References

E. H. L. Aarts and P. J. M. van Laarhoven. A new polynomial-time cooling schedule. In Proc. IEEE Intl. Conf. on Computer-Aided Design, pages 206–208, 1985. (p. 188)

B. Abramson. Expected-outcome: A general model of static evaluation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(2):182–193, February 1990. (pp. 167, 168)

D. H. Ackley and M. L. Littman. A case for distributed Lamarckian evolution. In C. Langton, C. Taylor, J. D. Farmer, and S. Ramussen, editors, *Artificial Life III: Santa Fe Institute Studies in the Sciences of Complexity*, volume 10, pages 487–509. Addison-Wesley, Redwood City, CA, 1993. (p. 170)

J. S. Albus. *Brains, Behavior, and Robotics.* Byte Books, Peterborough, NH, 1981. (p. 65)

D. Aldous and U. Vazirani. "Go with the winners" algorithms. In Proceedings of the 35th Symposium on Foundations of Computer Science, pages 492–501, 1994. (p. 162)

D. Applegate, R. Bixby, V. Chvatal, and B. Cook. Finding cuts in the TSP (A preliminary report). Technical Report 95-05, DIMACS, April 10, 1995. (p. 42)

C. G. Atkeson and J. C. Santamaria. A comparison of direct and model-based reinforcement learning. In International Conference on Robotics and Automation, 1997. (p. 145)

C. G. Atkeson. Using local trajectory optimizers to speed up global optimization in dynamic programming. In J. D. Cowan, G. Tesauro, and J. Alspector, editors, *Advances in Neural Information Processing Systems 6.* Morgan Kaufmann, 1994. (p. 25)

L. Baird. Residual algorithms: Reinforcement learning with function approximation. In Machine Learning: Proceedings of the Twelfth International Conference, 1995. (pp. 13, 23)

S. Baluja and R. Caruana. Removing the genetics from the standard genetic algorithm. In Proc. 12th International Conference on Machine Learning, pages 38–46. Morgan Kaufmann, 1995. (p. 170)

S. Baluja and S. Davies. Combining multiple optimization runs with optimal dependency trees. Technical Report CMU-CS-97-157, Carnegie Mellon University School of Computer Science, 1997. (pp. 74, 170)

R. S. Barr, B. L. Golden, J. P. Kelly, M. G. C. Resende, and W. R. Stewart. Designing and reporting on computational experiments with heuristic methods. Journal of Heuristics, 1(1):9–32, 1995. (pp. 70, 72)

A. Barto, R. Sutton, and C. Watkins. Learning and sequential decision making. Technical Report COINS 89-95, University of Massachusetts, 1989. (p. 20)

A. G. Barto, S. J. Bradtke, and S. P. Singh. Real-time learning and control using asynchronous dynamic programming. Artificial Intelligence, 1995. (pp. 20, 21)

J. Baxter, A. Tridgell, and L. Weaver. KnightCap: A chess program that learns by combining TD(λ) with minimax search. Technical report, Department of Systems Engineering, Australian National University, Canberra, Australia, November 1997. (p. 168)

R. Bellman, R. Kalaba, and B. Kotkin. Polynomial approximation—a new computational technique in dynamic programming: Allocation processes. Mathematics Of Computation, 17, 1963. (p. 20)

R. Bellman. *Dynamic Programming.* Princeton University Press, 1957. (pp. 13, 19)

R. Bellman. *An Introduction to Artificial Intelligence: Can Computers Think?* Boyd & Fraser Publishing Company, 1978. (p. 20)

D. A. Berry and B. Fristedt. *Bandit Problems: Sequential Allocation of Experiments.* Chapman and Hall, 1985. (p. 32)

D. Bertsekas and J. Tsitsiklis. *Neuro-Dynamic Programming.* Athena Scientific, Belmont, MA, 1996. (pp. 13, 19, 20, 21, 23, 24, 64, 66, 137, 138, 139, 141, 142, 143, 150)

D. Bertsekas, J. Tsitsiklis, and C. Wu. Rollout algorithms for combinatorial optimization. Technical Report LIDS-P 2386, MIT Laboratory for Information and Decision Systems, 1997. (p. 167)

D. Bertsekas. A counterexample to temporal differences learning. Neural Computation, 7:270–9, 1995. (p. 13)

J.R. Beveridge, C. Graves, and C. E. Lesher. Local search as a tool for horizon line matching. Technical Report CS-96-109, Colorado State University, 1996. (pp. 41, 70, 161)

K. D. Boese, A. B. Kahng, and S. Muddu. A new adaptive multi-start technique for combinatorial global optimizations. Operations Research Letters, 16:101–113, 1994. (pp. 162, 163, 170)

K. D. Boese. Cost versus distance in the traveling salesman problem. Technical Report CSD-950018, UCLA Computer Science Department, May 1995. (pp. 162, 163)

K. D. Boese. *Models for Iterative Global Optimization.* PhD thesis, UCLA Computer Science Department, 1996. (pp. 162, 188)

J. A. Boyan and A. W. Moore. Generalization in reinforcement learning: Safely approximating the value function. In G. Tesauro, D. S. Touretzky, and T. K. Leen, editors, *Advances In Neural Information Processing Systems 7.* MIT Press, 1995. (pp. 13, 14, 23, 25, 27)

J. A. Boyan and A. W. Moore. Learning evaluation functions for large acyclic domains. In L. Saitta, editor, *Machine Learning: Proceedings of the Thirteenth International Conference.* Morgan Kaufmann, 1996. (pp. 14, 25)

J. A. Boyan and A. W. Moore. Using prediction to improve combinatorial optimization search. In Proceedings of the Sixth International Workshop on Artificial Intelligence and Statistics (AISTATS), January 1997. (p. 14)

J. A. Boyan and A. W. Moore. Learning evaluation functions for global optimization and Boolean satisfiability. In Proceedings of the Fifteenth National Conference on Artificial Intelligence (AAAI), 1998. (Outstanding Paper Award). (p. 14)

J. A. Boyan, A. W. Moore, and R. S. Sutton, editors. *Proceedings of the Workshop on Value Function Approximation*, Machine Learning Conference, July 1995. CMU-CS-95-206. Internet resource available at `http://www.cs.cmu.edu/~reinf/ml95/`. (pp. 13, 19)

J. A. Boyan, D. Freitag, and T. Joachims. A machine learning architecture for optimizing web search engines. Technical Report AAAI Technical Report WS-96-06, Proceedings of the AAAI workshop on Internet-Based Information Systems, 1996. (p. 41)

J. A. Boyan. Modular neural networks for learning context-dependent game strategies. Master's thesis, University of Cambridge, Department of Engineering and Computer Laboratory, 1992. (pp. 13, 21, 168)

S. J. Bradtke and A. G. Barto. Linear least-squares algorithms for temporal difference learning. Machine Learning, 22(1/2/3):33–57, 1996. (pp. 15, 137, 143, 175)

L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone. Classification and regression trees. Technical report, Wadsworth International, Monterey, CA, 1984. (p. 65)

W. L. Buntine, L. Su, R. Newton, and A. Mayer. Adaptive methods for netlist partitioning. In Proceedings of ICCAD-97, San Jose, CA, November 1997. (p. 167)

M. S. Campbell. Deep Blue: The IBM chess machine. Talk given at Carnegie Mellon University, March 1998. (p. 12)

Y. Censor, M. D. Altschuler, and W. D. Powlis. A computational solution of the inverse problem in radiation-therapy treatment planning. Applied Mathematics and Computation, 25:57–87, 1988. (pp. 90, 91)

H-Y. Chao and M. P. Harper. An efficient lower bound algorithm for channel routing. Integration: The VLSI Journal, 1996. (pp. 79, 154, 196)

D. M. Chickering, D. Geiger, and D. Heckerman. Learning bayesian networks is NP-hard. Technical Report MSR-TR-94-17, Microsoft Research, November 1994. (pp. 84, 85)

J. Christensen and R. Korf. A unified theory of heuristic evaluation functions and its application to learning. In Proceedings of the 4th National Conference on Artificial Intelligence, pages 148–152, 1986. (p. 20)

J. Christensen. Learning static evaluation functions by linear regression. In T. Mitchell, J. Carbonell, and R. Michalski, editors, *Machine learning: A guide to current research*, pages 39–42. Kluwer, Boston, 1986. (pp. 20, 168)

W. S. Cleveland and S. J. Devlin. Locally weighted regression: An approach to regression analysis by local fitting. Journal of the American Statistical Association, 83(403):596–610, September 1988. (pp. 38, 65)

E. G. Coffman, M. R. Garey, and D. S. Johnson. Approximation algorithms for bin packing: a survey. In D. Hochbaum, editor, *Approximation Algorithms for NP-Hard Problems*. PWS Publishing, 1996. (pp. 42, 43, 76, 196)

J. M. Cohn. *Automatic Device Placement for Analog Cells in KOAN*. PhD thesis, Carnegie Mellon University Department of Electrical and Computer Engineering, February 1992. (pp. 47, 178)

T. H. Cormen, C. E. Leiserson, and R. L. Rivest. *Introduction to Algorithms*. MIT Press, 1990. (p. 24)

J. Crawford. Propositional versions of parity learning problems. Internet resource available by anonymous FTP, at `ftp://dimacs.rutgers.edu/pub/challenge/-sat/contributed/crawford/README`, July 1993. (pp. 96, 199)

R. Crites and A. Barto. Improving elevator performance using reinforcement learning. In D. Touretzky, M. Mozer, and M. Hasselmo, editors, *Advances in Neural Information Processing Systems 8*, 1996. (pp. 13, 21)

S. Davies and S. Baluja. Personal communication, April 1998. (p. 171)

J. S. de Bonet, C. L. Isbell Jr., and P. Viola. MIMIC: Finding optima by estimating probability densities. In Michael C. Mozer, Michael I. Jordan, and Thomas Petsche, editors, *Advances in Neural Information Processing Systems*, volume 9, page 424. The MIT Press, 1997. (p. 170)

R. Dearden, N. Friedman, and S. Russell. Bayesian Q-Learning. In Proceedings of the Fifteenth National Conference on Artificial Intelligence (AAAI), 1998. (p. 130)

F. D'Epenoux. A probabilistic production and inventory problem. Management Science, 10:98–108, 1963. (p. 19)

D. N. Deutsch. A 'dogleg' channel router. In Proceedings of the 13th ACM/IEEE Design Automation Conference, pages 425–433, 1976. (p. 79)

J. Doran and D. Michie. Experiments with the graph traverser program. Proceedings of the Royal Society of London, 294, Series A:235–259, 1966. (p. 169)

M. Dorigo and L. Gambardella. Ant-Q: A reinforcement learning approach to combinatorial optimization. Technical Report 95-01, Irdia, Université Libre de Bruxelles, 1995. (p. 167)

D. Dorling. Cartograms for visualizing human geography. In H. M. Hearnshaw and D. J. Unwin, editors, *Visualization in Geographical Information Systems*, pages 85–102. Wiley, 1994. (p. 93)

D. Dorling. *Area cartograms: their use and creation.* Number 59 in Concepts and Techniques in Modern Geography. University of East Anglia: Environmental Publications, 1996. (p. 93)

Q. Duan, S. Sorooshian, and V. Gupta. Effective and efficient global optimization of conceptual rainfall-runoff models. Water Resources Research, 28(4):1015–1031, 1992. (p. 41)

S. Fahlman and C. Lebiere. The Cascade-Correlation learning architecture. In D. Touretzky, editor, *Advances in Neural Information Processing Systems 2*. Morgan Kaufmann, 1990. (p. 65)

E. Falkenauer and A. Delchambre. A genetic algorithm for bin packing and line balancing. In Proc. of the IEEE 1992 International Conference on Robotics and Automation, pages 1186–1192, Nice, France, May 1992. (pp. 47, 74)

E. Falkenauer. A hybrid grouping genetic algorithm for bin packing. Journal of Heuristics, 2(1):5–30, 1996. (pp. 74, 76, 77, 196)

N. Friedman and Z. Yakhini. On the sample complexity of learning Bayesian networks. In Proc. 12th Conference on Uncertainty in Artificial Intelligence, 1996. (p. 84)

J. H. Friedman. Multivariate adaptive regression splines. Annals of Statistics, 19(1):1–67, 1991. (p. 65)

N. Friedman. Learning belief networks in the presence of missing values and hidden variables. In Proc. 14th International Conference on Machine Learning, pages 125–133. Morgan Kaufmann, 1997. (pp. 84, 85)

M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness.* Freeman, 1979. (pp. 42, 72, 95)

F. Glover and M. Laguna. Tabu search. In Modern Heuristic Techniques for Combinatorial Problems. Scientific Publications, Oxford, 1993. (p. 164)

D. E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning.* Addison-Wesley, Reading, Mass., 1989. (p. 169)

G. Gordon. Stable function approximation in dynamic programming. In Proceedings of the 12th International Conference on Machine Learning. Morgan Kaufmann, 1995. (pp. 13, 23, 27)

S. M. Gusein-Zade and V. S. Tikunov. A new technique for constructing continuous cartograms. Cartography and Geographic Information Systems, 20(3):167–173, 1993. (p. 93)

L. W. Hagen and A. B. Kahng. Combining problem reduction and adaptive multistart: A new technique for superior iterative partitioning. IEEE Transactions on CAD, 16(7):709–717, 1997. (p. 164)

B. Hajek. Cooling schedules for optimal annealing. Math. Oper. Res., 13(2):311–329, 1988. (p. 188)

M. Harmon, L. Baird, and A. H. Klopf. Advantage updating applied to a differential game. In G. Tesauro, D. S. Touretzky, and T. K. Leen, editors, *Advances In Neural Information Processing Systems 7.* MIT Press, 1995. (p. 18)

D. Heckerman, D. Geiger, and M. Chickering. Learning Bayesian networks: The combination of knowledge and statistical data. Technical Report MSR-TR-94-09, Microsoft Research, 1994. (pp. 84, 85)

G. E. Hinton and S. J. Nowlan. How learning can guide evolution. Complex Systems, 1(1):495–502, June 1987. (p. 170)

R. Howard. *Dynamic Programming and Markov Processes.* MIT Press and John Wiley & Sons, 1960. (pp. 19, 167)

F. Hsu, T. S. Anantharaman, M. S. Campbell, and A. Nowatzyk. A grandmaster chess machine. Scientific American, 263(4):44–50, October 1990. (p. 12)

M. D. Huang, F. Romeo, and A. Sangiovanni-Vincentelli. An efficient general cooling schedule for simulated annealing. In Proceedings of the International Conference on Computer-Aided Design, pages 381–384, November 1986. (p. 188)

A. Jagota, L. Sanchis, and R. Ganesan. Approximating maximum clique using neural network and related heuristics. In D. S. Johnson and M. A. Trick, editors, *DIMACS Series: Second DIMACS Challenge.* American Mathematical Society, 1996. (p. 164)

A. Jagota. An adaptive, multiple restarts neural network algorithm for graph coloring. European J. of Oper. Res., 93:257–270, 1996. (p. 164)

Y. Jiang, H. Kautz, and B. Selman. Solving problems with hard and soft constraints using a stochastic algorithm for MAX-SAT. In First International Joint Workshop on Artificial Intelligence and Operations Research, Timberline, Oregon, 1995. (p. 102)

D. S. Johnson and L. A. McGeoch. The traveling salesman problem: A case study in local optimization. In E. H. L. Aarts and J. K. Lenstra, editors, *Local Search in Combinatorial Optimization.* Wiley and Sons, 1995. (to appear). (pp. 161, 163, 164, 170)

D. S. Johnson, C. R. Aragon, L. A. Mcgeoch, and C. Schevon. Optimization by simulated annealing: an experimental evaluation; part I, graph partitioning. Oper. Res., 37:865–892, 1989. (p. 188)

D. S. Johnson, C. R. Aragon, L. A. McGeoch, and C. Schevon. Optimization by simulated annealing: an experimental evaluation; part II, graph colouring and number partitioning. Operations Research, 39:378–406, 1991. (p. 188)

S. A. Johnson, J. R. Stedinger, C. A. Shoemaker, Y. Li, and J. A. Tejada-Guibert. Numerical solution of continuous-state dynamic programs using linear and spline interpolation. Operations Research, 41(3):484–500, 1993. (p. 20)

D. S. Johnson. A theoretician's guide to the experimental analysis of algorithms. Unpublished manuscript, April 1996. (p. 70)

A. Juels and M. Wattenberg. Stochastic hillclimbing as a baseline mathod for evaluating genetic algorithms. In David S. Touretzky, Michael C. Mozer, and Michael E. Hasselmo, editors, *Advances in Neural Information Processing Systems*, volume 8, pages 430–436. The MIT Press, 1996. (p. 161)

L. P. Kaelbling. *Learning in Embedded Systems.* The MIT Press, Cambridge, MA, 1993. (p. 177)

R. Karp. Algorithms as a tool for molecular biology. Talk given at Carnegie Mellon University, archived on Internet at `http://ulserver.speech.cs.cmu.edu/-v/karp/`, September 1997. (p. 41)

H. Kautz. Personal communication, April 1998. (p. 99)

C. Kenyon. Best-fit bin-packing with random order. In Proceedings of the Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, pages 359–364, 1996. (p. 76)

S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimisation by simulated annealing. Science, 220:671–680, 1983. (pp. 44, 188)

L. Kuvayev and R. Sutton. Approximation in model-based learning. In C. G. Atkeson and G. J. Gordon, editors, *ICML-97 Workshop on Modelling in Reinforcement Learning*, 1997. Internet resource available at `http://www.cs.cmu.edu/~ggordon/-ml97ws/`. (p. 145)

W. Lam and F. Bacchus. Learning Bayesian belief networks: An approach based on the MDL principle. Computational Intelligence, 10(4), 1994. (p. 84)

J. Lam and J.-M. Delosme. Performance of a new annealing schedule. In ACM/IEEE, editor, *Proceedings of the 25th ACM/IEEE Design Automation Conference*, pages 306–311, Anaheim, CA, June 1988. IEEE Computer Society Press. (pp. 71, 188)

J. Lam. *An Efficient Simulated Annealing Schedule.* PhD thesis, Computer Science Department, Yale University, 1988. (p. 188)

K.-F. Lee and S. Mahajan. A pattern classification approach to evaluation function learning. Artificial Intelligence, 36, 1988. (pp. 20, 168)

S. Lin and B. W. Kernighan. An effective heuristic algorithm for the traveling salesman problem. Operations Research, 21:498–516, 1973. (pp. 12, 43)

Z.-M. Lin. An efficient optimum channel routing algorithm. In IEEE Proceedings of the SOUTHEASTCON, volume 2, pages 1179–1183, 1991. (p. 79)

L.-J. Lin. *Reinforcement Learning for Robots Using Neural Networks.* PhD thesis, Carnegie Mellon University, 1993. (pp. 142, 166)

M. L. Littman and C. Szepesvári. A generalized reinforcement-learning model: Convergence and applications. In L. Saitta, editor, *Machine Learning: Proceedings of the Thirteenth International Conference*. Morgan Kaufmann, 1996. (pp. 13, 18, 32)

M. L. Littman. Markov games as a framework for multi-agent reinforcement learning. In Proc. 11th International Conference on Machine Learning, pages 157–163. Morgan Kaufmann, 1994. (p. 18)

M. L. Littman. *Algorithms for Sequential Decision Making*. PhD thesis, Department of Computer Science, Brown University, Providence, RI, 1996. Also Technical Report CS-96-09. (p. 18)

S. Mahadevan, N. Marchalleck, T. K. Das, and A. Gosavi. Self-improving factory simulation using continuous-time average-reward reinforcement learning. In Proc. 14th International Conference on Machine Learning, pages 202–210. Morgan Kaufmann, 1997. (p. 18)

O. C. Martin and S. W. Otto. Combining simulated annealing with local search heuristics. Technical Report CS/E 94-016, Oregon Graduate Institute Department of Computer Science and Engineering, June 1994. (pp. 162, 163, 170)

D. McAllester, H. Kautz, and B. Selman. Evidence for invariants in local search. In Proceedings of AAAI-97, 1997. (p. 96)

C.J. Merz and P.M. Murphy. UCI repository of machine learning databases, 1998. Internet resource available at `http://www.ics.uci.edu/~mlearn/-MLRepository.html`. (p. 197)

T. M. Mitchell. *Machine Learning*. McGraw-Hill, 1997. (p. 84)

D. Mitra, F. Romeo, and A. Sangiovanni-Vincentelli. Convergence and finite-time behavior of simulated annealing. Adv. in Applied Probability, 18:747–771, 1986. (p. 188)

R. Moll, A. Barto, T. Perkins, and R. Sutton. Reinforcement and local search: A case study. Technical Report UM-CS-1997-044, University of Massachusetts, Amherst, Computer Science, October, 1997. (pp. 64, 106)

J. Moody and C. J. Darken. Fast learning in networks of locally-tuned processing units. Neural Computation, 1(2):281–294, 1989. (p. 65)

A. W. Moore and C. G. Atkeson. Prioritized sweeping: Reinforcement learning with less data and less time. Machine Learning, 13:103–130, 1993. (pp. 13, 19, 130, 145)

A. W. Moore and M. S. Lee. Cached sufficient statistics for efficient machine learning with large datasets. Journal of Artificial Intelligence Research, 8:67–91, 1998. (pp. 85, 86, 197)

A. W. Moore and J. Schneider. Memory-based stochastic optimization. In D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo, editors, *Neural Information Processing Systems 8*. MIT Press, 1996. (pp. 177, 180)

A. W. Moore, J. G. Schneider, J. A. Boyan, and M. S. Lee. Q2: Memory-based active learning for optimizing noisy continuous functions. In Proceedings of the Fifteenth Interational Conference on Machine Learning (ICML), 1998. (p. 180)

D. Moriarty, A. Schultz, and J. Grefenstette. Reinforcement learning through evolutionary computation. Technical Report NCARAI Report AIC-97-015, Naval Research Laboratory, 1997. (p. 179)

Y. Nakakuki and N. M. Sadeh. Increasing the efficiency of simulated annealing search by learning to recognize (un)promising runs. Technical Report CMU-RI-TR-94-30, CMU Robotics Institute, 1994. (p. 177)

A. Neumaier. Molecular modeling of proteins and mathematical prediction of protein structure. SIAM Rev., 39:407–460, 1997. (p. 41)

N.J. Nilsson. *Principles of Artificial Intelligence.* McGraw-Hill, 1980. (p. 12)

E. Ochotta. *Synthesis of High-Performance Analog Cells in ASTRX/OBLX.* PhD thesis, Carnegie Mellon University Department of Electrical and Computer Engineering, April 1994. (pp. 179, 188)

T. Perkins, R. Moll, and S. Zilberstein. Filtering to improve the performance of multi-trial optimization algorithms. Unpublished manuscript, November 1997. (p. 177)

J. Pollack, A. Blair, and M. Land. Coevolution of a backgammon player. In C.G. Langton, editor, *Proceedings of Artificial Life 5.* MIT Press, 1996. (p. 179)

W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing.* Cambridge University Press, second edition, 1992. (pp. 67, 142, 176, 179, 188)

M. L. Puterman and M. C. Shin. Modified policy iteration algorithms for discounted Markov decision processes. Management Science, 24:1127–1137, 1978. (p. 19)

M. L. Puterman. *Markov Decision Processes—Discrete Stochastic Dynamic Programming.* John Wiley & Sons, Inc., New York, NY, 1994. (pp. 12, 13)

L. A. Rendell. A new basis for state-space learning systems and a successful implementation. Artificial Intelligence, 20:369–392, July 1983. (p. 168)

D. Rumelhart, G. Hinton, and R. Williams. Learning internal representations by error propagation. In D. Rumelhart and J. McClelland, editors, *Parallel Distributed Processing*, volume 1, chapter 8. MIT Press, 1986. (p. 65)

G. A. Rummery and M. Niranjan. On-line Q-learning using connectionist systems. Technical Report CUED/F-INFENG/TR166, Cambridge University Engineering Department, 1994. (p. 19)

S. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach.* Prentice Hall, 1995. (p. 44)

J. Rust. Numerical dynamic programming in economics. In Handbook of Computational Economics. Elsevier, North Holland, 1996. (p. 20)

A. L. Samuel. Some studies in machine learning using the game of checkers. IBM Journal of Research and Development, 3:211–229, 1959. (pp. 20, 168)

A. L. Samuel. Some studies in machine learning using the game of checkers II—Recent progress. IBM Journal of Research and Development, 11(6):601–617, 1967. (pp. 20, 168)

J. G. Schneider, J. A. Boyan, and A. W. Moore. Value function based production scheduling. In Proceedings of the Fifteenth International Conference on Machine Learning (ICML), 1998. (p. 35)

N. Schraudolph, P. Dayan, and T. Sejnowski. Using TD(λ) to learn an evaluation function for the game of Go. In J. D. Cowan, G. Tesauro, and J. Alspector, editors, *NIPS-6*. Morgan Kaufmann, 1994. (p. 168)

C. Sechen and A. Sangiovanni-Vincentelli. TimberWolf 3.2: A new standard cell placement and global routing package. In Proceedings of the IEEE/ACM Design Automation Conference, pages 432–439, 1986. (p. 188)

B. Selman and H. A. Kautz. An empirical study of greedy local search for satisfiability testing. In Proceedings of the 11th National Conference on Artificial Intelligence, pages 46–53, Menlo Park, CA, USA, July 1993. AAAI Press. (p. 43)

B. Selman, H. Kautz, and B. Cohen. Local search strategies for satisfiability testing. In Cliques, Coloring, and Satisfiability: Second DIMACS Implementation Challenge, DIMACS Series in Discrete Mathematics and Theoretical Computer Science. American Mathematical Society, 1996. (pp. 44, 95, 96)

B. Selman, H. Kautz, and D. McAllester. Ten challenges in propositional reasoning and search. In Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence (IJCAI), 1997. (p. 96)

S. Singh and D. Bertsekas. Reinforcement learning for dynamic channel allocation in cellular telephone systems. In M. C. Mozer, M. I. Jordan, and T. Petsche, editors, *Advances in Neural Information Processing Systems*, volume 9, page 974. The MIT Press, 1997. (p. 150)

S. P. Singh and R. S. Sutton. Reinforcement learning with replacing eligibility traces. Machine Learning, 22:123–158, 1996. (pp. 19, 139)

S. P. Singh and R. Yee. An upper bound on the loss from approximate optimal-value functions (Technical Note). Machine Learning, 1994. (p. 23)

S. Singh, T. Jaakkola, M. L. Littman, and C. Szepesvári. Convergence results for single-step on-policy reinforcement learning algorithms. Machine Learning, 1998. To appear. (pp. 19, 21, 22)

L. Su, W. L. Buntine, R. Newton, and B. Peters. Learning as applied to stochastic optimization for standard cell placement. International Conference on Computer Design (Submitted), 1998. (p. 178)

R. Subramanian, R. P. Scheff Jr., J. D. Quillinan, D. S. Wiper, and R. E. Marsten. Coldstart: Fleet assignment at delta air lines. Interfaces, 24(1):104–120, Jan.-Feb. 1994. (pp. 41, 42)

R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction.* MIT Press, 1998. (p. 138)

R. S. Sutton and S. D. Whitehead. Online Learning with Random Representations. In Machine Learning: Proceedings of the Tenth International Conference, pages 314–321, San Mateo, CA, 1993. Morgan Kaufmann. (pp. 65, 66)

R. S. Sutton. Implementation details of the TD(λ) procedure for the case of vector predictions and backpropagation. Technical Note TN87-509.1, GTE Laboratories, May 1987. (p. 22)

R. S. Sutton. Learning to predict by the methods of temporal differences. Machine Learning, 3, 1988. (pp. 13, 19, 21, 137, 138, 139)

R. S. Sutton. Integrated architectures for learning, planning, and reacting based on approximating dynamic programming. In Proceedings of the Seventh International Conference on Machine Learning. Morgan Kaufmann, 1990. (pp. 19, 20, 145)

R. S. Sutton. Gain adaptation beats least squares. In Proceedings of the 7^{th} Yale Workshop on Adaptive and Learning Systems, pages 161–166, 1992. (p. 150)

R. S. Sutton. Generalization in reinforcement learning: Successful examples using sparse coarse coding. In D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo, editors, *Advances in Neural Information Processing Systems*, volume 8, pages 1038–1044. The MIT Press, 1996. (pp. 23, 27, 139)

W. Swartz and C. Sechen. New algorithms for the placement and routing of macro cells. In Satoshi Sangiovanni-Vincentelli, Alberto; Goto, editor, *Proceedings of the IEEE International Conference on Computer-Aided Design*, pages 336–339, Santa Clara, CA, November 1990. IEEE Computer Society Press. (pp. 71, 188, 189)

S. Szykman and J. Cagan. A simulated annealing-based approach to three-dimensional component packing. ASME Journal of Mechanical Design, 117, June 1995. (pp. 41, 47)

T. G. Szymanski. Dogleg channel routing is NP-complete. IEEE Transactions on Computer-Aided Design, CAD-4:31–40, 1985. (p. 79)

G. Tesauro and G. R. Galperin. On-line policy improvement using Monte-Carlo search. In M. C. Mozer, M. I. Jordan, and T. Petsche, editors, *Advances in Neural Information Processing Systems*, volume 9. MIT Press, 1997. (pp. 167, 168)

G. Tesauro. Practical issues in temporal difference learning. Machine Learning, 8(3/4), May 1992. (pp. 13, 20, 21, 168)

G. Tesauro. TD-Gammon, a self-teaching backgammon program, achieves master-level play. Neural Computation, 6(2):215–219, 1994. (p. 20)

S. Thrun. Efficient exploration in reinforcement learning. Technical Report CMU-CS-92-102, Carnegie Mellon University, March 1992. (pp. 21, 22, 130)

S. Thrun. Learning to play the game of chess. In G. Tesauro, D. Touretzky, and T. Leen, editors, *Advances in Neural Information Processing Systems 7*, pages 1069–1076. The MIT Press, Cambridge, MA, 1995. (p. 168)

J. N. Tsitsiklis and B. Van Roy. An analysis of temporal-difference learning with function approximation. Technical Report LIDS-P-2322, MIT, 1996. (pp. 13, 21, 23, 137, 139)

W. Tunstall-Pedoe. Genetic algorithms optimizing evaluation functions. International Computer Chess Association Journal, 14(3):119–128, 1991. (p. 179)

P. Utgoff and J. Clouse. Two kinds of training information for evaluation function learning. In Proceedings of the National Conference on Artificial Intelligence (AAAI), 1991. (p. 179)

P. E. Utgoff. Feature function learning for value function approximation. Technical Report UM-CS-1996-009, University of Massachusetts, Amherst, Computer Science, January, 1996. (p. 176)

C. Watkins. *Learning From Delayed Rewards*. PhD thesis, Cambridge University, 1989. (pp. 13, 19, 20, 178)

S. Webb. Optimization by simulated annealing of three-dimensional conformal treatment planning for radiation fields defined by a multileaf collimator. Phys. Med. Biol., 36:1201–1226, 1991. (pp. 41, 91)

S. Webb. Optimising the planning of intensity-modulated radiotherapy. Phys. Med. Biol., 39:2229–2246, 1994. (p. 91)

A. Wilk. Constraint satisfaction channel routing. Technical Report CS96-04, Department of Applied Mathematics and Computer Science, Weizmann Institute of Science, 1996. (pp. 79, 80)

D. F. Wong, H.W. Leong, and C.L. Liu. *Simulated Annealing for VLSI Design*. Kluwer, 1988. (pp. 12, 41, 46, 79, 80, 81, 82, 111, 114, 119, 170, 196)

K. Woolsey. Rollouts. Inside Backgammon, 1(5):4–7, September-October 1991. (p. 167)

T. Yoshimura and E. S. Kuh. Efficient algorithms for channel routing. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, CAD-1(1):21–35, January 1982. (pp. 79, 196)

W. Zhang and T. G. Dietterich. A reinforcement learning approach to job-shop scheduling. In Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), pages 1114–1120, 1995. (pp. 13, 21, 161, 164)

W. Zhang and T. G. Dietterich. Solving combinatorial optimization tasks by reinforcement learning: A general methodology applied to resource-constrained scheduling. Submitted for publication, 1998. (pp. 155, 165, 166)

W. Zhang. *Reinforcement Learning for Job-Shop Scheduling*. PhD thesis, Oregon State University, 1996. (p. 164)

M. Zweben, B. Daun, and M. Deale. Scheduling and rescheduling with iterative repair. In M. Zweben and M. S. Fox, editors, *Intelligent Scheduling*, chapter 8, pages 241–255. Morgan Kaufmann, 1994. (pp. 165, 166)