[108] Shepherd, G.M., Brayton, R.K., Miller, J.P., Segev, I., Rinzel, J., & Rall, W. Signal enhancement in distal cortical dendrites by means of interactions between active dendritic spines. *PNAS*, 1985, *82*, 2192-2195.

[109] Shepherd, G.M., Woolf, T.B., & Carnevale, N.T. Comparisons between active properties of distal dendritic branches and spines: implications for neuronal computations. *Cognitive Neurosci.*, 1989, *1*, 273-286.

[110] Singer, W. Effects of monocular deprivation on excitatory and inhibitory pathways in cat striate cortex. *Experimental Brain Res.*, 1977, *134*, 508-518.

[111] Singer, W. Central core control of visual cortex functions. In em The Neurosciences Fourth Study Program, F.O. Schmitt & F.G. Worden, (Eds.), Cambridge, MA: MIT Press, 1979.

[112] Softky, B. & Kammen, D. Can simple cells learn curves? A Hebbian model in a structured environment. In *Advances in neural information processing systems, vol. 2*, D.S. Touretzsky, (Ed.), San Mateo, CA: Morgan Kaufmann, 1990.

[113] Stanton, P.K. & Sejnowski, T.J. Associative long-term depression in the hippocampus: induction of synaptic plasticity by Hebbian covariance. *Nature*, 1989, *339*, 215-218.

[114] Stent, G.S. A physiological mechanism for Hebb's postulate of learning. *Proc. Nat. Acad. Sci. USA*, 1973, *70*, 997-1001.

[115] Suga, N. Functional organization of the auditory cortex. In *Cortical sensory organization, vol. 3, Multiple auditory areas*, C.N. Woolsey, (Ed.), New Jersey: Humana, 1982.

[116] Szentagothai, J. The neuron network of the cerebral cortex: a functional interpretation. (1977) *Proc. R. Soc. Lond. B.*, 201:219-248.

[117] Tesauro, G. & Sejnowski, T.J. A parallel network that learns to play backgammon. *Artificial Intelligence*, 1989, *39*, 357-390.

[118] Van Essen, D.C. Functional organization of primate visual cortex. In *Cerebral Cortex*, Vol. 3, A. Peters & E.G. Jones, (Eds.), Plenum, 1985.

[119] Van Essen, D.C. & Maunsell, J.H.R. Hierarchical organization and functional streams in the visual cortex. *TINS*, September 1983, 370-375.

[120] Volper, D.J. & Hampson, S.E. Learning and using specific instances. *Biol. Cybern.*, 1987, *57*, 57-71.

[121] Volper, D.J. & Hampson, S.E. Quadratic function nodes: use, structure and training. *Neural Networks.*, 1990, *3*, 93-107.

[122] Walters, D. Biological vs. artificial representations of variables in fine-grain parallel systems. In *Proc. IEEE Computer Soc. Workshop on Computer Vision*, pp. 298-300, 1987.

[123] White, B.L., Castle, P., & Held, R. Observations on the development of visually-directed reaching. *Child Development*, 1964, *35*, 349-364.

[124] Widrow, B. & Hoff, M.E. Adaptive switching circuits. *Institute of Radio Engineers, Western Electronic Show and Convention, Convention Record, Part 4*, 1960, 96-104.

[125] Willshaw, D.J. & von der Marlsburg, C. How patterned neural connections can be set up by self-organization. *Proc. R. Soc. Lond. B*, 1976, *194*, 431-445.

[90] Peters, A. Plenary address, 1989 Soc. Neurosc. Meeting, Phoenix, AZ.

[91] Poggio, T. & Girosi, F. Regularization algorithms for learning that are equivalent to multilayer networks. *Science*, 1990, *247*, 978-982.

[92] Poggio, T. & Torre, V. A new approach to synaptic interaction. In *Theoretical approaches to complex systems, vol. 21*, R. Heim & G. Palm, (Eds.), Berlin: Springer, 1978.

[93] Powell, M.J.D. Radial basis functions for multivariable interpolation: a review. In *IMA Conference on Algorithms for the Approximation of Functions and Data*, RMCS Shrivenham, 1985.

[94] Presson, J. & Gordon, B. *Soc. Neurosci. Abstr.*, 1982, *8*, 5.10.

[95] Purves, D. *Body and brain: A trophic theory of neural connections.* Cambridge, MA: Havard University Press, 1988.

[96] Purves, D. & Lichtman, J.W. *Principles of neural development.* Sunderland, MA: A. Sinauer Associates, 1985.

[97] Rall, W. & Segev, I. Functional possibilities for synapses on dendrites and on dendritic spines. In *Synaptic function*, G.E. Edelman, W.F. Gall, & W.M. Cowan, (Eds.), New York: Wiley, 1987.

[98] Reiter, H.O. & Stryker, M.P. Neural plasticity without postsynaptic action potentials: Less-active inputs become dominant when kitten visual cortical cells are pharmacologically inhibited. *Proc. Nat. Acad. Sci. USA*, 1988, *85*, 3623-3627.

[99] Rosenblatt, F. *Principles of neurodynamics.* New York: Spartan, 1962.

[100] Rumelhart, D.E. "On learning to do what you want." Talk given at 1st connectionist summer school, Carnegie-Mellon University, 1986.

[101] Rumelhart, D.E., Hinton, G.E., & McClelland, J.L. A general framework for parallel distributed processing. In *Parallel distributed processing: explorations in the microstructure of cognition, vol. 1*, D.E. Rumelhart, J.L. McClelland, (Eds.), Cambridge, MA: Bradford, 1986a.

[102] Rumelhart, D.E., Hinton, G.E., & Williams, R.J. Learning internal representations by error propagation. In *Parallel distributed processing: explorations in the microstructure of cognition, vol. 1*, D.E. Rumelhart, J.L. McClelland, (Eds.), Cambridge, MA: Bradford, 1986b.

[103] Rumelhart, D.E. & Zipser, D. Feature discovery by competitive learning. In *Parallel distributed processing: explorations in the microstructure of cognition, vol. 1*, D.E. Rumelhart, J.L. McClelland, (Eds.), Cambridge, MA: Bradford, 1986b.

[104] Segev, I., Fleshman, J.W., & Burke, R.E. Compartmental models of complex neurons. In *Methods in neuronal modeling: From synapses to networks*, C. Koch & I. Segev, (Eds.), Cambridge, MA: MIT Press, 1989

[105] Segev, I., Fleshman, J.W., Miller, J.P., & Bunow, B. Modeling the electrical properties of anatomically complex neurons using a network analysis program: passive membrane. *Biol. Cybern.*, 1985, *53*, 27-40.

[106] Sejnowski, T.J. & Tesauro, G. The Hebb rule for synaptic plasticity: algorithms and implementations. In *Neural models of plasticity*, J.H. Byrne & W. Berry, (Eds.), Cambridge, MA: Academic Press, 1989.

[107] Shepherd, G.M. & Brayton, R.K. Logic operations are properties of computer-simulated interactions between excitable dendritic spines. *Neuroscience*, 1987, *21*, 151-166.

[70] McCulloch, W.S., & Pitts, W. A logical calculus of the ideas immanent in nervous activity. (1943) *Bull. of Math. Biophys., 5*, 115-133.

[71] McMaughton, B.L., Douglas, R.M., & Goddard, G.V. Synaptic enhancement in fascia dentata: cooperativity among co-active afferents. *Brain Res.*, 1978, *157*, 277-293.

[72] Mel, B.W. MURPHY: A robot that learns by doing. In *Neural Information Processing Systems*, D.Z. Anderson, (Ed.), American Institute of Physics: New York, 1988.

[73] Mel, B.W. *Connectionist robot motion planning.* Cambridge, MA: Academic Press, 1990 (in press).

[74] Miller, K.D., Chapman, B., & Stryker, M.P. Visual responses in adult cat visual cortex depend on N-methyl-D-aspartate receptors . *Proc. Natl. Acad. Sci. USA*, 1989a, *86*, 5183-5187.

[75] Miller, K.D., Keller, J.B., & Stryker, M.P. Ocular domninance column development: analysis and simulation. *Science*, 1989b, *245*, 605-615.

[76] Miller W.T., Hewes, R.P., Glanz, F.H., & Kraft, L.G. Real time dynamic control of an industrial manipulator using a neural network based learning controller. Technical Report, Dept. of Electrical and Computer Engineering, University of New Hampshire, 1988.

[77] Minsky, M. & Papert, S. *Perceptrons.*, Cambridge, MA: MIT Press, 1969.

[78] Moody, J. & Darken, C. Fast learning in networks of locally-tuned processing units. *Neural Computation*, 1989, *1*, 281-294.

[79] Mountcastle, V.B., Lynch, J.C., Georgopoulos, A., Sakata, H., & Acuna, C. Posterior parietal association cortex of the monkey: command functions for operations within extrapersonal space. *J. Neurophysiol.*, 1975, *38*, 871-908.

[80] Mountcastle, V.B. An organization principle for cerebral function: the unit module and the distributed system. In *The mindful brain*, G.E. Edelman & V.B. Mountcastle, (Eds.), Cambridge, MA: MIT, 1978.

[81] Mower, G.D. et al. Dark rearing prolongs physiological but not anatomical plasticity of cat visual cortex. *J. Comp. Neurol.*, 1985, *235*, 448.

[82] Mower, G.D., White, W.F., & Rustad, R. [^3H] Muscimol binding of GABA receptors in the visual cortex of normal and monocularly deprived cats. *Brain Res.*, 1986, *380*, 253-260.

[83] Omohundro, S. Efficient algorithms with neural network behavior. *Complex Systems*, 1987, *1*, 273-347.

[84] Omohundro, S. Five balltree construction algorithms. Technical Report TR-89-063, International Computer Science Institute, Berkeley, CA, 1990.

[85] Palm, G. *Neural assemblies: an alternative approach to artificial intelligence.* NY: Springer Verlag, 1982.

[86] Parker, D. *Learning logic.* Technical Report 47, Cambridge, MA: MIT Center for Computational Research in Economics and Management Science, 1985.

[87] Perkel, D.H. & Mulloney, B. Electrotonic properties of neurons: steady-state compartmental model. *J. Neurophysiol.*, 1978, *41*, 621-639.

[88] Perkel, D.H., Mulloney, B., & Budelli, R.W. Quantitative methods for predicting neuronal behavior. *Neuroscience*, 1981, *6*, 823-837.

[89] Peters, A. & Kara, D.A. The neuronal composition of area 17 of rat visual cortex. IV. The organization of pyramidal cells. *J. Comp. Neurol.*, 1987, *260*, 573-590.

[49] Ito, M. Long-term depression. *Ann. Rev. Neurosci.*, 1989, *12*, 85-102.

[50] James, W. *The principles of psychology.* Cambridge, MA: Harvard University Press, 1983. (originally published in 1890).

[51] Jones, E.G. Anatomy of cerebral cortex: columnar input-output relations. In *The organization of cerebral cortex*, F.O. Schmitt, F.G. Worden, G. Adelman, & S.G. Dennis, (Eds.), MIT Press: Cambridge, MA, 1981.

[52] Jones, E.G. Varieties and distribution of non-pyramidal cells in the somatic sensory cortex of the squirrel monkey. *J. Comp. Neurol.*, 1975, *160*, 205-268.

[53] Kammen, D. & Yuille, A. Spontaneous symmetry-breaking energy functions and the emergence of orientation selective cortical cells. *Biol. Cybern.*, 1988, *59*, 23-31.

[54] Kauer, J.A., Malenka, R.C., & Nicholl, R.A. A persistent postsynaptic modification mediates long-term potentiation in the hippocampus. *Neuron*, 1988, *1*, 911-917.

[55] Kelso, S.R., Ganong, A.H., & Brown, T.H. Hebbian synapses in hippocampus. *PNAS* USA, 1986, *83*, 5326-5330.

[56] Kennedy, M.B. Regulation of synaptic transmission in the central nervous system: long-term potentiation. *Cell*, 1989, *59*, 777-787.

[57] Koch, C. The action of the corticofugal pathway on sensory thalamic nuclei: a hypothesis. *Neuroscience*, 1987, *23*, 399-406.

[58] Koch, C. & Poggio, T. Biophysics of computation: neurons, synapses, and membranes. In *Synaptic function*, G.E. Edelman, W.F. Gall, & W.M. Cowan, (Eds.), New York: Wiley, 1987.

[59] Koch, C., Poggio, T., & Torre, V. Retinal ganglion cells: a functional interpretation of dendritic morphology. *Phios. Trans. R. Soc. Lond. B*, 1982, *298*, 227-264.

[60] Koch, C., Poggio, T., & Torre, V. Nonlinear interactions in a dendritic tree: localization, timing, and role in information processing. *PNAS*, 1983, *80*, 2799-2802.

[61] Kohonen, T. *Self organization and associative memory.* Springer-Verlag: Berlin, 1984.

[62] Kuperstein, M. Neural model of adaptive hand-eye coordination for single postures. *Science*, March 1988, *239*, 1308-1311.

[63] Lapedes, A. & Farber, R. How neural nets work. In *Neural Information Processing Systems*, D.Z. Anderson, (Ed.), American Institute of Physics: New York, 1988.

[64] Le Cun, Y. Une procedure d'apprentissage pour reseau a seuil assymetrique [A learning procedure for an assymetric threshold network], *Proceedings of Cognitiva*, 1985, *85*, 599-604, Paris.

[65] Levy, W.B. & Desmond, W.B. The rules of elemental synaptic plasticity. In *Synaptic modification, neuron selectivity, and nervous system organization*, W.B. Levy, J. Anderson, & S. Lehmkuhle, (Eds.), Hillsdale, NJ: Lawrence Erlbaum, 1985.

[66] Linsker, R. Self-organization in a perceptual network. *Computer*, 1988, *March*, 105-117.

[67] Lund, J.S. Spiny stellate neurons. In *Cerebral Cortex, vol. 1*, A. Peters & E.G. Jones, (Eds.), Plenum: New York, 1984.

[68] Marr, D. A theory of cerebellar cortex. *J. Physiol.*, 1969, *202*, 437-470.

[69] Marr, D. A theory for cerebral neocortex. *Proc. Roy. Soc. Lond. B*, 1970, *176*, 161-234.

[29] Eccles, J.C. The modular operation of the cerebral cortex considered as the material basis for mental events. *Neuroscience*, 1981, *6*, 1839-1856.

[30] Edamatsu, H., Kawasaki, M., & Suga, N. Distrubution of combination-sensitive neurons in the ventral fringe area of the auditory cortex of the mustached bat. *J. Neurophysiol.*, 1989, *61*, 202-207.

[31] Fawcett, J.W. & O'Leary, D.D.M. The role of electrical activity in the formation of topographic maps in the nervous system. *TINS*, May 1985, 201-206.

[32] Feldman, M.L. Morphology of the neocortical pyramidal neuron. In *Cerebral Cortex, vol. 1*, A. Peters & E.G. Jones, (Eds.), Plenum: New York, 1984.

[33] Feldman, J.A. & Ballard, D.H. Connectionist models and their properties. *Cognitive Science*, 1982, *6*, 205-254.

[34] Ferster, D. A comparison of binocular depth mechanisms in areas 17 and 18 of the cat visual cortex. *J. Physiol.*, 1981, *311*, 623-655.

[35] Finkel, L.H. & Edelman, G.M. Population rules for synapses in networks. In *Synaptic function*, G.E. Edelman, W.F. Gall, & W.M. Cowan, (Eds.), New York: Wiley, 1987.

[36] Fraser, S.E. Cell interactions involved in neuronal patterning: an experimental and theoretical approach. In *Molecular bases of neural development*, G.M. Edelman, W.E. Gall, & W.M. Cowan, (Eds.), New York: Wiley, 1985.

[37] Gamble, E. & Koch, C. The dynamics of free calcium in dendritic spines in response to repetive synaptic input. *Science*, 1987, *236*, 1311-1315.

[38] Giles, C.L. & Maxwell, T. Learning, invariance, and generalization in high-order neural networks. *Applied Optics*, 1987, *26(23)*, 4972-4978.

[39] Goldman, P.S. & Nauta, W.J.H. Columnar distribution of cortico-cortical fibers in the frontal association, limbic and motor cortex of the developing rhesus monkey. *Brain Res.*, 1977, *122*, 393-414.

[40] Greenough, W.T. & Bailey, C.H. The anatomy of a memory: convergence of results across a diversity of tests. *TINS*, 1988, *11*, 142-147.

[41] Grossberg, S. Adaptive pattern classification and universal recoding. *Biol. Cybern.*, 1976, *23*, 121-134.

[42] Hebb, D.O. *The organization of behavior.* New York: Wiley, 1949.

[43] Herkenham, M. New perspectives on the organization and evolution of nonspecific thalamocortical projections. In *Cerebral Cortex, vol. 5*, A. Peters & E.G. Jones, (Eds.), Plenum: New York, 1986.

[44] Hinton, G.E. Connectionist learning procedures. Technical Report CMU-CS-87-115, Pittsburgh, PA: Carnegie Mellon University, 1987.

[45] Hinton, G.E. & Sejnowski, T.J. Learning and relearning in Boltzmann machines. In *Parallel distributed processing: explorations in the microstructure of cognition, vol. 1*, D.E. Rumelhart, J.L. McClelland, (Eds.), Cambridge, MA: Bradford, 1986b.

[46] Hubel, D.H. & Wiesel, T.N. The period of susceptibility to the physiological effects of unitaleral eye closure in kittens. *J. Physiol. London*, 1965, *206*, 419-436.

[47] Hyvarinen, J. *The parietal cortex of monkey and man.* New York: Springer-Verlag, 1982.

[48] Iriki, A., Pavlides, C., Keller, A., & Asanuma, H. Long-term potentiation in the motor cortex. *Science*, *245*, 1989, 1385-1387.

[11] Bear, M.F., Cooper, L.N., & Ebner, F.F. A physiolgical basis for a theory of synapse modification. *Science*, 1987, *237*, 42-48.

[12] Bear, M.F. Schmechel, D.E., & Ebner, F.F. Glutamic acid decarboxylase in the striate cortex of normal and monocularly deprived kittens. *J. Neurosci.*, 1985, *5*, 1262-1275.

[13] Beaulieu, C. & Colonnier, M. A laminar analysis of the number of round-asymmetrical and flat- symmetrical synapses on spines, dendritic trunks, and cell bodies in area 17 of the cat. *J. Comp. Neurol.*, 1985, *231*, 180-189.

[14] Bliss, T.V.P. & Lømo, T. Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J. Physiol.*, 1973, *232*, 331-356.

[15] Broomhead, D.S. & Lowe, D. Multivariable functional interpolation and adaptive networks. *Complex Systems*, 1988, *2*, 321-355.

[16] Brown, T.H., Chapman, P.F., Kairiss, E.W., & Keenan, C.L. Long-term synaptic potentiation. *Science*, 1988a, *242*, 724-728.

[17] Brown, T.H., Chang, V.C., Ganong, A.H., Keenan, C.L., & Kelso, S.R. Biophysical properties of dendrites and spines that may control the induction and expression of long-term synaptic potentiaion. In *Long-term potentiation,: from biophysics to behavior*. Alan R. Liss, Inc., 1988b.

[18] Brown, T.H., Fricke, R.A., & Perkel, D.H. Passive electrical constants in three classes of hippocampal neurons. *J. Neurophysiol.*, 1981a, *46*, 812-827.

[19] Brown, T.H., Perkel, D.H., Norris, J.C., & Peacock, J.H. Electrotonic structure and specific membrane properties of mouse dorsal root ganglion neurons. *J. Neurophysiol.*, 1981b, *45*, 1-15.

[20] Burton, H. Second somatosensory cortex and related areas. In *Cerebral Cortex, vol. 5*, A. Peters & E.G. Jones, (Eds.), Plenum: New York, 1986.

[21] Costanzo, R.M. & Gardner, E.P. Multiple-joint neurons in somatosensory cortex of awake monkeys. *Brain Res.*, 1981, *214*, 321-333.

[22] Cotman, C.W., Nieto-Sampedro, M., & Harris, E.W. Synapse replacement in the nervous system of adult vertebrates. Physiol. Rev., 1981, *61*, 684-785.

[23] Cotman, C.W., Monaghan, D.T., & Ganong, A.H. Excitatory amino acid neurotransmission: NMDA receptors and Hebb- type synaptic plasticity. *Ann. Rev. Neurosci.*, 1988, *11*, 61-80.

[24] Crick, F.H.C. & Asanuma, C. Certain aspects of the anatomy and physiology of the cerebral cortex. In *Parallel distributed processing: explorations in the microstructure of cognition, vol. 2*, D.E. Rumelhart, J.L. McClelland, (Eds.), Cambridge, MA: Bradford, 1986.

[25] Denker, J.S., Gardner, W.R., Graf, H.P., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D., Baird, H.S., & Guyon, I. Neural network recognizers for hand-written zip code digits. In *Advances in Neural Information Processing Systems vol. 1*, D.S. Touretzsky, (Ed.), San Mateo, CA: Morgan Kaufmann, 1989.

[26] Duda, R.O. & Hart, P.E. *Pattern classificiation and scene analysis.* New York: Wiley, 1973.

[27] Durbin, R. & Rumelhart, D.E. Product units: a computationally powerful and biologically plausible extension to backpropagation networks. *Complex Systems*, 1989, *1*, 133.

[28] Eccles, J.C. The cerebral neocortex: a theory of its operation. In *Cerebral Cortex, vol. 2*, A. Peters & E.G. Jones, (Eds.), Plenum: New York, 1984.

Anatomically speaking, how serious a simplification is the assumption that cortical associative learning can be viewed as a feedforward process, given the common pattern of reciprocal connectionis seen in all parts of the cortex? How may the widespread patchy structure of cortex relate to supervised associative learning scheme proposed here?

In continuing work, we are further considering the implications of the *sigma-pi* model neuron for cases of unsupervised learning, i.e. when the post-synaptic teacher signal is due solely to the association fibers themselves, rather than to an explicit, unconditioned unit-wide teacher signal. This type of learning may be directly relevant to the establishment of multi-dimensional stimulus selectivities of cortical neurons during development.

Acknowledgements

Thanks are due to Ojvind Bernander, Tom Brown, Rodney Douglas, Richard Durbin, Kamil Grajski, David MacKay, and John Moody for numerous helpful discussions. Thanks also to Dan Kammen and Ojvind Bernander for their many comments regarding the manuscript. Particular thanks to Christof Koch for his support, stimulating discussions, and careful reading of the manuscript. This research has been supported by the Office of Naval Research, the James S. McDonnell Foundation, and the Del Webb Foundation.

References

[1] Albus, J.S. A theory of cerebellar function. *Math. Biosci.*, 1971, *10*, 25-61.

[2] Andersen, R.A. Inferior parietal lobule function in spatial perception and visuomotor integration. In *Handbook of Physiology, Sec. 1 The Nervous System (vol. 5)*, F. Plum, (Ed.), Bethesda, MD: American Physiological Society, 1987.

[3] Artola, A. & Singer, W. Long-term potentiation and NMDA receptors in rat visual cortex. *Nature*, 1987, *330*, 649-652.

[4] Atkeson, C.G. Using local models to control movement. In *Advances in neural information processing systems, vol. 2*, D.S. Touretzsky, (Ed.), San Mateo, CA: Morgan Kaufmann, 1990.

[5] Ballard, D. Cortical connections and parallel processing: structure and function. *Behav. Brain Sci.*, 1986, *9*, 67-120.

[6] Ballard, D.H., Hinton, G.E., & Sejnowski, T.J. Parallel visual computation. *Nature*, 1983, *306*, 21-26.

[7] Ballard, D.H. Interpolation coding: a representation for numbers in neural models. *Biol. Cybern.*, 1987, *57*, 389-402.

[8] Baranyi, A. & Feher, O. Synaptic facilitation requires paired activation of convergent pathways in the neocortex. *Nature*, 1981, *290*, 413-415.

[9] Barron, A.R. & Barron, R.L. Statistical learning networks: a unifying view. Presented at the *1988 Symposium on the Interface: Statistics and Computing Science*, Reston, Virginia.

[10] Bear, M.F. & Cooper, L.N. Molecular mechanisms for synaptic modification in the visual cortex: interaction between theory and experiment. In *Neuroscience and connectioist theory*, M. Gluck & D.E. Rumelhart, (Eds.), Erlbaum, 1990.

depolarization, the current model suggests that (3) the ability to potentiate the NMDA (vs. the non-NMDA) signal during learning could allow a local cluster of synapses to be potentiated *as a whole entity*, without significantly increasing the ability of each constituent synapse to depolarize the post-synaptic cell as an *individual*.

3. The model gives an account for, and critically depends on, the complex 3-dimensional arborizations of axons and dendrites in layer 1. Within the *sigma-pi* learning scheme, this rich "combinatorial" input-output interface is needed in order to make it likely that any arbitrarily chosen small group of association fibers crosses near to each other *somewhere* in the 3-D volume of a post-synaptic dendrite. In conjunction with a Hebb-type learning rule, post-synaptic cells are thus given "virtual" access to a very large number of *possible* clusters of association inputs (and hence the higher-order interactions they represent) without dedicating the enormous neural resources needed to build them outright.

4. The model depends crucially on synaptogenesis and describes the necessary conditions for survival of a given association synapse. A synapse survives at its current dendritic locus when by chance it falls into a cluster that is strongly correlated with activation of the post-synaptic cell. Otherwise, the synapse retracts, making room for a different synapse.

5. The model gives an account for the gross asymmetry in number and patterns of termination of excitatory vs. inhibitory synapses onto cortical pyrmadial cells, is consistent with the fact that plasticity in inhibitory synapses has not yet been demonstrated. Within the *sigma-pi* learning scheme, only a single layer of excitatory synapses are used or modified (according to a simple Hebb-type rule), lending support to the conjecture (Bear et al., 1987) that inhibitory synapses play no direct role in cortical associative learning.

6. The model gives an interpretation to the replication of like-activated neurons beneath a single point in cerebral cortex, and in particular to the clumping of apical dendrites of pyramidal cells on their rise to the cortical surface (Jones, 1989). Within the *sigma-pi* learning scheme, this type of grouping allows a direct pooling of dendritic space for a corresponding increase in memory capacity.

A number of questions have remained entirely unaddressed in this work. Most prominently, detailed biophysical simulations of complex dendritic trees are needed to test many of the ideas in this paper. It is necessary to establish, as discussed, the degree to which clusters of active (i.e. pre-synaptically driven) synapses can function independently of one another when distributed across a pyramidal dendritic tree. How large a cluster is feasible while still maintaining significant AND-like selectivity within the cluster? What are the effects of cluster placement on dendritic branches of varying order?

7. Discussion

In summary, the current model of cortical associative learning predicts that the input-output pyramidal cell modules of association cortex can be modelled as *sigma-pi* columns, composed of units that can enforce AND-like selectivity among small groups of neighboring synapses. The *sigma-pi* column can learn supervised nonlinear associative maps in a way that is cortically plausible. Only a single layer of excitatory cortico-cortical association synapses is modified, with learning under the control of a Hebb-type learning rule. The function of the local multiplicative nonlinearity differs depending on the class of associative learning problem to which the column is applied. In the case of a smooth nonlinear learning problem over a low-dimensional problem space, typical of certain sensory-motor learning problems, a *sigma-pi* cluster can be viewed as a multi-dimensional receptive field "centered" on a specific input state, that may be weighted according to the average value of the post-synaptic teacher signal associated with that state. In this mode of operation, the *sigma-pi* unit is essentially a locally-smoothing multi-dimensional lookup table, with synaptic clusters acting as individual table entries. In the case of a high-dimensional associative learning problem, for which it may be necessary to discriminate among a large number of complex, unrelated input patterns (such as images), then each *sigma-pi* cluster may be viewed as a higher-order feature (i.e. polynomial term) computed over the original input space, helping to orthogonalize the set of input patterns for ease of discrimination. We believe the model presented here to be a good first guess as to the supervised associative learning paradigm at work in the cerebral cortex, on the basis of anatomical, physiological, biophysical, and ecological arguments.

A number of features of the model relevant to the neurobiology of learning in cerebral cortex are summarized and interpreted below.

1. The extrinsically-projecting pyramidal cells of layers 2, 3, and 5 of association cortex are modeled as *sigma-pi* units. A *sigma-pi* unit computes its activation level as a sum of contributions from a set of multiplicative (or locally-thresholded) clusters of synapses distributed throughout its dendritic tree. A dendritic tree is thus modeled as a two layer analog OR-of-AND processor, a somewhat simpler view of dendritic function than one which has been previously suggested, namely that branching patterns of dendritic trees may implement a complex hierarchy of "logic" operations including AND, OR, and AND-NOT operations (Koch & Poggio, 1987; Shepherd & Brayton, 1987).

2. Three distinct learning-related roles for the NMDA channel are discussed, one of them new. In addition to its ability to (1) enforce Hebb-type pre- and post-synaptic co-activation as a condition for the induction of LTP at a given synapse, and (2) encourage cooperativity among neighboring synapses during learning via shared post-synaptic

onto a post-synaptic dendrite when called for by a Hebb-type learning rule. This geometry of pre- and post-synaptic interface in the neocortex is to be strongly contrasted with the architecture of cerebellum, where the afferent "parallel" fibers have little or no possibility of clustering on the dendrites of Purkinje cells. Thus, the arborizations of cortical axons and dendrites can be viewed as an interface with extremely rich inherent combinatorics, obviating the advance construction of an exponentially large number of "higher order" features whose utility can only be determined during learning. As the *a priori* probability of a given cluster falls off exponentially with cluster size, however, this geometry of arborization can only be considered facilitatory to the establishment of relatively small clusters. Interestingly, this combinatorial constraint on cluster size would seem to parallel a likely *biophysical* constraint on cluster size limiting the length of dendritic membrane within which significant AND-like selectivity can be achieved among all contributing synapses.

The assumption that a large number of virtual clusters could be selected amongst via a Hebb-type learning rule is closely akin to the proposed mechanisms of target dependent axonal sprouting and growth cone guidance during neural development and regeneration in the adult (Purves, 1988). One scenario for *sigma-pi* cluster formation would entail that active association fibers continuously compete as individuals for post-synaptic sites, where the competition is mediated through a limited post-synaptic resource, such as a trophic factor. This type of competition among nerve terminals is well known in both the peripheral and central nervous system (see Lichtman & Purves, 1980; Fawcett & O'Leary, 1985). Evidence is also widespread for synaptogensis and associated growth and remodeling of axons and dendrites in the mature CNS in response to a variety of natural and experimentally-induced stimuli (Cotman et al., 1981, Purves, 1988), notably under experimental conditions associated with intensive behavioral learning (Greenough & Bailey, 1988). Following this scenario, once an association fiber has formed a synapse, its probability of survival depends on the degree to which the cluster of synapses in which it fortuitously sits is rewarded, as a whole, according to a learning rule such as that of equation 5.3. In this way, a very large number of possible clusters can be selected amongst, where only those that consistently fire with the post-synaptic teacher signal are maintained with strength. The proposal that conjoint activity of pre-synaptic axons should contribute to their mutual survival in competition with other axons on a post-synaptic target (Fawcett & O'Leary, 1985, Purves & Lichtman, 1985) is almost indistinguishable, at an abstract level, from the proposal that a cooperativity among inputs to a neuron should lead to a potentiation of existing synaptic connections, as discussed at length above (e.g. Stent, 1973). Purves (1988, p. 169) suggests that these two forms of "learning" may coexist in the adult nervous system.

binocular deprivation (see Bear et al.,1987). Deprived-eye visual afferents become functionally disconnected from their cortical targets within days. However, clear anatomical changes in the axonal arborizations of geniculo-striate afferents can be detected only if the deprivation occurs early in the critical period; late-critical-period deprivation (Presson & Gordon, 1982) or deprivation after a period of dark-rearing (Mower et al., 1985) results in functional disconnection with no apparent anatomical change. A selective, temporary loss of "hot" weight, e.g. NMDA channels, from deprived-eye inputs could account for this late-critical-period effect. In this view, the post-synaptic effects of local cluster activation among deprived-eye afferents would be strongly attenuated, while leaving the synapses physically intact (i.e. with their cold weight relatively unchanged) and with the possibility of rapidly regaining their hot weight should circumstances call for it. Fraser (1985) discusses the electrically-mediated motility of membrane proteins, such as channels, that could underlie such come-and-go among hot channels based on local patterns of synaptic activation.

In sum, we conclude that the NMDA channel, at once receptor and voltage gated, can play three distinct roles in learning by (1) enforcing the Hebbian condition of pre-and-postsynaptic coactivation, (2) underlying the crucial multiplicative or AND-like interaction needed for cluster selectivity, and (3) enabling the potentiation of synaptic clusters as whole entities, without destroying AND-like selectivity internal to the clusters.

6.3 A Virtual Cluster Reservoir

The abstract definition of *sigma-pi* learning specifies that all possible clusters c_j of size $1 < k < k_{max}$ pre-exist on the "dendrites" of each virtual *sigma-pi* unit (which we have previously proposed to consist of a vertically aggregated clump of about 100 pyramidal cells that receive the same focal teacher input). During learning, the weight on each cluster is governed by the Hebb rule of equation 5.3. Since the number of *possible* clusters of size k quickly overwhelms total available dendritic space for even small k, it must be possible to *create* a cluster when it is needed. For example, there are $\binom{28000}{3} > 10^{12}$ clusters of size 3 chosen from a set of 28,000 afferents (see section 6.1.2). However, if we assume 10,000 synapses per pyramidal cell, 2,000 of which are available for association fibers, then there is dendritic space in the *sigma-pi* column for at most 67,000 3-clusters, or only 1 in 15 million *possible* 3-clusters. Likewise, dendritic space exists for only 1 in 10^{11} possible 4-clusters.

We propose that the complex 3-dimensional structure of axonal and dendritic arborizations in layer 1 in cerebral cortex are ideal for maximizing the probability that arbitrary (small) subsets of association axons cross *near* to each other in space at some point in their intermeshing arborizations. According to this hypothesis, the tangle of axons within a dendrite's receptive field gives rise to an enormous set of *almost*-clusters, poised to "latch"

modest fixed value, will begin to degrade the AND-like selectivity of the cluster in which it sits. In contrast, the conductance of the NMDA-type channel is voltage-*sensitive*, and to a first approximation comes on line only when the post-synaptic membrane is already significantly depolarized from its resting level (MacDermott & Dale, 1987). The NMDA channel thereby fulfills the basic requirements of "hot" weight, since it does not significantly help to depolarize the membrane from its resting level of -70 mv, but once sufficient depolarization is achieved, it provides a large additional depolarizing current with relatively long time course (MacDermott & Dale, 1987). Thus, selective increases in NMDA channel density in a cluster of synapses could be one mechanism for potentiating the cluster as a whole during learning, while preserving significant AND-like selectivity within the cluster. (Note that standard active channels, such as the Na^+ channel that drives the action potential, do not qualify as hot weight, since they non-specifically amplify *any* depolarizing influence in the "neighborhood" unlike the NMDA channel which, in addition to its voltage-dependence, is linked to, and depends, on, an excitatory pre-synaptic event.)

In apparent contradiction to the hypothesis that the NMDA signal is potentiated in pyramidal dendrites during learning, Kauer et al. (1988) and Muller et al. (1988) have shown in hippocampus that synaptic potentials produced by NMDA channels are not altered after induction of LTP, whereas those produced by the kainate-quisqualate type receptors are strongly enhanced. However, the NMDA channel is known to contribute much more strongly to normal synaptic transmission onto neocortical (Miller et al, 1989) vs. hippocampal (Kauer et al., 1988) pyramidal cells, suggesting that the NMDA channel may play a different learning-related role in the two structures. Thus, if observed experimentally, the selective potentiation of hot, or supra-threshold, weight in a neural learning system may be a signpost indicating the importance of local clustering and AND-like cooperativity among neighboring synapses in that system. Otherwise, the potentiation of cold weight would suffice, although the presence of *fixed* hot weight may still be necessary to amplify otherwise passive dendritic inputs. Unfortunately, the prediction of selective learning-induced potentiation of the NMDA signal in neocortical pyramidal dendrites and/or spines cannot be made with strength, as at least one other obvious mechanism could contribute to increases in hot vs. cold weight: The hot weight of a cluster could also be increased by distributing multiple copies of the cluster (or its near functional equivalents) across the available dendritic space in a *sigma-pi* column. Thus, cluster weight w_j would in this view represent the probability of cluster formation or survival in competition with other clusters. The hot weight associated with any *single* cluster could in that case remain fixed.

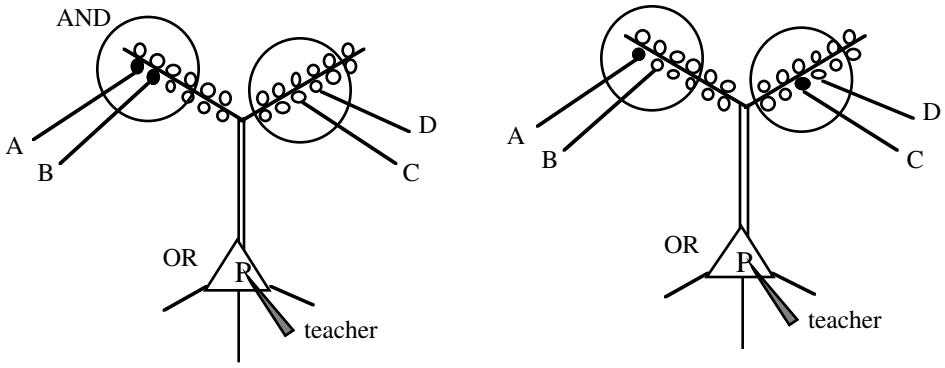
One suggestion that hot and cold weight exist and may be manipulated independently is the effect of "functional disconnection" of synapses seen in developmental studies in visual cortex under various forms of monocular and

NMDA receptor channels are concentrated in cortical structures, among relatively few other brain areas (Monaghan & Cotman, 1985). In the neocortex, receptor binding sites were concentrated in the superficial, synapse-enriched layers, even localized to post-synaptic densities (see Cotman et al., 1988). At minimum, this suggests a readily availabile local thresholding mechanism pervading the dendrites of cortical pyramidal modules. We reiterate here that within this perspective on dendritic function, individual synapses can be members of more than one cluster, such that no strict cluster boundaries need exist. In addition, use of the ligand-gated NMDA receptor channel as the dendritic nonlinearity underlying AND-like selectivity among neighboring synapses, in contrast to a simple voltage-gated channel (e.g. Na^+), offers the advantage that the excitable membrane distributed throughout the dendritic tree is *selective*, i.e. can only be "kindled" at the small fraction of synaptic sites that are pre-synaptically active. The dendritic tree as a whole is thus much less generically excitable, and lessens the risk of non-selectively "igniting" the entire dendritic arbor for arbitrary or irrelevant patterns of synaptic input.

6.2.4 Hot vs. Cold Synaptic Weight: A Third Role for NMDA

The assumption of a local thresholding nonlinearity, such as the NMDA channel, in place of a multiplicative nonlinearity, invites the suggestion that synaptic weight may be subdivided into two functionally distinct types. Consider the following quandry: In order to potentiate a cluster as a whole during learning (i.e. to increase cluster weight w_j), it is not sufficient to distribute the increase across the individual synaptic weights v_i to the cluster, since this procedure will eventually destroy cluster selectivity by allowing fewer than k_θ coactive inputs to drive the post-synaptic membrane to threshold; in the limit, a single synaptic input that is a member of a strongly weighted cluster could drive the membrane to threshold by itself, thereby completely eliminating the crucial AND-like property of the cluster. We therefore introduce the distinction of "hot" vs. "cold" synaptic weight. Cold weight, assigned to individual synaptic inputs, measures the *sub*-threshold efficacy of a single synaptic input, i.e. the efficacy with which it drives the post-synaptic membrane *toward* threshold. Hot weight, assigned to entire clusters, measures the post-synaptic "bang" associated with the *supra*-threshold activation of the cluster as a whole. The potentiation of hot weight is that which is governed by the w_j term in equation 5.3.

While theoretically motivated, this hot-cold dichotomy lies in tantalizing correspondence to the distinction of NMDA vs. non-NMDA channels (Cotman et al., 1988). The non-NMDA type channels (including kainate and quisqualate subtypes) are relatively simple voltage-*independent* membrane conductances, thereby fulfilling the requirements of "cold" synaptic weight. Potentiating this non-NMDA signal at a given synapse increases that synapse's ability to depolarize the membrane from its resting level, and above some



1. Response to AB >> AC.

2. LTP of AB >> AC.

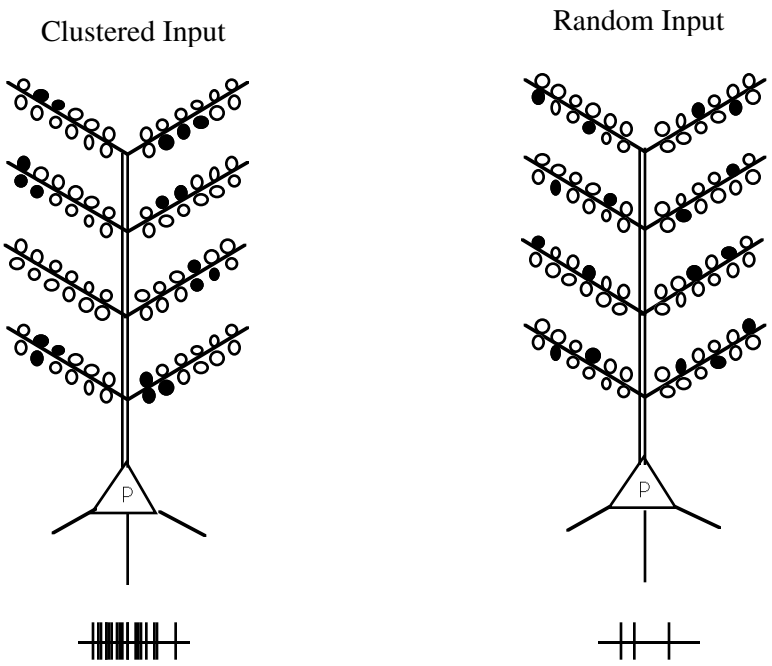


Figure 9: At top, a putative thresholding nonlinearity is seen to increase both the response and the potentiation of paired association inputs when they are spatially clustered, rather than separated, on the post-synaptic dendritic tree. Widespread distribution of NMDA channels in the dendrites of cortical pyramidal cells could, in this way, lead to cell responses that are highly selective for afferent vectors whose active inputs are spatially clumped (lower left) rather than uniformly distributed (lower right) across the dendritic arbor.

in all inactive synapses, regardless of their relative utilities.

In sum, the *sigma-pi* learning scheme predicts on *a priori* theoretical grounds that pre-without-postsynaptic activity should be the primary condition for long term depression at excitatory synapses onto cortical pyramidal cells. However, while the biological basis for such a condition for synaptic weakening has been experimentally observed during postnatal development in the visual cortex, strong conclusions as to the spectrum of synaptic depression (and enhancement) in neocortex are not yet warranted.

6.2.3 Cluster Multiplication: A Second Role for NMDA

Learning aside, the idea that dendritic computations could include local multiplicative nonlinearities has been repeatedly suggested for a number of years (Poggio & Torre, 1978; Koch et al., 1982; Shepherd et al., 1985; Shepherd & Brayton, 1987; Koch & Poggio, 1987, Rall & Segev, 1987), made all the more plausible given observed multiplicative interactions among inputs to sensory neurons in a variety of systems (Mountcastle et al., 1975; Ferster, 1981; Costanzo & Gardner, 1981; Hyvarinen, 1982; Edamatsu et al., 1989). Thus, in addition to its putative role in selectively potentiating like-activated clusters of neighboring synapses, the voltage-sensitive NMDA channel may be hypothesized to simultaneously underlie the crucial AND-like interaction among neighboring inputs needed for *sigma-pi* cluster-selectivity (Koch, 1987; Brown et al, 1988b). Pharmacological and biophysical studies suggest that the NMDA channel only opens to a large conductance state when strongly depolarized, to remove a Mg^{2+} block that is normally present at the membrane resting potential (see Cotman et al., 1988). A thresholding nonlinearity such as this applied to the sum of local synaptic input may be able to mimick the essential AND-like property of *sigma-pi* cluster multiplication for small clusters (see fig. 5). In this case, cluster response may be rewritten as,

$$c_j = \sigma(\sum_{i=1}^{k} v_i x_i - \theta_{local}),$$

where θ_{local} determines the number of fully activated inputs necessary to trigger a suprathreshold response, and σ is a thresholding nonlinearity. (It is important to note that the AND-like selectivity of a locally-thresholded cluster depends strongly on the assumption that the input weights v_i are small, such that no single synaptic input can be large enough to exceed the local threshold on its own. The physiological implications of this issue are taken up in more detail in the next section.)

As discussed above, if sufficiently endowed with NMDA channels in their dendritic spines then, cortical pyramidal cells could respond highly selectively to associative input "vectors" whose active afferents are spatially clustered, rather than scattered uniformly, across the dendritic arbor (fig. 9). This possibility for neocortical dendrites is in keeping with the finding that (in the rat)

recent models for development of modularity in cerebral cortex (personal communication).

We recall that the learning rule of equation 5.3 also prescribes non-associative conditions for synaptic weakening that were motivated by theoretical considerations. Long-term synaptic depression (LTD) has been observed only rarely in the CNS, primarily in the hippocampus (Levy & Steward, 1983; Bramham & Srebo, 1987; Stanton & Sejnowski, 1989), and its mechanisms are not yet understood. The best evidence for synaptic depression in neocortex comes from work in visual deprivation studies in the developing kitten (Hubel & Wiesel, 1965; Singer, 1979; Reiter & Stryker, 1988). Bear & Cooper (1990) have proposed a molecular mechanism inspired by earlier theoretical models (Stent, 1973; Cooper, Lieberman, & Oja, 1979; Bienenstock, Cooper, & Munroe, 1982) that can account for the intricate patterns of synaptic potentiation and depression under a variety of conditions of monocular and binocular deprivation. In this model, changes in synaptic strength are governed by a post-synaptic balance of signals from two types of ion channel (NMDA vs. non-NMDA). Synapses are strengthened when the NMDA signal exceeds the non-NMDA signal. Since the NMDA-type channel is voltage-dependent, this enhancement condition occurs only when the post-synaptic membrane is sufficiently depolarized, i.e. in a standard Hebb-type "pre-with-postsynaptic" condition. In the *absence* of sufficient membrane depolarization, NMDA channels are blocked, the balance shifts in favor of the non-NMDA signal, and the synapse is weakened (i.e. in a "pre-without-postsynaptic" condition). In both cases, synaptic change is dependent upon pre-synaptic activation, as in equation 5.3, while the sign of the change depends on the degree of induced post-synaptic depolarization relative to a plasticity threshold. Bienenstock et al. further postulate that the "modification threshold" is variable, depending on time-averaged post-synaptic activity. Direct evidence for a pre-without-postsynaptic condition for depression in neocortex comes from a study in which ocular dominance was shown to shift away from the more-active inputs and in favor of the less-active inputs (from a previously deprived eye) when the cortical cells to which they projected were pharmacologically inhibited (Reiter & Stryker, 1988). Evidence for the opposite, or post-without-presynaptic, condition for depression is currently uncertain in the visual deprivation data[10]. an adult animal it would moreover constitute a oddly non-selective plasticity mechanisms for associative learning, since each occurence of post-synaptic activation would bring about a uniform decrement

[10]Confusion exists in the interpretation of the result that ocular dominance shifts from a deprived eye to an open eye during monocular deprivation. This was interpreted by Stent (1973) as evidence for a "post-without-presynaptic" condition for weakening of the closed-eye afferents, since these afferents are relatively inactive while the post-syanptic cells continue to be driven by the open eye afferents. On the other hand, this condition is interpreted as a "pre-without-postsynaptic" condition for weakening by Bear et al. (1987) and Bear & Cooper (1990), since the closed-eye afferents still fire spontaneous action potentials, but in a fashion that is uncorrelated with the post-synaptic activity as driven through the open eye (i.e. pre-without-post).

Under the working assumption that the dendrites of cortical pyramidal cells are electrically inhomogeneous to an extent that permits a significant degree of independent dendritic processing, we examine in the following sections the neurobiological evidence for specific mechanisms that may underlie cluster multiplication and plasticity. Three distinct for the NMDA channel are discussed, one of them novel. Then, we examine possible mechanisms that could account for "online" synaptic cluster formation.

6.2.2 Hebbian Clustering: One Role for The NMDA Channel

The process of cluster formation in *sigma-pi* learning is driven by a local Hebb-type rule. Long term Hebb-type synaptic modification has been demonstrated in hippocampus (Bliss & Lømo, 1973) and a variety of neocortical areas including visual (Artola & Singer, 1987) and motor (Baranyi & Feher, 1981; Iriki et al., 1989) cortex (for reviews see Brown et al., 1988ab; Kennedy, 1989). It has been shown to be dependent upon local post-synaptic depolarization (Kelso et al., 1986), and mediated by a rise in post-synaptic calcium concentration gated by the voltage-dependent NMDA receptor channel (see Brown et al., 1988b). In addition to Hebb-type potentiation with pre- and post-synaptic coactivation, *sigma-pi* learning implicitly specifies cooperation among *pre*-synaptic inputs, in the sense that the largest increase in cluster weight w_j occurs when all inputs x_i to a cluster are simultaneously and fully active. The basic idea that cooperation among inputs to a neuron may contribute to synaptic plasticity is present implicitly in Hebb (1949) and developed further by others (Stent, 1973; McNaughton et al., 1978). The modeling of aspects of neural plasticity based on cooperative, "unsupervised"[9] Hebb-type learning rules has also become a vigorous industry (Grossberg, 1976; Kohonen, 1985; Rumelhart & Zipser, 1986), particularly for studies of the development of visual orientation selectivity (Linsker, 1988; Softky & Kammen, 1990) and ocular dominance (Miller et al., 1989b). The extension of this basic idea to *local* cooperativity among small groups of neighboring synapses (Fawcett & O'Leary, 1985; Gamble & Koch, 1987; Brown et al., 1988b; Mel, 1988) follows directly from the biological fact that local post-synaptic depolarization is the key mediator in the induction of LTP (Kelso et al., 1986). The central idea is that like-activated synaptic inputs should contribute to each other's enhancement during learning to the extent they are spatially clustered on a post-synaptic dendrite. This type of cooperativity gives key importance to *dendritic space* in neural learning, and has been shown to be plausible based on detailed biophysical modeling studies in progress (T. Brown and A. Zador, oral presentation, Snowbird, 1990). J. Moody has also pointed out that the mathematics that describe cluster formation on dendrites are closely akin to those embedded in several

[9]The term "unsupervised" denotes the absence of a designated teacher input that clamps the output neuron in a desired state. Instead, synaptic modification is governed by the level of postsynaptic activity generated directly by the normal weighted inputs to the neuron.

In an analysis of retinal ganglion cells, however, Koch et al. (1982) argue that dendrites could contain on the order of 30 subunits, where a subunit is a region of the dendritic tree that is relatively electrically homogeneous within, but relatively electrically isolated from other subunits. The number of dendritic subunits needed in order for a cortical pyramidal cell to function as a *sigma-pi* unit appears to be not significantly different from that shown to be plausible by Koch et al. (1982). Further, the number of predicted subunits from this earlier work would be expected to grow if transient, rather than D.C., synaptic potentials are considered.

While this question must be addressed through biophysical modeling studies of the kind pioneered by Rall (1959) and carried forward by others (Perkel & Mulloney, 1978; Perkel et al., 1981; Koch et al., 1982; Shepherd et al., 1985, Koch & Poggio, 1987; Rall & Segev, 1987; Shepherd & Brayton, 1987; Shepherd et al., 1989), we may nonetheless outline one set of assumptions under which a pyramidal dendrite could be expected to support *sigma-pi*-type cluster operations. First, we assume that on average only a small proportion of the association inputs to a cell is active, say 1-5%, and that this number of inputs is sufficient to fire the cell (if appropriately placed on the dendrites). For purposes of illustration, we assume this number to be 100 synapses, and consider two distinct conditions of synaptic activation. In condition 1, we activate 100 association synapses at random, representing an arbitrary input pattern to which this neuron has not been previously exposed. We then examine the spatial distribution of these 100 synapses on the dendritic tree, and compute the average electrotonic length L_{av} along the dendritic tree from each active synapse to its nearest active neighbor. In input condition 2, the 100 active synapses are spatially grouped in clusters of from two to several synapses. In this condition, definitionally, the average distance from each active synapse to its nearest active neighbor is sharply reduced relative to condition 1, i.e. $L_{av}^{(2)} \ll L_{av}^{(1)}$. The dendritic tree need only have the property that somatic depolarization in condition 2 be significantly larger than it is in condition 1 (fig. 9). This condition should exist if i) the dendritic tree is amply endowed with voltage-sensitive channels of some description, and ii) the average number of activated synapses per input pattern (e.g., 100), divided by the total number of subunits (e.g., 50), yields an average number of activated synapses per subunit (in this case, 2) which is less than the number needed within a subunit (on average) in order to fire the voltage-sensitive membrane nonlinearity. In other words, the argument goes, uniformly distributed inputs to a cell may be too spatially dilute in the dendritic tree to trigger the active membrane processes that can lead to significant somatic depolarization. On the other hand, input patterns that have been previously learned according the *sigma-pi* rule or equivalent, will activate synapses that are grouped in *clusters* on the post-synaptic dendrite, such that the pooled post-synaptic depolarization within each cluster is sufficient to fire the local local membrane nonlinearity.

moves through the field of view (Mel, 1990). The retinally-derived input to a visual neuron may functionally depend on the angles of all of the joints of the limb (though perhaps not equally); the visual neuron would therefore need access to the entire proprioceptive representation of the forelimb if "asked" to learn this functional relation. It is likely that a typical cortical pyramidal cell only has direct, monosynaptic access to an area of tissue in the pre-synaptic population that is at most 5-6 mm in diameter (Szentagothai, 1978), which, assuming 30μm centers, would comprise 25-30,000 pyramidal modules. If associations between cortical populations are, as suggested here, monosynaptically encoded at the layer 1 synapses onto pyramidal dendrites, then this would suggest evolutionary pressure limiting the size of any cortical population that gives rise to an association projection. In particular, all state information in the pre-synaptic population relevant to a post-synaptic neuron's unconditioned activation function should fall within a presynaptic area of no more than 5 mm in diameter. Interestingly, many higher order cortical sensory areas with association projections fit easily within these dimensions (for reviews, see Burton, 1985; Suga, 1982), though certain higher-order visual areas, while orders of magnitude smaller than V1, appear to exceed these general limitations (Van Essen & Maunsell, 1983; Andersen, 1987). Nonetheless, trends in modern neuroanatomy would seem to indicate significant evolutionary pressure toward fractionation of higher cortical representations into smaller, functionally specialized areas (Merzenich & Kaas, 1980).

6.2 Aspects of Sigma-Pi Clustering

6.2.1 Dendritic Subunits and Cluster Size

In the sections to follow, the question of cluster size vs. dendritic length constant is crucial. Simply put, what length of dendrite and/or number of synapses can be considered a cluster on a cortical pyramidal cell? Obviously, if the entire dendritic tree falls within one length constant of the soma, then it will be an essentially equipotential structure in which dendritic locality has little meaning. Equally obviously, this circumstance would make impossible the type of activity-dependent spatial clustering and locality dependence for the nonlinear AND-like operation central to the *sigma-pi* learning scheme. In general, the greater the degree of electrical isolation, and consequently, the greater the number of possible independent dendritic computations, the greater the computational advantage for a scheme such as *sigma-pi* learning. Apparently at odds with this need for a high degree of local dendritic processing, there are indications in a variety of types of mammalian neurons that dendrites are electrically compact structures (see Segev et al., 1989 for more complete references), including α-motoneurons (Lux et al., 1970; Fleshman et al., 1988), dorsal root ganglion cells (Brown et al., 1981b), three types of hippocampal cells (Brown et al., 1981a), and purkinje cells (Shelton, 1985).

We therefore tentatively define the "teacher" input to an association area to be those inputs that terminate primarily on granule cells in layer 4 (though perhaps including deep layer 3, Jones, 1981). With reference to the inherent modularity of the *sigma-pi* input-output relation (sec. 5.1), we recall that groups of *sigma-pi* units that are driven by the same teacher signal, and whose outputs are summed linearly at a common site of termination, may be viewed as a single, much larger "virtual" *sigma-pi* unit. We therefore propose that within each dendritic module, each laminar subgroup of pyramidal cells with a common site of projection (see Jones, 1981) can be considered as a single virtual *sigma-pi* unit, or *sigma-pi* column, whose collective dendritic space is interchangeable for purposes of learning (e.g. for cluster formation).

We note that the type of teacher input discussed here is similar to the suggested function of cerebellar climbing fibers (Albus, 1971; Marr, 1969), an idea that has also been incorporated into models of learning in cerebral cortex (Marr, 1970; Eccles, 1981). However, the idea of a common teacher signal to a *group* of output cells, whose dendritic space may then be shared during learning, is particular to the current model. Additionally, this cerebral cortical afferent system, if functionally akin to the system of cerebellar climbing fibers, is much less focused, such that a single pyramidal module is driven by a weighted average over a neighborhood several hundred microns in diameter (Szentagothai, 1978). The crucial feature of this pathway, for purposes of the current model, is that its action be unconditional (i.e. innately potent).

6.1.2 The Associational Input

A second major form of extrinsic excitatory input with access to the output pyramidal cells is the massive system of horizontal fibers in layer I. The primary source of these fibers is believed to be long range excitatory association fibers from both other cortical and subcortical areas (Szentagothai, 1978; Goldman & Nauta, 1977). In accordance with Marr (1970) and Eccles (1981), we interpret this system of horizontal fibers, which virtually permeates the dendritic "mat" of the pyramidal cell modules, as the primary *conditioned* input pathway at which cortical associative learning takes place. There is evidence that an individual layer 1 fiber can make excitatory synapses on apical dendrites of pyramidal cells across an area of cortex 5-6mm in diameter (Szentagothai, 1978). (Unlike the current emphasis on spatial clustering of synapses at the interface of association axons and post-synaptic dendrites, the proposals of Marr (1970) and Eccles (1981) assume this interface to be that of a standard Hebb-type associative memory, i.e. in which every synapse acts independently of every other.)

It may be recalled that a *sigma-pi* output unit y must have physical access to any and all of the afferents from X upon which its teacher signal functionally depends. For example, consider an associative map relating the proprioceptive state of a forelimb to the consequent visual activity as the limb