

# Bibliography

- [Acorn & Walden, 1992] Acorn, T., and Walden, S., "SMART: Support Management Automated Reasoning Technology for COMPAQ Customer Service," *Proc. Fourth Annual Conf. on Innovative Applications of Artificial Intelligence*, Menlo Park, CA: AAAI Press, 1992.
- [Aha, 1991] Aha, D., Kibler, D., and Albert, M., "Instance-Based Learning Algorithms," *Machine Learning*, 6, 37-66, 1991.
- [Anderson & Bower, 1973] Anderson, J. R., and Bower, G. H., *Human Associative Memory*, Hillsdale, NJ: Erlbaum, 1973.
- [Anderson, 1958] Anderson, T. W., *An Introduction to Multivariate Statistical Analysis*, New York: John Wiley, 1958.
- [Barto, Bradtke, & Singh, 1994] Barto, A., Bradtke, S., and Singh, S., "Learning to Act Using Real-Time Dynamic Programming," to appear in *Artificial Intelligence*, 1994.
- [Baum & Haussler, 1989] Baum, E., and Haussler, D., "What Size Net Gives Valid Generalization?" *Neural Computation*, 1, pp. 151-160, 1989.
- [Baum, 1994] Baum, E., "When Are  $k$ -Nearest Neighbor and Backpropagation Accurate for Feasible-Sized Sets of Examples?" in Hanson, S., Drastal, G., and Rivest, R., (eds.), *Computational Learning Theory and Natural Learning Systems, Volume 1: Constraints and Prospects*, pp. 415-442, Cambridge, MA: MIT Press, 1994.
- [Bellman, 1957] Bellman, R. E., *Dynamic Programming*, Princeton: Princeton University Press, 1957.
- [Blumer, *et al.*, 1987] Blumer, A., *et al.*, "Occam's Razor," *Info. Process. Lett.*, vol 24, pp. 377-80, 1987.

- [Blumer, *et al.*, 1990] Blumer, A., *et al.*, "Learnability and the Vapnik-Chervonenkis Dimension," *JACM*, 1990.
- [Bollinger & Duffie, 1988] Bollinger, J., and Duffie, N., *Computer Control of Machines and Processes*, Reading, MA: Addison-Wesley, 1988.
- [Brain, *et al.*, 1962] Brain, A. E., *et al.*, "Graphical Data Processing Research Study and Experimental Investigation," Report No. 8 (pp. 9-13) and No. 9 (pp. 3-10), Contract DA 36-039 SC-78343, SRI International, Menlo Park, CA, June 1962 and September 1962.
- [Breiman, *et al.*, 1984] Breiman, L., Friedman, J., Olshen, R., and Stone, C., *Classification and Regression Trees*, Monterey, CA: Wadsworth, 1984.
- [Brent, 1990] Brent, R. P., "Fast Training Algorithms for Multi-Layer Neural Nets," Numerical Analysis Project Manuscript NA-90-03, Computer Science Department, Stanford University, Stanford, CA 94305, March 1990.
- [Bryson & Ho 1969] Bryson, A., and Ho, Y.-C., *Applied Optimal Control*, New York: Blaisdell.
- [Buchanan & Wilkins, 1993] Buchanan, B. and Wilkins, D., (eds.), *Readings in Knowledge Acquisition and Learning*, San Francisco: Morgan Kaufmann, 1993.
- [Carbonell, 1983] Carbonell, J., "Learning by Analogy," in *Machine Learning: An Artificial Intelligence Approach*, Michalski, R., Carbonell, J., and Mitchell, T., (eds.), San Francisco: Morgan Kaufmann, 1983.
- [Cheeseman, *et al.*, 1988] Cheeseman, P., *et al.*, "AutoClass: A Bayesian Classification System," *Proc. Fifth Intl. Workshop on Machine Learning*, Morgan Kaufmann, San Mateo, CA, 1988. Reprinted in Shavlik, J. and Dietterich, T., *Readings in Machine Learning*, Morgan Kaufmann, San Francisco, pp. 296-306, 1990.
- [Cover & Hart, 1967] Cover, T., and Hart, P., "Nearest Neighbor Pattern Classification," *IEEE Trans. on Information Theory*, 13, 21-27, 1967.
- [Cover, 1965] Cover, T., "Geometrical and Statistical Properties of Systems of Linear Inequalities with Applications in Pattern Recognition," *IEEE Trans. Elec. Comp.*, EC-14, 326-334, June, 1965.

- [Dasarathy, 1991] Dasarathy, B. V., *Nearest Neighbor Pattern Classification Techniques*, IEEE Computer Society Press, 1991.
- [Dayan & Sejnowski, 1994] Dayan, P., and Sejnowski, T., "TD( $\lambda$ ) Converges with Probability 1," *Machine Learning*, 14, pp. 295-301, 1994.
- [Dayan, 1992] Dayan, P., "The Convergence of TD( $\lambda$ ) for General  $\lambda$ ," *Machine Learning*, 8, 341-362, 1992.
- [DeJong & Mooney, 1986] DeJong, G., and Mooney, R., "Explanation-Based Learning: An Alternative View," *Machine Learning*, 1:145-176, 1986. Reprinted in Shavlik, J. and Dietterich, T., *Readings in Machine Learning*, San Francisco: Morgan Kaufmann, 1990, pp 452-467.
- [Dietterich & Bakiri, 1991] Dietterich, T. G., and Bakiri, G., "Error-Correcting Output Codes: A General Method for Improving Multiclass Inductive Learning Programs," *Proc. Ninth Nat. Conf. on A.I.*, pp. 572-577, AAAI-91, MIT Press, 1991.
- [Dietterich, *et al.*, 1990] Dietterich, T., Hild, H., and Bakiri, G., "A Comparative Study of ID3 and Backpropagation for English Text-to-Speech Mapping," *Proc. Seventh Intl. Conf. Mach. Learning*, Porter, B. and Mooney, R. (eds.), pp. 24-31, San Francisco: Morgan Kaufmann, 1990.
- [Dietterich, 1990] Dietterich, T., "Machine Learning," *Annu. Rev. Comput. Sci.*, 4:255-306, Palo Alto: Annual Reviews Inc., 1990.
- [Duda & Fossum, 1966] Duda, R. O., and Fossum, H., "Pattern Classification by Iteratively Determined Linear and Piecewise Linear Discriminant Functions," *IEEE Trans. on Elect. Computers*, vol. EC-15, pp. 220-232, April, 1966.
- [Duda & Hart, 1973] Duda, R. O., and Hart, P.E., *Pattern Classification and Scene Analysis*, New York: Wiley, 1973.
- [Duda, 1966] Duda, R. O., "Training a Linear Machine on Mislabeled Patterns," SRI Tech. Report prepared for ONR under Contract 3438(00), SRI International, Menlo Park, CA, April 1966.
- [Efron, 1982] Efron, B., *The Jackknife, the Bootstrap and Other Resampling Plans*, Philadelphia: SIAM, 1982.

- [Ehrenfeucht, *et al.*, 1988] Ehrenfeucht, A., *et al.*, "A General Lower Bound on the Number of Examples Needed for Learning," in *Proc. 1988 Workshop on Computational Learning Theory*, pp. 110-120, San Francisco: Morgan Kaufmann, 1988.
- [Etzioni, 1991] Etzioni, O., "STATIC: A Problem-Space Compiler for PRODIGY," *Proc. of Ninth National Conf. on Artificial Intelligence*, pp. 533-540, Menlo Park: AAAI Press, 1991.
- [Etzioni, 1993] Etzioni, O., "A Structural Theory of Explanation-Based Learning," *Artificial Intelligence*, 60:1, pp. 93-139, March, 1993.
- [Evans & Fisher, 1992] Evans, B., and Fisher, D., *Process Delay Analyses Using Decision-Tree Induction*, Tech. Report CS92-06, Department of Computer Science, Vanderbilt University, TN, 1992.
- [Fahlman & Lebiere, 1990] Fahlman, S., and Lebiere, C., "The Cascade-Correlation Learning Architecture," in Touretzky, D., (ed.), *Advances in Neural Information Processing Systems, 2*, pp. 524-532, San Francisco: Morgan Kaufmann, 1990.
- [Fayyad, *et al.*, 1993] Fayyad, U. M., Weir, N., and Djorgovski, S., "SKI-CAT: A Machine Learning System for Automated Cataloging of Large Scale Sky Surveys," in *Proc. Tenth Intl. Conf. on Machine Learning*, pp. 112-119, San Francisco: Morgan Kaufmann, 1993. (For a longer version of this paper see: Fayyad, U. Djorgovski, G., and Weir, N., "Automating the Analysis and Cataloging of Sky Surveys," in Fayyad, U., *et al.*(eds.), *Advances in Knowledge Discovery and Data Mining*, Chapter 19, pp. 471ff., Cambridge: The MIT Press, March, 1996.)
- [Feigenbaum, 1961] Feigenbaum, E. A., "The Simulation of Verbal Learning Behavior," *Proceedings of the Western Joint Computer Conference*, 19:121-132, 1961.
- [Fikes, *et al.*, 1972] Fikes, R., Hart, P., and Nilsson, N., "Learning and Executing Generalized Robot Plans," *Artificial Intelligence*, pp 251-288, 1972. Reprinted in Shavlik, J. and Dietterich, T., *Readings in Machine Learning*, San Francisco: Morgan Kaufmann, 1990, pp 468-486.
- [Fisher, 1987] Fisher, D., "Knowledge Acquisition via Incremental Conceptual Clustering," *Machine Learning*, 2:139-172, 1987. Reprinted in Shavlik, J. and Dietterich, T., *Readings in Machine Learning*, San Francisco: Morgan Kaufmann, 1990, pp. 267-283.

- [Friedman, *et al.*, 1977] Friedman, J. H., Bentley, J. L., and Finkel, R. A., "An Algorithm for Finding Best Matches in Logarithmic Expected Time," *ACM Trans. on Math. Software*, 3(3):209-226, September 1977.
- [Fu, 1994] Fu, L., *Neural Networks in Artificial Intelligence*, New York: McGraw-Hill, 1994.
- [Gallant, 1986] Gallant, S. I., "Optimal Linear Discriminants," in *Eighth International Conf. on Pattern Recognition*, pp. 849-852, New York: IEEE, 1986.
- [Genesereth & Nilsson, 1987] Genesereth, M., and Nilsson, N., *Logical Foundations of Artificial Intelligence*, San Francisco: Morgan Kaufmann, 1987.
- [Gluck & Rumelhart, 1989] Gluck, M. and Rumelhart, D., *Neuroscience and Connectionist Theory*, The Developments in Connectionist Theory, Hillsdale, NJ: Erlbaum Associates, 1989.
- [Hammerstrom, 1993] Hammerstrom, D., "Neural Networks at Work," *IEEE Spectrum*, pp. 26-32, June 1993.
- [Haussler, 1988] Haussler, D., "Quantifying Inductive Bias: AI Learning Algorithms and Valiant's Learning Framework," *Artificial Intelligence*, 36:177-221, 1988. Reprinted in Shavlik, J. and Dietterich, T., *Readings in Machine Learning*, San Francisco: Morgan Kaufmann, 1990, pp. 96-107.
- [Haussler, 1990] Haussler, D., "Probably Approximately Correct Learning," *Proc. Eighth Nat. Conf. on AI*, pp. 1101-1108. Cambridge, MA: MIT Press, 1990.
- [Hebb, 1949] Hebb, D. O., *The Organization of Behaviour*, New York: John Wiley, 1949.
- [Hertz, Krogh, & Palmer, 1991] Hertz, J., Krogh, A., and Palmer, R., *Introduction to the Theory of Neural Computation*, Lecture Notes, vol. 1, Santa Fe Inst. Studies in the Sciences of Complexity, New York: Addison-Wesley, 1991.
- [Hirsh, 1994] Hirsh, H., "Generalizing Version Spaces," *Machine Learning*, 17, 5-45, 1994.

- [Holland, 1975] Holland, J., *Adaptation in Natural and Artificial Systems*, Ann Arbor: The University of Michigan Press, 1975. (Second edition printed in 1992 by MIT Press, Cambridge, MA.)
- [Holland, 1986] Holland, J. H., "Escaping Brittleness; The Possibilities of General-Purpose Learning Algorithms Applied to Parallel Rule-Based Systems." In Michalski, R., Carbonell, J., and Mitchell, T. (eds.), *Machine Learning: An Artificial Intelligence Approach, Volume 2*, chapter 20, San Francisco: Morgan Kaufmann, 1986.
- [Hunt, Marin, & Stone, 1966] Hunt, E., Marin, J., and Stone, P., *Experiments in Induction*, New York: Academic Press, 1966.
- [Jabbour, K., et al., 1987] Jabbour, K., et al., "ALFA: Automated Load Forecasting Assistant," *Proc. of the IEEE Power Engineering Society Summer Meeting*, San Francisco, CA, 1987.
- [John, 1995] John, G., "Robust Linear Discriminant Trees," *Proc. of the Conf. on Artificial Intelligence and Statistics*, Ft. Lauderdale, FL, January, 1995.
- [Kaelbling, 1993] Kaelbling, L. P., *Learning in Embedded Systems*, Cambridge, MA: MIT Press, 1993.
- [Kohavi, 1994] Kohavi, R., "Bottom-Up Induction of Oblivious Read-Once Decision Graphs," *Proc. of European Conference on Machine Learning (ECML-94)*, 1994.
- [Kolodner, 1993] Kolodner, J., *Case-Based Reasoning*, San Francisco: Morgan Kaufmann, 1993.
- [Koza, 1992] Koza, J., *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, Cambridge, MA: MIT Press, 1992.
- [Koza, 1994] Koza, J., *Genetic Programming II: Automatic Discovery of Reusable Programs*, Cambridge, MA: MIT Press, 1994.
- [Laird, et al., 1986] Laird, J., Rosenbloom, P., and Newell, A., "Chunking in Soar: The Anatomy of a General Learning Mechanism," *Machine Learning*, 1, pp. 11-46, 1986. Reprinted in Buchanan, B. and Wilkins, D., (eds.), *Readings in Knowledge Acquisition and Learning*, pp. 518-535, Morgan Kaufmann, San Francisco, CA, 1993.
- [Langley, 1992] Langley, P., "Areas of Application for Machine Learning," *Proc. of Fifth Int'l. Symp. on Knowledge Engineering*, Sevilla, 1992.

- [Langley, 1996] Langley, P., *Elements of Machine Learning*, San Francisco: Morgan Kaufmann, 1996.
- [Lavrač & Džeroski, 1994] Lavrač, N., and Džeroski, S., *Inductive Logic Programming*, Chichester, England: Ellis Horwood, 1994.
- [Lin, 1992] Lin, L., "Self-Improving Reactive Agents Based on Reinforcement Learning, Planning, and Teaching," *Machine Learning*, 8, 293-321, 1992.
- [Lin, 1993] Lin, L., "Scaling Up Reinforcement Learning for Robot Control," *Proc. Tenth Intl. Conf. on Machine Learning*, pp. 182-189, San Francisco: Morgan Kaufmann, 1993.
- [Littlestone, 1988] Littlestone, N., "Learning Quickly When Irrelevant Attributes Abound: A New Linear-Threshold Algorithm," *Machine Learning* 2: 285-318, 1988.
- [Maass & Turán, 1994] Maass, W., and Turán, G., "How Fast Can a Threshold Gate Learn?," in Hanson, S., Drastal, G., and Rivest, R., (eds.), *Computational Learning Theory and Natural Learning Systems, Volume 1: Constraints and Prospects*, pp. 381-414, Cambridge, MA: MIT Press, 1994.
- [Mahadevan & Connell, 1992] Mahadevan, S., and Connell, J., "Automatic Programming of Behavior-Based Robots Using Reinforcement Learning," *Artificial Intelligence*, 55, pp. 311-365, 1992.
- [Marchand & Golea, 1993] Marchand, M., and Golea, M., "On Learning Simple Neural Concepts: From Halfspace Intersections to Neural Decision Lists," *Network*, 4:67-85, 1993.
- [McCulloch & Pitts, 1943] McCulloch, W. S., and Pitts, W. H., "A Logical Calculus of the Ideas Immanent in Nervous Activity," *Bulletin of Mathematical Biophysics*, Vol. 5, pp. 115-133, Chicago: University of Chicago Press, 1943.
- [Michie, 1992] Michie, D., "Some Directions in Machine Intelligence," unpublished manuscript, The Turing Institute, Glasgow, Scotland, 1992.
- [Minton, 1988] Minton, S., *Learning Search Control Knowledge: An Explanation-Based Approach*, Kluwer Academic Publishers, Boston, MA, 1988.

- [Minton, 1990] Minton, S., "Quantitative Results Concerning the Utility of Explanation-Based Learning," *Artificial Intelligence*, 42, pp. 363-392, 1990. Reprinted in Shavlik, J. and Dietterich, T., *Readings in Machine Learning*, San Francisco: Morgan Kaufmann, 1990, pp. 573-587.
- [Mitchell, *et al.*, 1986] Mitchell, T., *et al.*, "Explanation-Based Generalization: A Unifying View," *Machine Learning*, 1:1, 1986. Reprinted in Shavlik, J. and Dietterich, T., *Readings in Machine Learning*, San Francisco: Morgan Kaufmann, 1990, pp. 435-451.
- [Mitchell, 1982] Mitchell, T., "Generalization as Search," *Artificial Intelligence*, 18:203-226, 1982. Reprinted in Shavlik, J. and Dietterich, T., *Readings in Machine Learning*, San Francisco: Morgan Kaufmann, 1990, pp. 96-107.
- [Moore & Atkeson, 1993] Moore, A., and Atkeson, C., "Prioritized Sweeping: Reinforcement Learning with Less Data and Less Time," *Machine Learning*, 13, pp. 103-130, 1993.
- [Moore, *et al.*, 1994] Moore, A. W., Hill, D. J., and Johnson, M. P., "An Empirical Investigation of Brute Force to Choose Features, Smoothers, and Function Approximators," in Hanson, S., Judd, S., and Petsche, T., (eds.), *Computational Learning Theory and Natural Learning Systems*, Vol. 3, Cambridge: MIT Press, 1994.
- [Moore, 1990] Moore, A., *Efficient Memory-based Learning for Robot Control*, PhD. Thesis; Technical Report No. 209, Computer Laboratory, University of Cambridge, October, 1990.
- [Moore, 1992] Moore, A., "Fast, Robust Adaptive Control by Learning Only Forward Models," in Moody, J., Hanson, S., and Lippman, R., (eds.), *Advances in Neural Information Processing Systems 4*, San Francisco: Morgan Kaufmann, 1992.
- [Mueller & Page, 1988] Mueller, R. and Page, R., *Symbolic Computing with Lisp and Prolog*, New York: John Wiley & Sons, 1988.
- [Muggleton, 1991] Muggleton, S., "Inductive Logic Programming," *New Generation Computing*, 8, pp. 295-318, 1991.
- [Muggleton, 1992] Muggleton, S., *Inductive Logic Programming*, London: Academic Press, 1992.
- [Muroga, 1971] Muroga, S., *Threshold Logic and its Applications*, New York: Wiley, 1971.



- [Natarajan, 1991] Natarajan, B., *Machine Learning: A Theoretical Approach*, San Francisco: Morgan Kaufmann, 1991.
- [Nilsson, 1965] Nilsson, N. J., "Theoretical and Experimental Investigations in Trainable Pattern-Classifying Systems," Tech. Report No. RADC-TR-65-257, Final Report on Contract AF30(602)-3448, Rome Air Development Center (Now Rome Laboratories), Griffiss Air Force Base, New York, September, 1965.
- [Nilsson, 1990] Nilsson, N. J., *The Mathematical Foundations of Learning Machines*, San Francisco: Morgan Kaufmann, 1990. (This book is a reprint of *Learning Machines: Foundations of Trainable Pattern-Classifying Systems*, New York: McGraw-Hill, 1965.)
- [Oliver, Dowe, & Wallace, 1992] Oliver, J., Dowe, D., and Wallace, C., "Inferring Decision Graphs using the Minimum Message Length Principle," *Proc. 1992 Australian Artificial Intelligence Conference*, 1992.
- [Pagallo & Haussler, 1990] Pagallo, G. and Haussler, D., "Boolean Feature Discovery in Empirical Learning," *Machine Learning*, vol.5, no.1, pp. 71-99, March 1990.
- [Pazzani & Kibler, 1992] Pazzani, M., and Kibler, D., "The Utility of Knowledge in Inductive Learning," *Machine Learning*, 9, 57-94, 1992.
- [Peterson, 1961] Peterson, W., *Error Correcting Codes*, New York: John Wiley, 1961.
- [Pomerleau, 1991] Pomerleau, D., "Rapidly Adapting Artificial Neural Networks for Autonomous Navigation," in Lippmann, P., *et al.* (eds.), *Advances in Neural Information Processing Systems*, 3, pp. 429-435, San Francisco: Morgan Kaufmann, 1991.
- [Pomerleau, 1993] Pomerleau, D., *Neural Network Perception for Mobile Robot Guidance*, Boston: Kluwer Academic Publishers, 1993.
- [Quinlan & Rivest, 1989] Quinlan, J. Ross, and Rivest, Ron, "Inferring Decision Trees Using the Minimum Description Length Principle," *Information and Computation*, 80:227-248, March, 1989.
- [Quinlan, 1986] Quinlan, J. Ross, "Induction of Decision Trees," *Machine Learning*, 1:81-106, 1986. Reprinted in Shavlik, J. and Dietterich, T., *Readings in Machine Learning*, San Francisco: Morgan Kaufmann, 1990, pp. 57-69.

- [Quinlan, 1987] Quinlan, J. R., "Generating Production Rules from Decision Trees," In *IJCAI-87: Proceedings of the Tenth Intl. Joint Conf. on Artificial Intelligence*, pp. 304-7, San Francisco: Morgan-Kaufmann, 1987.
- [Quinlan, 1990] Quinlan, J. R., "Learning Logical Definitions from Relations," *Machine Learning*, 5, 239-266, 1990.
- [Quinlan, 1993] Quinlan, J. Ross, *C4.5: Programs for Machine Learning*, San Francisco: Morgan Kaufmann, 1993.
- [Quinlan, 1994] Quinlan, J. R., "Comparing Connectionist and Symbolic Learning Methods," in Hanson, S., Drastal, G., and Rivest, R., (eds.), *Computational Learning Theory and Natural Learning Systems, Volume 1: Constraints and Prospects*, pp. 445-456,, Cambridge, MA: MIT Press, 1994.
- [Ridgway, 1962] Ridgway, W. C., *An Adaptive Logic System with Generalizing Properties*, PhD thesis, Tech. Rep. 1556-1, Stanford Electronics Labs., Stanford, CA, April 1962.
- [Rissanen, 1978] Rissanen, J., "Modeling by Shortest Data Description," *Automatica*, 14:465-471, 1978.
- [Rivest, 1987] Rivest, R. L., "Learning Decision Lists," *Machine Learning*, 2, 229-246, 1987.
- [Rosenblatt, 1958] Rosenblatt, F., *Principles of Neurodynamics*, Washington: Spartan Books, 1961.
- [Ross, 1983] Ross, S., *Introduction to Stochastic Dynamic Programming*, New York: Academic Press, 1983.
- [Rumelhart, Hinton, & Williams, 1986] Rumelhart, D. E., Hinton, G. E., and Williams, R. J., "Learning Internal Representations by Error Propagation," In Rumelhart, D. E., and McClelland, J. L., (eds.) *Parallel Distributed Processing*, Vol 1, 318-362, 1986.
- [Russell & Norvig 1995] Russell, S., and Norvig, P., *Artificial Intelligence: A Modern Approach*, Englewood Cliffs, NJ: Prentice Hall, 1995.
- [Samuel, 1959] Samuel, A., "Some Studies in Machine Learning Using the Game of Checkers," *IBM Journal of Research and Development*, 3:211-229, July 1959.

- [Schwartz, 1993] Schwartz, A., "A Reinforcement Learning Method for Maximizing Undiscounted Rewards," *Proc. Tenth Intl. Conf. on Machine Learning*, pp. 298-305, San Francisco: Morgan Kaufmann, 1993.
- [Sejnowski, Koch, & Churchland, 1988] Sejnowski, T., Koch, C., and Churchland, P., "Computational Neuroscience," *Science*, **241**: 1299-1306, 1988.
- [Shavlik, Mooney, & Towell, 1991] Shavlik, J., Mooney, R., and Towell, G., "Symbolic and Neural Learning Algorithms: An Experimental Comparison," *Machine Learning*, 6, pp. 111-143, 1991.
- [Shavlik & Dietterich, 1990] Shavlik, J. and Dietterich, T., *Readings in Machine Learning*, San Francisco: Morgan Kaufmann, 1990.
- [Sutton & Barto, 1987] Sutton, R. S., and Barto, A. G., "A Temporal-Difference Model of Classical Conditioning," in *Proceedings of the Ninth Annual Conference of the Cognitive Science Society*, Hillsdale, NJ: Erlbaum, 1987.
- [Sutton, 1988] Sutton, R. S., "Learning to Predict by the Methods of Temporal Differences," *Machine Learning* 3: 9-44, 1988.
- [Sutton, 1990] Sutton, R., "Integrated Architectures for Learning, Planning, and Reacting Based on Approximating Dynamic Programming," *Proc. of the Seventh Intl. Conf. on Machine Learning*, pp. 216-224, San Francisco: Morgan Kaufmann, 1990.
- [Taylor, Michie, & Spiegelhalter, 1994] Taylor, C., Michie, D., and Spiegelhalter, D., *Machine Learning, Neural and Statistical Classification*, Paramount Publishing International.
- [Tesauro, 1992] Tesauro, G., "Practical Issues in Temporal Difference Learning," *Machine Learning*, 8, nos. 3/4, pp. 257-277, 1992.
- [Towell & Shavlik, 1992] Towell G., and Shavlik, J., "Interpretation of Artificial Neural Networks: Mapping Knowledge-Based Neural Networks into Rules," in Moody, J., Hanson, S., and Lippmann, R., (eds.), *Advances in Neural Information Processing Systems*, 4, pp. 977-984, San Francisco: Morgan Kaufmann, 1992.
- [Towell, Shavlik, & Noordweier, 1990] Towell, G., Shavlik, J., and Noordweier, M., "Refinement of Approximate Domain Theories by Knowledge-Based Artificial Neural Networks," *Proc. Eighth Natl. Conf. on Artificial Intelligence*, pp. 861-866, 1990.

- [Unger, 1989] Unger, S., *The Essence of Logic Circuits*, Englewood Cliffs, NJ: Prentice-Hall, 1989.
- [Utgoff, 1989] Utgoff, P., "Incremental Induction of Decision Trees," *Machine Learning*, 4:161-186, Nov., 1989.
- [Valiant, 1984] Valiant, L., "A Theory of the Learnable," *Communications of the ACM*, Vol. 27, pp. 1134-1142, 1984.
- [Vapnik & Chervonenkis, 1971] Vapnik, V., and Chervonenkis, A., "On the Uniform Convergence of Relative Frequencies," *Theory of Probability and its Applications*, Vol. 16, No. 2, pp. 264-280, 1971.
- [Various Editors, 1989-1994] *Advances in Neural Information Processing Systems*, vols 1 through 6, San Francisco: Morgan Kaufmann, 1989-1994.
- [Watkins & Dayan, 1992] Watkins, C. J. C. H., and Dayan, P., "Technical Note: Q-Learning," *Machine Learning*, 8, 279-292, 1992.
- [Watkins, 1989] Watkins, C. J. C. H., *Learning From Delayed Rewards*, PhD Thesis, University of Cambridge, England, 1989.
- [Weiss & Kulikowski, 1991] Weiss, S., and Kulikowski, C., *Computer Systems that Learn*, San Francisco: Morgan Kaufmann, 1991.
- [Werbos, 1974] Werbos, P., *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*, Ph.D. Thesis, Harvard University, 1974.
- [Widrow & Lehr, 1990] Widrow, B., and Lehr, M. A., "30 Years of Adaptive Neural Networks: Perceptron, Madaline and Backpropagation," *Proc. IEEE*, vol. 78, no. 9, pp. 1415-1442, September, 1990.
- [Widrow & Stearns, 1985] Widrow, B., and Stearns, S., *Adaptive Signal Processing*, Englewood Cliffs, NJ: Prentice-Hall.
- [Widrow, 1962] Widrow, B., "Generalization and Storage in Networks of Adaline Neurons," in Yovits, Jacobi, and Goldstein (eds.), *Self-organizing Systems—1962*, pp. 435-461, Washington, DC: Spartan Books, 1962.
- [Winder, 1961] Winder, R., "Single Stage Threshold Logic," *Proc. of the AIEE Symp. on Switching Circuits and Logical Design*, Conf. paper CP-60-1261, pp. 321-332, 1961.

- [Winder, 1962] Winder, R., *Threshold Logic*, PhD Dissertation, Princeton University, Princeton, NJ, 1962.
- [Wnek, *et al.*, 1990] Wnek, J., *et al.*, "Comparing Learning Paradigms via Diagrammatic Visualization," in *Proc. Fifth Intl. Symp. on Methodologies for Intelligent Systems*, pp. 428-437, 1990. (Also Tech. Report MLI90-2, University of Illinois at Urbana-Champaign.)