

**UNIVERSITAT POLITÈCNICA DE CATALUNYA**

*Departament de Llenguatge i Sistemes Informàtics  
Ph.D. Programme: Artificial Intelligence*

**SYMBOLIC AND CONNECTIONIST  
LEARNING TECHNIQUES FOR  
GRAMMATICAL INFERENCE**

Autor: René Alquézar Mancho  
Director: Alberto Sanfeliu Cortés

March 1997

## References and Bibliography

- [AbDa:91] H. Abramson and V. Dahl, *Logic Grammars*, Springer-Verlag, 1991.
- [AhUl:72] A.V. Aho and J.D. Ullman, *The Theory of Parsing, Translation and Compiling*, Vol.1. Prentice-Hall, 1972.
- [AlDO:91] N. Alon, A.K. Dewdney, T.J. Ott, "Efficient simulation of finite automata by neural nets," *Journal of the ACM* **38** (2), pp.495-514, 1991.
- [Allen:87] J.F. Allen, *Natural Language Understanding*, Benjamin Cummings, 1987.
- [Allen:90] R.B. Allen, "Connectionist language users," *Connection Science* **2** (4), pp.279-311, 1990.
- [AlSa:93] R. Alquézar and A. Sanfeliu, "Representation and recognition of regular grammars by means of second-order recurrent neural networks," in *New Trends in Neural Computation, Proc. of the Int. Workshop on Artificial Neural Networks IWANN'93, Sitges, Spain*, J.Mira, J.Cabestany, A.Prieto (eds.), Springer-Verlag, pp.143-148, 1993.
- [AlSa:94a] R. Alquézar and A. Sanfeliu, "Inference and recognition of regular grammars by training recurrent neural networks to learn the next-symbol prediction task," in *Advances in Pattern Recognition and Applications, Selected Papers from the Vth Spanish Symposium on Pattern Recognition and Image Analysis, Valencia, Sept.1992*, F.Casacuberta and A.Sanfeliu (eds.), World Scientific Pub., Singapore, pp.48-59, 1994.
- [AlSa:94b] R. Alquézar and A. Sanfeliu, "A hybrid connectionist-symbolic approach to regular grammatical inference based on neural learning and hierarchical clustering," in *Grammatical Inference and Applications*, R.C.Carrasco and J.Oncina (eds.), Springer-Verlag, Lecture Notes in Artificial Intelligence 862, pp.203-211, 1994.

- [AlSa:95a] R. Alquézar and A. Sanfeliu, "Incremental grammatical inference from positive and negative data using unbiased finite state automata," in *Shape, Structure and Pattern Recognition*, Proc. Int. Workshop on Structural and Syntactic Pattern Recognition, SSPR'94, Nahariya (Israel), D.Dori and A.Bruckstein (eds.), World Scientific Pub., Singapore, pp.291-300, 1995.
- [AlSa:95b] R. Alquézar and A. Sanfeliu, "An algebraic framework to represent finite-state machines in single-layer recurrent neural networks," *Neural Computation* 7 (5), pp.931-949, 1995.
- [AlSa:95c] R. Alquézar and A. Sanfeliu, "Augmented regular expressions: a formalism to describe, recognize, and learn a class of context-sensitive languages," *Research Report LSI-95-17-R, Universitat Politècnica de Catalunya, Barcelona*, 1995.
- [AlSa:96] R. Alquézar and A. Sanfeliu, "Learning of context-sensitive languages described by augmented regular expressions," in *Proc. of the 13th International Conference on Pattern Recognition, ICPR'96, Vienna (Austria)*, IEEE Computer Society Press, Vol.IV, pp.745-749, 1996.
- [AlSa:97a] R. Alquézar and A. Sanfeliu, "Recognition and learning of a class of context-sensitive languages described by augmented regular expressions," *Pattern Recognition* 30 (1), pp.163-182, 1997.
- [AlSa:97b] R. Alquézar and A. Sanfeliu, "Experimental comparison of connectionist and symbolic methods for regular inference from positive and negative samples," submitted to *Int. Journal of Pattern Recognition and Artificial Intelligence*, 1997.
- [AlSC:96] R. Alquézar, A. Sanfeliu and J. Cueva, "Learning of context-sensitive language acceptors through regular inference and constraint induction," in *Grammatical Inference: Learning Syntax from Sentences*, Proc. of the 3rd Int. Colloquium, ICGI'96, Montpellier (France), Sept.1996, L.Miclet and C.de la Higuera (eds.), Springer-Verlag, Lecture Notes in Artificial Intelligence 1147, pp.134-145, 1996.
- [AlSo:94] R. Alquézar and J.M. Sopena, "Effect of unbounded activation functions on learning performance of recurrent networks," *Tech. Rep. IC-DT-9402, Institut de Cibernètica, UPC-CSIC, Barcelona, Spain*, January 1994.
- [AlSS:97] R. Alquézar, A. Sanfeliu and M. Sainz, "Experimental assessment of connectionist regular inference from positive and negative examples," in *Proc. of the 7th Spanish Symposium on Pattern Recognition and Image Analysis, Bellaterra (Spain)*, April 1997.

- [AmPu:64] V. Amar and G. Putzolu, "On a family of linear grammars," *Information and Control* **7**, pp.283-291, 1964.
- [Angl:78] D. Angluin, "On the complexity of minimum inference of regular sets," *Information and Control* **39**, pp.337-350, 1978.
- [Angl:80a] D. Angluin, "Inductive inference of formal languages from positive data," *Information and Control* **45**, pp.117-135, 1980.
- [Angl:80b] D. Angluin, "Finding patterns common to a set of strings," *J. Comput. System Science* **21**, pp.46-62, 1980.
- [Angl:81] D. Angluin, "A note on the number of queries needed to identify regular languages," *Information and Control* **51**, pp.76-87, 1981.
- [Angl:82] D. Angluin, "Inference of reversible languages," *Journal of the ACM* **29** (3), pp.741-765, 1982.
- [Angl:87] D. Angluin, "Learning regular sets from queries and counter-examples," *Information and Computation* **75**, pp.87-106, 1987.
- [Angl:88] D. Angluin, "Queries and concept learning," *Machine Learning* **2**, pp.319-342, 1988.
- [AnSm:83] D. Angluin and C.H. Smith, "Inductive inference: theory and methods," *ACM Computing Survey* **15** (3), pp.237-269, 1983.
- [Arden:60] D.N. Arden, "Delayed logic and finite state machines," in *Theory of Computing Machine Design*, Univ. of Michigan Press, Ann Arbor, pp.1-35, 1960.
- [BaJM:83] L.R. Bahl, F. Jelinek and R.L. Mercer, "A maximum likelihood approach to continuous speech recognition," *IEEE Trans. on PAMI* **5**, pp.179-190, 1983.
- [BiDu:84] G. Biswas and R.C. Dubes, "Some experiments in two-dimensional grammatical inference," *Pattern Recognition Letters* **2**, pp.173-177, 1984.
- [BiFe:72] A.W. Biermann and J. Feldman, "On the synthesis of finite-state machines from samples of their behavior," *IEEE Trans. Computers* **21**, pp.592-596, 1972.
- [Booth:67] T.L. Booth, *Sequential Machines and Automata Theory*. John Wiley, New York, 1967.
- [BrFu:76] J.M. Brayer and K.S. Fu, "Some multidimensional grammar inference methods," in *Pattern Recognition and Artificial Intelligence*, C.H.Chen (ed.), Academic-Press, 1976.

- [BuSa:90] H. Bunke and A. Sanfeliu (eds), *Syntactic and Structural Pattern Recognition: Theory and Applications*, World Scientific, 1990.
- [CaCa:96] M.A. Castaño and F. Casacuberta "Inference of regular grammars through neural networks: a review," submitted to *IEEE Trans. on Neural Networks*, 1996.
- [CaCh:92] G. Carroll and E. Charniak, "Learning probabilistic dependency grammars from labeled text," *Working Notes, Fall Symposium Series, AAAI*, pp.25-32, 1992.
- [CaCV:93] M.A. Castaño, F. Casacuberta and E. Vidal, "Simulation of stochastic regular grammars through simple recurrent networks," in *New Trends in Neural Computation, Proc. of the Int. Workshop on Artificial Neural Networks IWANN'93, Sitges, Spain*, J.Mira, J.Cabestany, A.Prieto (eds.), Springer-Verlag, pp.210-215, 1993.
- [CaFS:96] R. Carrasco, M. Forcada and L. Santamaría, "Inferring stochastic regular grammars with recurrent neural networks," in *Grammatical Inference: Learning Syntax from Sentences*, Proc. of the 3rd Int. Colloquium, ICGI'96, Montpellier (France), Sept.1996, L.Miclet and C.de la Higuera (eds.), Springer-Verlag, LNAI 1147, pp.274-281, 1996.
- [CaGV:94] A. Castellanos, I. Galiano and E. Vidal, "Application of OSTIA to machine translation tasks" in *Grammatical Inference and Applications*, R.C.Carrasco and J.Oncina (eds.), Springer-Verlag, Lecture Notes in Artificial Intelligence 862, pp.93-105, 1994.
- [Casa:94] F. Casacuberta, "Statistical estimation of stochastic context-free grammars using the inside-outside algorithm and a transformation on grammars," in *Grammatical Inference and Applications*, R.C.Carrasco and J.Oncina (eds.), Springer-Verlag, Lecture Notes in Artificial Intelligence 862, pp.119-129, 1994.
- [CaSm:83] J. Case and C. Smith, "Comparison of identification criteria for machine inductive inference," *Theor. Comput. Science* **25**, pp.193-220, 1983.
- [ChAc:80] C. Chirathamjaree and M.H. Ackroyd, "A method for the inference of non-recursive context-free grammars," *Int. J. Man-Machine Studies* **12**, pp.379-387, 1980.
- [ChFu:75] S.M. Chou and K.S.Fu, "Transition Networks for Pattern Recognition," *Tech. Rep. TR-EE-75-39, School of Electrical Engineering, Purdue University, W. Lafayette, Indiana*, December 1975.
- [ChFu:76] S.M. Chou and K.S.Fu, "Inference for transition network grammars," in *Proc. of Int. Joint Conf. on Pattern Recognition*, **3**, CA, 1976, pp.79-84.

- [Chom:59] N. Chomsky, "On certain formal properties of grammars," *Information and Control* **2** (2), pp.137-167, 1959.
- [Chom:65] N. Chomsky, *Aspects of the Theory of Syntax*, MIT Press, Cambridge MA, 1965.
- [ClSM:89] A. Cleeremans, D. Servan-Schreiber and J.L. McClelland, "Finite-state automata and simple recurrent networks," *Neural Computation* **1**, pp.372-381, 1989.
- [CoRA:76] C.M. Cook, A. Rosenfeld, and A.R. Aronson, "Grammatical inference by hill-climbing," *Information Science* **10**, pp.59-80, 1976.
- [Cres:72] S. Crespi-Reghizzi, "An effective model for grammatical inference," in "Information Processing-71", *Proc. of IFIP Congress, Ljubljana, 1971*, B.Gilchrist (ed.), Elsevier/North-Holland, New York, pp.524-529, 1972.
- [CrGM:78] S. Crespi-Reghizzi, G. Guida, and D. Mandrioli, "Noncounting context-free languages," *Journal of the ACM* **25**, pp.571-580, 1978.
- [Cybe:89] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of Control, Signals, and Systems* **2**, pp.303-314, 1989.
- [DaDa:91] S. Das and R. Das, "Induction of discrete-state machine by stabilizing a simple recurrent network using clustering," *Computer Science and Informatics* **21** (2), pp.35-40, 1991.
- [DaGS:92] S. Das, C.L. Giles and G.Z. Sun, "Learning context-free grammars: limitations of a recurrent neural network with an external stack memory," in *Proc. of the 14th Annual Conf. of the Cognitive Science Society*, Morgan Kaufmann Pub., San Mateo CA, pp.791-795, 1992.
- [DaGS:93] S. Das, C.L. Giles and G.Z. Sun, "Using hints to successfully learn context-free grammars with a neural network pushdown automaton," in *Advances in Neural Information Processing Systems 5*, S.J. Hanson, J.D. Cowan, C.L. Giles, (eds.), Morgan Kaufmann Pub., San Mateo CA, pp.65-, 1993.
- [DaMo:93] S. Das and M.C. Mozer, "A connectionist symbol manipulator that induces rewrite rules in context-free grammars," in *Grammatical Inference: Theory, Applications and Alternatives. Proc. of the First Int. Colloquium, Essex, UK*, S. Lucas (ed.), IEE Press, Digest No:1993/092, London, 1993.
- [DaMo:94] S. Das and M.C. Mozer, "A unified gradient-descent/clustering architecture for finite state machine induction," in *Advances in Neural Information Processing Systems 6*, J.D. Cowan, G. Tesauro, J. Alspector, (eds.), Morgan Kaufmann Pub., San Mateo CA, 1994.

- [DaSc:92] M.R.W. Dawson and D.P. Schopflocher, "Modifying the generalized delta rule to train networks of non-monotonic processors for pattern classification," *Connection Science* 4 (1), pp.19-31, 1992.
- [DeJo:90] K. DeJong, "Genetic-algorithm-based learning," in *Machine Learning: An Artificial Intelligence Approach, Vol.3*, Y. Kodratoff and R.S. Michalski (eds.), Morgan Kaufmann, San Mateo CA, pp.611-638, 1990.
- [DeKi:82] P.A. Devijver and J. Kittler, *Pattern Recognition: A Statistical Approach*, Prentice-Hall International, London, UK, 1982.
- [DelH:96] C. De la Higuera, "Characteristic sets for polynomial grammatical inference," in *Grammatical Inference: Learning Syntax from Sentences*, L.Miclet and C.De la Higuera (eds.), Springer-Verlag, LNAI 1147, pp.59-71, 1996.
- [DiLC:82] T.G. Dietterich, R. London, K. Clarkson and R. Dromey, "Learning and inductive inference," in *The Handbook of Artificial Intelligence*, P.Cohen and E.Feigenbaum (eds.), Kaufman, Los Altos CA, pp.323-512, 1982.
- [DuMV:94] P. Dupont, L. Miclet and E. Vidal, "What is the search space of the regular inference?" in *Grammatical Inference and Applications*, R.C.Carrasco and J.Oncina (eds.), Springer-Verlag, Lecture Notes in Artificial Intelligence 862, pp.25-37, 1994.
- [Dupo:94] P. Dupont, "Regular grammatical inference from positive and negative samples by genetic search: the GIG method," in *Grammatical Inference and Applications*, R.C.Carrasco and J.Oncina (eds.), Springer-Verlag, Lecture Notes in Artificial Intelligence 862, pp.236-245, 1994.
- [Dupo:96] P. Dupont, "Incremental regular inference," in *Grammatical Inference: Learning Syntax from Sentences*, L.Miclet and C.De la Higuera (eds.), Springer-Verlag, LNAI 1147, pp.222-237, 1996.
- [Elman:90] J.L. Elman, "Finding structure in time", *Cognitive Science* 14, pp.179-211, 1990.
- [Fahl:88] S.E. Fahlman, "Faster-learning variations on back-propagation: an empirical study," in *Proc. of the 1988 Connectionist Models Summer School*, Morgan Kaufmann, 1988.
- [Fahl:91a] S.E. Fahlman, "The recurrent cascade-correlation architecture," in *Advances in Neural Information Processing Systems 3*, R. Lippmann, J. Moody, D. Touretzky (eds.), Morgan Kaufmann Pub., San Mateo CA, pp.190-196, 1991.

- [Fahl:91b] S.E. Fahlman, "The recurrent cascade-correlation architecture," *Tech. Rep. CMU-CS-91-100, School of Computer Science, Carnegie Mellon Univ., Pittsburgh, PA*, May 1991.
- [FaLe:90] S.E. Fahlman and C. Lebiere, "The cascade-correlation learning architecture," in *Advances in Neural Information Processing Systems 2*, D.S.Touretzky (ed.), Morgan Kaufmann Pub., San Mateo CA, 1990.
- [FeGH:69] J.A. Feldman, J. Gips, J.J. Horning, and S. Reder, "Grammatical complexity and inference", *Tech. Rep. No. CS-125, Computer Science Dept., Stanford Univ., California*, 1969.
- [Feld:72] J.A. Feldman, "Some decidability results in grammatical inference," *Information and Control* **20**, pp.244-262, 1972.
- [FoCa:95] M. Forcada and R. Carrasco, "Learning the initial state of a second-order recurrent neural network during regular-language inference," *Neural Computation* **7** (5), pp.923-930, 1995.
- [FoSa:97] M. Font and A. Sanfeliu, "A new codification of boundaries by means of syntactic representation," in *Proc. of the 7th Spanish Symposium on Pattern Recognition and Image Analysis, Bellaterra (Spain)*, April 1997.
- [Fox:64] L. Fox, *An Introduction to Numerical Linear Algebra*, Oxford University Press, 1964.
- [Free:74] H. Freeman, "Computer processing of line drawing images," *Computer Surveys* **6**, pp.57-98, 1974.
- [FrGM:91] P. Frasconi, M. Gori, M. Maggini, G. Soda, "A unified approach for integrating explicit knowledge and learning by example in recurrent networks," in *Proc. Int. Joint Conf. on Neural Networks, IJCNN-91*, IEEE Press, Piscataway NJ, Vol.1, pp.811-816, 1991.
- [Fu:82] K.S. Fu, *Syntactic Pattern Recognition and Applications*. Prentice-Hall, New York, 1982.
- [FuBo:75] K.S. Fu and T.L. Booth, "Grammatical inference: introduction and survey," *IEEE Trans. on SMC* **5**, Part 1: pp.85-111, Part 2: pp.409-423, 1975.
- [Funa:89] K. Funahashi, "On the approximate realization of continuous mappings by neural networks," *Neural Networks* **2**, pp.183-192, 1989.
- [GaOn:93] P. García and J. Oncina, "Learning general context-free grammars from positive structural samples and negative strings," *DSIC Research Report, Universidad Politécnica de Valencia, Spain*, 1993.



- [GaSV:90] P. García, E. Segarra, E. Vidal and I. Galiano, "On the use of the morphic generator grammatical inference (MGGI) methodology in automatic speech recognition," *Int. J. Pattern Recogn. Artif. Intell.* **4**, pp.667-685, 1990.
- [GaVC:87] P. García, E. Vidal and F. Casacuberta, "Local languages, the successor method, and a step towards a general methodology for the inference of regular grammars," *IEEE Trans. on PAMI* **9**, pp.841-845, 1987.
- [GaVi:90] P. García and E. Vidal, "Inference of K-testable languages in the strict sense and applications to syntactic pattern recognition," *IEEE Trans. on PAMI* **12** (9), pp.920-925, 1990.
- [GiMC:92] C.L. Giles, C.B. Miller, D. Chen, H.H. Chen, G.Z. Sun, Y.C. Lee, "Learning and extracting finite state automata with second-order recurrent neural networks," *Neural Computation* **4**, pp.393-405, 1992.
- [GiOm:92] C.L. Giles and C.W. Omlin, "Inserting rules into recurrent neural networks," *Neural Networks for Signal Processing II, Proc. of the 1992 IEEE Workshop*, S.Y. Kung, F. Fallside, J.A. Sorenson, C.A. Kamm, Eds., Piscataway NJ: IEEE Press, pp.13-22, 1992.
- [GiOm:93] C.L. Giles and C.W. Omlin, "Extraction, insertion and refinement of symbolic rules in dynamically-driven recurrent neural networks," *Connection Science* **5** (3/4), pp.307-337, 1993.
- [Gior:94] J-Y. Giordano, "Inference of context-free grammars by enumeration: structural containment as an ordering bias," in *Grammatical Inference and Applications*, R.C. Carrasco and J. Oncina (eds.), Springer-Verlag, Lecture Notes in Artificial Intelligence 862, pp.212-221, 1994.
- [GiSC:90] C.L. Giles, G.Z. Sun, H.H. Chen, Y.C. Lee, and D. Chen, "Higher order recurrent networks and grammatical inference," in *Advances in Neural Information Processing Systems 2*, D.S. Touretzky, (ed.), Morgan Kaufmann Pub., San Mateo CA, pp.380-387, 1990.
- [GoGC:94] M.W. Goudreau, C.L. Giles, S.T. Chakradhar, D. Chen, "First-order vs. second-order single layer recurrent neural networks," *IEEE Trans. on Neural Networks* **5** (3), pp.511-513, 1994.
- [GoGi:93] M.W. Goudreau and C.L. Giles, "On recurrent neural networks and representing finite state recognizers," *Proc. Third Int. Conf. on Artificial Neural Networks*, IEE, London, UK, 1993.
- [Gold:67] E.M. Gold, "Language identification in the limit," *Information and Control* **10**, pp.447-474, 1967.

- [Gold:78] E.M. Gold, "Complexity of automaton identification from given data," *Information and Control* **37**, pp.302-320, 1978.
- [Goldb:89] D.E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, New York, 1989.
- [Greg:94] J. Gregor, "Data-driven inductive inference of finite-state automata," *Int. J. of Pattern Recognition and Artificial Intelligence* **8** (1), pp.305-322, 1994.
- [GrJu:92] T. Grönfors and M. Juhola, "Experiments and comparison of inference methods of regular grammars," *IEEE Trans. on SMC* **22** (4), pp.821-826, 1992.
- [Gros:86] S. Grossberg, *The Adaptive Brain I: Cognition, Learning, Reinforcement, and Rhythm*, and *The Adaptive Brain II: Vision, Speech, Language, and Motor Control*. Elsevier/North-Holland, Amsterdam, 1986.
- [HaKK:90] E. Hartman, J.D. Keeler and J.M. Kowalski, "Layered neural networks with Gaussian hidden units as universal approximations," *Neural Computation* **2**, pp.210-215, 1990.
- [Harr:78] M.A. Harrison, *Introduction to Formal Language Theory*. Addison-Wesley, Reading MA, 1978.
- [HeKP:91] J. Hertz, A. Krogh, R.G. Palmer *Introduction to the Theory of Neural Computation*. Addison-Wesley, Redwood City CA, 1991.
- [Hopf:82] J.J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proc. National Academy of Sciences, USA*, **79**, pp.2554-2558, 1982.
- [Horn:69] J.J. Horning, *A study of grammatical inference*. Tech. Rep. No. CS-139, Computer Science Dept., Stanford Univ., Stanford CA, 1969.
- [Horn:91] K. Hornik, "Approximation capabilities of multilayer feedforward networks," *Neural Networks* **4**, pp.251-257, 1991.
- [HoSW:89] K. Hornik, M. Stinchcombe, H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks* **2**, pp.359-366, 1989.
- [HoUl:79] J.E. Hopcroft and J.D. Ullman, *Introduction to Automata Theory, Languages and Computation*. Addison-Wesley, Reading MA, 1979.
- [Ishi:89] I. Ishizaka, "Learning simple deterministic languages," *Proc. of the 2nd Annual Conference on Computational Learning Theory, COLT'89*, ACM, pp.162-174, 1989.

- [Itoga:81] S.Y. Itoga, "A new heuristic for inferring regular grammars," *IEEE Trans. on PAMI* **3**, pp.191-197, 1981.
- [JaAd:94] M. Jardino and G. Adda, "Automatic determination of a stochastic bigram class language model" in *Grammatical Inference and Applications*, R.C.Carrasco and J.Oncina (eds.), Springer-Verlag, Lecture Notes in Artificial Intelligence 862, pp.57-65, 1994.
- [Jant:84] K.P. Jantke, "Polynomial time inference of general pattern languages," in *Proc. Symposium of Theoretical Aspects of Computer Science*, Lecture Notes in Computer Science 166, Springer-Verlag, Berlin, pp.314-325, 1984.
- [Jeli:91] F. Jelinek, "Up from trigrams! The struggle for improved language models," *Proc. of EUROSPEECH-91*, pp.1037-1039, 1991.
- [KeVa:89] M. Kearns and L. Valiant, "Cryptographic limitations on learning boolean formulae and finite automata," *Proc. of the 21st ACM Symposium on Theory of Computing*, pp.433-444, 1989.
- [Knuu:96] T. Knuutila, "Inductive inference from positive data: from heuristic to characterizing methods," in *Grammatical Inference: Learning Syntax from Sentences*, L.Miclet and C.De la Higuera (eds.), Springer-Verlag, LNAI 1147, pp.22-47, 1996.
- [Kodr:92] Y. Kodratoff, "Recent advances in machine learning," *Int. Journal of Pattern Recognition and Artificial Intelligence* **6** (4), pp.469-511, 1992.
- [Koha:78] Z. Kohavi, *Switching and Finite Automata Theory*, (2nd edition). Tata McGraw-Hill, New Delhi, India, 1978.
- [Koho:84] T. Kohonen, *Self-Organization and Associative Memory*, Springer-Verlag, Berlin, Germany, 1984.
- [KoHu:87] K. Ko and C. Hua, "A note on the two-variable pattern finding problem," *J. Comput. System Science* **34**, pp.75-86, 1987.
- [KoMi:90] Y. Kodratoff and R.S. Michalski (eds.), *Machine Learning: An Artificial Intelligence Approach, Vol.3*, Morgan Kaufmann, San Mateo CA, 1990.
- [KoSS:85] J.L. Kolodner, R.L. Simpson and K. Sycara, "A process model of case-based reasoning in problem solving," *Proc. 9th Int. Joint Conf. on Artificial Intelligence*, pp.284-290, 1985.
- [Kupi:92] J. Kupiec, "Hidden Markov estimation for unrestricted stochastic context-free grammars," *Proc. Int. Conference on Acoustic, Speech and Signal Processing, Vol.I*, pp.177-180, 1992.

- [KuSh:88] M. Kudo, M. Shimbo, "Efficient regular grammatical inference techniques by the use of partial similarities and their logical relationships", *Pattern Recognition* **21** (4), pp.401-409, 1988.
- [LaFa:87] A. Lapedes, R. Farber, *Nonlinear signal processing using neural networks: prediction and system modelling*. Tech. Rep. LA-UR-87-2662, Los Alamos National Laboratory, Los Alamos NM, 1987.
- [LaFa:88] A. Lapedes, R. Farber, *How neural networks work*. Tech. Rep. LA-UR-88-418, Los Alamos National Laboratory, Los Alamos NM, 1988.
- [Lang:92] K.J. Lang, "Random DFA's can be approximately learned from sparse uniform examples," *Proc. of the 5th ACM Workshop on Computational Learning Theory*, pp.45-52, 1992.
- [LaWe:90] S. Lange and R. Weihagen, "Polynomial time inference of pattern languages," *Proc. of the First Int. Workshop on Algorithmic Learning Theory*, pp.289-308, 1990.
- [LaYo:90] K. Lari, S.J. Young, "The estimation of stochastic context-free grammars using the Inside-Outside algorithm," *Computer Speech and Language* **4** (1), pp.35-36, 1990.
- [LaYo:91] K. Lari, S.J. Young, "Applications of stochastic context-free grammars using the Inside-Outside algorithm," *Computer Speech and Language* **5**, pp.237-257, 1991.
- [LeLP:93] M. Leshno, V.Y. Lin, A. Pinkus and S. Schocken, "Multilayer feedforward networks with a nonpolynomial activation function can approximate any function," *Neural Networks* **6**, pp.861-867, 1993.
- [Levi:85] S.E. Levinson, "Structural methods in automatic speech recognition," *Proc. of the IEEE* **73** (11), pp.1625-1650, 1985.
- [Lipp:87] R.P. Lippmann, "An introduction to computing with neural nets," *IEEE ASSP Magazine* **2**, April, pp.4-22, 1987.
- [LoRe:80] B. Lowerre and R. Reddy, "The HARPY speech understanding system," in *Trends in Speech Recognition*, W. Lea (ed.), Prentice-Hall, Englewood Cliffs NJ, pp.340-346, 1980.
- [Lucas:90] S. Lucas, R. Dampier, "Syntactic neural networks," *Connection Science* **2**, pp.199-225, 1990.
- [Luen:84] D.G. Luenberger, *Linear and Nonlinear Programming*, 2nd ed., Addison-Wesley, Reading MA, 1984.

- [LuVA:94] S. Lucas, E. Vidal, A. Amiri, S. Hanlon and J.C. Amengual, "A comparison of syntactic and statistical techniques for off-line OCR" in *Grammatical Inference and Applications*, R.C.Carrasco and J.Oncina (eds.), Springer-Verlag, Lecture Notes in Artificial Intelligence 862, pp.168-179, 1994.
- [MaCo:92] S. Mahadevan and J. Connell, "Automatic programming of behavior-based robots using reinforcement learning," *Artificial Intelligence* **55**, pp.311-365, 1992.
- [MaFa:94] P. Manolios and R. Fanelli, "First-order recurrent neural networks and deterministic finite state automata," *Neural Computation* **6**, pp.1155-1173, 1994.
- [Mage:94] D.M. Magerman, *Natural Language Parsing as Statistical Pattern Recognition*. Ph.D. thesis dissertation, Stanford Univ., Stanford CA, 1994.
- [Mage:96] D.M. Magerman, "Learning grammatical structure using statistical decision-trees," in *Grammatical Inference: Learning Syntax from Sentences*, L.Miclet and C.De la Higuera (eds.), Springer-Verlag, LNAI 1147, pp.1-21, 1996.
- [Maki:92] E. Makinen, "Remarks on the structural grammatical inference problem for context-free grammars," *Information Processing Letters* **44**, pp.125-127, 1992.
- [MaKo:87] A. Marron and K. Ko, "Identification of pattern languages from examples and queries," *Information and Computation* **74**, pp.91-112, 1987.
- [MaNo:92] A. Maskara and A. Noetzel, "Forcing simple recurrent neural networks to encode context," in *Proc. of the 1992 Long Island Conference on Artificial Intelligence and Computer Graphics*, 1992.
- [MaRW:82] H.A. Maurer, G. Rozenberg and E. Welzl, "Using string languages to describe picture languages," *Information and Control* **54**, pp.155-185, 1982.
- [McPi:43] W.S. McCulloch and W. Pitts, "A logical calculus of the ideas imminent in nervous activity," *Bulletin of Mathematical Biophysics* **5**, pp.115-133, 1943.
- [Mich:83] R.S. Michalski, "A theory and methodology of inductive learning," *Artificial Intelligence* **20**, pp.111-161, 1983.
- [Micl:76] L. Miclet, "Inference of regular expressions," *Proc. of the 3rd Int. Conf. on Pattern Recognition*, pp.100-105, 1976.

- [Micl:80] L. Miclet, "Regular inference with a tail-clustering method," *IEEE Trans. Syst. Man Cybern.* **10**, pp.737-743, 1980.
- [Micl:90] L. Miclet, "Grammatical inference," in *Syntactic and Structural Pattern Recognition: Theory and Applications*, H.Bunke and A.Sanfeliu, Eds., World Scientific, 1990.
- [MiCM:83] R.S. Michalski, J.G. Carbonell and T.M. Mitchell (eds.), *Machine Learning: An Artificial Intelligence Approach*, Morgan Kaufmann, San Mateo CA, 1983.
- [MiCM:86] R.S. Michalski, J.G. Carbonell and T.M. Mitchell (eds.), *Machine Learning: An Artificial Intelligence Approach, Vol.2*, Morgan Kaufmann, San Mateo CA, 1986.
- [MiDH:96] L.Miclet and C.De la Higuera (eds.), *Grammatical Inference: Learning Syntax from Sentences*, Proc. of the 3rd Int. Colloquium, ICGI'96, Montpellier (France), Sept.1996, Springer-Verlag, Lecture Notes in Artificial Intelligence 1147, 1996.
- [MiGe:94] L. Miclet and C. de Gentile, "Inférence grammaticale à partir d'exemples et de contre-exemples: deux algorithmes optimaux (BIG et RIG) et une version heuristique (BRIG)," in *Actes des JFA-94, Strasbourg, France*, pp.F1-F13, 1994.
- [MiGi:93] C.B. Miller and C.L. Giles, "Experimental comparison of the effect of order in recurrent neural networks," *Int. Journal of Pattern Recognition and Artificial Intelligence* **7** (4), pp.849-872, 1993.
- [MiKK:86] T.M. Mitchell, R.M. Keller and S.T. Kedar-Cabelli, "Explanation based learning: A unifying view," *Machine Learning* **1**, pp.47-80, 1986.
- [Mins:67] M. Minsky, *Computation: Finite and Infinite Machines*, Chap.3, Prentice-Hall, Englewood Cliffs NJ, 1967.
- [MiPa:69] M. Minsky and S. Papert, *Perceptrons: an Introduction to Computational Geometry*. MIT Press, Cambridge MA, 1969.
- [Mitc:82] T.M. Mitchell, "Generalization as search," *Artificial Intelligence* **18**, pp.203-266, 1982.
- [MoBa:90] M.C. Mozer and J. Bachrach, "Discovering the structure of a reactive environment by exploration," *Neural Computation* **2** (4), pp.447-, 1990.
- [MoDa:93] M.C. Mozer and S. Das, "A connectionist symbol manipulator that discover the structure of context-free languages," in *Advances in Neural Information Processing Systems 5*, S.J. Hanson, J.D. Cowan, C.L. Giles, (eds.), Morgan Kaufmann Pub., San Mateo CA, pp.863-, 1993.

- [Moisl:92] H. Moisl, "Connectionist finite state natural language processing," *Connection Science* 4 (2), pp.67-91, 1992.
- [OmGi:92] C.W. Omlin and C.L. Giles, "Training second-order recurrent neural networks using hints," *Proc. Ninth Int. Conf. on Machine Learning*, D. Sleeman and P. Edwards, Eds., San Mateo CA: Morgan Kaufmann Pub., pp.363-368, 1992.
- [OmGi:96] C.W. Omlin and C.L. Giles, "Stable encoding of large finite-state automata in recurrent neural networks with sigmoid discriminants," *Neural Computation* 8, pp.675-696, 1996.
- [OnGa:92a] J. Oncina and P. García, "Inferring regular languages in polynomial update time," in *Pattern Recognition and Image Analysis*, N.Pérez de la Blanca, A.Sanfeliu, E.Vidal (eds.), Series in Machine Perception and Artificial Intelligence, Vol.1, World-Scientific, pp.49-61, 1992.
- [OnGa:92b] J. Oncina and P. García, "Identifying regular languages in polynomial time," in *Advances in Structural and Syntactic Pattern Recognition*, H.Bunke (ed.), World-Scientific, Singapore, pp.99-108, 1992.
- [OnGV:93] J. Oncina, P. García and E. Vidal, "Learning subsequential transducers for pattern recognition interpretation tasks," *IEEE Trans. on PAMI* 15 (5), pp.448-458, 1993.
- [OsBr:94] M. Osborne and D. Bridge, "Learning unification-based grammars using the Spoken English Corpus" in *Grammatical Inference and Applications*, R.C.Carrasco and J.Oncina (eds.), Springer-Verlag, LNAI 862, pp.260-270, 1994.
- [OsSW:86] D.N. Osherson, M. Stob and S. Weinstein (eds.), *Systems that Learn: an Introduction to Learning Theory for Cognitive and Computer Scientists*. MIT Press, Cambridge MA, 1986.
- [PaCa:78] T.W. Pao and J.W. Carr III, "A solution of the syntactical induction-inference problem for regular languages," *Computer Languages* 3, pp.53-64, 1978.
- [Pear:89] B.A. Pearlmutter, "Learning state space trajectories in recurrent neural networks," *Neural Computation* 1, pp.263-269, 1989.
- [Pitt:89] L. Pitt, "Inductive inference, DFA's, and computational complexity," in *Analogical and Inductive Inference, Proc. of the Int. Workshop AII'89*, K.P.Jantke (ed.), Lecture Notes in Artificial Intelligence 397, Springer-Verlag, Berlin, pp.18-44, 1989.
- [PoFe:91] S. Porat and J.A. Feldman, "Learning automata from ordered examples," *Machine Learning* 7, pp.109-138, 1991.

- [Poll:90] J.B. Pollack, "Recursive distribute representation," *Journal of Artificial Intelligence* **46**, pp.77-105, 1990.
- [Poll:91] J.B. Pollack, "The induction of dynamical recognizers," *Machine Learning* **7**, pp.227-252, 1991.
- [Quin:86] J.R. Quinlan, "Induction of decision trees," *Machine Learning* **1**, pp.81-106, 1986.
- [Rabi:89] L.R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. of the IEEE* **77** (2), pp.257-286, 1989.
- [RaNa:87] V. Radhakrishnan and G. Nagaraja, "Inference of regular grammars via skeletons," *IEEE Trans. Syst. Man Cybern.* **17**, pp.982-992, 1987.
- [RaNa:88] V. Radhakrishnan and G. Nagaraja, "Inference of even linear grammars and its application to picture description languages," *Pattern Recognition* **21** (1), pp.55-62, 1988.
- [RaWJ:86] L.R. Rabiner, J.G. Wilpon and B.H. Juang, "A model-based connected digit recognition system using either hidden Markov models or templates," *Computer, Speech, and Language* **1** (2), pp.167-197, 1986.
- [Reed:93] R. Reed, "Pruning algorithms - A survey," *IEEE Trans. on Neural Networks* **4** (5), pp.740-747, 1993.
- [RiVe:84] M. Richetin and F. Vernadat, "Efficient regular grammatical inference for pattern recognition," *Pattern Recognition* **17**, pp.245-250, 1984.
- [Rose:62] F. Rosenblatt, *Principles of Neurodynamics*. Academic Press, New York, 1962.
- [RuGa:94] J. Ruiz and P. García, "The algorithms RT and k-TTI: a first comparison," in *Grammatical Inference and Applications*, R.C.Carrasco and J.Oncina (eds.), Springer-Verlag, Lecture Notes in Artificial Intelligence 862, pp.180-188, 1994.
- [RuHW:86] 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, 1. Foundations*, D.E. Rumelhart, J.L. McClelland and the PDP Research Group (eds.), MIT Press, Cambridge MA, 1986.
- [RuMc:86] D.E. Rumelhart, J.L. McClelland and the PDP Research Group (eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. MIT Press, Cambridge MA, 1986.



- [RuPV:89] H. Rulot, N. Prieto and E. Vidal, "Learning accurate finite-state structural models of words through the ECGI algorithm," in *Proc. ICASSP'89*, Vol.1, pp.643-646, 1989.
- [RuVi:87] H. Rulot and E. Vidal, "Modelling (sub)string-length-based constraints through a grammatical inference method," in *Pattern Recognition Theory and Applications*, P.A. Devijver and J. Kittler (eds.), Springer-Verlag, pp.451-459, 1987.
- [SaAl:92] A. Sanfeliu and R. Alquézar, "Understanding neural networks for grammatical inference and recognition," in *Advances in Structural and Syntactic Pattern Recognition*, H.Bunke (ed.), World Scientific, pp.75-98, 1992.
- [SaAl:95] A. Sanfeliu and R. Alquézar, "Active grammatical inference: a new learning methodology," in *Shape, Structure and Pattern Recognition*, Proc. SSPR'94, Nahariya (Israel), D.Dori and A.Bruckstein (eds.), World Scientific Pub., Singapore, pp.191-200, 1995.
- [SaAl:96] A. Sanfeliu and R. Alquézar, "Efficient recognition of a class of context-sensitive languages described by augmented regular expressions," in *Advances in Structural and Syntactical Pattern Recognition*, Proc. SSPR'96, Leipzig (Germany), P.Perner, P.Wang and A.Rosenfeld (eds.), Springer-Verlag, Lecture Notes in Computer Science 1121, pp.1-10, 1996.
- [Saka:88] Y. Sakakibara, "Learning context-free grammars from structural data in polynomial time," in *Proc. 1st Workshop on Computational Learning Theory*, pp.330-344, Kaufmann, San Mateo, CA, 1988; also in *Theoretical Comput. Science* 76, pp.223-242.
- [Saka:92] Y. Sakakibara, "Efficient learning of context-free grammars from positive structural examples," *Information and Computation* 97, pp.23-60, 1992.
- [Salo:73] A. Salomaa, *Formal Languages*. Academic Press, New York, 1973.
- [SaSa:96a] A. Sanfeliu and M. Sainz, "Automatic recognition of bidimensional models learned by grammatical inference in outdoor scenes," in *Advances in Structural and Syntactical Pattern Recognition*, Proc. SSPR'96, Leipzig (Germany), P.Perner, P.Wang and A.Rosenfeld (eds.), Springer-Verlag, Lecture Notes in Computer Science 1121, pp.160-169, 1996.
- [SaSa:96b] M. Sainz and A. Sanfeliu, "Learning bidimensional context dependent models using a context-sensitive language," in *Proc. of the 13th International Conference on Pattern Recognition, ICPR'96, Vienna (Austria)*, IEEE Computer Society Press, Vol.IV, pp.565-569, 1996.

- [SaSa:97] M. Sainz and A. Sanfeliu, "A syntactical approach to learn and identify bidimensional image models," in *Proc. of the 7th Spanish Symposium on Pattern Recognition and Image Analysis, Bellaterra (Spain)*, April 1997.
- [Schm:92] J. Schmidhuber, "A fixed size storage  $O(n^3)$  time complexity learning algorithm for fully recurrent continually running networks", *Neural Computation* 4, pp.243-248, 1992.
- [SeCM:88] D. Servan-Schreiber, A. Cleeremans and J.L. McClelland, *Encoding Sequential Structure in Simple Recurrent Networks*, Technical Report CMU-CS-88-183, Carnegie-Mellon University, Pittsburgh, 1988.
- [SeCM:91] D. Servan-Schreiber, A. Cleeremans and J.L. McClelland, "Graded state machines: the representation of temporal contingencies in simple recurrent networks," *Machine Learning* 7, pp.161-193, 1991.
- [SeGa:93] J.M. Sempere and P. García, "A new regular language learning algorithm from lexicographically ordered complete samples," in *Grammatical Inference: Theory, Applications and Alternatives. Proc. of the First Int. Colloquium, Essex, UK*, S. Lucas (ed.), IEE Press, Digest No:1993/092, London, 1993.
- [SeGa:94] J.M. Sempere and P. García, "A characterization of even linear languages and its application to the learning problem," in *Grammatical Inference and Applications*, R.C.Carrasco and J.Oncina (eds.), Springer-Verlag, Lecture Notes in Artificial Intelligence 862, pp.38-44, 1994.
- [SeRo:87] T.J. Sejnowski and C.R. Rosenberg, "Parallel networks that learn to pronounce English text," *Complex Systems* 1, pp.145-168, 1987.
- [SiSG:92] H.T. Siegelmann, E.D. Sontag and C.L. Giles, "The complexity of language recognition by neural networks," *Information Processing* 92, Vol.1, Elsevier/North-Holland, pp.329-335, 1992.
- [SiSM:92] R. Siromoney, K.G. Subramanian and L. Mathew, "Learning of pattern and picture languages," *Int. J. of Pattern Recognition and Artificial Intelligence* 6 (2-3), pp.275-284, 1992.
- [SiSo:91] H.T. Siegelmann and E.D. Sontag, "Turing computability with neural nets," *Applied Mathematics Letters* 4 (6), pp.77-80, 1991.
- [SiSo:92] H.T. Siegelmann and E.D. Sontag, "On the computational power of neural nets," in *Proc. of the Fifth ACM Workshop on Computational Learning Theory*, ACM Press, New York, pp.440-449, 1992.
- [SmZi:89] A.W. Smith and D. Zipser, "Learning sequential structure with the real-time recurrent learning algorithm," *Int. Journal of Neural Systems* 1 (2), pp.125-131, 1989.

- [SoAl:94] J.M. Sopena and R. Alquézar, "Improvement of learning in recurrent networks by substituting the sigmoid activation function," *ICANN'94, Proc. of the International Conference on Artificial Neural Networks*, Sorrento, Italy, Springer-Verlag, Vol.1, pp.417-420, 1994.
- [Solo:64] R.J. Solomonoff, "A formal theory of inductive inference," *Information and Control* 7, pp.1-22, 224-254, 1964.
- [Sope:91] J.M. Sopena, "ERSP: A distributed connectionist parser that use embedded sequences to represent structure", *Tech. Rep. UB-PB-91-1, Dep. Psicologia Basica, Univ. de Barcelona*, 1991.
- [StMi:83] R.E. Stepp and R.S. Michalski, "Learning from observations: Conceptual clustering," in *Machine Learning: An Artificial Intelligence Approach*, R.S. Michalski, J.G. Carbonell and T.M. Mitchell (eds.), Morgan Kaufmann, San Mateo CA, pp.331-363, 1983.
- [SuCG:90] G.Z. Sun, H.H. Chen, C.L. Giles, Y.C. Lee, and D. Chen, , "Connectionist pushdown automata that learn context-free grammars," in *Proc. Int. Joint Conf. on Neural Networks, IJCNN-90*, Lawrence Erlbaum, Hillsdale NJ, Vol.1, pp.577-580, 1990.
- [SuGC:93] G.Z. Sun, C.L. Giles, H.H. Chen and Y.C. Lee, *The Neural Network Pushdown Automaton: Model, Stack and Learning Simulations*, Technical Report UMIACS-TR-93-77, Institute for Advanced Computer Studies, University of Maryland, 1993.
- [Taka:88] Y. Takada, "Grammatical inference for even linear languages based on control sets," *Information Processing Letters* 28 (4), pp.193-199, 1988.
- [Taka:94] Y. Takada, "A hierarchy of language families learnable by regular language learners," in *Grammatical Inference and Applications*, R.C. Carrasco and J. Oncina (eds.), Springer-Verlag, Lecture Notes in Artificial Intelligence 862, pp.16-24, 1994.
- [Tana:95] E. Tanaka, "Theoretical aspects of syntactic pattern recognition," *Pattern Recognition* 28, pp.1053-1061, 1995.
- [ThGB:86] M.G. Thomason, E. Granum and R.E. Blake, "Experiments in dynamic programming inference of Markov networks with strings representing speech data," *Pattern Recognition* 19 (5), pp.343-352, 1986.
- [ThGr:86] M.G. Thomason and E. Granum, "Dynamic programming inference of Markov networks from finite sets of sample strings," *IEEE Trans. on PAMI* 8, pp.491-501, 1986.

- [Tomi:82] M. Tomita, "Dynamic construction of finite-state automata from examples using hill-climbing," in *Proc. Fourth Annual Cognitive Science Conf.*, Ann Arbor, MI, pp.105-108, 1982.
- [TrBa:73] B. Trakhtenbrot and Y. Barzdin, *Finite Automata: Behavior and Synthesis*, North Holland Pub. Comp., Amsterdam, 1973.
- [VaBa:87] K. Vanlehn and W. Ball, "A version space approach to learning context-free grammars," *Machine Learning* 2, pp.39-74, 1987.
- [Vali:84] L.G. Valiant, "A theory of the learnable," *Communications of the ACM* 27, pp.1134-1142, 1984.
- [VaWa:78] A. Van der Mude and A. Walker, "On the inference of stochastic regular grammars," *Information and Control* 38, pp.310-329, 1978.
- [VeRi:84] F. Vernadat and M. Richetin, "Regular inference for syntactic pattern recognition: a case study," *Proc. of the 7th Int. Conf. on Pattern Recognition*, pp.1370-1372, 1984.
- [Vidal:94] E. Vidal, "Grammatical inference: an introductory survey," in *Grammatical Inference and Applications, Proc. of the Second Int. Colloquium, ICGI'94, Alicante, Spain*, R.C.Carrasco and J.Oncina (eds.), Springer-Verlag, Lecture Notes in Artificial Intelligence 862, pp.1-4, 1994.
- [ViLl:96] E. Vidal and D. Llorens, "Using knowledge to improve N-gram language modelling through the MGCI methodology," in *Grammatical Inference: Learning Syntax from Sentences*, L.Miclet and C.De la Higuera (eds.), Springer-Verlag, LNAI 1147, pp.179-190, 1996.
- [ViRV:92] E. Vidal, H. Rulot, J.M. Valiente and G. Andreu, "Font independent mixed-size digit recognition through error-correcting grammatical inference (ECGI)," in *Proc. ICPR'92, La Hague, The Netherlands*, pp.334-337, 1992.
- [ViRV:93] E. Vidal, H. Rulot, J.M. Valiente and G. Andreu, "Application of the error-correcting grammatical inference algorithm (ECGI) to planar shape recognition," in *Grammatical Inference: Theory, Applications and Alternatives. Proc. of the First Int. Colloquium, Essex, UK*, S. Lucas (ed.), IEE Press, Digest No:1993/092, London, 1993.
- [WaDa:90] P.S.P. Wang and X.W. Dai, "An algorithm for inferring context-free array grammars," in *Syntactic and Structural Pattern Recognition: Theory and Applications*, H. Bunke and A. Sanfeliu (eds.), World Scientific, 1990.
- [WaHH:89] A. Waibel, T. Hanazawa, G. Hiitoinen, K. Shikano and K. Lang, "Phoneme recognition using time delay neural networks," *IEEE Trans. on ASSP* 37 (3), pp.328-340, 1989.

- [WaKu:92] R.L. Watrous and G.M. Kuhn, "Induction of finite state languages using second-order recurrent networks," *Neural Computation* 4, pp.406-414, 1992.
- [Wang:92] P.S.P. Wang, "Three-dimensional object pattern representation by array grammars," *Int. J. Pattern Recogn. Artif. Intell.* 6 (2-3), pp.227-240, 1992.
- [Whar:74] R.M. Wharton, "Approximate language identification," *Information and Control* 26, pp.236-255, 1974.
- [WiHo:60] B. Widrow and M.E. Hoff, "Adaptive switching circuits," *1960 IRE WESCON Conv. Record, Part 4*, pp.96-104, August 1960.
- [WiPe:90] R.J. Williams and J. Peng, "An efficient gradient-based algorithm for on-line training of recurrent network trajectories," *Neural Computation* 2, pp.491-501, 1990.
- [WiZi:89] R.J. Williams and D. Zipser, "A learning algorithm for continually running fully recurrent neural networks," *Neural Computation* 1 (2), pp.270-280, 1989.
- [Woods:70] W.A. Woods, "Transition networks grammars for natural language analysis," *CACM* 13, pp.591-606, 1970.
- [Wyard:94] P. Wyard, "Representational issues for context-free grammar induction using genetic algorithms," in *Grammatical Inference and Applications*, R.C.Carrasco and J.Oncina (eds.), Springer-Verlag, Lecture Notes in Artificial Intelligence 862, pp.222-235, 1994.
- [WyNi:90] P.J. Wyard and C. Nightingale, "A single layer higher order neural net and its application to context free grammar recognition," *Connection Science* 2 (4), pp.347-370, 1990.
- [YoFu:79] K.C. You and K.S. Fu, "A syntactic approach to shape recognition using attributed grammars," *IEEE Trans. on SMC* 9 (6), pp.334-345, 1979.
- [ZeGS:93] Z. Zeng, R.M. Goodman, and P. Smyth, "Learning finite state machines with self-clustering recurrent networks," *Neural Computation* 5, pp.976-990, 1993.

