References

[1] N. Abe. Polynomial learning of semilinear sets. In *Proc. 2nd Annu. Workshop on Comput. Learning Theory*, pages 25–40, San Mateo, CA, 1989. Morgan Kaufmann.

[2] N. Abe. Learning commutative deterministic finite state automata in polynomial time. *New Generation Computing*, 8:319–335, 1991.

[3] N. Abe and J. Takeuchi. The 'lob-pass' problem and an on-line learning model of rational choice. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 422–428. ACM Press, New York, NY, 1993.

[4] N. Abe, J. Takeuchi, and M. K. Warmuth. Polynomial learnability of probabilistic concepts with respect to the Kullback-Leibler divergence. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 277–289, San Mateo, CA, 1991. Morgan Kaufmann.

[5] N. Abe and M. K. Warmuth. On the computational complexity of approximating distributions by probabilistic automata. *Machine Learning*, 1991. Special issue for COLT90, to appear.

[6] N. Abe and O. Watanabe. Polynomially sparse variations and reducibility among prediction problems. *IEICE Trans. Inf. & Syst.*, E75-D(4):449–458, July 1992.

[7] Y. S. Abu-Mostafa. The Vapnik-Chervonenkis dimension: Information verses complexity in learning. *Neural Computation*, 1(3):312–317, 1989.

[8] D. H. Ackley, G. E. Hinton, and T. J. Sejnowski. A learning algorithm for Boltzmann machines. *Cognitive Science*, 9:147–169, 1985.

[9] R. T. Adams. Inference of LISP programs from examples. Bachelor's thesis, MIT EECS Department, June 1990.

[10] L. Adleman and M. Blum. Inductive inference and unsolvability. *J. Symbolic Logic*, 56(3):891–900, 1991.

[11] S. Ahmad. A study of scaling and generalization in neural networks. Technical Report UIUCDCS-R-88-1454, University of Illinois at Urbana-Champaign, Sept. 1988.

[12] W. Aiello and M. Mihail. Learning the fourier spectrum of probabilistic lists and trees. In *Proceedings SODA 91*, pages 291–299, Boston, MA, Jan. 1991. AAAI Press/The MIT Press.

[13] H. Aizenstein, L. Hellerstein, and L. Pitt. Read-thrice DNF is hard to learn with membership and equivalence queries. In *Proc. of the 33rd Symposium on the Foundations of Comp. Sci.*, pages 523–532. IEEE Computer Society Press, Los Alamitos, CA, 1992.

[14] H. Aizenstein and L. Pitt. Exact learning of read-twice DNF formulas. In *Proc. 32th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 170–179. IEEE Computer Society Press, 1991.

[15] H. Aizenstein and L. Pitt. Exact learning of read-k disjoint DNF and not-so-disjoint DNF. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 71–76. ACM Press, New York, NY, 1992.

[16] D. Aldous and U. Vazirani. A Markovian extension of Valiant's learning model. In *Proc. of the 31st Symposium on the Foundations of Comp. Sci.*, pages 392–396. IEEE Computer Society Press, Los Alamitos, CA, 1990.

[17] J. Alspector and R. B. Allen. Performance of a stochastic learning chip. In *Proc. 1988 Neural Information Processing Conference*, pages 748–760, 1989.

[18] K. Amakawa. Self-learning reaching motion of a multi-joint arm using a trial-and-error heuristic and a neural network. Master's thesis, UC Santa Cruz, 1991.

[19] S. Amari, N. Fujita, and S. Shinomoto. Four types of learning curves. *Neural Computation*, 4(4):605–618, 1992.

[20] F. Ameur, P. Fischer, K. U. Höffgen, and F. M. auf der Heide. Trial and error: a new approach to space-bounded learning. In *Computational Learning Theory: Eurocolt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 133–144, Oxford, 1994. Oxford University Press.

[21] J. Amsterdam. Extending the Valiant learning model. In *Proceedings of the Fifth International Conference on Machine Learning*, pages 381–394, San Mateo, CA, 1988. Morgan Kaufmann.

[22] J. Amsterdam. Some philosophical problems with formal learning theory. In *Proceedings AAAI-88*, pages 580–584, Saint Paul, Minn., Aug. 1988. American Association for Artificial Intelligence.

[23] J. Amsterdam. The valiant learning model: Extensions and assessment. Master's thesis, MIT Department of Electrical Engineering and Computer Science, Jan. 1988.

[24] J. A. Anderson and E. Rosenfeld, editors. *Neurocomputing: Foundations of Research.* MIT Press, 1988.

[25] P. M. Andreae. Justified generalization: Acquiring procedures from examples. Technical Report AI-TR-834, MIT Artificial Intelligence Laboratory, Jan. 1985.

[26] D. Angluin. On the complexity of minimum inference of regular sets. *Inform. Control*, 39(3):337–350, 1978.

[27] D. Angluin. Finding patterns common to a set of strings. *J. of Comput. Syst. Sci.*, 21:46–62, 1980.

[28] D. Angluin. Inductive inference of formal languages from positive data. *Inform. Control*, 45(2):117–135, May 1980.

[29] D. Angluin. Inference of reversible languages. *J. ACM*, 29(3):741–765, July 1982.

[30] D. Angluin. A note on the number of queries needed to identify regular languages. *Inform. Control*, 51:76–87, 1982.

[31] D. Angluin. Types of queries for concept learning. Technical Report YALEU/DCS/TR-479, Yale University Department of Computer Science, June 1986.

[32] D. Angluin. Debugging a DAG efficiently. Technical Report 591, Yale University, 1987.

[33] D. Angluin. Learning k-bounded context-free grammars. Technical Report YALEU/DCS/RR-557, Department of Computer Science, Yale University, Aug. 1987.

[34] D. Angluin. Learning k-term DNF formulas using queries and counterexamples. Technical Report YALEU/DCS/RR-559, Department of Computer Science, Yale University, Aug. 1987.

[35] D. Angluin. Learning propositional Horn sentences with hints. Technical report, Yale University, YALE/DCS/RR-590, 1987.

[36] D. Angluin. Learning regular sets from queries and counterexamples. *Inform. Comput.*, 75(2):87–106, Nov. 1987.

[37] D. Angluin. Learning Propositional Horn sentences with hints. Technical Report YALEU/DCS/RR-590, Department of Computer Science, Yale University, Dec. 1987.

[38] D. Angluin. Equivalence queries and DNF formulas. Technical Report YALEU/DCS/RR-659, Yale University Department of Computer Science, Nov. 1988.

[39] D. Angluin. Identifying languages from stochastic examples. Technical Report YALEU/DCS/RR-614, Yale University Dept. of Computer Science, New Haven, CT, 1988.

[40] D. Angluin. Learning with hints. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 167–181, San Mateo, CA, 1988. Morgan Kaufmann.

[41] D. Angluin. Queries and concept learning. *Machine Learning*, 2(4):319–342, Apr. 1988.

[42] D. Angluin. Requests for hints that return no hints. Technical report, Yale University, YALE/DCS/RR-647, 1988.

[43] D. Angluin. Equivalence queries and approximate fingerprints. In *Proc. 2nd Annu. Workshop on Comput. Learning Theory*, pages 134–145, San Mateo, CA, 1989. Morgan Kaufmann.

[44] D. Angluin. Using queries to identify μ-formulas. Technical report, Yale University, YALE/DCS/RR-694, 1989.

[45] D. Angluin. Negative results for equivalence queries. *Machine Learning*, 5:121–150, 1990.

[46] D. Angluin. Computational learning theory: survey and selected bibliography. In *Proc. 24th Annu. ACM Sympos. Theory Comput.*, pages 351–369. ACM Press, New York, NY, 1992.

[47] D. Angluin, L. Birnbaum, J. Feldman, R. Rivest, and L. Valiant. Summary of the panel discussion. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 398–401, San Mateo, CA, 1988. Morgan Kaufmann.

[48] D. Angluin, M. Frazier, and L. Pitt. Learning conjunctions of Horn clauses. *Machine Learning*, 9:147–164, 1992.

[49] D. Angluin, W. Gasarch, and C. Smith. Training sequences. *Theoret. Comput. Sci.*, 66:255–272, 1989.

[50] D. Angluin, L. Hellerstein, and M. Karpinski. Learning read-once formulas with queries. *J. ACM*, 40:185–210, 1993.

[51] D. Angluin and M. Kharitonov. When won't membership queries help? In *Proc. of the 23rd Symposium on Theory of Computing*, pages 444–454. ACM Press, New York, NY, 1991.

[52] D. Angluin and M. Kriķis. Learning with malicious membership queries and exceptions. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 57–66. ACM Press, New York, NY, 1994.

[53] D. Angluin and P. Laird. Learning from noisy examples. *Machine Learning*, 2(4):343–370, 1988.

[54] D. Angluin and D. K. Slonim. Learning monotone DNF with an incomplete membership oracle. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 139–146, San Mateo, CA, 1991. Morgan Kaufmann.

[55] D. Angluin and C. Smith. Inductive inference. In *Encyclopedia of Artificial Intelligence*, pages 409–418. J. Wiley and Sons, New York, 1987.

[56] D. Angluin and C. H. Smith. A survey of inductive inference: Theory and methods. *ACM Comput. Surv.*, 15(3):237–269, Sept. 1983.

[57] S. Anoulova, P. Fischer, S. Pölt, and H. U. Simon. PAB-decisions for Boolean and real-valued features. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 353–362. ACM Press, New York, NY, 1992.

[58] S. Anoulova and S. Pölt. Using kullback-leibler divergence in learning theory. In *Computational Learning Theory: Eurocolt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 145–156, Oxford, 1994. Oxford University Press.

[59] M. Anthony and N. Biggs. *Computational Learning Theory.* Cambridge Tracts in Theoretical Computer Science (30). Cambridge University Press, 1992.

[60] M. Anthony, N. Biggs, and J. Shawe-Taylor. The learnability of formal concepts. In *Proc. 3rd Annu. Workshop on Comput. Learning Theory*, pages 246–257, San Mateo, CA, 1990. Morgan Kaufmann.

[61] M. Anthony, G. Brightwell, D. Cohen, and J. Shawe-Taylor. On exact specification by examples. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 311–318. ACM Press, New York, NY, 1992.

[62] M. Anthony and S. Holden. On the power of polynomial discriminators and radial basis function networks. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 158–164. ACM Press, New York, NY, 1993.

[63] M. Anthony and J. Shawe-Taylor. A result of Vapnik with applications. Technical Report CSD-TR-628, University of London, Surrey, England, 1990.

[64] M. Anthony and J. Shawe-Taylor. Valid generalisation of functions from close approximations on a sample. In *Computational Learning Theory: Eurocolt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 94–108, Oxford, 1994. Oxford University Press.

[65] S. Ar, R. J. Lipton, R. Rubinfeld, and M. Sudan. Reconstructing algebraic functions from mixed data. In *Proc. of the 33rd Symposium on the Foundations of Comp. Sci.*, pages 503–512. IEEE Computer Society Press, Los Alamitos, CA, 1992.

[66] A. Arapostathis and S. I. Marcus. Analysis of an identification algorithm arising in the adaptive estimation of markov chains. *Math. Control Signals Systems*, 3:1–29, 1990.

[67] S. Arikawa, S. Goto, S. Ohsuga, and T. Yokomori, editors. *Proceedings of the First International Workshop on Algorithmic Learning Theory.* Japanese Society for Artificial Intelligence, Tokyo, 1990.

[68] S. Arikawa, A. Maruoka, and T. Sato, editors. *Proceedings of the Second International Workshop on Algorithmic Learning Theory.* Japanese Society for Artificial Intelligence, Tokyo, 1991.

[69] S. Arikawa, T. Shinohara, and A. Yamamoto. Elementary formal system as a unifying framework for language learning. In *Proc. 2nd Annu. Workshop on Comput. Learning Theory*, pages 312–327, San Mateo, CA, 1989. Morgan Kaufmann.

[70] H. Arimura, H. Ishizaka, and T. Shinohara. Polynomial time inference of a subclass of context-free transformations. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 136–143. ACM Press, New York, NY, 1992.

[71] J. A. Aslam. Inferring graphs from walks. Master's thesis, Massachusetts Institute of Technology, Jan. 1992.

[72] J. A. Aslam and S. E. Decatur. General bounds on statistical query learning and PAC learning with noise via hypothesis boosting. In *Proc. 35th Annu. IEEE Sympos. Found. Comput. Sci.*, Nov. 1993.

[73] J. A. Aslam and A. Dhagat. Searching in the presence of linearly bounded errors. In *Proc. of the 23rd Symposium on Theory of Computing*, pages 486–493. ACM Press, New York, NY, 1991.

[74] J. A. Aslam and R. L. Rivest. Inferring graphs from walks. In *Proc. 3rd Annu. Workshop on Comput. Learning Theory*, pages 359–370, San Mateo, CA, 1990. Morgan Kaufmann.

[75] P. Assouad. Densité et dimension. *Ann. Inst. Fourier*, pages 233–282, 1983.

[76] C. G. Atkeson. Learning arm kinematics and dynamics. In *Annual Review of Neuroscience*, volume 12, pages 157–183. Annual Reviews, 1989.

[77] P. Auer. On-line learning of rectangles in noisy environments. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 253–261. ACM Press, New York, NY, 1993.

[78] P. Auer, P. Long, W. Maass, and G. Woeginger. On the complexity of function learning. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 392–401. ACM Press, New York, NY, 1993.

[79] J. R. Bachrach. *Connectionist Modeling and Control of Finite-State Environments.* PhD thesis, University of Massachusetts, Feb. 1992.

[80] J. L. Balcázar, J. Díaz, R. Gavaldà, and O. Watanabe. An optimal parallel algorithm for learning DFA. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 208–217. ACM Press, New York, NY, 1994.

[81] G. Baliga and J. Case. Learnability: Admissible, co–finite and hypersimple sets. In *Proceedings of the 20th International Colloquium on Automata, Languages and Programming, Lund, Sweden*, 1993. July.

[82] G. Baliga, J. Case, and S. Jain. Language learning with some negative information. *J. of Comput. Syst. Sci.*, 1993. Accepted.

[83] G. Baliga, J. Case, S. Jain, and M. Suraj. Machine learning of higher order programs. *J. Symbolic Logic*, 1993. To Appear.

[84] G. Baliga, S. Jain, and A. Sharma. Learning from multiple sources of inaccurate data. In *Proceedings of the Third International Workshop on Analogical and Inductive Inference, Dagstuhl Castle, Germany*, pages 108–128, Oct. 1992.

[85] R. B. Banerji. The logic of learning: a basis for pattern recognition and for improvement of performance. *Advances in Computers*, 24:177–216, 1985.

[86] R. B. Banerji. Learning theories in a subset of a polyadic logic. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 267–278, San Mateo, CA, 1988. Morgan Kaufmann.

[87] I. Barland. Some ideas on learning with directional feedback. Master's thesis, UC Santa Cruz, June 1992.

[88] A. R. Barron. Approximation and estimation bounds for artificial neural networks. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 243–249, San Mateo, CA, 1991. Morgan Kaufmann.

[89] P. Bartlett. Lower bounds on the Vapnik-Chervonenkis dimension of multi-layer threshold networks. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 144–150. ACM Press, New York, NY, 1993.

[90] P. L. Bartlett. Learning with a slowly changing distribution. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 243–252. ACM Press, New York, NY, 1992.

[91] P. L. Bartlett, P. Fischer, and K. U. Höffgen. Exploiting random walks for learning. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 318–327. ACM Press, New York, NY, 1994.

[92] P. L. Bartlett, P. M. Long, and R. C. Williamson. Fat-shattering and the learnability of real-valued functions. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 299–310. ACM Press, New York, NY, 1994.

[93] P. L. Bartlett and R. C. Williamson. Investigating the distribution assumptions in the PAC learning model. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 24–32, San Mateo, CA, 1991. Morgan Kaufmann.

[94] J. M. Barzdin. On decoding automata in the absence of an upper bound on the number of states. *Dokl. Akad. Nauk. SSSR*, pages 1048–1051, 1970.

[95] J. M. Barzdin. Prognostication of automata and functions. *Information Processing*, 1:81–84, 1971.

[96] J. M. Barzdin. Prognostication of automata and functions. In *Information Processing 71*, pages 81–84. North-Holland, Amsterdam, 1972.

[97] J. M. Barzdin. Limiting synthesis of t numbers. *Latvijas Valsts Univ. Zimatm. Raksti*, 210:112–116, 1974.

[98] J. M. Barzdin. On synthesizing programs given by examples. In *Lecture Notes in Computer Science 5*, pages 53–63. Springer-Verlag, 1974.

[99] J. M. Barzdin. Two theorems on the limiting synthesis of functions. *Theory of Algorithms and Programs, Latvian State University, Riga*, 210:82–88, 1974. In Russian.

[100] J. M. Barzdin. Uber eine eigenschaft limitar berech eubarer funktionale. *Theory of Algorithms and Programs, Latvian State University, Riga*, 1:20–24, 1974.

[101] J. M. Barzdin. Inductive inference of automata, functions and programs. In *Amer. Math. Soc. Transl.*, pages 107–122, 1977.

[102] J. M. Barzdin and R. Freivalds. Prediction and limiting synthesis of recursively enumerable classes of functions. *Latvijas Valsts Univ. Zimatm. Raksti*, 210:101–111, 1974.

[103] J. M. Barzdin and R. V. Frievald. On the prediction of general recursive functions. *Soviet Math. Doklady*, 13:1224–1228, 1972.

[104] J. M. Barzdin, E. B. Kinber, and K. M. Podnieks. Concerning synthesis and prediction of functions. *Theory of Algorithms and Programs, Latvian State University, Riga*, pages 117–128, 1974.

[105] J. M. Barzdin and K. M. Podnieks. The theory of inductive inference. In *Proceedings of the Mathematical Foundations of Computer Science*, pages 9–15. Math. Inst. of the Slovak Academy of Sciences, 1973.

[106] K. Basye, T. Dean, and J. Vitter. Coping with uncertainty in map learning. In *Proceedings of IJCAI 11*, pages 663–668. IJCAI, 1989.

[107] R. Battiti. Accelerated backpropagation learning: Two optimization methods. *Complex Systems*, 3(4):331–342, Aug. 1989.

[108] E. Baum. A polynomial time algorithm that learns two hidden net units. *Neural Computation*, 2:510–522, 1990.

[109] E. Baum. Neural net algorithms that learn in polynomial time from examples and queries. *IEEE Transactions on Neural Networks*, 2:5–19, 1991.

[110] E. Baum and D. Haussler. What size net gives valid generalization? *Neural Computation*, 1(1):151–160, 1989.

[111] E. B. Baum. On learning a union of half spaces. *Journal of Complexity*, 6(1):67–101, Mar. 1990.

[112] E. B. Baum. The perceptron algorithm is fast for nonmalicious distributions. *Neural Computation*, 2:248–260, 1990.

[113] E. B. Baum. Polynomial time algorithms for learning neural nets. In *Proc. 3rd Annu. Workshop on Comput. Learning Theory*, pages 258–272, San Mateo, CA, 1990. Morgan Kaufmann.

[114] M. F. Bear, L. N. Cooper, and F. F. Ebner. A physiological basis for a theory of synapse modification. *Science*, 237:42–48, July 3 1987.

[115] D. Beaver. Polynomially sized boolean circuits are not learnable. Technical Report TR-13-87, Aiken Computation Laboratory, Harvard University, Dec. 1987.

[116] H. R. Beick. Zur konvergenzgeschwindigkeit von strategien der induktiven inferenz. *Elektronische Informationsverarbeitung und Kybernetik*, 18:163–172, 1982.

[117] M. Bellare. A technique for upper bounding the spectral norm with applications to learning. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 62–70. ACM Press, New York, NY, 1992.

[118] S. Ben-David. Can finite samples detect singularities of real-valued functions? In *Proc. 24th Annu. ACM Sympos. Theory Comput.*, pages 390–399. ACM Press, New York, NY, 1992.

[119] S. Ben-David and G. M. Benedek. Measurability constraints on PAC learnability. Technical report, Technion, Haifa, 1991.

[120] S. Ben-David, G. M. Benedek, and Y. Mansour. A parametrization scheme for classifying models of learnability. In *Proc. 2nd Annu. Workshop on Comput. Learning Theory*, pages 285–302, San Mateo, CA, 1989. Morgan Kaufmann.

[121] S. Ben-David, N. Cesa-Bianchi, and P. M. Long. Characterizations of learnability for classes of $\{0, ..., n\}$-valued functions. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 333–340. ACM Press, New York, NY, 1992.

[122] S. Ben-David and E. Dichterman. Learning with restricted focus of attention. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 287–296. ACM Press, New York, NY, 1993.

[123] S. Ben-David, A. Itai, and E. Kushilevitz. Learning by distances. In *Proc. 3rd Annu. Workshop on Comput. Learning Theory*, pages 232–245, San Mateo, CA, 1990. Morgan Kaufmann.

[124] S. Ben-David and M. Jacovi. On learning in the limit and non-uniform (ϵ, δ)- learning. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 209–217. ACM Press, New York, NY, 1993.

[125] S. Ben-David and M. Lindenbaum. Localization vs. identification of semi-algebraic sets. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 327–336. ACM Press, New York, NY, 1993.

[126] G. Benedek and A. Itai. Learnability with respect to fixed distributions. *Theoret. Comput. Sci.*, 86(2):377–389, 1991.

[127] G. M. Benedek and A. Itai. Learnability by fixed distributions. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 80–90, San Mateo, CA, 1988. Morgan Kaufmann.

[128] G. M. Benedek and A. Itai. Nonuniform learnability. In *ICALP*, pages 82–92, July 1988.

[129] G. M. Benedek and A. Itai. Dominating distributions and learnability. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 253–264. ACM Press, New York, NY, 1992.

[130] F. Bergadano and L. Saitta. On the error probabilty of boolean concept descriptions. In *Proceedings of the 1989 European Working Session on Learning*, pages 25–35, 1989.

[131] B. Berger, J. Rompel, and P. W. Shor. Efficient NC algorithms for set cover with applications to learning and geometry. In *Proc. 30th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 54–59. IEEE Computer Society Press, Los Alamitos, CA, 1989.

[132] J. Berger and C. Pair. Inference for regular bilanguages. *J. of Comput. Syst. Sci.*, 16:100–122, 1978.

[133] U. Berggren. Linear time deterministic learning of k-term DNF. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 37–40. ACM Press, New York, NY, 1993.

[134] P. Berman and R. Roos. Learning one-counter languages in polynomial time. In *Proc. 28th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 61–67. IEEE Computer Society Press, Los Alamitos, CA, 1987.

[135] E. J. Bernstein. Absolute error bounds for learning linear functions on line. *Proc. 6th Annu. Workshop on Comput. Learning Theory*, 1992.

[136] A. Bertoni, P. Campadelli, A. Morpurgo, and S. Panizza. Polynomial uniform convergence and polynomial-sample learnability. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 265–271. ACM Press, New York, NY, 1992.

[137] R. Berwick. Learning from positive-only examples. In *Machine Learning, Vol. II*, pages 625–645. Morgan Kaufmann, 1986.

[138] R. C. Berwick. Language learning by automata induction. *Machine Learning*, 2:9–38, July 3 1987.

[139] M. Betke. Algorithms for exploring an unknown graph. Master's thesis, MIT Department of Electrical Engineering and Computer Science, Feb. 1992. (Published as MIT Laboratory for Computer Science Technical Report MIT/LCS/TR-536 (March, 1992)).

[140] M. Betke, R. Rivest, and M. Singh. Piecemeal learning of an unknown environment. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 277–286. ACM Press, New York, NY, 1993.

[141] A. W. Biermann. On the inference of turing machines from sample computations. *Art. Int.*, 3:181–198, 1972.

[142] A. W. Biermann. The inference of regular LISP programs from examples. *IEEE Trans. on Systems, Man, and Cybernetics*, SMC-8:585–600, 1978.

[143] A. W. Biermann and J. Feldman. On the synthesis of finite-state machines from samples of their behavior. *IEEE Trans. Computers*, 21:592–597, 1972.

[144] A. W. Biermann and J. A. Feldman. A survey of results in grammatical inference. In *Frontiers of Pattern Recognition*. Academic Press, N. Y., 1972.

[145] A. W. Biermann and R. Krishnaswamy. Constructing programs from example computations. *IEEE Trans. on Software Engineering*, SE-2:141–153, 1976.

[146] A. W. Biermann and D. R. Smith. The hierarchical synthesis of LISP scanning functions. In *Information Processing 77*, pages 41–45. North-Holland, Amsterdam, 1977.

[147] A. W. B. R. I. Biermann and F. E. Petry. Speeding up the synthesis of programs from traces. *IEEE Trans. on Computers*, C-24:122–136, 1975.

[148] A. Blum. On the computational complexity of training simple neural networks. Master's thesis, MIT Department of Electrical Engineering and Computer Science, May 1989. (Published as Laboratory for Computer Science Technical Report MIT/LCS/TR-445 (May, 1989)).

[149] A. Blum. Separating distribution-free and mistake-bound learning models over the Boolean domain. In *Proc. of the 31st Symposium on the Foundations of Comp. Sci.*, pages 211–218. IEEE Computer Society Press, Los Alamitos, CA, 1990.

[150] A. Blum. Separating PAC and mistake-bound learning models over the boolean domain. *Proc. 31th Annu. IEEE Sympos. Found. Comput. Sci.*, 1990.

[151] A. Blum. Learning boolean functions in an infinite attribute space. *Machine Learning*, 9(4):373–386, Oct. 1992.

[152] A. Blum and P. Chalasani. Learning switching concepts. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 231–242. ACM Press, New York, NY, 1992.

[153] A. Blum and P. Chalasani. An on-line algorithm for improving performance in navigation. In *Proc. 34th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 2–11. IEEE Computer Society Press, Los Alamitos, CA, 1993.

[154] A. Blum, P. Chalasani, and J. Jackson. On learning embedded symmetric concepts. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 337–346. ACM Press, New York, NY, 1993.

[155] A. Blum, L. Hellerstein, and N. Littlestone. Learning in the presence of finitely or infinitely many irrelevant attributes. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 157–166, San Mateo, CA, 1991. Morgan Kaufmann.

[156] A. Blum, R. Khardon, E. Kushilevitz, L. Pitt, and D. Roth. On learning read-k-satisfy-j DNF. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 110–117. ACM Press, New York, NY, 1994.

[157] A. Blum, P. Raghavan, and B. Schieber. Navigating in unfamiliar geometric terrain. In *Proc. 23th Annu. ACM Sympos. Theory Comput.*, pages 494–504. ACM, 1991.

[158] A. Blum and R. L. Rivest. Training a 3-node neural net is NP-Complete. In *Advances in Neural Information Processing Systems I*, pages 494–501. Morgan Kaufmann, 1989.

[159] A. Blum and M. Singh. Learning functions of k terms. In *Proc. 3rd Annu. Workshop on Comput. Learning Theory*, pages 144–153, San Mateo, CA, 1990. Morgan Kaufmann.

[160] L. Blum and M. Blum. Toward a mathematical theory of inductive inference. *Inform. Control*, 28(2):125–155, June 1975.

[161] A. Blumer, A. Ehrenfeucht, D. Haussler, and M. K. Warmuth. Occam's razor. *Inform. Proc. Lett.*, 24:377–380, Apr. 1987.

[162] A. Blumer, A. Ehrenfeucht, D. Haussler, and M. K. Warmuth. Learnability and the Vapnik-Chervonenkis dimension. *J. ACM*, 36(4):929–965, 1989.

[163] A. Blumer and N. Littlestone. Learning faster than promised by the Vapnik-Chervonenkis dimension. *Discrete Applied Mathematics*, 24:47–53, 1989.

[164] R. Board and L. Pitt. On the necessity of Occam algorithms. *Theoret. Comput. Sci.*, 100:157–184, 1992.

[165] R. A. Board and L. Pitt. Semi-supervised learning. *Machine Learning*, 4(1):41–65, 1989.

[166] D. Boneh and R. Lipton. Amplification of weak learning under the uniform distribution. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 347–351. ACM Press, New York, NY, 1993.

[167] B. E. Boser, I. M. Guyon, and V. N. Vapnik. A training algorithm for optimal margin classifiers. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 144–152. ACM Press, New York, NY, 1992.

[168] S. Boucheron and J. Sallantin. Some remarks about space-complexity of learning, and circuit complexity of recognizing. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 125–138, San Mateo, CA, 1988. Morgan Kaufmann.

[169] Y. Brandman, J. Hennessy, and A. Orlitsky. A spectral lower bound technique for the size of decision trees and two level circuits. *IEEE Trans. Computers*, 39(2):282–287, 1990.

[170] J. M. Brayer and K. S. Fu. A note on the k-tail method of tree grammar inference. *IEEE Trans. on Systems, Man, and Cybernetics*, SMC-7:293–300, 1977.

[171] A. Brāzma. Efficient identification of regular expressions from representative examples. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 236–242. ACM Press, New York, NY, 1993.

[172] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone. *Classification and Regression Trees.* Wadsworth International Group, 1984.

[173] R. Brooks. The role of learning in autonomous robots. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 5–10, San Mateo, CA, 1991. Morgan Kaufmann.

[174] D. Bshouty and N. H. Bshouty. On learning arithmetic read-once formulas with exponentiation. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 311–317. ACM Press, New York, NY, 1994.

[175] N. Bshouty, S. Goldman, T. Hancock, and S. Matar. Asking questions to minimize errors. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 41–50. ACM Press, New York, NY, 1993.

[176] N. Bshouty, L. Hellerstein, T. Hancock, and M. Karpinski. Read-once threshold formulas, justifying assignments, and transformations. *Journal of Complexity Theory.* To appear.

[177] N. H. Bshouty and R. Cleve. On the exact learning of formulas in parallel. In *Proc. of the 33rd Symposium on the Foundations of Comp. Sci.*, pages 513–522. IEEE Computer Society Press, Los Alamitos, CA, 1992.

[178] N. H. Bshouty, R. Cleve, S. Kannan, and C. Tamon. Oracles and queries that are sufficient for exact learning. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 130–139. ACM Press, New York, NY, 1994.

[179] N. H. Bshouty, T. R. Hancock, and L. Hellerstein. Learning arithmetic read-once formulas. In *Proc. 24th Annu. ACM Sympos. Theory Comput.*, pages 370–381. ACM Press, New York, NY, 1992.

[180] N. H. Bshouty, T. R. Hancock, and L. Hellerstein. Learning Boolean read-once formulas with arbitrary symmetric and constant fan-in gates. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 1–15. ACM Press, New York, NY, 1992.

[181] K. Buescher and P. R. Kumar. Simultaneous learning of concepts and simultaneous estimation of probabilities. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 33–42, San Mateo, CA, 1991. Morgan Kaufmann.

[182] K. L. Buescher and P. R. Kumar. Learning stochastic functions by smooth simultaneous estimation. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 272–279. ACM Press, New York, NY, 1992.

[183] W. J. Bultman and W. Maass. Fast identification of geometric objects with membership queries. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 337–353, San Mateo, CA, 1991. Morgan Kaufmann.

[184] A. Bundy and B. Silver. A critical survey of rule learning programs. Technical report, Edinburgh Univ. Dept. of AI, No. 169, 1981.

[185] W. Buntine. A critique of the Valiant model. In *Proceedings of the 11th International Joint Conference on Artificial Intelligence*, pages 837–842, 1989.

[186] W. L. Buntine. *A Theory of Learning Classification Rules.* PhD thesis, University of Technology, Sydney, 1990.

[187] A. Burago. Learning structurally reversible context-free grammars from queries and counterexamples in polynomial time. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 140–146. ACM Press, New York, NY, 1994.

[188] T. Bylander. Polynomial learnability of linear threshold approximations. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 297–302. ACM Press, New York, NY, 1993.

[189] T. Bylander. Learning linear threshold functions in the presence of classification noise. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 340–347. ACM Press, New York, NY, 1994.

[190] J. G. Carbonell and Y. Gil. Learning by experimentation. In *Proceedings of the Fourth International Workshop on Machine Learning*, pages 256–266. Morgan Kaufmann, June 1987.

[191] J. Case. Periodicity in generations of automata. *Math. Syst. Theory*, 8:15–32, 1974.

[192] J. Case. Problems in inductive inference. *Bulletin of the European Association for Theoretical Computer Science*, 18:117–118, 1982.

[193] J. Case. Learning machines. In *Language Learning and Concept Acquisition.* Ablex Publishing Company, 1986.

[194] J. Case. The power of vacillation. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 196–205, San Mateo, CA, 1988. Morgan Kaufmann.

[195] J. Case. Infinitary self-reference in learning theory. *J. of Experimental and Theoret. Artif. Intell.*, 4:281–293, 1993. To appear in the special issue from AII 92.

[196] J. Case. Infinitary self reference in learning theory. Technical Report TR-93-15, University of Delaware, Jan. 1993.

[197] J. Case, K. J. Chen, and S. Jain. Strong separation of learning classes. *J. of Experimental and Theoret. Artif. Intell.*, 4:281–293, 1992.

[198] J. Case and H. Chi. Machine learning of nearly minimal size grammars. Unpublished Manuscript, 1986.

[199] J. Case and M. Fulk. Solution to inductive inference problem P3. *Bulletin of the European Association for Theoretical Computer Science*, 19:8, 1983.

[200] J. Case, S. Jain, and S. N. Manguelle. Refinements of inductive inference by popperian and reliable machines. *Kybernetika*, 1994. To appear.

[201] J. Case, S. Jain, and A. Sharma. Anomalous learning helps succinctness. In *Proceedings of the First Workshop on Algorithmic Learning Theory, Tokyo, Japan*, pages 282–288. Japanese Society for Artificial Intelligence, Oct. 1990.

[202] J. Case, S. Jain, and A. Sharma. Convergence to nearly minimal size grammars by vacillating learning machines. Technical Report 90-10, University of Delaware, Newark, Delaware, 1990.

[203] J. Case, S. Jain, and A. Sharma. On learning limiting programs. *Internat. J. of Found. of Comput. Sci.*, 3(1):93–115, 1992.

[204] J. Case, S. Jain, and A. Sharma. Complexity issues for vacillatory function identification. *Inform. Comput.*, 1993. To appear.

[205] J. Case, S. Jain, and A. Sharma. Vacillatory learning of nearly minimal size grammers. *J. of Comput. Syst. Sci.*, 1993. To Appear.

[206] J. Case and C. Lynes. Inductive inference and language identification. In *Proc. ICALP 82*, pages 107–115. Springer Verlag, 1982.

[207] J. Case and C. Lynes. Machine inductive inference and language identification. In *Proceedings of the 9th International Colloquium on Automata, Languages and Programming*, volume 140, pages 107–115. Springer-Verlag, Berlin, 1982.

[208] J. Case, S. Mandayam, and S. Jain. Not-so-nearly-minimal-size program inference. In *Proceedings of the Fourth Workshop on Algorithmic Learning Theory, Tokyo, Japan.* Japanese Society for Artificial Intelligence, Oct. 1993. To Appear.

[209] J. Case and S. N. Manguelle. Refinements of inductive inference by Popperian machines. Technical Report 152, SUNY/Buffalo, 1979.

[210] J. Case and S. Ngo-Manguelle. Refinements of inductive inference by popperian machines. Technical report, SUNY Buffalo, Dept. of Computer Science, 1979.

[211] J. Case and C. Smith. Comparison of identification criteria for machine inductive inference. *Theoret. Comput. Sci.*, 25:193–220, 1983.

[212] N. Cesa-Bianchi. Learning the distribution in the extended PAC model. In *Proceedings of the First International Workshop on Algorithmic Learning Theory*, pages 236–246. Japanese Society for Artificial Intelligence, 1990.

[213] N. Cesa-Bianchi, Y. Freund, D. P. Helmbold, D. Haussler, R. E. Schapire, and M. K. Warmuth. How to use expert advice. In *Proc. 25th Annu. ACM Sympos. Theory Comput.*, pages 382–391. ACM Press, New York, NY, 1993.

[214] N. Cesa-Bianchi, Y. Freund, D. P. Helmbold, and M. Warmuth. On-line prediction and conversion strategies. In *Computational Learning Theory: Eurocolt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 205–216, Oxford, 1994. Oxford University Press.

[215] N. Cesa-Bianchi, P. Long, and M. Warmuth. Worst-case quadratic loss bounds for a generalization of the Widrow-Hoff rule. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 429–438. ACM Press, New York, NY, 1993.

[216] I. K. Cestnik, B and I. Bratko. Assistant 86: A knowledge-elicitation tool for sophisticated users. In *Progress in Machine Learning–Proceedings of EWSL 87: 2nd European Working Session on Learning*, pages 31–45, Bled, Yogoslavia, May 1987.

[217] P. C. Cheeseman. A method of computing generalized bayesian probability values for expert systems. In *Proceedings Eighth International Conference on Artificial Intelligence (Karlruhe, West Germany)*, pages 198–202, Aug. 1983.

[218] P. C. Cheeseman. Learning of expert systems from data. In *Proceedings of the Workshop on Principles of Knowledge-Based Systems*, pages 115–122, Dec. 1984.

[219] K. Chen. *Tradeoffs in Machine Inductive Inference.* PhD thesis, SUNY at Buffalo, 1981.

[220] K. J. Chen. Tradeoffs in the inductive inference of nearly minimal size programs. *Inform. Control*, 52:68–86, 1982.

[221] Z. Chen. Learning unions of two rectangles in the plane with equivalence queries. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 243–252. ACM Press, New York, NY, 1993.

[222] Z. Chen and S. Homer. On learning counting fuctions with queries. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 218–227. ACM Press, New York, NY, 1994.

[223] Z. Chen and W. Maass. On-line learning of rectangles. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 16–28. ACM Press, New York, NY, 1992.

[224] J. Cherniavsky and C. Smith. Using telltales in developing program test sets. Technical report, Georgetown University, Dept. of Computer Science, TR-4, 1986.

[225] J. Cherniavsky and R. Statman. Testing and inductive inference: abstract approaches. Preprint, Georgetown University, 1987.

[226] J. C. Cherniavsky, M. Velauthapillai, and R. Statman. Inductive inference: an abstract approach. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 251–266, San Mateo, CA, 1988. Morgan Kaufmann.

[227] H. Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Annals of Math. Stat.*, 23:493–509, 1952.

[228] P. Cholak, R. Downey, L. Fortnow, W. Gasarch, E. Kinber, M. Kummer, S. Kurtz, and T. Slaman. Degrees of inferability. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 180–192. ACM Press, New York, NY, 1992.

[229] T. H. Chung. Approximate methods for sequential decision making using expert advice. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 183–189. ACM Press, New York, NY, 1994.

[230] P. Clark and T. Niblett. The CN2 induction algorithm. , 1987.

[231] P. Clark and T. Niblett. Induction in noisy domains. In *Progress in Machine Learning–Proceedings of EWSL 87: 2nd European Working Session on Learning*, pages 11–30, Bled, Yogoslavia, May 1987.

[232] P. R. Cohen and E. A. Feigenbaum, editors. *The Handbook of Artificial Intelligence*, volume 3, chapter XIV: *Learning and Inductive Inference*, pages 324–511. William Kaufman, Inc., Los Altos, California, 1968.

[233] W. W. Cohen and H. Hirsh. Learnability of description logics. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 116–127. ACM Press, New York, NY, 1992.

[234] C. M. Cook, A. Rosenfeld, and A. R. Aronson. Grammatical inference by hill-climbing. *Inform. Sci.*, 10:59–80, 1976.

[235] D. Coulon and D. Kayser. Construction of natural language sentence acceptors by a supervised learning technique. *IEEE Trans. Pattern Anal. Mach. Intell.*, PAMI-1:94–99, 1979.

[236] T. Cover. Behavior of sequential predictors of binary sequences. In *Proceedings of the 4th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes*, pages 263–272. Publishing House of the Czechoslovak Academy of Sciences, 1965.

[237] T. Cover and J. Thomas. *Elements of Information Theory.* Wiley, 1991.

[238] T. M. Cover. Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE Trans. on Electronic Computers*, EC-14:326–334, 1965.

[239] T. M. Cover. Learning in pattern recognition. In *Methodologies of Pattern Recognition*, pages 111–132. Academic Press, 1969.

[240] T. M. Cover and P. E. Hart. Nearest neighbor pattern classification. *IEEE Transactions in Information Theory*, IT-13(1):21–27, Jan. 1967.

[241] T. M. Cover and T. J. Wagner. Topics in statistical pattern recognition. In *Communication and Cybernetics: Digital Pattern Recognition*, volume 10, chapter 2. Springer-Verlag, 1976.

[242] S. Crespi-Reghizzi. An effective model for grammar inference. In *Information Processing 71*, pages 524–529. North-Holland Publishing Co., 1972.

[243] S. Crespi-Reghizzi and D. Mandrioli. Inferring grammars by means of profiles: a unifying view. Internal report, Istituto di Elettrotechnica ed Elettronica del Politechnico di Milano, 1980.

[244] M. M. A. Crespi-Reghizzi, S and L. Lichten. The use of grammatical inference for designing programming languages. *Commun. ACM*, 16:83–90, 1973.

[245] G. Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of Control*, 2:303–314, 1989.

[246] B. Daley and D. Haussler, editors. *Proc. 5th Annu. Workshop on Comput. Learning Theory.* Morgan Kaufmann, Inc., San Mateo, CA, 1992.

[247] R. Daley. On the inference of optimal descriptions. *Theoret. Comput. Sci.*, 4:301–319, 1977.

[248] R. Daley. On the error correcting power of pluralism in BC-type inductive inference. *Theoret. Comput. Sci.*, 24:95–104, 1983.

[249] R. Daley. Towards the development of an analysis of learning algorithms. *Lecture Notes in Computer Science*, 265:1–18, 1986.

[250] R. Daley. Transformation of probabilistic learning strategies into deterministic learning strategies. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 220–226, San Mateo, CA, 1988. Morgan Kaufmann.

[251] R. Daley and B. Kalyanasundaram. Capabilities of probabilistic learners with bounded mind changes. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 182–191. ACM Press, New York, NY, 1993.

[252] R. Daley, B. Kalyanasundaram, and M. Velauthapillai. Breaking the probability 1/2 barrier in FIN-type learning. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 203–217. ACM Press, New York, NY, 1992.

[253] R. Daley, B. Kalyanasundaram, and M. Velauthapillai. Capabilities of fallible FINite learning. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 199–208. ACM Press, New York, NY, 1993.

[254] R. Daley, L. Pitt, M. Velauthapillia, and T. Will. Relations between probabilistic and team one-shot learners. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 228–239, San Mateo, CA, 1991. Morgan Kaufmann.

[255] R. Daley and C. Smith. On the complexity of inductive inference. *Inform. Control*, 69:12–40, 1986.

[256] R. P. Daley. Inductive inference hierarchies: Probabilistic vs pluralistic. *Lecture Notes in Computer Science*, 215:73–82, 1986.

[257] C. Darken, M. Donahue, L. Gurvits, and E. Sontag. Rate of approximation results motivated by robust neural network learning. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 303–309. ACM Press, New York, NY, 1993.

[258] B. DasGupta, H. T. Siegelmann, and E. Sontag. On a learnability question associated to neural networks with continuous activations. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 47–56. ACM Press, New York, NY, 1994.

[259] T. Dean, D. Angluin, K. Basye, S. Engelson, L. Kaelbling, E. Kokkevis, and O. Maron. Inferring finite automata with stochastic output functions and an application to map learning. In *Proceedings of AAAI-92*, pages 208–214. AAAI, 1992.

[260] S. E. Decatur. Statistical queries and faulty PAC oracles. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 262–268. ACM Press, New York, NY, 1993.

[261] R. Dechter. On the design of networks with hidden variables. Technical report, Technion, Haifa, July 1990.

[262] X. Deng, T. Kameda, and C. Papadimitriou. How to learn in an unknown environment. In *Proc. of the 32nd Symposium on the Foundations of Comp. Sci.*, pages 298–303. IEEE Computer Society Press, Los Alamitos, CA, 1991.

[263] X. Deng and C. H. Papadimitriou. Exploring an unknown graph. In *Proc. 31th Annu. IEEE Sympos. Found. Comput. Sci.*, volume I, pages 355–361, 1990.

[264] J. Denker, D. Schwartz, B. Wittner, S. Solla, R. Howard, L. Jackel, and J. Hopfield. Automatic learning, rule extraction and generalization. *Complex Systems*, 1:877–922, 1987.

[265] A. V. der Mude and A. Walker. On the inference of stochastic regular grammars. *Inform. Control*, 38:310–329, 1978.

[266] A. DeSantis, G. Markowsky, and M. N. Wegman. Learning probabilistic prediction functions. In *Proc. 29th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 110–119. IEEE Computer Society Press, Los Alamitos, CA, 1988.

[267] L. Devroye. Automatic pattern recognition: A study of the probability of error. *IEEE Trans. Pattern Anal. Mach. Intell.*, 10(4):530–543, July 1988.

[268] L. P. Devroye and T. J. Wagner. A distribution-free performance bound in error estimation. *IEEE Trans. Inform. Theory*, IT-22:586–587, 1976.

[269] L. P. Devroye and T. J. Wagner. Distribution-free performance bounds with the resubstitution error estimate. *IEEE Trans. Inform. Theory*, IT-25:208–210, 1979.

[270] T. G. Dietterich. Learning about systems that contain state variables. In *Proceedings of the National Conference on Artificial Intelligence*, pages 96–100, Aug. 1984.

[271] T. G. Dietterich and R. S. Michalski. Learning and generalization of characteristic descriptions: evaluation criteria and comparative review of selected methods. In *Proc. 6th IJCAI*, pages 223–231. IJCAI, 1979.

[272] L. R. C. K. Dietterich, T and R. Dromey. Learning and inductive inference. In *The Handbook of Artificial Intellegence*, volume 3, pages 323–512. William Kaufman, Inc., 1982.

[273] C. P. Dolan. *Tensor Manipulation Networks: Connectionist and Symbolic Approaches to Comprehension, Learning, and Planning.* PhD thesis, UCLA Computer Science Department, June 1989.

[274] P. G. Doucet. The syntactic inference problem for D0L sequences. In *L-systems*, pages 146–161. Springer-Verlag, 1974.

[275] R. Downey, P. Evans, and M. Fellows. Parameterized learning complexity. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 51–57. ACM Press, New York, NY, 1993.

[276] J. Doyle. On rationality and learning. Technical report, Carnegie Mellon University, 1988.

[277] G. L. Drescher. Suggestions for genetic A.I. Technical Report 198, MIT Artificial Intelligence Laboratory, Feb. 1980.

[278] G. L. Drescher. The schema mechanism: A conception of constructivist intelligence. Master's thesis, MIT Department of Electrical Engineering and Computer Science, Feb. 1985.

[279] G. L. Drescher. Genetic AI—translating Piaget into Lisp. Technical Report 890, MIT Artificial Intelligence Laboratory, Feb. 1986.

[280] G. L. Drescher. A mechanism for early Piagetian learning. In *Proceedings of AAAI-87: Sixth National Conference on Artificial Intelligence*, pages 290–294, Seattle, Washington, July 1987.

[281] G. L. Drescher. *Made-up Minds: A Constructivist Approach to Artificial Intelligence.* PhD thesis, MIT EECS Department, Sept. 1989. (Reprinted by MIT Press.).

[282] H. Drucker, R. Schapire, and P. Simard. Improving performance in neural networks using a boosting algorithm. In *Advances in Neural Information Processing Systems.* Morgan Kaufmann, 1992.

[283] R. O. Duda and P. E. Hart. *Pattern Classification and Scene Analysis.* Wiley, 1973.

[284] R. Dudley, S. Kulkarni, T. Richardson, and O. Zeituni. A metric entropy bound is not sufficient for learnability. Technical Report LIDS-P-2144, MIT LIDS, Oct. 1992.

[285] R. M. Dudley. Central limit theorems for empirical measures. *Annals of Probability*, 6(6):899–929, 1978.

[286] G. M. Edelman. *Neural Darwinism – The Theory of Neuronal Group Selection.* Basic Books, 1987.

[287] J. C. (Editor). *Machine Learning: Paradigms and Methods.* MIT Press, 1990.

[288] A. Ehrenfeucht and D. Haussler. Learning decision trees from random examples. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 182–194, San Mateo, CA, 1988. Morgan Kaufmann.

[289] A. Ehrenfeucht and D. Haussler. A general lower bound on the number of examples needed for learning. *Inform. Comput.*, 82(3):247–251, Sept. 1989.

[290] A. Ehrenfeucht, D. Haussler, M. Kearns, and L. Valiant. A general lower bound on the number of examples needed for learning. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 139–154, San Mateo, CA, 1988. Morgan Kaufmann.

[291] B. Eisenberg and R. L. Rivest. On the sample complextity of PAC-learning using random and chosen examples. In *Proc. 3rd Annu. Workshop on Comput. Learning Theory*, pages 154–162, San Mateo, CA, 1990. Morgan Kaufmann.

[292] B. B. Eisenberg. On the sample complexity of PAC-learning using random and chosen examples. Master's thesis, MIT EECS Department, 1992. (LCS Technical report MIT/LCS/TR-535, March 1992.).

[293] W. Evans, S. Rajagopalan, and U. Vazirani. Choosing a reliable hypothesis. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 269–276. ACM Press, New York, NY, 1993.

[294] V. Faber and J. Mycielski. Applications of learning theorems. *Fundamenta Informaticae*, 15(2):145–167, 1991.

[295] S. E. Fahlman and C. Lebiere. The cascade-correlation learning architecture. In *Advances in Neural Information Processing Systems 2*. Morgan Kaufmann, 1990.

[296] S. Fang and S. Venkatesh. On the average tractability of binary integer programming and the curious transition to perfect generalization in learning majority functions. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 310–316. ACM Press, New York, NY, 1993.

[297] J. D. Farmer and J. J. Sidorowich. Exploiting chaos to predict the future and reduce noise. Technical Report LA-UR-88-901, Los Alamos National Laboratory, Mar. 1988.

[298] L. Fass. Inference of skeletal automata. Technical report, Georgetown University Computer Science Dept. TR-2, 1984.

[299] U. Feige and A. Shamir. Learning in permutation groups (extended abstract). Preprint, Applied Mathematics Dept., The Weizmann Institute of Science., 1987.

[300] J. Feldman. Some decidability results on grammatical inference and complexity. *Inform. Control*, 20:244–262, 1972.

[301] J. A. Feldman. First thoughts on grammatical inference. Technical report, Stanford University Artificial Intelligence Memo 55, 1967.

[302] J. A. Feldman and P. Shields. Total complexity and the inference of best programs. *Math. Syst. Theory*, 10:181–191, 1977.

[303] J. A. G. J. H. J. J. Feldman and S. Reder. Grammatical complexity and inference. Technical report, Computer Science Dept., Stanford University, 1969.

[304] H. Feliciangeli and G. Herman. Algorithms for producing grammars from sample derivations: a common problem of formal language theory and developmental biology. *J. of Comput. Syst. Sci.*, 7:97–118, 1977.

[305] W. Feller. *An Introduction to Probability and its Applications*, volume 2. John Wiley and Sons, second edition, 1971.

[306] A. Fiat, S. Moses, A. Shamir, I. Shimshoni, and G. Tardos. Planning and learning in permutation groups. In *Proc. 30th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 274–279. IEEE Computer Society Press, Los Alamitos, CA, 1989.

[307] C. N. Fiechter. Efficient reinforcement learning. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 88–97. ACM Press, New York, NY, 1994.

[308] P. Fischer. Learning unions of convex polygons. In *Computational Learning Theory: Eurocolt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 61–67, Oxford, 1994. Oxford University Press.

[309] P. Fischer, S. Pölt, and H. U. Simon. Probably almost Bayes decisions. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 88–94, San Mateo, CA, 1991. Morgan Kaufman.

[310] P. Fischer and H. Simon. On learning ring-sum expansions. *SIAM J. Comput.*, 21:181–192, 1992.

[311] M. Flammini, A. Marchetti-Spaccamela, and L. K. Cera. Learning DNF formulae under classes of probability distributions. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 85–92. ACM Press, New York, NY, 1992.

[312] P. A. Flanagan. A discrete model of semantic learning. Technical report, Brown University Computer Science Dept., CS-79, 1981.

[313] M. J. Flannagan, L. S. Fried, and K. J. Holyoak. Distributional expectations and the induction of category structure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12(2):241–256, 1986.

[314] S. Floyd. Space-bounded learning and the Vapnik-Chervonenkis dimension. In *Proc. 2nd Annu. Workshop on Comput. Learning Theory*, pages 349–364, San Mateo, CA, 1989. Morgan Kaufmann.

[315] S. Floyd. *Space-bounded learning and the Vapnik-Chervonenkis Dimension (Ph.D)*. PhD thesis, U.C. Berkeley, Dec. 1989. ICSI Tech Report TR-89-061.

[316] L. Fortnow, W. Gasarch, S. Jain, E. Kinber, M. Kummer, S. Kurtz, M. Pleszkoch, T. Slaman, R. Solovay, and F. Stephan. Extremes in the degrees of inferability. *Annals of Pure and Applied Logic*, 66:21–276, 1994.

[317] M. Frazier, S. Goldman, N. Mishra, and L. Pitt. Learning from a consistently ignorant teacher. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 328–339. ACM Press, New York, NY, 1994.

[318] M. Frazier and L. Pitt. CLASSIC learning. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 23–34. ACM Press, New York, NY, 1994.

[319] M. Frazier and L. Pitt. Some new directions in computational learning theory. In *Computational Learning Theory: Eurocolt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 19–32, Oxford, 1994. Oxford University Press.

[320] E. Fredkin. Techniques using LISP for automatically discovering interesting relations in data. In *The Programming Language LISP*, pages 108–124. MIT Press, 1964.

[321] J. A. Freeman. Neural networks in mathematica. *AI Expert: the Magazine of Aritificial Intelligence in Practice*, 7(11):26–35, Nov. 1992.

[322] R. Freivalds. On the complexity and optimality of computation in the limit. *Theory of Algorithms and Programs, Latvian State University, Riga*, pages 155–173, 1974. (in Russian).

[323] R. Freivalds. Uniform and non uniform predictability. *Theory of Algorithms and Programs, Latvian State University, Riga*, 1:89–100, 1974.

[324] R. Freivalds. Functions computable in the limit by probabilistic machines. *Mathematical Foundations of Computer Science*, 1975.

[325] R. Freivalds. On the principle capabilities of probabilistic algorithms in inductive inference. *Semiotika Inform*, 12:137–140, 1979.

[326] R. Freivalds. Recursiveness of the enumerating functions increases the inferrability of recursively enumerable sets. *Bulletin of the European Association for Theoretical Computer Science*, 27:35–40, 1985.

[327] R. Freivalds. Inductive inference of minimal programs. In *Proc. 3rd Annu. Workshop on Comput. Learning Theory*, pages 3–20, San Mateo, CA, 1990. Morgan Kaufmann.

[328] R. Freivalds, M. Karpinski, and C. H. Smith. Co-learning of total recursive functions. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 190–197. ACM Press, New York, NY, 1994.

[329] R. Freivalds, E. Kinber, and C. Smith. On the impact of forgetting on learning machines. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 165–174. ACM Press, New York, NY, 1993.

[330] R. Freivalds and E. B. Kinber. Limit identification of minimal Gödel numbers. *Theory of Algorithms and Programs, Latvian State University, Riga*, 3:3–34, 1977.

[331] R. Freivalds, E. B. Kinber, and R. Wiehagen. Inductive inference and computable one-one numberings. *Zeitschr. f. math. Logik und Grundlagen d. Math. Bd.*, 28:463–479, 1982.

[332] R. Freivalds, E. B. Kinber, and R. Wiehagen. Connections between identifying functionals, standardizing operations, and computable numberings. *Zeitschr. f. math. Logik und Grundlagen d. Math. Bd.*, 30:145–164, 1984.

[333] R. Freivalds, E. B. Kinber, and R. Wiehagen. Probabilistic versus deterministic inductive inference in nonstandard numberings. *Zeitschr. f. math. Logik und Grundlagen d. Math. Bd.*, 34:531–539, 1988.

[334] R. Freivalds, E. B. Kinber, and R. Wiehagen. Inductive inference from good examples. In *Proceedings of the Second International Workshop on Analogical and Inductive Inference*, pages 1–17, 1989.

[335] R. Freivalds and C. H. Smith. On the role of procrastination for machine learning. *Inform. Comput.*, 107(2):237–271, 1993.

[336] R. Freivalds, C. H. Smith, and M. Velauthapillai. Trade-off among parameters affecting inductive inference. *Inform. Comput.*, 82:323–349, 1989.

[337] R. V. Freivalds. On the limit synthesis of numbers of general recursive functions in various computable numerations. *Soviet Math. Doklady*, 15:1681–1683, 1974.

[338] R. V. Freivalds. Minimal godel numbers and their identification in the limit. In *Lecture Notes in Computer Science 32*, pages 219–225. Springer-Verlag, 1975.

[339] R. V. Freivalds. Finite identification of general recursive functions by probabilistic strategies. In *Proceedings of the Conference on Algebraic, Arithmetic, and Categorial Methods in Computation Theory*, pages 138–145, 1979.

[340] R. V. Freivalds and R. Wiehagen. Inductive inference with additional information. *Elektronische Informationsverarbeitung und Kybernetik*, 15:179–185, 1979.

[341] R. V. K. E. B. Freivalds and R. Wiehagen. Inductive inference and computable one-one numberings. *Zeitschr. f. Logik und Grundlagen d. Math.*, 23:463–479, 1982.

[342] Y. Freund. Boosting a weak learning algorithm by majority. In *Proc. 3rd Annu. Workshop on Comput. Learning Theory*, pages 202–216, San Mateo, CA, 1990. Morgan Kaufmann.

[343] Y. Freund. An improved boosting algorithm and its implications on learning complexity. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 391–398. ACM Press, New York, NY, 1992.

[344] Y. Freund and D. Haussler. Unsupervised learning of distributions on binary vectors using two layer networks. Technical Report UCSC-CRL-91-20, Univ. of Calif. Computer Research Lab, Santa Cruz, CA, 1991.

[345] Y. Freund, M. Kearns, D. Ron, R. Rubinfeld, R. Schapire, and L. Sellie. Efficient learning of typical finite automata from random walks. In *Proc. 25th Annu. ACM Sympos. Theory Comput.*, pages 315–324. ACM Press, New York, NY, 1993.

[346] J. H. Friedman and W. Stuetzle. Projection pursuit regression. *J. American Statistical Association*, 76(376):817–823, Dec. 1981.

[347] K. Fu and T. L. Booth. Grammatical inference: Introduction and survey – part 1. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-5(1):95–111, Jan. 1975.

[348] K. Fu and T. L. Booth. Grammatical inference: Introduction and survey – part 2. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-5(4):409–423, July 1975.

[349] K. S. Fu. *Syntactic Methods in Pattern Recognition.* Academic Press, N. Y., 1975.

[350] K. S. Fu. *Syntactic Pattern Recognition, Applications.* Springer-Verlag, N. Y., 1977.

[351] K. S. Fu. *Syntactic Pattern Recognition and Applications.* Prentice-Hall, Englewood Cliffs, N. J., 1982.

[352] M. Fulk. *A Study of Inductive Inference machines.* PhD thesis, SUNY at Buffalo, 1985.

[353] M. Fulk. The disjunctive learning problem. *20th Asilomar Conference on Signals, Systems and Computers*, pages 343–345, 1986.

[354] M. Fulk. Saving the phenomenon: Requirements that inductive machines not contradict known data. *Inform. Comput.*, 79:193–209, 1988.

[355] M. Fulk. Regressiveness. unpublished, 1989.

[356] M. Fulk. Inductive inference with additional information. *J. of Comput. Syst. Sci.*, 1990. To appear.

[357] M. Fulk. Prudence and other conditions on formal language learning. *Inform. Comput.*, 85:1–11, 1990.

[358] M. Fulk and J. Case, editors. *Proc. 3rd Annu. Workshop on Comput. Learning Theory*. Morgan Kaufmann, Inc., San Mateo, CA, 1990.

[359] M. Fulk, S. Jain, and D. Osherson. Open problems in systems that learn. *J. of Comput. Syst. Sci.*, 1993. To appear.

[360] M. A. Fulk. Robust separations in inductive inference. In *Proc. of the 31st Symposium on the Foundations of Comp. Sci.*, pages 405–410. IEEE Computer Society Press, Los Alamitos, CA, 1990.

[361] M. A. Fulk and S. Jain. Learning in the presence of inaccurate information. In *Proceedings of the Second Annual Workshop on Computational Learning Theory, Santa Cruz, California*, pages 175–188. Morgan Kaufmann, Inc., Aug. 1989.

[362] M. A. Fulk and S. Jain. Approximate inference and scientific method. *Inform. Comput.*, 1992. To Appear.

[363] M. L. Furst, J. C. Jackson, and S. W. Smith. Improved learning of AC^0 functions. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 317–325. Morgan Kaufmann, 1991.

[364] B. R. Gaines. Maryanski's grammatical inferencer. *IEEE Trans. on Computers*, C-28:62–64, 1978.

[365] B. R. Gaines. Maryanski's grammatical inferencer. *IEEE Trans. Computers*, C-27(1):62–66, Jan. 1979.

[366] M. Gams and N. Lavrac. Review of five empirical learning systems within a proposed schemata. In *Progress in Machine Learning–Proceedings of EWSL 87: 2nd European Working Session on Learning*, pages 46–66, Bled, Yogoslavia, May 1987.

[367] Q. Gao and M. Li. An application of minimum description length principle to online r ecognition of handprinted alphanumerals. In *Proc. 11th International Joint Conference on Artificial Intellige nce*, pages 843–848. Morgan Kaufmann, 1989.

[368] W. Gasarch, M. Pleszkoch, and R. Solovay. Learning via queries to $[+,<]$. *J. Symbolic Logic*, 57(1):53–81, Mar. 1992.

[369] W. Gasarch and C. Smith. On the inference of sequences of functions. *Lecture Notes in Computer Science*, 265:23–41, 1986.

[370] W. Gasarch and C. H. Smith. Learning via queries. *J. ACM*, 39(3):649–675, July 1992. A shorter version is in 29th FOCS conference, 1988, pp. 130-137.

[371] W. I. Gasarch, R. K. Sitaraman, C. H. Smith, and M. Velauthapillai. Learning programs with an easy to calculate set of errors. *Fundamentca Informaticae*, pages 355–370, 1992.

[372] R. Gavaldà. On the power of equivalence queries. In *Computational Learning Theory: Eurocolt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 193–203, Oxford, 1994. Oxford University Press.

[373] S. Geisser. The inferential use of predictive distributions. In *Foundations of Statistical Inference*. Holt, Rinehart, and Winston, 1970.

[374] S. Geman. Stochastic relaxation methods for image restoration and expert systems. In *Automated Image Analysis: Theory and Experiments*. Academic Press, 1986.

[375] F. H. George. Logical networks and probability. *Bulletin of Mathematical Biophysics*, 19:187–199, 1957.

[376] F. H. George. Inductive machines and the problem of learning. *Cybernetica*, II:109–126, 1959.

[377] M. P. Georgeff and C. S. Wallace. A general selection criterion for inductive inference. In *ECAI 84: Advances in Artificial Intelligence*, pages 473–482. Elsevier Science, 1984.

[378] D. Gillman and M. Sipser. Inference and minimization of hidden Markov chains. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 147–158. ACM Press, New York, NY, 1994.

[379] G. J. Goetsch. CONSENSUS: A statistical learning procedure in a connectionist network. Technical Report CMU-CS-86-131, Carnegie-Mellon Computer Science Department, May 1986.

[380] E. M. Gold. Limiting recursion. *J. Symbolic Logic*, 30:28–48, 1965.

[381] E. M. Gold. Language identification in the limit. *Inform. Control*, 10:447–474, 1967.

[382] E. M. Gold. Universal goal seekers. *Inform. Control*, 18:395–403, 1971.

[383] E. M. Gold. System identification via state characterization. *Automatica*, 8:621–636, 1972.

[384] E. M. Gold. Complexity of automaton identification from given data. *Inform. Control*, 37:302–320, 1978.

[385] M. Gold. *Models of Goal Seeking and Learning*. PhD thesis, University of California, Los Angeles, 1965.

[386] D. E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, 1989.

[387] D. E. Goldberg and J. H. Holland. Special issue on genetic algorithms. *Machine Learning*, 3(2/3), Oct. 1988.

[388] P. Goldberg and M. Jerrum. Bounding the Vapnik-Chervonenkis dimension of concept classes parameterized by real numbers. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 361–369. ACM Press, New York, NY, 1993.

[389] P. W. Goldberg and S. A. Goldman. Learning one-dimensional geometric patterns under one-sided random misclassification noise. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 246–255. ACM Press, New York, NY, 1994.

[390] P. W. Goldberg, S. A. Goldman, and H. D. Mathais. Learning unions of boxes with membership and equivalence queries. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 198–207. ACM Press, New York, NY, 1994.

[391] S. Goldman. *Learning Binary Relations, Total Orders, and Read-Once Formulas.* PhD thesis, MIT Dept. of Electrical Engineering and Computer Science, Sept. 1990. (MIT Laboratory for Computer Science Technical Report MIT/LCS/TR-483, July 1990).

[392] S. Goldman and D. Mathias. Teaching a smarter learner. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 67–76. ACM Press, New York, NY, 1993.

[393] S. Goldman and R. Sloan. Can PAC learning algorithms tolerate random attribute noise? Technical Report WUCS-92-25, Washington University Department of Computer Science, July 1992. To appear, *Algorithmica.*

[394] S. Goldman and R. Sloan. The difficulty of random attribute noise. Technical Report WUCS-92-25, Washington University Department of Computer Science, July 1992. To appear, *Algorithmica.*

[395] S. Goldman and M. Warmuth. Learning binary relations using weighted majority voting. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 453–462. ACM Press, New York, NY, 1993.

[396] S. A. Goldman and M. J. Kearns. On the complexity of teaching. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 303–314, San Mateo, CA, 1991. Morgan Kaufmann. To appear, *Journal of Computer and System Sciences.*

[397] S. A. Goldman, M. J. Kearns, and R. E. Schapire. On the sample complexity of weak learning. In *Proc. 3rd Annu. Workshop on Comput. Learning Theory*, pages 217–231, San Mateo, CA, 1990. Morgan Kaufmann. To appear, *Information and Computation.*

[398] S. A. Goldman, M. J. Kearns, and R. E. Schapire. Exact identification of circuits using fixed points of amplification functions. *SIAM J. Comput.*, 22(4):705–726, Aug. 1993.

[399] S. A. Goldman and H. D. Mathias. Learning k-term DNF formulas with an incomplete membership oracle. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 77–84. ACM Press, New York, NY, 1992.

[400] S. A. Goldman and R. L. Rivest. A non–iterative maximum entropy algorithm. In *Uncertainty in Artificial Intelligence*, volume 2, pages 133–148. Elsevier Science B.V. North-Holland, 1988.

[401] S. A. Goldman, R. L. Rivest, and R. E. Schapire. Learning binary relations and total orders. *SIAM J. Comput.*, 22(5):1006–1034, Oct. 1993.

[402] S. A. Goldman and R. H. Sloan. The power of self-directed learning. Technical Report WUCS-92-49, Washington University in St. Louis, Nov. 1992. To appear, *Machine Learning.*

[403] S. A. Goldman and R. H. Sloan. The power of self-directed learning. *Machine Learning*, 14(3):271–294, 1994.

[404] M. Golea and M. Marchand. Average case analysis of the clipped Hebb rule for nonoverlapping perceptron networks. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 151–157. ACM Press, New York, NY, 1993.

[405] M. Golea and M. Marchand. On learning simple deterministic and probabilistic neural concepts. In *Computational Learning Theory: Eurocolt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 47–60, Oxford, 1994. Oxford University Press.

[406] R. C. Gonzalez and M. G. Thomason. *Syntactic Pattern Recognition, An Introduction.* Addison-Wesley, Reading, Mass., 1978.

[407] R. C. E. J. J. Gonzalez and M. G. Thomason. An algorithm for the inference of tree grammars. *Internat. J. Comput. Info. Sci.*, 5:145–164, 1976.

[408] J. Gordon and E. H. Shortliffe. A method for managing evidential reasoning in a hierarchical hypothesis space. *Artificial Intelligence*, 26(3):323–357, July 1985.

[409] R. P. Gorman and T. J. Sejnowski. Analysis of hidden units in a layered network trained to classify sonar targets. *Neural Networks*, 1:75–89, 1988.

[410] J. Grabowski. Inductive inference of functions from noised observations. *Lecture Notes in Computer Science*, 265:55–60, 1986.

[411] H. P. Graf, S. Satyanarayana, and Y. Tsividis. A reconfigurable analog (VLSI neural network chip. In *Proc. 1989 Neural Information Processing Conference*, pages 758–768, 1989.

[412] J. Grefenstette. Genetic algorithms and machine learning. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 3–4. ACM Press, New York, NY, 1993.

[413] U. Grenander. *Abstract Inference.* John WIley & Sons, Inc., 1981.

[414] G. Györgi and N. Tishby. Statistical theory of learning a rule. In *Proceedings of the STATPHYS-17 Workshop on Neural Networks and Spin Glasses*, Singapore, Teaneck NJ, Hong Kong, 1989. World Scientific Publishing Co.

[415] G. Gyorgyi. Inference of a rule by a neural network with thermal noise. *Phys. Rev. Lett.*, 64:2957, 1990.

[416] G. Györgyi and N. Tishby. Statistical theory of learning a rule. In *Neural Networks and Spin Glasses.* World Scientific, 1990.

[417] T. Hagerup and C. Rub. A guided tour of Chernov bounds. *Inform. Proc. Lett.*, 33:305–308, 1990.

[418] R. Hall, E. Wenger, D. Kibler, and P. Langley. The effect of multiple knowledge sources on learning and teaching. Technical Report 85-11, University of California, Irvine, 1985.

[419] H. Hamburger and K. Wexler. Identifiability of a class of transformational grammars. In *Approaches to Natural Language.* Dordecht, D. reidel, 1973.

[420] S. Hampson and D. Kibler. A boolean complete neural model of adaptive behavior. *Biol. Cybern.*, 49:9–19, 1983.

[421] S. E. Hampson and D. J. Volper. Linear function neurons: Structure and training. *Biol. Cybern.*, 53:203–217, 1986.

[422] T. Hancock. Identifying decision trees with equivalence queries. Manuscript, Harvard University, 1989.

[423] T. Hancock. Identifying μ-decision trees and μ-formulas with constrained instance queries. Manuscript, Harvard University, 1989.

[424] T. Hancock. Learning $k\mu$ decision trees on the uniform distribution. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 352–360. ACM Press, New York, NY, 1993.

[425] T. Hancock, M. Golea, and M. Marchand. Learning nonoverlapping perceptron networks from examples and membership queries. Technical report, Harvard University Center for Research in Computing Technology, TR-26-91, 1991.

[426] T. Hancock and L. Hellerstein. Learning read-once formulas over fields and extended bases. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 326–336. Morgan Kaufmann, San Mateo, CA, 1991.

[427] T. Hancock and Y. Mansour. Learning monotone $k\mu$-DNF formulas on product distributions. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 179–183, San Mateo, CA, 1991. Morgan Kaufmann.

[428] T. R. Hancock. Identifying μ-formula decision trees with queries. In *Proc. 3rd Annu. Workshop on Comput. Learning Theory*, pages 23–37, San Mateo, CA, 1990. Morgan Kaufmann.

[429] T. R. Hancock. Learning 2μ-DNF formulas and $k\mu$ decision trees. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 199–209, San Mateo, CA, 1991. Morgan Kaufmann.

[430] D. Hansel and H. Sompolinsky. Learning from examples in a single-layer neural network. *Europhys. Lett.*, 11(7):687–692, 1990.

[431] S. J. Hanson and D. J. Burr. What connectionist models learn: Learning and representation in connectionist networks. *Behavioral and Brain Sciences*, 13:471–518, 1990.

[432] S. J. Hanson and W. Remmele, editors. *Machine Learning: From Theory to Applications; Cooperative Research at Siemens and MIT*. Springer-Verlag, 1993.

[433] S. Hardy. Synthesis of LISP programs from examples. In *Proc. 4th IJCAI*, pages 268–273. IJCAI, 1975.

[434] G. W. Hart. Prototype nonintrusive appliance load monitor. Technical Report Progress Report #2, MIT Energy Laboratory, Sept. 1985.

[435] G. W. Hart. *Minimum Information Estimation of Structure*. PhD thesis, MIT Dept. of Electrical Engineering and Computer Science, Apr. 1987. Appears as LIDS-TH-1664.

[436] D. Haussler. Applying valiant's learning framework to AI concept learning problems. Technical Report UCSC-CRL-87-11, UCSC, 1987.

[437] D. Haussler. Bias, version spaces, and Valiant's learning framework. In *Proceedings of the 4th International Workshop on Machine Learning*, pages 324–336, San Mateo, CA, June 1987. Morgan Kaufmann.

[438] D. Haussler. Quantifying inductive bias: AI learning algorithms and Valiant's learning framework. *Artificial Intelligence*, 36:177–221, 1988.

[439] D. Haussler. Space efficient learning algorithms. Technical Report UCSC-CRL-88-2, University of Calif. Computer Research Laboratory, Santa Cruz, CA, 1988.

[440] D. Haussler. Generalizing the PAC model for neural net and other learning applications. Technical Report UCSC-CRL-89-30, University of California Santa Cruz, Sept. 1989.

[441] D. Haussler. Generalizing the PAC model: sample size bounds from metric dimension-based uniform convergence results. In *Proc. 30th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 40–45. IEEE Computer Society Press, Los Alamitos, CA, 1989.

[442] D. Haussler. Learning conjunctive concepts in structural domains. *Machine Learning*, 4(1):7–40, 1989.

[443] D. Haussler. Probably approximately correct learning. In *Proc. of the 8th National Conference on Artificial Intelligence*, pages 1101–1108. Morgan Kaufmann, 1990.

[444] D. Haussler. Sphere packing numbers for subsets of the Boolean n-cube with bounded Vapnik-Chervonenkis dimension. Technical Report UCSC-CRL-91-41, University of Calif. Computer Research Laboratory, Santa Cruz, CA, 1991.

[445] D. Haussler. Decision theoretic generalizations of the PAC model for neural net and other learning applications. *Inform. Comput.*, 100(1):78–150, Sept. 1992.

[446] D. Haussler, M. Kearns, N. Littlestone, and M. K. Warmuth. Equivalence of models for polynomial learnability. *Inform. Comput.*, 95(2):129–161, Dec. 1991.

[447] D. Haussler, M. Kearns, and R. E. Schapire. Bounds on the sample complexity of Bayesian learning using information theory and the VC dimension. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 61–74, San Mateo, CA, 1991. Morgan Kaufmann.

[448] D. Haussler, M. Kearns, H. S. Seung, and N. Tishby. Rigorous learning curve bounds from statistical mechanics. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 76–87. ACM Press, New York, NY, 1994.

[449] D. Haussler, N. Littlestone, and M. K. Warmuth. Predicting {0,1} functions on randomly drawn points. In *Proceedings of the 29th Annual IEEE Symposium on Foundations of Computer Science*, pages 100–109. IEEE Computer Society Press, 1988.

[450] D. Haussler and P. Long. A generalization of sauer's lemma. Technical Report UCSC-CRL-90-15, U.C. Santa Cruz Computer Research Laboratory, Apr. 1990.

[451] D. Haussler and L. Pitt, editors. *Proc. 1st Annu. Workshop on Comput. Learning Theory.* Morgan Kaufmann, Inc., San Mateo, CA, 1988.

[452] D. Haussler and E. Welzl. Epsilon-nets and simplex range queries. *Discrete Computational Geometry*, 2:127–151, 1987.

[453] F. Hayes-Roth and J. McDermott. An interference matching technique for inducing abstractions. *Commun. ACM*, 21:401–410, 1978.

[454] T. Hegedüs. Geometrical concept learning and convex polytopes. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 228–236. ACM Press, New York, NY, 1994.

[455] T. Hegedüs. On training simple neural networks and small-weight neurons. In *Computational Learning Theory: Eurocolt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 69–82, Oxford, 1994. Oxford University Press.

[456] L. Hellerstein. *On Characterizing and Learning Some Classes of Read-once Functions.* PhD thesis, University of California, Berkeley, 1989.

[457] L. Hellerstein and M. Karpinski. Learning read-once formulas using membership queries. In *Proc. of the Second Annual Workshop on Computational Learning Theory*, pages 146–161. Morgan Kaufmann, 1989.

[458] L. Hellerstein and M. Karpinski. Computational complexity of learning read-once formulas over different bases. Technical report, International Computer Science Institute, Berkeley, CA, TR-91-014, 1991.

[459] D. Helmbold, R. Sloan, and M. K. Warmuth. Learning nested differences of intersection-closed concept classes. In *Proc. 2nd Annu. Workshop on Comput. Learning Theory*, pages 41–56, San Mateo, CA, 1989. Morgan Kaufmann.

[460] D. Helmbold, R. Sloan, and M. K. Warmuth. Learning integer lattices. *SIAM J. Comput.*, 21(2):240–266, 1992.

[461] D. P. Helmbold, N. Littlestone, and P. M. Long. Apple tasting and nearly one-sided learning. In *Proc. of the 33rd Symposium on the Foundations of Comp. Sci.*, pages 493–502. IEEE Computer Society Press, Los Alamitos, CA, 1992.

[462] D. P. Helmbold and P. M. Long. Tracking drifting concepts by minimizing disagreements. Technical Report UCSC-CRL-91-26, UC Santa Cruz, 1991. To appear, *Machine Learning*.

[463] D. P. Helmbold and P. M. Long. Tracking drifting concepts using random examples. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 13–23, San Mateo, CA, 1991. Morgan Kaufmann.

[464] D. P. Helmbold and M. K. Warmuth. Some weak learning results. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 399–412. ACM Press, New York, NY, 1992.

[465] R. J. Hendel. Mathematical learning theory: a formalized, axiomatic, abstract approach. *Inform. Control*, 41:67–117, 1979.

[466] G. Hinton and D. van Camp. Keeping neural networks simple by minimizing the description length of the weights. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 5–13. ACM Press, New York, NY, 1993.

[467] G. E. Hinton. Learning distributed representations of concepts. In *Proceedings of the Eighth Annual Conference of the Cognitive Science Society*, Amherst, Mass, Aug. 1986.

[468] G. E. Hinton. Deterministic Boltzmann learning performs steepest descent in weight-space. *Neural Computation*, 1(1):143–150, 1989.

[469] G. E. Hinton, T. J. Sejnowski, and D. H. Ackley. Boltzmann machines: Constraint satisfaction networks that learn. Technical Report CMS-CS-84-119, CMU Computer Science Department, May 1984.

[470] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, Mar. 1963.

[471] K. Höffgen. Learning and robust learning of product distributions. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 77–83. ACM Press, New York, NY, 1993.

[472] K. Höffgen and H. Simon. Robust trainability of single neurons. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 428–439, New York, New York, 1992. ACM Press.

[473] J. H. Holland. *Adaptation in Natural and Artificial Systems.* The University of Michigan Press, 1975.

[474] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad Sci. USA*, 79:2554–2558, Apr. 1982.

[475] K. Hornik, M. Stinchcombe, and H. White. Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. *Neural Networks*, 3:551–560, 1990.

[476] J. J. Horning. *A Study of Grammatical Inference*. PhD thesis, Stanford University, 1969.

[477] P. J. Huber. Projection pursuit. *Ann. of Statist.*, 13(2):435–475.

[478] O. H. Ibarra and T. Jiang. Learning regular languages from counterexamples. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 371–385, San Mateo, CA, 1988. Morgan Kaufmann.

[479] B. Indurkhya. On the role of interpretive analogy in learning. *New Generation Computing*, 8:385–402, 1991.

[480] N. Intrator. Feature extraction using an unsupervised neural network. In *Proceedings of the 1990 Connectionist Models Summer School*, pages 310–318. Morgan Kaufmann., 1990.

[481] H. Ishizaka. Learning simple deterministic languages. In *Proc. 2nd Annu. Workshop on Comput. Learning Theory*, pages 162–174, San Mateo, CA, 1989. Morgan Kaufmann.

[482] H. Ishizaka. Polynomial time learnability of simple deterministic languages. *Machine Learning*, 5:151–164, 1990.

[483] H. Ito, S. Amari, and K. Kobayashi. Identifiability of hidden Markov information sources and their minimum degrees of freedom. *IEEE Trans. Information Theory*, 38(2):324–333, Mar. 1992.

[484] S. Y. Itoga. A new heuristic for inferring regular grammars. *IEEE Trans. Pattern Anal. Mach. Intell.*, TPAMI-3:191–197, 1981.

[485] J. Jackson and A. Tomkins. A computational model of teaching. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 319–326. ACM Press, New York, NY, 1992.

[486] R. A. Jacobs, M. A. Jordan, and A. G. Barto. Task decomposition through competition in a modular connectionist architecture: The what and where vision tasks. Technical Report 90-27, COINS, Mar. 1990.

[487] S. Jain. *Learning in the Presence of Additional Information and Inaccurate Information.* PhD thesis, University of Rochester, 1990.

[488] S. Jain and A. Sharma. Knowledge of an upper bound on grammar size helps language learning. Technical Report 283, University of Rochester, 1989.

[489] S. Jain and A. Sharma. Recursion theoretic characterizations of language learning. Technical Report 281, University of Rochester, 1989.

[490] S. Jain and A. Sharma. Finite learning by a team. In *Proc. 3rd Annu. Workshop on Comput. Learning Theory*, pages 163–177, San Mateo, CA, 1990. Morgan Kaufmann.

[491] S. Jain and A. Sharma. Hypothesis formation and language acquisition with an infinitely often correct teacher. In *Proceedings of the Third Conference on Theoretical Aspects of Reasoning About Knowledge*, pages 225–239. Morgan Kaufmann, Inc., Mar. 1990.

[492] S. Jain and A. Sharma. Language learning by a team. In *Proceedings of the 17th International Colloquium on Automata, Languages and Programming*, pages 153–166. Springer-Verlag, July 1990.

[493] S. Jain and A. Sharma. Program size restrictions in inductive learning. Technical Report 90-06, University of Delaware, Newark, Delaware, 1990.

[494] S. Jain and A. Sharma. Learning in the presence of partial explanations. *Inform. Comput.*, 95-2:162–191, 1991.

[495] S. Jain and A. Sharma. Restrictions on grammar size in language identification. In *Proceedings MLNLO'91, Machine Learning of Natural Language and Ontology, Stanford University, California. Document* **D91-09**, *DFKI: Kaiserslautern FRG, 1991*, pages 87–92, Mar. 1991.

[496] S. Jain and A. Sharma. Characterizing language learning by standardizing operations. *J. of Comput. Syst. Sci.*, 1993. To Appear.

[497] S. Jain and A. Sharma. Computational limits on team identification of languages. Technical Report 9301, School of Computer Science and Engineering; University of New S outh Wales, 1993.

[498] S. Jain and A. Sharma. Learning with the knowledge of an upper bound on program size. *Inform. Comput.*, 102–1:118–166, 1993. Presented at the Workshop on Computational Learning Theory and Natural Learning Systems.

[499] S. Jain and A. Sharma. On aggregating teams of learning machines. In *Proceedings of the Fourth Workshop on Algorithmic Learning Theory, Tokyo, Japan.* Japanese Society for Artificial Intelligence, Oct. 1993. To Appear.

[500] S. Jain and A. Sharma. On the non-existence of maximal inference degrees for language identification. *Inform. Proc. Lett.*, 1993. To appear.

[501] S. Jain and A. Sharma. Probability is more powerful than team for language identification from positive data. In *Proc. 6th Annu. Conference on Comput. Learning Theory*, pages 192–198, New York, New York, 1993. ACM Press.

[502] S. Jain and A. Sharma. Prudence in vacillatory language identification. *Math. Syst. Theory*, 1993. To Appear.

[503] S. Jain and A. Sharma. On the intrinsic complexity of language identification. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 278–286. ACM Press, New York, NY, 1994.

[504] S. Jain and A. Sharma. Program size restrictions in computational learning. *Theoret. Comput. Sci. A*, 1994. To appear.

[505] S. Jain, A. Sharma, and J. Case. Convergence to nearly minimal size grammars by vacillating learning machines. In *Proc. 2nd Annu. Workshop on Comput. Learning Theory*, pages 189–199, San Mateo, CA, 1989. Morgan Kaufmann.

[506] K. Jantke. Automatic synthesis of programs and inductive inference of functions. In *Int. Conf. Fundamentals of Computations Theory*, pages 219–225, 1979.

[507] K. Jantke. *Leistungsfahigkeit und Kmpliziertheit universeller Verfahren zur Erkennung allgemein-rekursiver Funktinen.* PhD thesis, Humboldt-Universitat, Berlin, 1979.

[508] K. Jantke and H. Beick. Combining postulates of naturalness in inductive inference. *Elektronische Informationsverarbeitung und Kybernetik*, 17:465–484, 1981.

[509] K. P. Jantke. Natural properties of strategies identifying recursive functions. *Elektronische Informationsverarbeitung und Kybernetik*, 15:487–496, 1979.

[510] K. P. Jantke. Polynomial-time inference of general pattern languages. In *Proceedings of the Symposium of Theoretical Aspects of Computer Science; Lecture Notes in Computer Science*, volume 166, pages 314–325. Springer, 1984.

[511] K. P. Jantke. An algebraic framework for inductive program synthesis. *Lecture Notes in Computer Science*, 215:101–122, 1986.

[512] K. P. Jantke, editor. *Analogical and Inductive Inference, International Workshop AII '86. Wendisch-Rietz, GDR*, volume 265 of *Lecture Notes in Computer Science*. Springer-Verlag, Berlin, Oct. 1986.

[513] K. P. Jantke. Monotonic and nonmonotonic inductive inference of functions and patterns. In *Nonmonotonic and Inductive Logic, 1st International Workshop, Karlsruhe, Germany*, pages 161–177. Springer Verlag, 1990. Lecture Notes in Computer Science 543.

[514] K. P. Jantke. Monotonic and non-monotonic inductive inference. *New Generation Computing*, 8:349–360, 1991.

[515] K. P. Jantke. Case based learning in inductive inference. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 218–223. ACM Press, New York, NY, 1992.

[516] M. Jerrum. Simple translation-invariant concepts are hard to learn. Technical report, University of Edinburgh, Department of Computer Science, CSR-12-91, 1991.

[517] C. Ji and D. Psaltis. The VC-dimension vs. the statistical capacity for two layer networks with binary weights. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 250–256, San Mateo, CA, 1991. Morgan Kaufmann.

[518] T. Jiang and M. Li. On the complexity of learning strings and sequences. *Theoret. Comput. Sci.*, 119:363–371, 1993.

[519] M. Joardan and D. Rumelhart. Forward models: Supervised learning with a distal teacher. (Unpublished), July 1990.

[520] L. K. Jones. A simple lemma on greedy approximation in hilbert space and convergence rates for projection pursuit regression and neural network training. To appear in Annals of Statistics., 1990.

[521] K. D. Jong. Special issue on genetic algorithms. *Machine Learning*, 5(4), Oct. 1990.

[522] M. I. Jordan. Supervised learning and systems with excess degrees of freedom. Technical Report 88-27, COINS, May 1988.

[523] M. I. Jordan. A statistical approach to decision tree modeling. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 13–20. ACM Press, New York, NY, 1994.

[524] J. P. Jouannaud and G. Guiho. Inference of functions with an interactive system. In *Machine Intelligence 9*, pages 227–250. John Wiley and Sons, N. Y., 1979.

[525] J. P. Jouannaud and Y. Kodratoff. Characterization of a class of functions synthesized by a summers-like method using a B.M.W. matching technique. In *Proc. 6th IJCAI*, pages 440–447. IJCAI, 1979.

[526] J. P. Jouannaud and Y. Kodratoff. An automatic construction of LISP programs by transformations of functions synthesized from their input-output behavior. *International J. of Policy Analysis and Information Systems*, 4:331–358, 1980.

[527] J. P. G. G. Jouannaud and T. P. Treuil. SISP/1, an interactive system able to synthesize functions from examples. In *Proc. 5th IJCAI*, pages 412–418. IJCAI, 1977.

[528] B. Juang. On the hidden markov model and dynamic time warping for speech recognition – A unified view. *AT&T Bell Laboratories Technical Journal*, 63(7):1213–1242, Sept. 1985.

[529] J. S. Judd. Complexity of connectionist learning with various node functions. Technical Report 87-60, Department of Computer and Information Science, University of Massachusetts at Amherst, July 1987. Also presented at the First IEEE International Conference on Neural Networks, June 21–24, 1987, San Diego, California.

[530] J. S. Judd. *Neural Network Design and the Complexity of Learning.* MIT Press, 1990.

[531] S. Judd. Learning in neural networks. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 2–8, San Mateo, CA, 1988. Morgan Kaufmann.

[532] H. Jung. *Zur Untersuchung von Abstrakten Interaktiven Erkennungssystemen.* PhD thesis, Humbolt University, Berlin, 1977.

[533] Y. Kabashima and S. Shinomoto. Acceleration of learning in binary choice problems. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 446–452. ACM Press, New York, NY, 1993.

[534] P. Kanerva. *Sparse Distributed Memory.* MIT Press, 1988.

[535] S. Kannan. On the query complexity of learning. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 58–66. ACM Press, New York, NY, 1993.

[536] M. Kao, J. H. Reif, and S. R. Tate. Searching in an unknown environment: An optimal randomized algorithm for the cow-path problem. In *Proceedings SODA 93*, pages 441–447, Jan. 1993.

[537] S. Kapur and G. Bilardi. Language learning from stochastic input. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 303–310. ACM Press, New York, NY, 1992.

[538] M. Kearns. Thoughts on hypothesis boosting. (Unpublished), Dec. 1988.

[539] M. Kearns. *The Computational Complexity of Machine Learning.* PhD thesis, Harvard University Center for Research in Computing Technology, May 1989. Technical Report TR-13-89. Also published by MIT Press as an ACM Distinguished Dissertation.

[540] M. Kearns. Efficient noise-tolerant learning from statistical queries. In *Proc. 25th Annu. ACM Sympos. Theory Comput.*, pages 392–401. ACM Press, New York, NY, 1993.

[541] M. Kearns and M. Li. Robust learning in polynomial time: possibilities and limitations. Manuscript., 1987.

[542] M. Kearns and M. Li. Learning in the presence of malicious errors. *SIAM J. Comput.*, 22:807–837, 1993.

[543] M. Kearns, M. Li, L. Pitt, and L. Valiant. On the learnability of Boolean formulae. In *Proc. 19th Annu. ACM Sympos. Theory Comput.*, pages 285–294. ACM Press, New York, NY, 1987.

[544] M. Kearns, M. Li, L. Pitt, and L. Valiant. Recent results on boolean concept learning. In *Proceedings of the Fourth International Workshop on Machine Learning*, pages 337–352, Irvine, California, June 1987. (published by Morgan Kaufmann, Los Altos California).

[545] M. Kearns and L. Pitt. A polynomial-time algorithm for learning k-variable pattern languages from examples. In *Proc. 2nd Annu. Workshop on Comput. Learning Theory*, pages 57–71, San Mateo, CA, 1989. Morgan Kaufmann.

[546] M. Kearns and H. S. Seung. Learning from a population of hypotheses. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 101–110. ACM Press, New York, NY, 1993.

[547] M. Kearns and L. G. Valiant. Learning boolean formulae or finite automata is as hard as factoring. Technical Report TR 14-88, Harvard University Aiken Computation Laboratory, 1988.

[548] M. Kearns and L. G. Valiant. Cryptographic limitations on learning Boolean formulae and finite automata. In *Proc. of the 21st Symposium on Theory of Computing*, pages 433–444. ACM Press, New York, NY, 1989.

[549] M. Kearns and U. Vazirani. Topics in learning theory. (Prelinary draft), 1992.

[550] M. J. Kearns and R. E. Schapire. Efficient distribution-free learning of probabilistic concepts. In *Proc. of the 31st Symposium on the Foundations of Comp. Sci.*, pages 382–391. IEEE Computer Society Press, Los Alamitos, CA, 1990.

[551] M. J. Kearns, R. E. Schapire, and L. M. Sellie. Toward efficient agnostic learning. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 341–352. ACM Press, New York, NY, 1992.

[552] K. Kelly and C. Glymour. On convergence to the truth and nothing but the truth. Technical report, Carnegie Mellon University, Laboratory for Computational Linguistics, CMU-LCL-87-4, 1987.

[553] K. T. Kelly. Induction from the general to the more general. In *Proc. 2nd Annu. Workshop on Comput. Learning Theory*, pages 334–348, San Mateo, CA, 1989. Morgan Kaufmann.

[554] K. T. Kelly and C. Glymour. Inductive inference from theory laden data. Technical Report CMU-LCL-89-5, CMU Laboratory for Computational Linguistics, Oct. 1989.

[555] M. Kharitonov. Cryptographic lower bounds on learnability of Boolean functions on the uniform distribution. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 29–36. ACM Press, New York, NY, 1992.

[556] M. Kharitonov. Cryptographic hardness of distribution-specific learning. In *Proc. 25th Annu. ACM Sympos. Theory Comput.*, pages 372–381. ACM Press, New York, NY, 1993.

[557] J. Kilian, K. J. Lang, and B. A. Pearlmutter. Playing the matching-shoulders lob-pass game with logarithmic regret. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 159–164. ACM Press, New York, NY, 1994.

[558] J. Kilian and H. Siegelmann. On the power of sigmoid neural networks. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 137–143. ACM Press, New York, NY, 1993.

[559] J. H. Kim. *CONVINCE: A Conversational Inference Consolidation Engine.* PhD thesis, University of California, Los Angeles, 1983.

[560] W. M. Kim. Learning by smoothing: a morphological approach. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 43–57, San Mateo, CA, 1991. Morgan Kaufmann.

[561] D. Kimber and P. Long. The learning complexity of smooth functions of a single variable. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 153–159. ACM Press, New York, NY, 1992.

[562] E. Kinber, C. Smith, M. Velauthapillai, and R. Wiehagen. On learning multiple concepts in parallel. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 175–181. ACM Press, New York, NY, 1993.

[563] E. B. Kinber. On the synthesis in the limit of almost minimal Gödel numbers. *Theory of Algorithms and Programs, Latvian State University, Riga*, 1:221–223, 1974.

[564] E. B. Kinber. On comparison of limit identification and limit standardization of general recursive functions. *Uch. zap. Latv. univ.*, 233:45–56, 1975.

[565] E. B. Kinber. On a theory of inductive inference. *Lecture Notes in Computer Science*, 56:435–440, 1977.

[566] E. B. Kinber. On limit identification of minimal Gödel numbers for functions from enumerable classes. *Theory of Algorithms and Programs, Latvian State University, Riga*, 3:35–56, 1977.

[567] E. B. Kinber. A note on limit identification of c-minimal indices. *Elektronische Informationsverarbeitung und Kybernetik*, 19:459–463, 1983.

[568] E. B. Kinber. Inductive syntactical synthesis of programs from sample computations. *Lecture Notes in Computer Science*, pages 73–84, 1988.

[569] E. B. Kinber. Some problems of learning with an oracle. In *Proc. 3rd Annu. Workshop on Comput. Learning Theory*, pages 178–186, San Mateo, CA, 1990. Morgan Kaufmann.

[570] E. B. Kinber, W. I. Gasarch, T. Zeugmann, M. K. Pleszkoch, and C. H. Smith. Learning via queries with teams and anomalies. In *Proc. 3rd Annu. Workshop on Comput. Learning Theory*, pages 327–337, San Mateo, CA, 1990. Morgan Kaufmann.

[571] E. B. Kinber and T. Zeugmann. Inductive inference of almost everywhere correct programs by reliably working strategies. *Elektronische Informationsverarbeitung und Kybernetik*, 21:91–100, 1985.

[572] E. B. Kinber and T. Zeugmann. Monte-carlo inference and its relations to reliable frequency identification. In *Proc. Fundamentals of Computation Theory*, pages 257–266. Springer–Verlag, 1989. Lecture Notes in Computer Science 380.

[573] E. B. Kinber and T. Zeugmann. Refined query inference. In *Proc. 2nd Internat. Workshop on Analogical and Inductive Inference*, pages 148–160. Springer–Verlag, 1989. Lecture Notes in Artificial Intelligence 397.

[574] E. B. Kinber and T. Zeugmann. One-sided error probabilistic inductive inference and reliable frequency identification. *Inform. Comput.*, 92:253–284, 1991.

[575] J. Kivinen. Reliable and useful learning. In *Proc. 2nd Annu. Workshop on Comput. Learning Theory*, pages 365–380, San Mateo, CA, 1989. Morgan Kaufmann.

[576] J. Kivinen, H. Mannila, and E. Ukkonen. Learning heirarchical rule sets. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 37–44. ACM Press, New York, NY, 1992.

[577] J. Kivinen, H. Mannila, and E. Ukkonen. Learning rules with local exceptions. In *Computational Learning Theory: Eurocolt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 35–46, Oxford, 1994. Oxford University Press.

[578] J. Kivinen and M. Warmuth. Using experts for predicting continuous outcomes. In *Computational Learning Theory: Eurocolt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 109–120, Oxford, 1994. Oxford University Press.

[579] P. J. Klaus. Zur kompliziertheit der konsistenten erkennung von klassen allgemeinrekursiver funktionen. *International Symposium on Algorithmische Kompliziertheit, lern, und Erkennungs-prozesse*, pages 45–52, 1976.

[580] R. Klette. Unterklassen in der familie NUM aller effektiv numerierbaren mengen von eistelligen allgemein rekursiven funktionen. *Kompliziertheit von Lern und Erkennungsprozessen*, 1:62–66, 1975.

[581] R. Klette. Indexmengen und erkennung rekursiver functionen. *Zeitschr. f. math. Logik und Grundlagen d. Math. Bd.*, 22:231–238, 1976.

[582] R. Klette. Recognizing algorithms for recursive functions. *Elektronische Informationsverarbeitung und Kybernetik*, 12:227–243, 1976.

[583] R. Klette. A few results on the complexity of classes of identifiable recursive function sets. *Sektion Mathematik report 4 (M) 77, Friedrich Schiller Universitat, Jena*, pages 1–16, 1977.

[584] R. Klette and B. Goetze. Limes-rekursive funktionen. *Elektronische Informationsverarbeitung und Kybernetik*, 11:586–589, 1975.

[585] R. Klette and R. Wiehagen. Research in the theory of inductive inference by GDR mathematicians – A survey. *Inform. Sci.*, 22:149–169, 1980.

[586] B. Knobe and K. Knobe. A method for inferring context-free grammars. *Inform. Control*, 31:129–146, 1976.

[587] K. Ko. On the complexity of learning minimum time-bounded Turing machines. In *Proc. 3rd Annu. Workshop on Comput. Learning Theory*, pages 82–96, San Mateo, CA, 1990. Morgan Kaufmann.

[588] K. Ko and T. W. Marron, Assaf. Learning string patterns and tree patterns from examples. In *Machine Learning: Proceedings of the seventh international conference*, pages 384–391, 1990.

[589] Y. Kodratoff. A class of functions synthesized from a finite number of examples and a LISP program scheme. *Internat. J. Comput. Info. Sci.*, 8:489–521, 1979.

[590] Y. Kodratoff and J. Fargues. A sane algorithm for the synthesis of LISP functions from example problems. In *Proc. of AISB/GI Conference on Artificial Intelligence*, pages 169–175. AISB and GI, 1978.

[591] Y. Kodratoff and R. S. Michalski, editors. *Machine Learning: An Artificial Intelligence Approach*, volume III. Morgan Kaufmann, Los Altos, California, 1990.

[592] P. Koiran. Efficient learning of continuous neural networks. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 348–355. ACM Press, New York, NY, 1994.

[593] Kramer. A note on the self-consistency definitions of generalization and inductive inference. *J. ACM*, pages 280–281, 1962.

[594] P. Kugel. Induction, pure and simple. *Inform. Control*, 35:276–336, 1977.

[595] A. Kuh, T. Petsche, and R. L. Rivest. Learning time-varying concepts. In *Advances in Neural Information Processing Systems 3*, pages 183–189. Morgan Kaufmann, 1991.

[596] T. Kuh, T. Petsche, and R. Rivest. Learning time varying concepts. In *NIPS 3*. Morgan Kaufmann, 1990.

[597] T. Kuh, T. Petsche, and R. Rivest. Mistake bounds of incremental learners when concepts drift with applications to feedforward networks. In *NIPS 4*. Morgan Kaufmann, 1991.

[598] T. P. A. Kuh and R. L. Rivest. Incrementally learning time-varying half-planes. In *Advances in Neural Information Processing Systems 4*, pages 920–927. Morgan Kaufmann, 1992.

[599] S. Kulkarni. On metric entripy, Vapnik-Chervonenkis dimension, and learnability for a class of distributions. Technical Report LIDS-P-1910, Center for Intelligent Control Systems, MIT, 1989.

[600] S. Kulkarni and O. Zeitouni. On probably correct classification of concepts. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 111–116. ACM Press, New York, NY, 1993.

[601] S. R. Kulkarni. *Problems of computational and information complexity in machine vision and learning*. PhD thesis, MIT EECS, 1991. (Also CICS report CICS-TH-298.).

[602] S. R. Kulkarni, S. K. Mitter, and J. N. Tsitsiklis. Active learning using arbitrary binary valued queries. *Machine Learning*, 11(1), 1993.

[603] S. R. Kulkarni, S. K. Mitter, J. N. Tsitsiklis, and O. Zeitouni. PAC learning with generalized samples and an application to stochastic geometry. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 172–179. ACM Press, New York, NY, 1992.

[604] M. Kummer. The strength of noninclusions for teams of finite learners. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 268–277. ACM Press, New York, NY, 1994.

[605] M. Kummer and F. Stephan. On the structure of degrees of inferability. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 117–126. ACM Press, New York, NY, 1993.

[606] M. Kummer and F. Stephan. Inclusion problems in parallel learning and games. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 287–298. ACM Press, New York, NY, 1994.

[607] S. A. Kurtz and J. S. Royer. Prudence in language learning. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 143–156. Morgan Kaufmann, Inc., 1988.

[608] S. A. Kurtz and C. H. Smith. On the role of search for learning. In *Proc. 2nd Annu. Workshop on Comput. Learning Theory*, pages 303–311, San Mateo, CA, 1989. Morgan Kaufmann.

[609] E. Kushilevitz and Y. Mansour. Learning decision trees using the Fourier spectrum. In *Proc. of the 23rd Symposium on Theory of Computing*, pages 455–464. ACM Press, New York, NY, 1991.

[610] E. Kushilevitz and D. Roth. On learning visual concepts and DNF formulae. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 317–326. ACM Press, New York, NY, 1993.

[611] K. H. Kwon, K. Kang, and J. H. Oh. Generalization in partially connected layered neural networks. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 356–361. ACM Press, New York, NY, 1994.

[612] J. Laird, P. Rosenbloom, and A. Newell. Towards chunking as a general learning mechanism. In *Proceedings AAAI-84*, pages 188–192, Aug. 1984.

[613] J. Laird, P. Rosenbloom, and A. Newell. Chunking in Soar: the anatomy of a general learning mechanism. *Machine Learning*, 1(1):11–46, 1986.

[614] J. E. Laird, A. Newell, and P. S. Rosenbloom. SOAR: An architecture for general intelligence. *Artificial Intelligence*, 33(1):1–64, Sept. 1987.

[615] P. Laird. Inductive inference by refinement. In *Proc. of AAAI-86*, pages 472–476. AAAI, 1986.

[616] P. Laird. Learning by making models. Technical Report RIA-88-4-12-0, NASA-Ames Reasarch Center, 1988.

[617] P. Laird. A survey of computational learning theory. In *Formal Techniques in Artificial Intelligence: A Sourcebook*, pages 173–215. Elsevier Science, 1990.

[618] P. Laird and E. Gamble. Learning a probability distribution efficiently and reliably. Technical report, NASA Ames Research Center, Oct. 1988.

[619] P. Laird, S. R, and P. Dunning. A model of sequence extrapolation. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 84–93. ACM Press, New York, NY, 1993.

[620] P. D. Laird. Efficient unsupervised learning. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 297–311, San Mateo, CA, 1988. Morgan Kaufmann.

[621] K. J. Lang. Random DFA's can be approximately learned from sparse uniform examples. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 45–52. ACM Press, New York, NY, 1992.

[622] S. Lange. A note on polynomial-time inference of k-variable pattern language. In *Nonmonotonic and Inductive Logic, 1st International Workshop, Karlsruhe, Germany*, pages 178–183. Springer Verlag, 1990. Lecture Notes in Computer Science 543.

[623] S. Lange. Polynomial time inference of arbitrary pattern languages. *New Generation Computing*, 8:361–370, 1991.

[624] S. Lange. The representation of recursive languages and its impact on the efficiency of learning. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 256–267. ACM Press, New York, NY, 1994.

[625] S. Lange and T. Zeugmann. Learning recursive languages with bounded mind changes. Technical Report 16/92, GOSLER-Report, FB Mathematik und Informatik, TH Lepzig, 1992.

[626] S. Lange and T. Zeugmann. On the power of monotonic language learning. Technical Report 5/92, GOSLER-Report, FB Mathematik und Informatik, TH Lepzig, 1992.

[627] S. Lange and T. Zeugmann. Types of monotonic learning and their characterization. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 377–390. ACM Press, New York, NY, 1992.

[628] S. Lange and T. Zeugmann. A unifying approach to monotonic language learning on informant. In *Proc. 3rd Internat. Workshop on Analogical and Inductive Inference*, pages 244–259. Springer–Verlag, 1992. Lecture Notes in Artificial Intelligence 642.

[629] S. Lange and T. Zeugmann. Language learning in dependence on the space of hypotheses. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 127–136. ACM Press, New York, NY, 1993.

[630] S. Lange and T. Zeugmann. Language learning with a bounded number of mind changes. In *Proc. 10th Annual Symposium on Theoretical Aspects of Computer Science*, pages 682–691. Springer–Verlag, 1993. Lecture Notes in Computer Science 665.

[631] S. Lange and T. Zeugmann. The learnability of recursive languages in dependence on the hypothesis space. Technical Report 20/93, GOSLER-Report, FB Mathematik und Informatik, TH Lepzig, 1993.

[632] S. Lange and T. Zeugmann. Learning recursive languages with a bounded number of mind changes. *Internat. Journal of Foundations of Computer Science*, 4:157–178, 1993.

[633] S. Lange and T. Zeugmann. Monotonic versus non–monotonic language learning. In *Proc. 2nd Internat. Workshop on Nonmonotonic and Inductive Logic*, pages 254–269. Springer–Verlag, 1993. Lecture Notes in Artificial Intelligence 659.

[634] S. Lange and T. Zeugmann. On the impact of order independence to the learnability of recursive languages. Technical Report ISIS-RR-93-17E, Research Report, FUJITSU Laboratories Ltd., Numazu, 1993.

[635] S. Lange and T. Zeugmann. Charcterization of language learning from informant under various monotonicity constraints. *Journal of Experimental and Artificial Intelligence*, 6:73–94, 1994.

[636] P. Langley. A general theory of discrimination learning. Technical Report 86-18, University of California, Irvine, 1986.

[637] P. Langley and B. Nordhausen. A framework for empirical discovery. Technical Report 86-19, University of California, Irvine, 1986.

[638] R. H. Lathrop. *Efficient Methods for Massively Parallel Symbolic Induction: Algorithms and Implementation.* PhD thesis, MIT Department of Electrical Engineering and Computer Science, June 1990.

[639] K. Lee and S. Mahajan. A pattern classification approach to evaluation function learning. *Artificial Intelligence*, 36(1):1–26, Aug. 1988.

[640] W. S. Lee, P. L. Bartlett, and R. C. Williamson. Lower bounds on the VC-dimension of smoothly parametrized function classes. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 362–367. ACM Press, New York, NY, 1994.

[641] Y. C. Lee, S. Qian, R. D. Jones, C. W. Barnes, G. W. Flake, M. K. O'Rourke, K. Lee, H. H. Chen, G. Z. Sun, Y. Q. Zhang, D. Chen, and C. L. Giles. Adaptive stochastic cellular automata: Theory. *Phyisca D*, 666, 1990.

[642] H. C. Leung. *The Use of Artificial Neural Networks for Phonetic Recognition.* PhD thesis, MIT Dept. of Electrical Engineering and Computer Science, May 1989.

[643] E. Levin, N. Tishby, and S. Solla. A statistical approach to learning and generalization in neural networks. In *Proc. 3rd Annu. Workshop on Comput. Learning Theory.* Morgan Kaufmann, 1989.

[644] E. Levin, N. Tishby, and S. A. Solla. A statistical approach to learning and generalization in layered neural networks. *Proceedings of the IEEE – Special issue on Neural Networks*, 1990.

[645] B. Levine. Derivatives of tree sets with applications to grammatical inference. *IEEE Trans. Pattern Anal. Mach. Intell.*, PAMI-3:285–293, 1981.

[646] B. Levine. The use of tree derivatives and a sample support parameter for inferring tree systems. *IEEE Trans. Pattern Anal. Mach. Intell.*, PAMI-4:25–34, 1982.

[647] S. E. Levinson, L. R. Rabiner, and M. M. Sondhi. An introduction to the application of the theory of probabilistic functions of a markov process to automatic speech recognition. *Bell System Technical Journal*, 62(4):1035–1074, Apr. 1983.

[648] M. Li. Towards a DNA sequencing theory (learning a string). In *Proc. of the 31st Symposium on the Foundations of Comp. Sci.*, pages 125–134. IEEE Computer Society Press, Los Alamitos, CA, 1990.

[649] M. Li and U. Vazirani. On the learnability of finite automata. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 359–370, San Mateo, CA, 1988. Morgan Kaufmann.

[650] M. Li and P. Vitanyi. Inductive reasoning and kolmogorov complexity. *J. Comput. Syst. Sci.*, 44:343–384, 1992.

[651] M. Li and P. Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications.* Text and Monographs in Computer Science. Springer-Verlag, 1993.

[652] M. Li and P. M. B. Vitanyi. A theory of learning simple concepts under simple distributions and average case complexity for the universal distribution. In *Proc. 30th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 34–39. IEEE Computer Society Press, Los Alamitos, CA, 1989.

[653] M. Li and P. M. B. Vitanyi. Learning simple concepts under simple distributions. *SIAM J. Comput.*, 20:911–935, 1991.

[654] J. Lin and J. S. Vitter. Complexity issues in learning by neural nets. In *Proc. 2nd Annu. Workshop on Comput. Learning Theory*, pages 118–133, San Mateo, CA, 1989. Morgan Kaufmann.

[655] J. Lin and J. S. Vitter. A theory for memory-based learning. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 103–115. ACM Press, New York, NY, 1992.

[656] L. Lin. Self-improving reactive agents: case studies of reinforcement learning frameworks. Technical Report CMU-CS-90-109, Carnegie Mellon Computer Science Department, Aug. 1990.

[657] N. Lincoln and J. Skrzypek. Synergy of clustering multiple backpropagation networks. In *NIPS 2*. Morgan Kaufmann, 1989.

[658] R. Lindner. *Algorithmische Erkennung*. PhD thesis, Friedrich-Schiller-Universitat, 1972.

[659] N. Linial, Y. Mansour, and N. Nisan. Constant depth circuits, Fourier transform, and learnability. In *Proc. of the 31st Symposium on the Foundations of Comp. Sci.*, pages 574–579. IEEE Computer Society Press, 1989.

[660] N. Linial, Y. Mansour, and R. L. Rivest. Results on learnability and the Vapnik-Chervonenkis dimension. *Inform. Comput.*, 90(1):33–49, Jan. 1991.

[661] R. P. Lippmann. An introduction to computing with neural nets. *IEEE ASSP Magazine*, pages 4–22, Apr. 1987.

[662] N. Littlestone. Learning in a layered network with many fixed-function hidden nodes. In *Proceedings of the 1st Annual IEEE International Conference on Neural Networks*, volume II, pages 471–478, Washington, D.C., June 1987. IEEE Computer Society Press.

[663] N. Littlestone. Learning quickly when irrelevant attributes abound: A new linear-threshold algorithm. *Machine Learning*, 2:285–318, 1988.

[664] N. Littlestone. Learning when irrelevant attributes abound: A new linear-threshold algorithm. *Machine Learning*, 2:285–318, 1988.

[665] N. Littlestone. From on-line to batch learning. In *Proc. 2nd Annu. Workshop on Comput. Learning Theory*, pages 269–284, San Mateo, CA, 1989. Morgan Kaufmann.

[666] N. Littlestone. *Mistake Bounds and Logarithmic Linear-threshold Learning Algorithms*. PhD thesis, Technical Report UCSC-CRL-89-11, University of California Santa Cruz, 1989.

[667] N. Littlestone. Redundant noisy attributes, attribute errors, and linear threshold learning using Winnow. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 147–156, San Mateo, CA, 1991. Morgan Kaufmann.

[668] N. Littlestone and P. Long. On-line learning with linear loss constraints. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 412–421. ACM Press, New York, NY, 1993.

[669] N. Littlestone, P. M. Long, and M. K. Warmuth. On-line learning of linear functions. In *Proc. of the 23rd Symposium on Theory of Computing*, pages 465–475. ACM Press, New York, NY, 1991.

[670] N. Littlestone and M. K. Warmuth. The weighted majority algorithm. In *Proc. 30th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 256–261. IEEE Computer Society Press, Los Alamitos, CA, 1989.

[671] P. M. Long. *Towards a more comprehensive theory of learning in computers*. PhD thesis, UC Santa Cruz, 1992.

[672] P. M. Long and M. K. Warmuth. Composite geometric concepts and polynomial predictability. *Inform. Comput.*, 1993. To appear.

[673] W. Maass. On-line learning with an oblivious environment and the power of randomization. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 167–175, San Mateo, CA, 1991. Morgan Kaufmann.

[674] W. Maass. Bounds for the computational power and learning complexity of analog neural nets. In *Proc. 25th Annu. ACM Sympos. Theory Comput.*, pages 335–344. ACM Press, New York, NY, 1993.

[675] W. Maass. Efficient agnostic PAC-learning with simple hypotheses. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 67–75. ACM Press, New York, NY, 1994.

[676] W. Maass. On the complexity of learning on neural nets. In *Computational Learning Theory: Eurocolt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 1–17, Oxford, 1994. Oxford University Press.

[677] W. Maass, G. Schnitger, and E. D. Sontag. On the computational power of sigmoid versus Boolean threshold circuits. In *Proc. of the 32nd Symposium on the Foundations of Comp. Sci.*, pages 767–776. IEEE Computer Society Press, Los Alamitos, CA, 1991.

[678] W. Maass and G. Turán. On the complexity of learning from counterexamples. In *Proc. 30th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 262–267. IEEE Computer Society Press, Los Alamitos, CA, 1989.

[679] W. Maass and G. Turán. On the complexity of learning from counterexamples and membership queries. In *Proc. of the 31st Symposium on the Foundations of Comp. Sci.*, pages 203–210. IEEE Computer Society Press, Los Alamitos, CA, 1990.

[680] W. Maass and G. Turán. How fast can a threshold gate learn? In *Computational Learning Theory and Natural Learning Systems: Constraints and Prospects*. MIT Press, 1992. Previous versions appeared in FOCS89 and FOCS90.

[681] W. Maass and G. Turán. Lower bound methods and separation results for on-line learning models. *Machine Learning*, 9:107–145, 1992.

[682] A. Macintyre and E. D. Sontag. Finiteness results for sigmoid. In *Proc. 25th Annu. ACM Sympos. Theory Comput.*, pages 325–334. ACM Press, New York, NY, 1993.

[683] D. MacKay. *Bayesian Methods for Adaptive Models*. PhD thesis, California Institute of Technology, 1992.

[684] P. Maes. How to do the right thing. *Connection Science Journal, Special Issue on Hybrid Systems*, 1, 1990.

[685] P. Maes and R. A. Brooks. Learning to coordinate behaviors. In *Proceedings of AAAI-90*, volume 2, pages 796–802, Boston, MA, July 1990. AAAI Press/The MIT Press.

[686] S. Mahadevan and J. Connell. Automatic programming of behavior-based robots using reinforcement learning. Technical report, IBM Research at Yorktown Heights, Dec. 1990.

[687] J. Makhoul, S. Roucos, and H. Gish. Vector quantization in speech coding. *Proceedings of the IEEE*, 73(11):1551–1588, Nov. 1985.

[688] O. Maler and A. Pneuli. On the learnability of infinitary regular sets. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 128–136, San Mateo, CA, 1991. Morgan Kaufmann.

[689] Y. Mansour. Learning via fourier transform. (see KushilevitzMa91), Apr. 1990.

[690] A. Marron. Learning pattern languages from a single initial example and from queries. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 345–358, San Mateo, CA, 1988. Morgan Kaufmann.

[691] A. Marron and K. Ko. Identification of pattern languages from examples and queries. *Inform. Comput.*, 74(2):91–112, Aug. 1987.

[692] O. Maruyama and S. Miyano. Inferring a tree from walks. Technical Report RIFIS-TR-CS-51, Kyushu University, Dec. 1991.

[693] F. J. Maryanski and T. L. Booth. Inference of finite-state probabilistic grammars. *IEEE Trans. on Computers*, C-26:521–536, 1977.

[694] R. J. Matthews and W. Demopoulos. *Learnability and Linguistic Theory.* Kluwer, 1989.

[695] J. L. McClelland and D. E. Rumelhart, editors. *Explorations in Parallel Distributed Processing: A Handbook of Models, Programs, and Exercises.* MIT Press, 1988.

[696] J. L. McClelland, D. E. Rumelhart, and t. P. R. Group. *Parallel Distributed Processing — Explorations in the Microstructure of Cognition*, volume 2. MIT Press, 1986.

[697] R. Meir and J. F. Fontanari. On learning noisy theshold functions with finite precision weights. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 280–286. ACM Press, New York, NY, 1992.

[698] N. Merhav and M. Feder. Universal sequential learning and decisions from individual data sequences. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 413–427. ACM Press, New York, NY, 1992.

[699] A. Meyer. Program size in restricted programming languages. *Inform. Control*, 21:382–394, 1972.

[700] R. Michalski. Pattern recognition as rule guided inductive inference. *IEEE Trans. Pattern Anal. Mach. Intell.*, PAMI-2:349–361, 1980.

[701] R. M. Michalski. *Understanding the Nature of Learning: Issues and Research Directions*, pages 3–26. Morgan Kaufmann, Los Altos, CA, 1986.

[702] R. S. Michalski. A theory and methodology of inductive inference. In *Machine Learning Vol2*, pages 215–244. Morgan Kaufman, 1986.

[703] R. S. Michalski, J. G. Carbonell, and T. M. Mitchell, editors. *Machine Learning: An Artificial Intelligence Approach 1*, volume I. Morgan Kaufmann, Los Altos, California, 1983.

[704] R. S. Michalski, J. G. Carbonell, and T. M. Mitchell, editors. *Machine Learning: An Artificial Intelligence Approach 2*, volume II. Morgan Kaufmann, Los Altos, California, 1986.

[705] C. J. Michalski, R and T. Mitchell. *Machine Learning.* Tioga Publishing Co., Palo Alto, Calif., 1983.

[706] D. Michie. Inductive rule generation in the context of the fifth generation. In *Proceedings of the First International Workshop on Machine Learning*, pages 66–70, Monticello, Illinois, June 1983.

[707] L. Miclet. Regular inference with a tail-clustering method. *IEEE Trans. on Systems, Man, and Cybernetics*, SMC-10:737–743, 1980.

[708] W. T. Miller, F. H. Glanz, and L. G. Kraft. Application of a general learning algorithm to the control of robotic manipulators. *International Journal of Robotics Research*, 6(2):84–98, 1987.

[709] A. Milosavljevi'c, D. Haussler, and J. Jurka. Informed parsimonious inference of prototypical genetic sequences. In *Proc. 2nd Annu. Workshop on Comput. Learning Theory*, pages 102–117, San Mateo, CA, 1989. Morgan Kaufmann.

[710] E. Minicozzi. Some natural properties of strong identification in inductive inference. *Theoret. Comput. Sci.*, 2:345–360, 1976.

[711] M. Minsky and S. Papert. *Perceptrons.* MIT Press, Cambridge, MA, 1969.

[712] T. M. Mitchell. Version spaces: A candidate elimination approach to rule learning. In *Proceedings IJCAI-77*, pages 305–310, Cambridge, Mass., Aug. 1977. International Joint Committee for Artificial Intelligence.

[713] T. M. Mitchell. The need for biases in learning generalizations. Technical Report CBM-TR-117, Rutgers University, New Brunswick, NJ, 1980.

[714] T. M. Mitchell. Generalization as search. *Art. Int.*, 18:203–226, 1982.

[715] S. Miyano, A. Shinohara, and T. Shinohara. Which classes of elementary formal systems are polynomial-time learnable? In *Proceedings of the Second Workshop on Algorithmic Learning Theory*, pages 139–150. Japanese Society for Artificial Intelligence, 1991.

[716] J. Moody. Fast learning in multi-resolution hierarchies. In *Advances in Neural Information Processing Systems I*, pages 29–39, San Mateo, CA, 1989. Morgan Kaufmann.

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

[718] T. Motoki. Inductive inference from all positive and some negative data. Unpublished, 1992.

[719] T. Motoki, T. Shinohara, and K. Wright. The correct definition of finite elasticity: corrigendum to Identification of unions. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, page 375, San Mateo, CA, 1991. Morgan Kaufmann.

[720] M. C. Mozer and J. Bachrach. Discovering the structure of a reactive environment by exploration. Technical report, University of Colorado Department of Computer Science, Nov. 1989. Report CU-CS-451-89.

[721] S. Muggleton. Inductive logic programming. *New Generation Computing*, 8:295–318, 1991.

[722] S. Muggleton. Bayesian inductive logic programming. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 3–11. ACM Press, New York, NY, 1994.

[723] S. Muggleton. Recent advances in inductive logic programming. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 3–11. ACM Press, New York, NY, 1994.

[724] Y. Mukouchi. Characterization of pattern languages. In *Proceedings 2nd Workshop on Algorithmic Learning Theory*, pages 93–104, 1991.

[725] Y. Mukouchi. Characterization of finite identification. In *Proceedings of the Third International Workshop on Analogical and Inductive Inference, Dagstuhl Castle, Germany*, pages 260–267, Oct. 1992.

[726] Y. Mukouchi and S. Arikawa. Inductive inference machines that can refute hypothesis spaces. Technical Report RIFIS-TRCS-67, RIFIS, Kyushu University, 1993.

[727] K. S. Narendra and M. A. L. Thathachar. Learning automata – A survey. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-4(4):323–334, 1974.

[728] K. S. Narendra and M. A. L. Thathachar. *Learning Automata – An Introduction.* Prentice Hall, 1989.

[729] B. K. Natarajan. Learning functions from examples. Technical Report CMU-RI-TR-87-19, Robotics Institute, Carnegie-Mellon University, 1987.

[730] B. K. Natarajan. On learning Boolean functions. In *Proc. 19th Annu. ACM Sympos. Theory Comput.*, pages 296–304. ACM Press, New York, NY, 1987.

[731] B. K. Natarajan. On learning from exercises. In *Proc. 2nd Annu. Workshop on Comput. Learning Theory*, pages 72–87, San Mateo, CA, 1989. Morgan Kaufmann.

[732] B. K. Natarajan. On learning sets and functions. *Machine Learning*, 4(1), 1989.

[733] B. K. Natarajan. Probably-approximate learning over classes of distributions. Technical Report HPL-SAL-89-29, Hewlett Packard Labs, Palo Alto, CA, 1989.

[734] B. K. Natarajan. Some results on learning. Technical Report CMU-RI-TR-89-6, Carnegie Mellon, 1989.

[735] B. K. Natarajan. *Machine Learning: A Theoretical Approach.* Morgan Kaufmann, San Mateo, CA, 1991.

[736] B. K. Natarajan. Probably approximate learning of sets and functions. *SIAM J. Comput.*, 20(2):328–351, Apr. 1991.

[737] B. K. Natarajan. Occam's razor for functions. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 370–376. ACM Press, New York, NY, 1993.

[738] B. K. Natarajan and P. Tadepalli. Two new frameworks for learning. In *Proceedings of the 5th International Conference on Machine Learning*, pages 402–415, San Mateo, CA, June 1988. published by Morgan Kaufmann.

[739] R. M. Neal. Learning stochastic feedforward networks. Technical report, Department of Computer Science, University of Toronto, Nov. 1990.

[740] R. M. Neal. Bayesian training of backpropagation networks by the hybrid monte carlo method. Technical Report CRG-TR-92-1, Department of Computer Science, U. of Toronto, Apr. 1992.

[741] T. Niblett. Constructing decision trees in noisy domains. In *Progress in Machine Learning–Proceedings of EWSL 87: 2nd European Working Session on Learning*, pages 67–78, Bled, Yugoslavia, May 1987.

[742] N. J. Nilsson. *Learning Machines.* McGraw-Hill, New York, 1965.

[743] N. J. Nilsson. *The Mathematical Foundations of Learning Machines.* Morgan Kaufmann, 1990. (Republication of 1965 edition, with a new introduction by T.J. Sejnowski and H. White.).

[744] R. Nix. *Editing By Example (Ph.D Thesis).* PhD thesis, Yale University, 1983. Computer Science TR-280.

[745] R. Nix. Editing by example. In *Proc. 11th ACM Symposium on Principles of Programming Languages*, pages 186–195. ACM Press, 1984.

[746] A. Nobel. Evaluating the performance of a simple inductive procedure in the presence of overfitting error. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 267–274, San Mateo, CA, 1991. Morgan Kaufmann.

[747] E. Oja. Neural networks, principle components, and subspaces. *Internat. J. Neural Syst.*, 1(1):61–68, 1989.

[748] S. Omohundro. Efficient algorithms with neural networks behavior. *Complex Systems*, 1:273–347, 1987.

[749] M. Opper and D. Haussler. Calculation of the learning curve of Bayes optimal classification algorithm for learning a perceptron with noise. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 75–87, San Mateo, CA, 1991. Morgan Kaufmann.

[750] P. Orponen and R. Greiner. On the sample complexity of finding good search strategies. In *Proc. 3rd Annu. Workshop on Comput. Learning Theory*, pages 352–358, San Mateo, CA, 1990. Morgan Kaufmann.

[751] D. Osherson. Computer output. *Cognition*, 20:261–264, 1985.

[752] D. Osherson, M. Stob, and S. Weinstein. Ideal learning machines. *Cognitive Science*, 6:277–290, 1982.

[753] D. Osherson, M. Stob, and S. Weinstein. Learning strategies. *Inform. Control*, 53:32–51, 1982.

[754] D. Osherson, M. Stob, and S. Weinstein. Environments for learning. *Inform. Control*, 1983.

[755] D. Osherson, M. Stob, and S. Weinstein. Note on a central lemma of learning theory. *J. of Math. Psych.*, 27:86–92, 1983.

[756] D. Osherson, M. Stob, and S. Weinstein. Aggregating inductive expertise. *Inform. Control*, 70:69–95, 1986.

[757] D. Osherson, M. Stob, and S. Weinstein. An analysis of a learning paradigm. In *Language Learning and Concept Acquisition.* Ablex Publ. Co., 1986.

[758] D. Osherson, M. Stob, and S. Weinstein. *Systems That Learn.* MIT Press, Cambridge, MA, 1986.

[759] D. Osherson, M. Stob, and S. Weinstein. Synthesising inductive expertise. *Inform. Comput.*, 77:138–161, 1988.

[760] D. Osherson and S. Weinstein. A note on formal learning theory. *Cognition*, 11:77–88, 1982.

[761] D. Osherson and S. Weinstein. Formal learning theory. In *Handbook of Cognitive Neurology,New York, Plenum.* 1983.

[762] D. Osherson and S. Weinstein. Models of language acquisition. In *The Biology of Learning*. 1984.

[763] D. Osherson and S. Weinstein. On the study of first language acquisition. Manuscript, 1992.

[764] D. N. Osherson, M. Stob, and S. Weinstein. Learning theory and natural language. *Cognition*, 17:1–28, 1984.

[765] D. N. Osherson, M. Stob, and S. Weinstein. *Systems that Learn: An Introduction to Learning Theory for Cognitive and Computer Scientists.* MIT Press, 1986.

[766] D. N. Osherson, M. Stob, and S. Weinstein. Mathematical/mechanical? learners pay a price for Bayesianism. *J. Symbolic Logic*, 53(4):1245–1251, 1988.

[767] D. N. Osherson, M. Stob, and S. Weinstein. On approximate truth. In *Proc. 2nd Annu. Workshop on Comput. Learning Theory*, pages 88–101, San Mateo, CA, 1989. Morgan Kaufmann.

[768] D. N. Osherson, M. Stob, and S. Weinstein. A mechanical method of successful ascientific inquiry. In *Proc. 3rd Annu. Workshop on Comput. Learning Theory*, pages 187–201, San Mateo, CA, 1990. Morgan Kaufmann.

[769] D. N. Osherson and S. Weinstein. Criteria of language learning. *Inform. Control*, 52:123–138, 1982.

[770] D. N. Osherson and S. Weinstein. Identification in the limit of first order structures. *Journal of Philosophical Logic*, 15:55–81, 1986.

[771] G. Pagallo and D. Haussler. A greedy method for learning μ-DNF functions under the uniforn distribution. Technical report, University of California at Santa Cruz, UCSC-CRL-89-12, 1989.

[772] G. Pagallo and D. Haussler. Boolean feature discovery in empirical learning. *Machine Learning*, 5(1):71–100, 1990.

[773] T. W. Pao. A solution of the syntactical induction-inference problem for regular languages. *Comput. Lang.*, 3:53–64, 1978.

[774] T. W. Pao and J. W. I. Carr. A solution of the syntactical induction-inference problem for regular languages. *Comput. Lang.*, 3:53–64, 1978.

[775] G. Pask and H. V. Foerster. A predictive model for self organizing systems, parts I and II. *Cybernetica*, III and IV:258–300 and 20–55, 1960 and 1961.

[776] R. Paturi, S. Rajasekaran, and J. Reif. The light bulb problem. In *Proc. 2nd Annu. Workshop on Comput. Learning Theory*, pages 261–268, San Mateo, CA, 1989. Morgan Kaufmann.

[777] R. Paturi and M. E. Saks. On threshold circuits for parity. In *Proc. of the 31st Symposium on the Foundations of Comp. Sci.*, pages 397–404. IEEE Computer Society Press, Los Alamitos, CA, 1990.

[778] J. Pearl. On the connection between the complexity and credibility of inferred models. *Journal of General Systems*, 4:255–264, 1978.

[779] J. Pearl. Capacity and error estimates for boolean classifiers with limited complexity. *IEEE Trans. Pattern Anal. Mach. Intell.*, PAMI-1(4):350–355, Oct. 1979.

[780] J. Pearl. Jeffrey's rule and the problem of autonomous inference agents. Class notes for 274A, Fall 1984, 1984.

[781] J. Pearl. Bayes decision methods. Technical Report CSD-850023, UCLA Computer Science Department, June 1985.

[782] J. Pearl. Bayesian networks: A model of self-activated memory for evidential reasoning. Technical Report CSD-850021, R-43, UCLA Computer Science Department, June 1985.

[783] J. Pearl. How to do with probabilities what people say you can't. Technical Report CSD-850031, R-49, UCLA Computer Science Department, Sept. 1985.

[784] J. Pearl. On evidential reasoning in A hierarchy of hypothesis. Technical Report CSD-850032, UCLA Computer Science Department, Sept. 1985.

[785] J. Pearl and R. Dechter. Learning structure from data: a survey. In *Proc. 2nd Annu. Workshop on Comput. Learning Theory*, pages 230–244, San Mateo, CA, 1989. Morgan Kaufmann.

[786] J. Pearl and A. Paz. GRAPHOIDS: A graph-based logic for reasoning about relevance relations, or when would x tell you more about y if you already know z. Technical report, UCLA Computer Science Department, Dec. 1985.

[787] J. Pearl and A. Paz. On the logic of representing dependencies by graphs. In *Proceedings 1986 Canadian AI Conference*, May 1986.

[788] J. Pearl and T. S. Verma. A formal theory of inductive causation. Technical Report R-1555 (I), UCLA Cognitive Systems Laboratory, Oct. 1990.

[789] F. Pereira. Frequencies vs. biases: machine learning problems in natural language processing. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, page 12. ACM Press, New York, NY, 1994.

[790] S. Persson. *Some Sequence Extrapolating Programs: a Study of Representation and Modeling in Inquiring Systems*. PhD thesis, Stanford University, 1966.

[791] N. Perugini. Neural network learning: Effects of network and training set size. Master's thesis, MIT Department of Electrical Engineering and Computer Science, June 1989.

[792] C. Peterson and J. R. Anderson. A mean field theory learning algorithm for neural networks. *Complex Systems*, 1:995–1019, 1987.

[793] K. Pillaipakkamnatt and V. Raghavan. On the limits of proper learnability of subclasses of DNF formulas. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 118–129. ACM Press, New York, NY, 1994.

[794] K. Pillaipakkamnatt and V. Raghavan. Read-twice DNF formulas are properly learnable. In *Computational Learning Theory: Eurocolt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 121–132, Oxford, 1994. Oxford University Press.

[795] S. Pinker. Formal models of language learning. *Cognition*, 7:217–283, 1979.

[796] S. Pinker. Language acquisition. In *Foundations of Cognitive Science*. The MIT Press, Cambridge, 1990.

[797] S. Pinker and A. Prince. On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. Technical Report Occasional Paper #33, MIT Center for Cognitive Science, 1987.

[798] L. Pitt. A characterization of probabilistic inference. In *Proc. 26th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 485–494. IEEE Computer Society Press, 1984.

[799] L. Pitt. *A characterization of probabilistic inference (Ph.D Thesis).* PhD thesis, Yale University, 1984.

[800] L. Pitt. *Probabilistic Inductive Inference (Ph.D Thesis).* PhD thesis, Yale University, 1985. Computer Science Dept. TR-400.

[801] L. Pitt. Inductive Inference, DFAs, and Computational Complexity. In *Proceedings of AII-89 Workshop on Analogical andInductive Inference; Lecture Notes in Artificial Intelligence 397*, pages 18–44, Heidelberg, Oct. 1989. Springer-Verlag.

[802] L. Pitt. Probabilistic inductive inference. *J. ACM*, 36(2):383–433, 1989.

[803] L. Pitt and R. Board. Semi-supervised learning. Technical report, University of Illinois, UIUCDCS-R-87-1372, 1987.

[804] L. Pitt and R. Board. On the necessity of Occam algorithms. In *Proc. 23th Annu. ACM Sympos. Theory Comput.*, pages 54–63. ACM Press, 1990.

[805] L. Pitt and R. E. Reinke. Criteria for polynomial-time (conceptual) clustering. *Machine Learning*, 2:371–396, 1988.

[806] L. Pitt and R. E. Reinke. Criteria for polynomial time (conceptual) clustering. *Machine Learning*, 2:371–396, 1988.

[807] L. Pitt and C. Smith. Probability and plurality for aggregations of learning machines. *Inform. Comput.*, 77:77–92, 1988.

[808] L. Pitt and L. Valiant. Computational limitations on learning from examples. *J. ACM*, 35:965–984, 1988.

[809] L. Pitt and M. Warmuth. The minimum consistent DFA problem cannot be approximated within any polynomial. *J. ACM*, 40(1):95–142, 1993.

[810] L. Pitt and M. K. Warmuth. Prediction preserving reducibility. *J. of Comput. Syst. Sci.*, 41(3):430–467, Dec. 1990. Special issue of the for the *Third Annual Conference of Structure in Complexity Theory* (Washington, DC., June 88).

[811] M. Pivar and M. Finkelstein. Automation, using LISP, of inductive inference on sequences. In *The Programming Language LISP*, pages 125–136. MIT Press, 1964.

[812] M. Pivar and E. Gord. The LISP program for inductive inference on sequences. In *The Programming Language LISP*, pages 260–289. MIT Press, 1964.

[813] G. D. Plotkin. A note on inductive generalization. In *Machine Intelligence 5*, pages 153–163. American Elsevier, 1970.

[814] G. D. Plotkin. *Automatic Methods of Inductive Inference*. PhD thesis, Edinburgh University, 1971.

[815] G. D. Plotkin. A further note on inductive generalization. In *Machine Intelligence 6*, pages 101–124. American Elsevier, 1971.

[816] K. Podnieks. Comparing various concepts of function prediction, Part 1. *Theory of Algorithms and Programs, Latvian State University, Riga*, 210:68–81, 1974.

[817] K. Podnieks. Comparing various concepts of function prediction, Part 2. *Theory of Algorithms and Programs, Latvian State University, Riga*, 233, 1975.

[818] K. Podnieks. Probabilistic program synthesis. *Theory of Algorithms and Programs, Latvian State University, Riga*, 16:57–88, 1977.

[819] K. M. Podnieks. Probabilistic synthesis of enumerated classes of functions. *Soviet Math. Doklady*, 16:1042–1045, 1975.

[820] T. Poggio and F. Girosi. Networks and the best approximation property. Technical Report AIM1164, MIT Artificial Intelligence Laboratory, Oct. 1989.

[821] T. Poggio and F. Girosi. Representation propoerties of networks: Kolmogorov's theorm is irrelevant. *Neural Computation*, 1:465–469, 1989.

[822] T. Poggio and F. Girosi. A theory of networks for approximation and learning. Technical Report AIM1140, MIT Artificial Intelligence Laboratory, July 1989.

[823] T. Poggio and F. Girosi. Extensions of a theory of networks for approximation and learning: dimensionality and reduction and clustering. Technical Report AIM1167, MIT Artificial Intelligence Laboratory, Apr. 1990.

[824] T. Poggio and F. Girosi. Regularization algorithms for learning that are equivalent to multilayer networks. *Science*, 247:978–982, Feb. 23 1990.

[825] D. Pollard. *Convergence of Stochastic Processes*. Springer-Verlag, 1984.

[826] S. Pölt. Improved sample size bounds for PAB-decisions. In *Computational Learning Theory: Eurocolt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 229–239, Oxford, 1994. Oxford University Press.

[827] S. Porat. Stability and looping in connectionist models with assymmetric weights. Technical Report TR 210, University of Rochester Computer Science Department, Mar. 1987.

[828] S. Porat and J. A. Feldman. Learning automata from ordered examples. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 386–396, San Mateo, CA, 1988. Morgan Kaufmann.

[829] S. E. Posner and S. R. Kulkarni. On-line learning of functions of bounded variation under various sampling schemes. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 439–445. ACM Press, New York, NY, 1993.

[830] M. J. D. Powell. Radial basis functions for multivariable interpolation: A review. In *Algorithms for Approximation*, pages 143–167, Oxford, 1987. Clarendon Press.

[831] P. Pudlak and F. N. Springsteel. Complexity of mechanized hypothesis formation. *Theoret. Comput. Sci.*, 8:203–225, 1979.

[832] H. Putnam. Probability and confirmation. In *Mathematics, Matter and Method*. Cambridge University Press, 1975.

[833] S. Qian, Y. C. Lee, R. D. Jones, C. W. Barnes, G. W. Flake, M. K. O'Rourke, K. Lee, H. H. Chen, G. Z. Sun, Y. Q. Zhang, D. Chen, and C. L. Giles. Adaptive stochastic cellular automata: Theory. *Phyisca D*, 666, 1990.

[834] J. R. Quinlan. Inferno: A cautious approach to uncertain inference. *The Computer Journal*, 26(3):255–269, 1983.

[835] J. R. Quinlan. The effect of noise on concept learning. In *Machine Learning, An Artificial Intelligence Approach (Volume II)*, chapter 6, pages 149–166. Morgan Kaufmann, 1986.

[836] J. R. Quinlan. Induction of decision trees. *Machine Learning*, 1:81–106, 1986.

[837] J. R. Quinlan. Simplifying decision trees. *International Journal of Man-Machine Studies*, 1987. (To appear.).

[838] J. R. Quinlan and R. L. Rivest. Inferring decision trees using the Minimum Description Length Principle. *Inform. Comput.*, 80(3):227–248, Mar. 1989. (An early version appeared as MIT LCS Technical report MIT/LCS/TM-339 (September 1987).).

[839] L. R. Rabiner and B. H. Juang. An introduction to hidden markov models. *IEEE ASSP Magazine*, 3(1):4–16, Jan. 1986.

[840] L. R. Rabiner, S. E. Levinson, and M. M. Sondhi. On the application of vector quantization and hidden markov models to speaker-independent, isolated word recognition. *Bell System Technical Journal*, 62(4):1075–1105, Apr. 1983.

[841] P. Raghavan. Learning in threshold networks. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 19–27, San Mateo, CA, 1988. Morgan Kaufmann.

[842] V. Raghavan. Bounded degree graph inference from walks. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 354–366. Morgan Kaufmann, 1991.

[843] V. Raghavan and S. R. Schach. Learning switch configurations. In *Proc. 3rd Annu. Workshop on Comput. Learning Theory*, pages 38–51, San Mateo, CA, 1990. Morgan Kaufmann.

[844] V. Raghavan and D. Wilkins. Learning μ-branching programs with queries. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 27–36. ACM Press, New York, NY, 1993.

[845] N. S. V. Rao, E. M. Oblow, C. W. Glover, and G. E. Liepins. N-learners problem: Fusion of concepts. Technical Report ORNL/TM-11904, CESAR-91/23, Oak Ridge National Laboratory, Sept. 1991.

[846] G. N. Reeke, Jr, O. Sporns, and G. M. Edelman. Synthetic neural modelling: Comparisons of population and connectionist approaches. In *Connectionism in Perspective*. Elsevier Science B. V. (North-Holland), 1989.

[847] J. Rissanen. Modeling by shortest data description. *Automatica*, 14:465–471, 1978.

[848] J. Rissanen. A universal prior for integers and estimation by minimum description length. *Ann. of Statist.*, 11(2):416–431, 1983.

[849] J. Rissanen. Stochastic complexity and modeling. *Ann. of Statist.*, 14(3):1080–1100, 1986.

[850] J. Rissanen. Stochastic complexity and sufficient statistics. Technical report, IBM Research Laboratory (San Jose), 1986.

[851] J. Rissanen. *Stochastic Complexity in Statistical Inquiry*, volume 15 of *Series in Computer Science*. World Scientific, 1989.

[852] R. Rivest, D. Haussler, and M. K. Warmuth, editors. *Proc. 2nd Annu. Workshop on Comput. Learning Theory*. Morgan Kaufmann, Inc., San Mateo, CA, 1989.

[853] R. Rivest and R. Schapire. Inference of visible simple assignment automata with planned experiments. Preprint, MIT Laboratory for Computer Science., 1987.

[854] R. L. Rivest. Learning decision lists. *Machine Learning*, 2(3):229–246, 1987.

[855] R. L. Rivest. Theory of learning: What's hard and what's easy to learn. In *Research Directions in Computer Science: An MIT Perspective*, chapter 11, pages 217–227. MIT Press, 1991.

[856] R. L. Rivest and W. Remmele. Machine learning: the human connection. *Siemens Review*, 55(2/88):33–38, Mar./Apr. 1988.

[857] R. L. Rivest and W. Remmele. Machine learning. In *Applied Computer Science and Software: Turing Theory into Practice*, pages 186–200. Springer-Verlag, 1991.

[858] R. L. Rivest and R. E. Schapire. Diversity-based inference of finite automata. In *Proc. 28th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 78–87. IEEE Computer Society Press, Los Alamitos, CA, 1987.

[859] R. L. Rivest and R. E. Schapire. A new approach to unsupervised learning in deterministic environments. In *Proceedings of the 4th International Workshop on Machine Learning*, pages 364–375, San Mateo, CA, June 1987. Morgan Kaufmann.

[860] R. L. Rivest and R. E. Schapire. A new approach to unsupervised learning in deterministic environments (reprint). In *Machine Learning, An Artificial Intelligence Approach*, volume III, pages 670–684. Morgan Kaufmann, 1990.

[861] R. L. Rivest and R. E. Schapire. Inference of finite automata using homing sequences. *Inform. Comput.*, 103(2):299–347, Apr. 1993.

[862] R. L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public key cryptosytems. *Commun. ACM*, 21(2):120–126, 1978.

[863] R. L. Rivest and R. Sloan. Learning complicated concepts reliably and usefully. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 69–79, San Mateo, CA, 1988. Morgan Kaufmann.

[864] R. L. Rivest and R. Sloan. A new model for inductive inference. In *Proceedings of the Second Conference on Theoretical Aspects of Reasoning about Knowledge*, pages 13–27. Morgan Kaufmann, Mar. 1988.

[865] R. L. Rivest and P. Winston. Machine learning research at MIT. *Siemens Review*, pages 12–14, 1990.

[866] R. L. Rivest and Y. Yin. Simulation results for a new two-armed bandit heuristic. In *Proceedings 1990 Conference on Computation Learning and Natural Learning (Princeton)*, 1990. To appear.

[867] K. Romanik. Testing as A dual to learning. Technical Report UMIACS-TR-91.93, CS-TR-2704, Dept. of Computer Science, University of Maryland, June 1991.

[868] K. Romanik. Approximate testing and learnability. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 327–332. ACM Press, New York, NY, 1992.

[869] D. Ron and R. Rubinfeld. Learning fallible finite state automata. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 218–227. ACM Press, New York, NY, 1993.

[870] D. Ron, Y. Singer, and N. Tishby. Learning probabilistic automata with variable memory length. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 35–46. ACM Press, New York, NY, 1994.

[871] F. Rosenblatt. The perceptron: A probabilistic model for information storage and organization in the brain. *Psych. Rev.*, 65:386–407, 1958. (Reprinted in *Neurocomputing* (MIT Press, 1988).).

[872] V. Roychowdhury, K. Siu, A. Orlitsky, and T. Kailath. A geometric approach to threshold circuit complexity. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 97–111, San Mateo, CA, 1991. Morgan Kaufmann.

[873] J. Royer. On machine inductive inference of approximations. Technical Report 85-005, University of Chicago, 1985.

[874] J. Royer. Inductive inference of approximations. *Inform. Control*, 70:156–178, 1986.

[875] S. Rudich. Inferring the structure of a Markov chain from its output. In *Proc. 27th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 321–326, Oct. 1985.

[876] D. E. Rumelhart, G. E. Hinton, and J. L. McClelland. A general framework for parallel distributed processing. In *Parallel Distributed Processing (Volume I: Foundations)*, chapter 2, pages 45–76. MIT Press, 1986.

[877] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning internal representations by error propagation. In *Parallel Distributed Processing – Explorations in the Microstructure of Cognition*, chapter 8, pages 318–362. MIT Press, 1986.

[878] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning representations by back-propagating errors. *Nature*, 323:533–536, 1986.

[879] D. E. Rumelhart and J. L. McClelland, editors. *Parallel Distributed Processing (Volume I: Foundations)*. MIT Press, 1986.

[880] D. E. Rumelhart and D. Zipser. Feature discovery by competitive learning. *Cognitive Science*, 9:75–112, 1985.

[881] L. Saitta and F. Bergadano. Pattern recognition and valiant's learning framework. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15(2):145–155, Feb. 1993.

[882] Y. Sakai and A. Maruoka. Learning monotone log-term DNF formulas. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 165–172. ACM Press, New York, NY, 1994.

[883] Y. Sakakibara. Inductive inference of logic programs based on algebraic semantics. Technical report, IIAS-SIS, Fujitsu Limited, Research Report No. 79, 1987.

[884] Y. Sakakibara. Inferring parsers of context free languages from structural examples. Technical Report 81, Fujitsu Ltd., 1987.

[885] Y. Sakakibara. Inferring parsers of context-free languages from structural examples. Technical report, IIAS-SIS, Fujitsu Limited, Research Report No. 81, 1987.

[886] Y. Sakakibara. Learnability in the presence of classification noise. In *Fujitsu IIAS-SIS Workshop on Computational Learning Theory*, pages 1–27, Nov. 1989.

[887] Y. Sakakibara. An efficient robust algorithm for learning decision lists. Technical Report No. 105, International Institute for Advanced Study of Social Information Science, Fujitsu Laboratories Ltd, Aug. 1990.

[888] Y. Sakakibara. Learning context-free grammars from structural data in polynomial time. *Theoret. Comput. Sci.*, pages 223–242, 1990.

[889] Y. Sakakibara. *Algorithmic Learning of Formal Languages and Decision Trees*. PhD thesis, Tokyo Institute of Technology, Oct. 1991. (International Institute for Advanced Study of Social Information Science, Fujitsu Laboratories Ltd, Research Report IIAS-RR-91-22E).

[890] Y. Sakakibara. On learning from queries and counterexamples in the presence of noise. *Inform. Proc. Lett.*, 37(5):279–284, Mar. 1991.

[891] Y. Sakakibara and R. Siromoney. A noise model on learning sets of strings. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 295–302. ACM Press, New York, NY, 1992.

[892] S. Salzberg. Exemplar-based learning: theory and implementation. Center for Research in Computing Technology TR-10-88, Harvard University, Cambridge, Mass., Oct. 1988.

[893] C. Sammut and R. Banerji. Learning concepts by asking questions. In *Machine Learning : An AI Approach, Vol2*. Morgan Kaufman, 1986.

[894] A. L. Samuel. *Computers and thought*, chapter : Some studies in machine learning using the game of checkers. McGraw Hill, 1959.

[895] A. L. Samuel. Some studies in machine learning using the game of checkers. *IBM J. of Res. Develop.*, 3:211–229, July 1959. (Reprinted in *Computers and Thought*, (eds. E. A. Feigenbaum and J. Feldman), McGraw-Hill, 1963, pages 39–70).

[896] T. D. Sanger. Optimal unsupervised learning in a single-layer linear feedforward neural network. *Neural Networks*, 2:459–473, 1989.

[897] A. Saoudi. Learning local and recognizable ω-languages and monadic logic programs. In *Computational Learning Theory: Eurocolt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 157–169, Oxford, 1994. Oxford University Press.

[898] G. Schäfer-Richter. *Uber Eingabeabhangigkeit und Komplexitat von Inferenzstrategien*. PhD thesis, RWTH Aachen, 1984.

[899] G. Schäfer-Richter. Some results in the theory of effective program synthesis - learning by defective information. *Lecture Notes in Computer Science*, 215:219–225, 1986.

[900] R. Schapire and L. Sellie. Learning sparse multivariate polynomials over a field with queries and counterexamples. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 17–26. ACM Press, New York, NY, 1993.

[901] R. E. Schapire. Diversity-based inference of finite automata. Master's thesis, MIT Lab. for Computer Science, May 1988. Technical Report MIT/LCS/TR-413.

[902] R. E. Schapire. Pattern languages are not learnable. In *Proc. 3rd Annu. Workshop on Comput. Learning Theory*, pages 122–129, San Mateo, CA, 1990. Morgan Kaufmann.

[903] R. E. Schapire. The strength of weak learnability. *Machine Learning*, 5(2):197–227, 1990.

[904] R. E. Schapire. Learning probabilistic read-once formulas on product distributions. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 184–198, San Mateo, CA, 1991. Morgan Kaufmann.

[905] R. E. Schapire. *The Design and Analysis of Efficient Learning Algorithms.* MIT Press, Cambridge, MA, 1992.

[906] J. C. Schlimmer and R. H. Granger. Incremental learning from noisy data. *Machine Learning*, 1:317–354, 1986.

[907] D. B. Schwartz, R. E. Howard, and W. E. Hubbard. Adaptive neural networks using MOS charge storage. In *Proc. 1988 Neural Information Processing Conference*, pages 761–768, 1989.

[908] T. J. Sejnowski and C. R. Rosenberg. Parallel networks that learn to pronounce english text. *Journal of Complex Systems*, 1(1):145–168, Feb. 1987.

[909] H. S. Seung, M. Opper, and H. Sompolinsky. Query by committee. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 287–294. ACM Press, New York, NY, 1992.

[910] G. Shackelford and D. Volper. Learning k-DNF with noise in the attributes. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 97–103, San Mateo, CA, 1988. Morgan Kaufmann.

[911] E. Shamir. A remark on discovery algorithms for grammars. *Inform. Control*, 5:246–251, 1962.

[912] E. Shapiro. A general incremental algorithm that infers theory from facts. In *Internat. Joint Conf. on Artif. Intell.*, 1981.

[913] E. Y. Shapiro. Inductive inference of theories from facts. Technical Report Research Report 192, Yale University Department of Computer Science, Feb. 1981.

[914] N. Shapiro. Review of "Limiting recursion" by E.M. Gold and "Trial and error predicates and the solution to a problem of Mostowski" by H. Putnam. *J. Symbolic Logic*, 36:342, 1971.

[915] A. Sharma. *Program Size and Teams for Computational Learning.* PhD thesis, SUNY at Buffalo, 1990.

[916] S. W. Shaw, D and C. Green. Inferring LISP programs from example problems. In *Proc. 4th IJCAI*, pages 260–267. IJCAI, 1975.

[917] J. Shawe-Taylor, M. Anthony, and R. L. Biggs. Bounding sample size with the vapnik-chervonenkis dimension. Technical Report CSD-TR-618, University of London, Royal Halloway and New Bedford College, 1989.

[918] A. Shinohara and S. Miyano. Teachability in computational learning. *New Generation Computing*, 8:337–347, 1991.

[919] T. Shinohara. Polynomial time inference of extended regular pattern languages. In *RIMS Symposia on Software Science and Engineering, Kyoto, Japan*, pages 115–127. Springer Verlag, 1982. Lecture Notes in Computer Science 147.

[920] T. Shinohara. Polynomial time inference of pattern languages and its applications. Proceedings, 7th IBM Symposium on Mathematical Foundations of Computer Science, 1982.

[921] T. Shinohara. Some problems on inductive inference from positive data. *Lecture Notes in Computer Science*, 215:41–58, 1986.

[922] T. Shinohara. *Studies on Inductive Inference from Positive Data.* PhD thesis, Kyushu University, Kyushu, Japan, 1986.

[923] T. Shinohara. Inductive inference from positive data is powerful. In *Proc. 3rd Annu. Workshop on Comput. Learning Theory*, pages 97–110. Morgan Kaufmann, 1990.

[924] T. Shinohara. Inductive inference of monotonic formal systems from positive data. *New Generation Computing*, 8:371–384, 1991.

[925] T. Shinohara. Teachability in computational learning. *New Generation Computing*, 8:337–347, 1991.

[926] H. Shvaytser. Linear manifolds are learnable from positive examples. Unpublished manuscript, 1988.

[927] H. Shvaytser. Non-learnable classes of Boolean formulae that are closed under variable permutation. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 155–166, San Mateo, CA, 1988. Morgan Kaufmann.

[928] H. T. Siegelmann and E. D. Sontag. On the computational power of neural nets. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 440–449. ACM Press, New York, NY, 1992.

[929] L. Siklossy and D. Sykes. Automatic program synthesis for example problems. In *Proc. 4th IJCAI*, pages 268–273. IJCAI, 1975.

[930] H. A. Simon. Why should machines learn? In *Machine Learning, An Artificial Intelligence Approach.* Tioga, Palo Alto, California, 1983.

[931] H. A. Simon and K. Kotovsky. Human acquisition of concepts for sequential patterns. *Psych. Rev.*, 70:534–546, 1963.

[932] H. U. Simon. On the number of examples and stages needed for learning decision trees. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 303–313. Morgan Kaufmann, 1990.

[933] H. U. Simon. General bounds on the number of examples needed for learning probabilistic concepts. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 402–411. ACM Press, New York, NY, 1993.

[934] H. U. Simon. Bounds on the number of examples needed for learning functions. In *Computational Learning Theory: Eurocolt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 83–94, Oxford, 1994. Oxford University Press.

[935] T. A. Slaman and R. M. Solovay. When oracles do not help. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 379–383. Morgan Kaufmann, 1991.

[936] R. Sloan. Types of noise in data for concept learning. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 91–96, San Mateo, CA, 1988. Morgan Kaufmann.

[937] R. H. Sloan. Some notes on Chernoff bounds. (Unpublished), 1987.

[938] R. H. Sloan. *Computational Learning Theory: New Models and Algorithms.* PhD thesis, MIT, 1989. Issued as MIT/LCS/TR-448.

[939] R. H. Sloan. Corrigendum to Types of noise in data for concept learning. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, page 450. ACM Press, New York, NY, 1992.

[940] R. H. Sloan and G. Turán. Learning with queries but incomplete information. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 237–245. ACM Press, New York, NY, 1994.

[941] C. Smith. The power of pluralism for automatic program synthesis. *J. ACM*, 29:1144–1165, 1982.

[942] C. Smith and M. Velauthapillai. On the inference of approximate programs. Technical Report 1427, University of Maryland, 1985.

[943] C. Smith and M. Velauthapillai. On the inference of programs approximately computing the desired function. *Lecture Notes in Computer Science*, 265:164–176, 1986.

[944] C. H. Smith. *Hierarchies of Identification Criteria for Mechanized Inductive Inference.* PhD thesis, SUNY Buffalo, 1979.

[945] C. H. Smith. An inductive inference bibliography. Technical report, Purdue University Computer Science Dept., CSD TR 323, 1979.

[946] D. R. Smith. A class of synthesizeable LISP programs. Technical report, Duke University Computer Science Dept., CS-1977-4, 1977.

[947] D. R. Smith. A survey of the synthesis of LISP programs from examples. In *Automatic Program Construction Techniques.* MacMillan Co., 1982.

[948] R. J. Solomonoff. A formal theory of inductive inference: Part 1. *Inform. Control*, 7:1–22, 1964.

[949] R. J. Solomonoff. A formal theory of inductive inference: Part 2. *Inform. Control*, 7:224–254, 1964.

[950] R. J. Solomonoff. Inductive inference theory - a unified approach to problems in pattern recognition and artificial intelligence. In *Proc. 4th IJCAI*, pages 274–280. IJCAI, 1975.

[951] R. J. Solomonoff. Complexity-based induction systems: comparisons and convergence theorems. *IEEE Trans. Inform. Theory*, IT-24:422–432, 1978.

[952] H. Sompolinsky, H. S. Seung, and N. Tishby. Learning curves in large neural networks. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 112–127, San Mateo, CA, 1991. Morgan Kaufmann.

[953] H. Sompolinsky, N. Tishby, and H. S. Seung. Learning from examples in large neural networks. *Phys. Rev. Lett.*, 65:1683–1686, 1990.

[954] S. J. Stolfo. *Automatic Discovery Heuristics for Nondeterministic Programs From Sample Execution Traces.* PhD thesis, New York University, 1979.

[955] S. J. Stolfo and M. C. Harrison. Automatic discovery of heuristics for nondeterministic programs. Technical report, New York University Computer Science Dept., No. 7, 1979.

[956] P. D. Summers. *Program Construction From Examples.* PhD thesis, Yale University, 1976.

[957] P. D. Summers. A methodology for LISP program construction from examples. *J. ACM*, 24(1):161–175, Jan. 1977.

[958] R. S. Sutton. Learning to predict by the methods of temporal differences. *Machine Learning*, 3(1):9–44, Aug. 1988.

[959] R. S. Sutton. Reinforcement learning architectures for animats. In *First International Conference on Simulation of Adaptive Behavior*, 1991.

[960] R. S. Sutton and A. G. Barto. Toward a modern theory of adaptive networks: Expectation and prediction. *Psych. Rev.*, 88(2):135–170, 1981.

[961] R. S. Sutton and A. G. Barto. An adaptive network that constructs and uses an internal model of its world. , 1985.

[962] Z. S. Szabo. Stratified inductive hypothesis generation. *Lecture Notes in Computer Science*, 265:177–195, 1986.

[963] E. Takimoto and A. Maruoka. Conservativeness and monotonicity for learning algorithms. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 377–383. ACM Press, New York, NY, 1993.

[964] G. Tesauro. Scaling relationships in back-propagation learning: dependence on training set size. *Complex Systems*, 1:367–372, 1987.

[965] G. Tesauro. Neurogammon: A neural-network backgammon program. Technical Report RC 15619 (#69436), IBM T.J. Watson Research Center, Mar. 1990.

[966] G. Tesauro and D. Cohn. Can neural networks do better than the Vapnik-Chervonenkis bounds? In *Advances in Neural Information Processing, Vol. 3*, pages 911–917. Morgan Kaufmann, 1991.

[967] G. Tesauro and B. Janssens. Scaling relationships in back-propagation learning. *Complex Systems*, 2:39–44, 1988.

[968] G. Tesauro and T. J. Sejnowski. A 'neural' network that learns to play backgammon. *Artificial Intelligence*, 39(3):357–390, July 1989.

[969] G. Tesauro and T. J. Sejnowski. A parallel network that learns to play backgammon. *Artificial Intelligence*, in press, 1989.

[970] H. Thiele. Lernverfahren zur erkennung formaler sprachen. *Kybernetik-Forschung*, 3:11–93, 1973.

[971] H. Thiele. Zur charakterisierung von erkennungssytemen mit einbettendem konvergenzbegriff. *Kompliziertheit von Lern und Erkennungsprozessen*, 2:188–207, 1975. Jena.

[972] Y. Tikochinsky, N. Z. Tishby, and R. D. Levine. Consistent inference of probabilities for reproducible experiments. *Phys. Rev. Lett.*, 52(16):1357–1360, 1984.

[973] N. Tishby, E. Levin, and S. Solla. Consistent inference of probabilities in layered networks: predictions and generalizations. In *IJCNN International Joint Conference on Neural Networks*, volume II, pages 403–409. IEEE, 1989.

[974] J. N. Tsitsiklis. A lemma on the multiarmed bandit problem. *IEEE Transactions on Automatic Control*, AC-31(6), June 1986.

[975] G. Turán. Lower bounds for PAC learning with queries. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 384–391. ACM Press, New York, NY, 1993.

[976] W. Tzeng. The equivalence and learning of probabilistic automata. In *Proc. 30th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 268–273. IEEE Computer Society Press, Los Alamitos, CA, 1989.

[977] P. Utgoff. Shift of bias for inductive concept learning. In *Machine Learning Vol2*. Morgan Kaufman, 1987.

[978] P. E. Utgoff. *Machine Learning of Inductive Bias*. Kluwer Academic, 1986.

[979] L. Valiant. Deductive learning. *Phil. Trans. Roy. Soc. Lond. A*, 312:441–446, 1984.

[980] L. Valiant. Functionality in neural nets (at AAAI). In *AAAI*, pages 629–634, 1988.

[981] L. Valiant. A view of computational learning theory. In *Computation & Cognition: Proceedings of the First NEC Research Symposium*, pages 32–51. SIAM, 1991.

[982] L. Valiant. A neuroidal model for cognitive functions. In *Computational Learning Theory: Eurocolt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, page 2, Oxford, 1994. Oxford University Press.

[983] L. Valiant and M. Warmuth, editors. *Proc. 4th Annu. Workshop on Comput. Learning Theory*. Morgan Kaufmann, Inc., San Mateo, CA, 1991.

[984] L. G. Valiant. A theory of the learnable. *Commun. ACM*, 27(11):1134–1142, Nov. 1984.

[985] L. G. Valiant. Learning disjunctions of conjunctions. In *Proceedings of the 9th International Joint Conference on Artificial Intelligence, vol. 1*, pages 560–566, Los Angeles, California, 1985. International Joint Committee for Artificial Intelligence.

[986] L. G. Valiant. Functionality in neural networks. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 28–39, San Mateo, CA, 1988. Morgan Kaufmann.

[987] L. G. Valiant. A view of computational learning theory. In *Proceedings of the First NEC Research Symposium*, pages 32–51. Society for Industrial and Applied Mathematics, Philadelphia, Morgan Kaufmann, 1989.

[988] K. VanLehn. Learning one subprocedure per lesson. *Artificial Intelligence*, 31(1):1–40, Jan. 1987.

[989] V. N. Vapnik. *Theorie der Zeichenerkennung*. Akademie-Verlag, 1979.

[990] V. N. Vapnik. *Estimation of Dependences Based on Empirical Data*. Springer-Verlag, New York, 1982.

[991] V. N. Vapnik. Inductive principles of the search for empirical dependences (methods based on weak convergence of probability measures). In *Proc. 2nd Annu. Workshop on Comput. Learning Theory*, pages 3–21, San Mateo, CA, Aug. 1989. Morgan Kaufmann.

[992] V. N. Vapnik and A. Y. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probab. and its Applications*, 16(2):264–280, 1971.

[993] L. P. J. Veelenturf. Inference of sequential machines from sample computations. *IEEE Trans. on Computers*, C-27:167–170, 1978.

[994] M. Velauthapillai. *On the Inductive Inference of Programs with Anomalies.* PhD thesis, University of Maryland, 1986.

[995] M. Velauthapillai. Inductive inference with bounded number of mind changes. In *Proc. 2nd Annu. Workshop on Comput. Learning Theory*, pages 200–213, San Mateo, CA, 1989. Morgan Kaufmann.

[996] S. S. Venkatesh. On learning binary weights for majority functions. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 257–266. Morgan Kaufmann, 1991.

[997] S. S. Venkatesh, R. R. Snapp, and D. Psaltis. BELLMAN STRIKES AGAIN! The growth rate of sample complexity with dimension for the nearest neighbor classifier. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 93–102. ACM Press, New York, NY, 1992.

[998] K. Verbeurgt. Learning DNF under the uniform distribution in quasi-polynomial time. In *Proc. 3rd Annu. Workshop on Comput. Learning Theory*, pages 314–326, San Mateo, CA, 1990. Morgan Kaufmann.

[999] S. A. Vere. Multilevel counterfactuals for generalizations of relational concepts and productions. *Art. Int.*, 14:139–164, 1980.

[1000] S. A. Vere. Constrained N-to-1 generalization. preprint, Jet Prop. Lab., 1981.

[1001] J. Vitter and J. Lin. Learning in parallel. *Inform. Comput.*, pages 179–202, 1992.

[1002] V. Vovk. Aggregating strategies. In *Proc. 3rd Annu. Workshop on Comput. Learning Theory*, pages 371–383. Morgan Kaufmann, 1990.

[1003] V. Vovk. Universal forecasting algorithms. *Inform. Comput.*, 96(2):245–277, 1992.

[1004] V. G. Vovk. An optimal-control application of two paradigms of on-line learning. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 98–109. ACM Press, New York, NY, 1994.

[1005] A. Waibel, T. Hanazawa, G. Ginton, and K. Shikano. Phoneme recognition using time-delay neural networks. Technical report, ATR Interpreting Telephony Laboratories, 1987.

[1006] C. S. Wallace and D. M. Boulton. An information measure for classification. *The Computer Journal*, 11(2):185–194, 1968.

[1007] R. S. Wallace. *Finding Natural Clusters Through Entropy Minimization.* PhD thesis, Carnegie Mellon Computer Science Department, June 1989. Tech. Report CMU-CS-89-183.

[1008] W. G. Walter. A machine that learns. *Scientific American*, pages 60–63, Aug. 1951.

[1009] M. Warmuth. Towards representation independence in PAC learning. In *Lecture Notes in Artificial Intelligence 397: Proceedings of the Second Workshop on Analogical and Inductive Inference*, pages 78–103. Springer-Verlag, 1989.

[1010] M. K. Warmuth. Towards representation independence in PAC-learning. In *Proceedings of AII-89 Workshop on Analogical and Inductive Inference; Lecture Notes in Artificial Intelligence 397*, pages 78–103, Heidelberg, Oct. 1989. Springer-Verlag.

[1011] O. Watanabe. A formal study of learning via queries. In *Lecture Notes in Computer Science 443: Proceedings of the 17th International Colloquium on Automata, Languages and Programming*, pages 139–152. Springer-Verlag, 1990.

[1012] S. Watanabe. Information-theoretical aspects of inductive and deductive inference. *IBM J. of Res. Develop.*, 4:208–231, 1960.

[1013] D. A. F. W. S. K. P. R. S. J. Waterman and R. Wesson. Design issues for exemplary programming. Technical report, Rand Corporation, N-1484-RC, 1980.

[1014] C. J. C. H. Watkins. Combining cross–validation and search. In *Progress in Machine Learning–Proceedings of EWSL 87: 2nd European Working Session on Learning*, pages 79–90, Bled, Yugoslavia, May 1987.

[1015] C. J. C. H. Watkins. *Learning from Delayed Rewards.* PhD thesis, King's College, Oxford, May 1989. (To be reprinted by MIT Press.).

[1016] A. Weigend, B. Huberman, and D. Rumelhart. Predicting the future: A connectionist approach. *Internat. J. Neural Syst.*, 1:193–209, 1990.

[1017] S. Weiss and C. Kulikowski. *Computer Systems that Learn.* Morgan Kaufmann, San Mateo, CA, 1991.

[1018] R. S. Wenocur and R. M. Dudley. Some special Vapnik-Chervonenkis classes. *Discrete Mathematics*, 33:313–318, 1981.

[1019] K. Wexler. On extensional learnability. *Cognition*, 11:89–95, 1982.

[1020] K. Wexler and P. Culicover. *Formal Principles of Language Acquisition.* MIT Press, Cambridge, Mass., 1980.

[1021] K. Wexler and H. Hamburger. On the sufficiency of surface data for the learning of transformational languages. In *Approaches to Natural Language.* Dordecht, D. reidel, 1973.

[1022] R. M. Wharton. Approximate language identification. *Inform. Control*, 26:236–255, 1974.

[1023] R. M. Wharton. Grammar enumeration and inference. *Inform. Control*, 33:253–272, 1977.

[1024] H. White. Connectionist nonparametric regression: Multilayer feedforward networks can learn arbitrary mappings. *Neural Networks*, 3:535–549, 1990.

[1025] H. White. Learning in artificial neural networks: a statistical perspective. *Neural Computation*, 1(4):425–464, 1990.

[1026] B. Widrow and M. E. Hoff. Adaptive switching circuits. *1960 IRE WESCON Convention Record*, pages 96–104, 1960. Reprinted in *Neurocomputing* (MIT Press, 1988).

[1027] R. Wiehagen. Inductive inference of recursive functions. In *Lecture Notes in Computer Science 32, Mathematical Foundations of Computer Science 75*, pages 462–464. Springer-Verlag, 1975.

[1028] R. Wiehagen. Limes-erkennung rekursiver funktionen durch spezielle strategien. *Elektronische Informationsverarbeitung und Kybernetik*, 12:93–99, 1976.

[1029] R. Wiehagen. Identification of formal languages. In *Lecture Notes in Computer Science 53*, pages 571–579. Springer-Verlag, 1977.

[1030] R. Wiehagen. Characterization problems in the theory of inductive inference. In *Proc. of 5th Colloquium on Automata, Languages, and Programming*, pages 494–508. Springer-Verlag, 1978.

[1031] R. Wiehagen. *Zur Theorie der Algorithmischen Erkennung.* PhD thesis, Humboldt-Universitat, Berlin, 1978.

[1032] R. Wiehagen. How fast is program synthesis from examples. *Lecture Notes in Computer Science*, 215:231–239, 1986.

[1033] R. Wiehagen. On the complexity of effective program synthesis. *Lecture Notes in Computer Science*, 265:209–219, 1986.

[1034] R. Wiehagen. On the complexity of program synthesis from examples. *Elektronische Informationsverarbeitung und Kybernetik*, 22:305–323, 1986.

[1035] R. Wiehagen. A thesis in inductive inference. In *Nonmonotonic and Inductive Logic, 1st International Workshop, Karlsruhe, Germany*, pages 184–207. Springer Verlag, 1990. Lecture Notes in Computer Science 543.

[1036] R. Wiehagen, R. Freivalds, and E. B. Kinber. On the power of probabilistic strategies in inductive inference. *Theoret. Comput. Sci.*, 28:111–133, 1984.

[1037] R. Wiehagen and H. Jung. Rekursionstheoretische charakterisierung von erkennbaren klassen rekursiver funktionen. *Elektronische Informationsverarbeitung und Kybernetik*, 13:385–397, 1977.

[1038] R. Wiehagen and W. Liepe. Characteristic properties of recognizable classes of recursive functions. *Elektronische Informationsverarbeitung und Kybernetik*, 12:421–438, 1976.

[1039] R. Wiehagen and W. Liepe. Charakteristche eigenschaften von erkennbaren klassen rekursiver funktionen. *Elektronische Informationsverarbeitung und Kybernetik*, 12:421–438, 1976.

[1040] R. Wiehagen and C. H. Smith. Generalization versus classification. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 224–230. ACM Press, New York, NY, 1992.

[1041] R. Wiehagen, C. H. Smith, and T. Zeugmann. Classifying recursive predicates and languages. Technical Report 21/93, GOSLER-Report, FB Mathematik und Informatik, TH Lepzig, 1993.

[1042] R. Wiehagen, C. H. Smith, and T. Zeugmann. Classification of predicates and languages. In *Computational Learning Theory: Eurocolt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 171–181, Oxford, 1994. Oxford University Press.

[1043] R. Wiehagen and T. Zeugmann. Too much information can be too much for efficient learning. In *Proc. 3rd Internat. Workshop on Analogical and Inductive Inference*, pages 72–86. Springer–Verlag, 1992. Lecture Notes in Artificial Intelligence 642.

[1044] R. Wiehagen and T. Zeugmann. Ingnoring data may be the only way to learn efficiently. *Journal of Experimental and Artificial Intelligence*, 6:131–144, 1994.

[1045] H. Wiklicky. The neural network loading problem is undecidable. In *Computational Learning Theory: Eurocolt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 183–192, Oxford, 1994. Oxford University Press.

[1046] S. W. Wilson. Knowledge growth in an artificial animal. In *Proceedings of an International Conference on Genetic Algorithms and their Applications*, pages 16–23, July 1985.

[1047] P. H. Winston. Learning structural descriptions from examples. In *The Psychology of Computer Vision*. McGraw-Hill, New York, 1975.

[1048] P. H. Winston. Learning new principles from precedents and exercises: The details. Technical Report AIM632, MIT Artificial Intelligence Laboratory, May 1981.

[1049] I. H. Witten. An adaptive optimal controller for discrete-time markov environments. *Inform. Control*, 34:286–295, 1977.

[1050] P. C. Woodland and S. G. Smyth. An experimental comparison of connectionist and conventional classification systems on natural data. *Speech Communication*, 9:73–82, Oct. 1989.

[1051] K. Wright. Identification of unions of languages drawn from an identifiable class. In *Proc. 2nd Annu. Workshop on Comput. Learning Theory*, pages 328–333. Morgan Kaufmann, 1989. (See also the correction by Motoki, Shinohara and Wright in the Proceedings of the Fourth Annual Workshop on Computational Learning Theory, page 375, 1991).

[1052] K. Wright. Inductive identification of pattern languages with restricted substitutions. In *Proc. 3rd Annu. Workshop on Comput. Learning Theory*, pages 111–121, San Mateo, CA, 1990. Morgan Kaufmann.

[1053] K. Yamanishi. A loss bound model for on-line stochastic prediction strategies. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 290–302, San Mateo, CA, 1991. Morgan Kaufmann.

[1054] K. Yamanishi. A loss-bound model for on-line stochastic prediction strategies. *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 290–302, 1991.

[1055] K. Yamanishi. A learning criterion for stochastic rules. *Machine Learning*, 1992. Special Issue on the Proceedings of the 3nd Workshop on Computational Learning Theory, to appear.

[1056] K. Yamanishi. Probably almost discriminative learning. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 164–171. ACM Press, New York, NY, 1992.

[1057] K. Yamanishi. On polynomial-time probably almost discriminative learnability. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 94–100. ACM Press, New York, NY, 1993.

[1058] K. Yamanishi. Learning non-parametric smooth rules by stochastic rules with finite partitioning. In *Computational Learning Theory: Eurocolt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 217–227, Oxford, 1994. Oxford University Press.

[1059] K. Yamanishi. The minimum l-complexity algorithm and its applications to learning non-parametric rules. In *Proc. 7th Annu. ACM Workshop on Comput. Learning Theory*, pages 173–182. ACM Press, New York, NY, 1994.

[1060] M. Yannakakis and D. Lee. Testing finite state machines. In *Proc. of the 23rd Symposium on Theory of Computing*, pages 476–485. ACM Press, New York, NY, 1991.

[1061] T. Yokomori. Polynomial-time learning of very simple grammars from posistive data. In *Proc. 4th Annu. Workshop on Comput. Learning Theory*, pages 213–227, San Mateo, CA, 1991. Morgan Kaufmann.

[1062] T. Yokomori. Learning two-tape automata from queries and counterexamples. In *Proc. 6th Annu. Workshop on Comput. Learning Theory*, pages 228–235. ACM Press, New York, NY, 1993.

[1063] S. D. Zeroski, S. Muggleton, and S. Russell. PAC-learnability of determinate logic programs. In *Proc. 5th Annu. Workshop on Comput. Learning Theory*, pages 128–135. ACM Press, New York, NY, 1992.

[1064] T. Zeugmann. A-psoteriori characterizations in inductive inference of recursive functions. *Elektronische Informationsverarbeitung und Kybernetik*, 19:559–594, 1983.

[1065] T. Zeugmann. On the synthesis of fastest programs in inductive inference. *Elektronische Informationsverarbeitung und Kybernetik*, 19:625–642, 1983.

[1066] T. Zeugmann. On the nonboundability of total effective operators. *Zeitschr. f. math. Logik und Grundlagen d. Math. Bd.*, 30:169–172, 1984.

[1067] T. Zeugmann. On barzdin's conjecture. In *Proc. 1st Internat. Workshop on Analogical and Inductive Inference*, pages 117–125. Springer–Verlag, 1986. Lecture Notes in Computer Science 265.

[1068] T. Zeugmann. On the power of recursive optimizers. *Theoret. Comput. Sci.*, 62:289–310, 1988.

[1069] T. Zeugmann. Inductive inference of optimal programs: A survey and open problems. In *Nonmonotonic and Inductive Logic, 1st International Workshop, Karlsruhe, Germany*, pages 208–222. Springer Verlag, 1990. Lecture Notes in Computer Science 543.

[1070] T. Zeugmann, S. Lange, and S. Kapur. Characterizations of class preserving monotonic and dual monotonic language learning. Technical Report IRCS 92 - 24, Technical Report, Institute for Research in Cognitive Science, Univ. Pennsylvania, Philadelphia, 1992.

[1071] Y. Zhuravlev and Y. Kogan. Realization of boolean functions with a small number of zeros by disjunctive normal forms, and related problems. *Soviet Math. Doklady*, 32:771–775, 1985.

[1072] M. Zloof. Query-by-example: a data base language. *IBM Syst. J.*, 16:324–343, 1977.