

EPILOGUE

Throughout this monograph, I have mainly concerned with the learning behaviors of stochastic automata operating in the multi-teacher environment. In our daily life, we often encounter the problem in which we need an intelligent behavior in the situation where one action elicits multi-responses from the unknown multi-criteria environment. Although this problem cannot be easily tackled, active research efforts should be directed to solve it. It is my hope that this monograph will play an important role to stimulate discussions in that direction.

This monograph has not been intended to be an encyclopedic treatise in the area of learning automata. If you are interested in the literature of learning automata, you should consult the book written by Lakshmivarahan [L6] (or the survey papers by Narendra et al ([N3], [N4], and [N9])).

REFERENCES

- A1 A.O. Allen, Probability, Statistics, and Queueing Theory, Academic Press, 1978.
- A2 H. Aso and M. Kimura, "The structures of automata to adapt to an unknown environment", IEEE Trans. Systems, Man, and Cybernetics, Vol. 6, pp. 494-504, 1976.
- A3 H. Aso, "A characterization of learning automata", Proceedings of the International Conference on Cybernetics and Society, Tokyo, pp. 1487-1491, 1978.
- A4 H. Aso and M. Kimura, "Absolute expediency of learning automata", Information Sciences, Vol. 17, pp. 91-112, 1979.
- A5 R.C. Atkinson, G.H. Bower, and E.J. Crothers, An Introduction to Mathematical Learning Theory, Wiley, 1965.
- A6 R.J. Aumann and M. Maschler, "Repeated games with incomplete information. The zero-sum extensive case", Report to the U.S. Arms Control and Disarmament Agency, Washington, D.C.: final report on contract ACDA/ST-143 prepared by MATHEMATICA, Princeton, pp. 25-108, 1968.
- B1 N. Baba and Y. Sawaragi, "On the learning behavior of stochastic automata under a nonstationary random environment", IEEE Trans. Systems, Man, and Cybernetics, Vol. 5, pp. 273-275, 1975.
- B2 N. Baba, "Learning behavior of stochastic automata in the last stage of learning", Information Sciences, Vol. 9, pp. 315-322, 1975.
- B3 N. Baba, T. Shoman, and Y. Sawaragi, "A modified convergence theorem for a random optimization theorem", Information Sciences, Vol. 13, pp. 159-166, 1977.

- B4 N. Baba, "Theoretical considerations of the parameter self-optimization by stochastic automata", *International Journal of Control*, Vol. 27, pp. 271-276, 1978.
- B5 N. Baba, T. Soeda, T. Shoman, and Y. Sawaragi, "An application of the stochastic automaton to the investment game", *International Journal of Systems Science*, Vol. 11, pp. 1447-1457, 1980.
- B6 N. Baba, "Convergence of a random optimization method for constrained optimization problems", *Journal of Optimization Theory and Applications*, Vol. 33, pp. 451-461, 1981.
- B7 N. Baba, "The absolutely expedient nonlinear reinforcement schemes under the unknown multi-teacher environment", *IEEE Trans. Systems, Man, and Cybernetics*, Vol. 13, pp. 100-108, 1983.
- B8 N. Baba, "On the learning behaviors of variable-structure stochastic automaton in the general n-teacher environment", *IEEE Trans. Systems, Man, and Cybernetics*, Vol. 13, pp. 224-231, 1983.
- B9 N. Baba, "An absolutely expedient nonlinear reinforcement scheme under a nonstationary multi-teacher environment and its applications to practical problems", *Proceedings of the Third Yale Workshop on Applications of Adaptive Systems Theory*, pp. 110-113, 1983.
- B10 N. Baba, "Learning behaviors of stochastic automata and some applications", *Working Paper, WP-83-119, IIASA, Austria*, pp. 1-26, 1983.
- B11 Y. Bar-Shalom and E. Tse, "Caution, probing, and the value of information in the control of uncertain systems", *Annals of Economic and Social Measurement*, pp. 323-337, 1976.
- B12 R. Bellman, Dynamic Programming, Princeton Univ. Press, 1957.
- B13 V.E. Benes, Mathematical Theory of Connecting Networks and Telephone Traffic, Academic Press, 1965.

- B14 J.O. Berger, Statistical Decision Theory, Springer-Verlag, 1980.
- B15 D.P. Bersekas, Dynamic Programming and Stochastic Control, 1976.
- C1 B. Chandrasekaran and D.W.C. Shen, "On expediency and convergence in variable-structure automata", IEEE Trans. Systems, Man, and Cybernetics, Vol. 4, pp. 52-60, 1968.
- C2 B. Chandrasekaran and D.W.C. Shen, "Adaptation of stochastic automata in nonstationary environments", Proc. NEC, Vol. 23, pp. 33-44, 1967.
- C3 B. Chandrasekaran and D.W.C. shen, "Stochastic automata games", IEEE Trans. Systems, Science and Cybernetics, Vol. 5, pp. 145-149, 1969.
- C4 V.K. Chichinadze, "Random search to determine the extremum of the function of several variables", Eng. Cybernetics, Vol. 1, pp. 115-123, 1967.
- C5 L.D. Cockrell and K.S. Fu, "On search techniques in switching environment", Proceedings of the 9th Symposium Adaptive Processes, Austin, Tex., 1970.
- C6 T.M. Cover and M.E. Hellman, "Two armed bandit problem with time-invariant finite memory", IEEE Trans. Information Theory, Vol. 14, pp. 185-195, 1970.
- D1 C. Derman, Finite State Markovian Decision Processes, Academic Press, 1970.
- D2 L.P. Devroye, "On the convergence of statistical search", IEEE Trans. Systems, Man, and Cybernetics, Vol. 6, pp. 46-56, 1976.
- D3 L.P. Devroye, "Probabilistic search as a search selection procedure", IEEE Trans. Systems, Man and Cybernetics, Vol. 6, pp. 315-321, 1976.
- D4 L.P. Devroye, "A class of optimal performance directed probabilistic automata", IEEE Trans. Systems, Man, and Cybernetics, Vol. 6, pp. 777-783, 1976.
- D5 L.P. Devroye, "Progressive global random search of continuous functions", Mathematical Programming, Vol. 15, pp. 330-342, 1978.

- D6 J. Dieudonne, Foundations of Modern Analysis, Academic Press, 1969.
- D7 I.C.W. Dixon and G.P. Szego, Editors, Towards Global Optimization, North-Holland, 1975.
- D8 I.C.W. Dixon and G.P. Szego, Editors, Towards Global Optimization 2, North-Holland, 1978
- D9 A.V. Dobrovidov and R.L. Stratonovich, "Construction of optimal automata functioning in random media", *Automation and Remote Control*, Vol. 25, pp. 1289-1296, 1964.
- D10 J.L. Doob, Stochastic Processes, Wiley, 1953.
- D11 A. Dvoretzky, "On stochastic approximation", in *Proc. 3rd Berkeley Symp. on Math. Stat. and Probability*, Vol.1, pp. 39-55, 1956.
- D12 E.B. Dynkin, Markov Processes, Springer-Verlag, 1965.
- E1 Y.M. El-Fattah and C. Foulard, Learning Systems: Decision, Simulation, and Control, Springer-Verlag, 1978.
- E2 Y.M. El-Fattah, "Stochastic automata modelling of certain problems of collective behavior", *IEEE Trans. Systems, Man, and Cybernetics*, Vol. 10, pp. 304-314, 1980.
- F1 Yu.A. Flerov, "Some class of multi-input automata", *Journal of Cybernetics*, Vol. 2, pp. 112-122, 1972.
- F2 K.S. Fu and G.J. McMurtry, "A study of stochastic automata as models of adaptive and learning controllers", *Purdue Univ., Tech. Rep., TR-EE 65-8*, 1965.
- F3 K.S. Fu and R.W. McLaren, "An application of stochastic automata to the synthesis of learning systems", *Purdue Univ., Tech. Rep., TR-EE 65-17*, 1965.

- F4 K.S. Fu and Z.J. Nikolic, "On some reinforcement techniques and their relation to the stochastic approximation", IEEE Trans. Automatic Control, Vol. 11, pp. 756-758, 1966.
- F5 K.S. Fu, Sequential Methods in Pattern Recognition and Machine Learning, Academic Press, 1968.
- F6 K.S. Fu and T.J. Li, "Formulation of learning automata and automata games", Information Sciences, pp. 237-256, 1969.
- F7 K.S. Fu, "Learning control systems - Review and outlook", IEEE Trans. Automatic Control, Vol. 15, pp. 210-221, 1970.
- F8 K.S. Fu, Editor, Learning Systems, The American Society of Mechanical Engineers, 1973.
- F9 K. Fukunaga, Introduction to Statistical Pattern Recognition, Academic Press, 1972.
- G1 I.I. Gihman and A.V. Skorohod, Stochastic Differential Equations, Springer-Verlag, 1972.
- G2 S.L. Ginsburg, V.Y. Krylov, and M.L. Tsetlin, "On one example of a game of many identical automata", Automation and Remote Control, Vol. 25, pp. 608-612, 1964.
- G3 S.L. Ginsburg and M.L. Tsetlin, "Some examples of simulation of the collective behavior of automata", Probl. Peredachi Informatsii, Vol. 1, pp. 54-62, 1965.
- G4 A. Ginzburg, Algebraic Theory of Automata, Academic Press, 1968.
- G5 E.G. Gladyshev, "On stochastic approximation", Theory of Probability and its Applications, Vol. 10, pp. 275-278, 1965.

- G6 R.M. Glorioso, G.R. Grueneich, and J.C. Dunn, "Self organization and adaptive routing for communication networks", 1969 EASCON Rec., pp. 243-250.
- G7 R.M. Glorioso and G.R. Grueneich, "A training algorithm for systems described by stochastic transition matrices", IEEE Trans. Systems, Man, and Cybernetics, Vol. 1, pp. 86-87, 1971.
- G8 M. Grauer, A. Lewandowski, and A.P. Wierzbicki, Editors, Multiobjective and Stochastic Optimization, CP-82-S12, IIASA, AUSTRIA, 1982.
- H1 Y.Y. Haimes, W.A. Hall, and H.T. Friedman, Multiobjective Optimization in Water Resources Systems, The Surrogate Worth Trade-off Method, Elsevier Scientific, 1975.
- H2 J.C. Harsanyi, "Games with incomplete information played by "Bayesian" players, I", Management Science, Vol. 14, pp. 159-182, 1967.
- H3 J.C. Harsanyi, "Games with incomplete information played by "Bayesian" players, part II. Bayesian equilibrium points", Ibid, Vol. 14, pp. 320-334, 1967.
- H4 J.C. Harsanyi, "Games with incomplete information played by "Bayesian" players, part III. The basic probability distribution of the game", Ibid, Vol. 14, pp. 486-502, 1968.
- H5 U. Herkenrath, D. Kalin, and S. Lakshmivarahan, "On a general of absorbing-barrier learning algorithms", Information Sciences, Vol. 24, pp. 255-263, 1981.
- H6 R.A. Howard, Dynamic Programming and Markov Processes, M.I.T. Press, 1960.
- I1 M. Iosifescu and R. Theodorescu, Random Processes and Learning, Springer-Verlag, 1969.
- I2 K. Ito, Probability Theory, in Japanese, Iwanami, 1952.
- I3 S. Ito, An Introduction to Lebesgue Integral, in Japanese, Shokabo, 1963

- J1 R.A. Jarvis, "Adaptive global search in a time-invariant environment using a probabilistic automaton", Proc. IREE, Australia, pp. 210-226, 1969.
- J2 R.A. Jarvis, "Adaptive global search in a time-variant environment using a probabilistic automaton with pattern recognition supervision", IEEE Trans. Systems, Science and Cybernetics, Vol. 6, pp. 209-217, 1970.
- K1 R.L. Kashyap, "Application of stochastic approximation" in Adaptive, Learning and Pattern Recognition Systems, J.M. Mendel and K.S. Fu, Editors, Academic Press, New York, 1970.
- K2 R.L. Kashyap, "Syntactic decision rules for recognition of spoken words and phrases using a stochastic automaton", IEEE Trans. Pat. Analys. and Mach. Intel., Vol. PAMI-1, No. 2, pp.154-163, 1979.
- K3 V.Y. Katkovnik and I.V. Antonov, "Generalization of the concept of statistical gradient", Automation and Remote Control, No. 6, pp. 26-33, 1972.
- K4 J.G. Kemeny and J.L. Snell, Finite Markov Chains, Springer-Verlag, 1976.
- K5 D.E. Koditschek and K.S. Narendra, "Fixed structure automata in multi-teacher environment", IEEE Trans. Systems, Man, and Cybernetics, Vol. 7, pp. 616-624, 1977.
- K6 E. Kohlberg, "On the nucleolus of a characteristic game", SIAM J. Appl. Math., Vol. 20, pp. 62-66, 1971.
- K7 E. Kohlberg, "Optimal strategies in repeated games with incomplete information", International Journal of Game Theory, Vol. 4, pp. 7-24, 1974.
- K8 V.I. Krinskii, "An asymptotically optimal automaton with exponential convergence", Bio Physics, Vol. 9, pp. 484-487, 1964.

- K9 V.Y. Krylov, "On one stochastic automaton which is asymptotically optimal in a random media", Automation and Remote Control, Vol. 24, pp. 1114-1116, 1963.
- K10 V.Y. Krylov and M.L. Tsetlin, "Games between automata", Automation and Remote Control, Vol. 24, pp. 889-900, 1963.
- K11 H.J. Kushner, Stochastic Stability and Control, Academic Press, 1967.
- K12 H.J. Kushner, Introduction to Stochastic Control, Holt, Rinehart and Winston, 1971.
- K13 H.J. Kushner, M.A.L. Thathachar, and S. Lakshmivarahan, "Two-state automaton with linear reward-inaction reinforcement scheme - A counter example", IEEE Trans. Systems, Man and Cybernetics, Vol. 2, pp. 292-294, 1972.
- K14 H.J. Kushner, "Stochastic approximation algorithms for the local optimization of functions with nonunique stationary points", IEEE Trans. Automatic Control, Vol. 17, pp. 646-654, 1972.
- K15 H.J. Kushner, "Convergence of recursive adaptive and identification procedures via weak convergence theory", IEEE Trans. Automatic Control, Vol. 22, pp. 921-930, 1977.
- L1 S. Lakshmivarahan and M.A.L. Thathachar, "Absolutely expedient learning algorithms for stochastic automata", IEEE Trans. Systems, Man, and Cybernetics, Vol. 3, pp. 281-286, 1973.
- L2 S. Lakshmivarahan, "Learning algorithms for stochastic automata acting in nonstationary random environments", Journal of Cybernetics, Vol. 4, pp. 73-85, 1974.

- L3 S. Lakshmivarahan and M.A.L. Thathachar, "Absolute expediency of Q- and S- model learning algorithms", IEEE Trans. Systems, Man and Cybernetics, Vol. 6, pp. 222-226, 1976.
- L4 S. Lakshmivarahan and M.A.L. Thathachar, "Bounds on the probability of convergence of learning automata", IEEE Trans. Systems, Man and Cybernetics, Vol. 6, pp. 756-763, 1976.
- L5 S. Lakshmivarahan and K.S. Narendra, "Learning algorithms for two person zero sum stochastic games with incomplete information", Mathematics of Operations Research, Vol. 6, 1981.
- L6 S. Lakshmivarahan, Learning Algorithms Theory and Applications, Springer-Verlag, 1981.
- L7 L. Ljung, "Analysis of recursive stochastic algorithms", IEEE Trans. Automatic Control, Vol. 22, pp. 551-575, 1977.
- L8 M. Loeve, Probability Theory, 4th Edition, Springer-Verlag, 1977.
- L9 R.D. Luce and H. Raiffa, Games and Decisions, Wiley, 1957.
- M1 P. Mars and M.S. Chrystall, "Real-time telephone traffic simulation using learning automata routing", S & IS Report No. 7909, Dept. of Eng. and Applied Science, Yale University, 1979.
- M2 P. Mars, K.S. Narendra, and M. Crystall, "Learning automata control of computer communication networks", Proceedings of the Third Yale Workshop on Applications of Adaptive Systems Theory, pp. 114-119, 1983.
- M3 L.G. Mason, "Self-optimizing allocation systems", Ph.D. University of Saskatchewan, Canada, 1972.

- M4 L.G. Mason, "An optimal learning algorithm for S-model environments", IEEE Trans. Automatic Control, Vol. 18, pp. 493-496, 1973.
- M5 R.W. McLaren, "A stochastic automaton model for synthesis of learning systems", IEEE Trans. Systems, Science and Cybernetics, Vol. 2, pp. 109-114, 1966.
- M6 G.J. McMurtry and K.S. Fu, "A variable-structure automaton used as a multimodal search technique", IEEE Trans. Automatic Control, Vol. 11, pp. 379-387, 1966.
- M7 N. Megiddo, "On repeated games with incomplete information played by non-bayesian players", International Journal of Game Theory, Vol. 9, pp. 157-167, 1979.
- M8 J.M. Mendel and K.S. Fu, Editors, Adaptive, Learning and Pattern Recognition Systems, Academic Press, 1970
- M9 J.M. Mendel, Discrete Techniques of Parameter Estimation, Dekker, 1973.
- M10 J.M. Mendel, "Reinforcement learning models and their applications to control problems: Learning Systems", 1973 Joint Automatic Control Conference Proceedings.
- M11 J.F. Mertens and S. Zamir, "The value of two-person zero-sum repeated games with lack of information on both sides", International Journal of Game Theory, Vol. 1, pp. 39-64, 1971.
- M12 J.F. Mertens and S. Zamir, "Minmax and maxmin of repeated games with incomplete information", International Journal of Game Theory, Vol. 9, pp. 201-215, 1979.
- M13 M.R. Meybodi and S. Lakshmivarahan, "A learning approach to priority assignment in a two class M/M/1 queueing system with unknown parameters", Proceedings of the Third Yale Workshop on Applications of Adaptive Systems Theory, pp. 106-109, 1983.

- N1 K.S. Narendra and R. Viswanathan, "A two-level system of stochastic automata for periodic random environments", IEEE Trans. Systems, Man, and Cybernetics, Vol. 2, pp. 285-289, 1972.
- N2 K.S. Narendra and R. Viswanathan, "Learning models using stochastic automata", in Proc. 1972 Int. Conf. Cybernetics and Society, Washington, D.C., pp. 9-12, 1972.
- N3 K.S. Narendra and M.A.L. Thathachar, "Learning automata - a survey", IEEE Trans. Systems, Man, and Cybernetics, Vol. 4, pp. 323-334, 1974.
- N4 K.S. Narendra and S. Lakshmivarahan, "Learning automata - a critique", Tech. Rep. 7703, Yale University, 1977.
- N5 K.S. Narendra, E. Wright, and L.G. Mason, "Application of learning automata to telephone traffic routing", IEEE Trans. Systems, Man, and Cybernetics, pp. 785-792, 1977.
- N6 K.S. Narendra and M.A.L. Thathachar, "On the behavior of learning automaton in a changing environment with routing applications", IEEE Trans. Systems, Man, and Cybernetics, Vol. 10, pp. 262-269, 1980.
- N7 K.S. Narendra, "The use of learning algorithms in telephone traffic routing - a methodology", Tech. Rep. 8203, Yale University, 1982.
- N8 K.S. Narendra and R.M. Wheeler, "An n-player sequential stochastic game with identical payoffs", Tech. Rep. 8209, Yale University, 1982.
- N9 K.S. Narendra, "Recent developments in learning automata - theory and applications", Proceedings of the Third Yale Workshop on Applications of Adaptive Systems Theory, pp. 90-99, 1983.
- N10 O.V. Nedzelnitsky, "Learning automata routing in data communication networks", Proceedings of the Third Yale Workshop on Applications of Adaptive Systems Theory, pp. 142-147, 1983.

- N11 J. Neumann and O. Morgenstern, Theory of Games and Economic Behavior, Princeton University Press, 1953.
- N12 M.B. Nevelson and R.Z. Has'minskii, Stochastic Approximation and Recursive Estimation, Translation of the American Mathematical Society, 1973.
- N13 N.J. Nilsson, Learning Machines, McGraw-Hill, 1965.
- N14 M.F. Norman, "Some convergence theorems for stochastic learning models with distance diminishing operators", *Journal of Mathematical Psychology*, Vol. 5, pp. 61-101, 1968.
- N15 M.F. Norman, "On linear models with two absorbing barriers", *Journal of Mathematical Psychology*, Vol. 5, pp. 225-241, 1968.
- N16 M.F. Norman, "Slow learning", *The British Journal of Mathematical and Statistical Psychology*, Vol. 21, pp. 141-159, 1968.
- N17 M.F. Norman, Markov Processes and Learning Models, Academic Press, 1972.
- N18 M.F. Norman, "A central limit theorem for Markov processes that move by small steps", *The Annals of Probability*, Vol. 2, pp. 1065-1074, 1974.
- N19 M.F. Norman, "Markovian learning process", *SIAM Review*, Vol. 16, pp. 143-162, 1974.
- N20 M.F. Norman, "Approximation of stochastic processes by Gaussian diffusions and applications to Wright-Fisher generic model.", *SIAM Journal of Applied Mathematics*, Vol. 29, pp. 225-242, 1975.
- P1 K.R. Parthasarathy, Probability Measures in Metric Spaces, Academic Press, 1965.
- P2 B.T. Poljak, "Nonlinear programming methods in the presence of noise", *Mathematical Programming*, pp. 87-97, 1978.

- P3 V.A. Ponomarev, "A construction of an automaton which is asymptotically optimal in a stationary random media", *Bio Physics*, Vol. 9, pp.104-110, 1964.
- P4 J.P. Ponsard and S. Zamir, "Zero-sum sequential games with incomplete information", *International Journal of Game Theory*, Vol. 2, pp. 99-110, 1974.
- P5 J.P. Ponsard, "Zero-sum games with "almost" perfect information", *Management Science*, Vol. 21, pp. 794-805, 1975.
- P6 A.S. Poznyak, "Investigation of convergence of algorithms for learning stochastic automata", *Automation and Remote Control*, pp. 77-91, 1973.
- P7 A.S. Poznyak, "Learning automata in stochastic programming problem", *Automation and Remote Control*, pp. 1608-1619, 1973.
- R1 K.R. Ramakrishnan, "Hierarchical Systems and Co-operative Games of Learning Automata", Ph.D. Thesis, Indian Institute of Science, Bangalore, India, 1982.
- R2 J.S. Riordon, "Optimal feedback characteristics from stochastic automaton models", *IEEE Trans. Automatic Control*, Vol. 14, pp. 89-92, 1969.
- R3 J.S. Riordon, "An adaptive automaton controller for discrete-time Markov processes", *Automatica*, Vol. 5, pp. 721-730, 1969.
- R4 H. Robbins and S. Monro, "A stochastic approximation method", *Annals of Mathematical Statistics*, Vol. 22, pp. 400-407, 1951.
- R5 H. Robbins, "Sequential decision problem with finite memory", *Proceedings of the National Academy of Sciences*, Vol. 42, pp. 920-923, 1956.
- R6 H.L. Royden, Real Analysis, Macmillan Pub. Co. Inc., 1963.

- S1 S.M. Samuels, "Randomized rules for the two-armed bandit with finite memory", *Annals of Mathematical Statistics*, Vol. 39, pp. 2103-2107, 1968.
- S4 G.N. Saridis, Self-Organizing Control of Stochastic Systems, Marcel Dekker Inc., 1978.
- S3 Y. Sawaragi and N. Baba, "A note on the learning behavior of variable-structure stochastic automata", *IEEE Trans. Systems, Man, and Cybernetics*, Vol. 3, pp. 644-647, 1973.
- S4 Y. Sawaragi and N. Baba, "Two ϵ -optimal nonlinear reinforcement schemes for stochastic automata", *IEEE Trans. Systems, Man, and Cybernetics*, Vol. 4, pp. 126-131, 1974.
- S5 Y. Sawaragi, N. Baba, and T. Soeda, "New topics of learning automata", *Journal of Cybernetics and Information Science*, Vol. 1, pp. 112-120, 1977.
- S6 I.J. Shapiro and K.S. Narendra, "Use of stochastic automata for parameter self-optimization with multimodal performance criteria", *IEEE Trans. Systems Science and Cybernetics*, Vol. 5, pp. 352-360, 1969.
- S7 J. Sklansky, "Learning systems for automatic control", *IEEE Trans. Automatic Control*, Vol. 11, pp. 6-19, 1966.
- S8 C.V. Smith and R. Pyke, "The Robbins-Isbell two armed bandit problem with finite memory", *Annals of Mathematical Statistics*, Vol. 36, pp. 1375-1386, 1965.
- S9 V.G. Sragovich, "Automata with multivalued input and their behavior in random environments", *Journal of Cybernetics*, Vol. 2, pp. 79-108, 1972.
- S10 P.R. Srikantakumar and K.S. Narendra, "A learning model for routing in telephone networks", *SIAM Journal on Control and Optimization*, Vol. 20, pp. 34-57, 1982.
- S11 P.R. Srikantakumar, "Application of learning theory to communication networks control", *Proceedings of the Third Yale Workshop on Applications of Adaptive Systems Theory*, pp. 135-141, 1983.

- T1 M.A.L. Thathachar and R. Bhakthavathsalam, "Learning automaton operating in parallel environments", *Journal of Cybernetics and Information Science*, Vol. 1, pp. 121-127, 1978.
- T2 M.A.L. Thathachar and K.R. Ramakrishnan, "An automaton model of a hierarchical learning system", *IFAC Congress*, Kyoto, Japan, 1981.
- T3 M.A.L. Thathachar and P.S. Sastry, "A new approach to the design of reinforcement schemes for learning automata", *Tech. Rep.*, Indian Institute of Science, 1983.
- T4 M.L. Tsetlin, "On behavior of finite automata in random media", *Automation and Remote Control*, Vol. 22, pp. 1345-1354, 1961.
- T5 M.L. Tsetlin, Automaton Theory and Modelling of Biological Systems, Academic Press, 1963.
- T6 H. Tsuji, M. Mizumoto, J. Toyoda, and K. Tanaka, "An automaton in the non-stationary random environment", *Information Sciences*, Vol. 6, pp. 123-142, 1973.
- T7 Y.Z. Tsypkin, Adaptation and Learning in Automatic Systems, Academic Press, 1971.
- T8 Y.Z. Tsypkin and A.S. Poznyak, "Finite learning automata", *Engineering Cybernetics*, Vol. 10, pp. 478-490, 1972.
- T9 Y.Z. Tsypkin, Foundations of the Theory of Learning Systems, Academic Press, 1973.
- VI V.I. Varshavskii and I.P. Vorontsova, "On the behavior of stochastic automata with variable structure", *Automation and Remote Control*, Vol. 24, pp. 327-333, 1963.

- V1 E.M. Vaisbord, "Game of two automata with differing memory depths", Automation and Remote Control, Vol. 29, pp. 440-451, 1968.
- V2 E.M. Vaisbord, "Game of many automata with various depths of memory", Automation and Remote Control, Vol. 29, pp. 1938-1943, 1968.
- V3 V.I. Varshavskii and I.P. Vorontsova, "On the behavior of stochastic automata with variable structure", Automation and Remote Control, Vol. 24, pp. 327-333, 1963.
- V4 V.I. Varshavskii, M.V. Meleshina, and M.L. Tsetlin, "Behavior of automata in periodic random media and the problem of synchronization in the presence of noise", Probl. Peredachĭ Informatsii, Vol. 1, pp. 65-71, 1965.
- V5 V.I. Varshavskii, "Some effects in the collective behavior of automata", Machine Intelligence, B. Meltzer and D. Michie, Editors, Edinburgh: Edinburgh Univ., 1969.
- V6 V.I. Varshavskii, "Automata games and control problems", IFAC Congress, Paris, 1972.
- V7 R. Viswanathan and K.S. Narendra, "Application of stochastic automata models to learning systems with multimodal performance criteria", Tech. Rep. CT 40, Yale University, 1971.
- V8 R. Viswanathan and K.S. Narendra, "A note on linear reinforcement scheme for variable structure stochastic automata", IEEE Trans. Systems, Man, and Cybernetics, Vol. 2, pp. 292-294, 1972.
- V9 R. Viswanathan and K.S. Narendra, "Competitive and cooperative games of variable-structure stochastic automata", Joint Automatic Control Conf., 1972.

- V10 R. Viswanathan and K.S. Narendra, "Stochastic automata models with application to learning systems", IEEE Trans. Systems, Man, and Cybernetics, Vol. 3, pp. 107-111, 1973.
- V11 R. Viswanathan and K.S. Narendra, "Games of stochastic automata", IEEE Trans. Systems, Man, and Cybernetics, Vol. 4, pp. 131-135, 1974.
- V12 I.P. Vorontsova "Algorithms for changing automaton transition probabilities", Problemi Peredachi Informatsii, Vol. 1, pp. 122-126, 1965.
- W1 M.D. Waltz and K.S. Fu, "A heuristic approach to reinforcement learning control system", IEEE Trans. Automatic Control, Vol. 10, pp. 390-398, 1965.
- W2 M.T. Wasan, Stochastic Approximations, Cambridge University Press, 1969.
- W3 S. Watanabe, Knowing and Guessing, Wiley, 1969.
- W4 S. Watanabe, "Creative learning and propensity automaton", IEEE Trans. Systems, Man, and Cybernetics, Vol. 5, pp. 603-610, 1975.
- W5 W.G. Wee and K.S. Fu, "A heuristic approach to reinforcement learning control system", IEEE Trans. Systems, Man, and Cybernetics, Vol. 5, pp. 215-223, 1969.
- W6 I.H. Witten, "Finite time performance of some two-armed bandit controller", IEEE Trans. Systems, Man, and Cybernetics, Vol. 3, pp. 194-197, 1973.
- W7 I.H. Witten, "The apparent conflict between estimation and control - A survey of two-armed bandit problem", Journal of Franklin Institute, Vol. 301, pp. 161-189.
- W8 I.H. Witten, "An adaptive optimal controller for discrete-time Markov environments", Information and Control, Vol. 34, pp. 286-295, 1977.

- Z1 L.A. Zadeh, "Fuzzy sets", Information and Control, Vol. 8, pp. 338-353, 1965.
- Z2 S. Zamir, "On the notion of the value for games with infinitely many stages" , Annals of Statistics, Vol. 1, pp. 791-796, 1973.

INDEX

| | |
|--|---------------|
| absolutely expedient | 2, 3, 8 |
| absolutely expedient algorithm | 10 |
| absolute expediency in the general n-teacher environment | 3, 23, 27 |
| a priori information | 1 |
| asymptotically optimal | 2 |
| average penalty | 6 |
| average sum of the penalty probabilities | 107 |
| average weighted reward in the general n-teacher environment | 22, 29 |
| Basic space | 12, 56 |
| Borel field | 12, 56 |
| coalition | 4, 91, 94, 96 |
| conditional expectation | 13, 29, 60 |
| converges with probabilities | 1, 16, 34 |
| cooperative game | 90, 94 |
| discrete parameter Semi-Martingale | 14 |
| distribution function | 13, 60, 74 |
| expedient | 7 |
| expedient in the general n-teacher environment | 23 |
| fixed-structure automata | 18, 21, 36 |

| | |
|---|------------------------|
| GAE reinforcement scheme | 3, 25, 26, 27, 37 |
| Gaussian white noise | 72, 78 |
| general n-teacher environment | 3, 18, 21, 55 |
| GL_{R-I} scheme | 26, 37, 64, 65 |
| GNA scheme | 26, 37 |
| hierarchical structure automata | 4, 90, 93, 104 |
| incomplete information | 90 |
| learning automaton | 1, 56 |
| learning behavior | 1, 17, 55, 71, 104 |
| learning control | 1 |
| L_{R-I} scheme | 9 |
| majority decision | 36 |
| mathematical expectation | 13, 16, 32 |
| MGAE scheme | 4, 55, 59, 76, 78 |
| multi-teacher environment | 1, 17, 25, 55, 71, 104 |
| noise-corrupted, multi-objective function | 4, 71, 78 |
| nonstationary multi-teacher environment | 3, 55, 56, 71, 88 |
| nonstationary random environment | 6 |
| optimal control | 1 |
| optimality | 7, 8 |
| ε -optimality | 2, 3, 7 |

| | |
|--|----------------------------|
| optimal in the general n-teacher environment | 24 |
| ϵ -optimal in the general n-teacher environment | 3, 24, 27 |
| optimal in NMT | 58 |
| ϵ -optimal in NMT | 58 |
| parameter self-optimization | 4, 71 |
| Pareto-optimal parameter | 77 |
| payoff | 91 |
| penalty response | 4, 9, 18, 25, 74 |
| penalty strength | 58 |
| perfect information | 1 |
| P-model | 4, 20, 55, 63, 69, 73, 104 |
| probability density function | 13 |
| probability measure space | 12, 15, 56, 74 |
| Q-model | 5, 20 |
| random environment | 1, 5, 6 |
| random variable | 12 |
| reinforcement scheme | 3, 6 |
| repeated game | 90, 91 |
| reward-inaction scheme | 9 |
| reward response | 4, 18, 25 |
| routing of messages in communication network | 17 |
| Semi-Martingale | 14, 16, 32 |

| | |
|---|---------------------------|
| single teacher environment | 4, 6, 23 |
| smallest Borel field | 15, 56, 74 |
| S-model | 5, 20, 56, 64 |
| stationary random environment | 1, 6 |
| stochastic automaton | 1, 4, 17, 55, 72, 91, 104 |
| stochastic process | 14 |
| sum of the penalty strength | 58 |
| super regular function | 34 |
| telephone network routing | 70 |
| two person zero-sum game | 71 |
| uniformly distributed random variable | 64 |
| variable-structure stochastic automaton | 6 |