UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Rules and associations : hidden Markov models and neural networks in the psychology of learning

Visser, I.

Publication date
2002

Link to publication

Citation for published version (APA):
Visser, I. (2002). *Rules and associations : hidden Markov models and neural networks in the psychology of learning*. [Thesis, fully internal, Universiteit van Amsterdam].

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:17 Oct 2026

Bibliography

ABRAMOWITZ, M. and SEGUN, I. A. (eds.) (1964). *Handbook of mathematical functions*. New York: Dover Publications.

AKAIKE, H. (1973). Information theory and an extension of the maximum likelihood principle. In *Second international symposium on information theory* (B. N. Petrov and F. Csaki, eds.). Budapest: Academiai Kiado, 267–281.

ARBIB, M. A. (ed.) (1995). *The Handbook of Brain Theory and Neural Networks*. Cambridge: MIT Press.

ARMINGER, G., CLOGG, C. C. and SOBEL, M. E. (eds.) (1995). *Handbook of statistical modeling for the social and behavioral sciences*. New York: Plenum press.

BADDELEY, R., HANCOCK, P. and FÖLDIÁK, P. (eds.) (1999). *Information theory and the brain*. Cambridge University Press.

BAUM, L. E. and PETRIE, T. (1966). Statistical inference for probabilistic functions of finite state Markov chains. *Annals of Mathematical Statistics* **67** 1554–40.

BECKER, J., HONERKAMP, J., HIRSCH, J., FRÖBE, U., SCHLATTER, E. and GREGER, R. (1994). Analysing ion channels with hidden Markov models. *European Journal of Physiology* **426** 328–332.

BERKELEY, I. S. N. (2000). What the #$% ! is a subsymbol? *Minds and Machines* **10** 1–13.

BORSBOOM, D., MELLENBERGH, G. J. and VAN HEERDEN, J. (2001). Philosophy of science and psychometrics: Reflections on the theoretical status of the latent variable. *Submitted for publication* .

BOWER, J. M. (1993). The modulation of learning state in a biological associative memory: An in vitro, in vivo, and in computo study of object recognition in mammalian olfactory cortex. *Artificial Intelligence Review* **7(5)** 261–269.

BOZDOGAN, H. (2000). Akaike's information criterion and recent developments in informational complexity. *Journal of Mathematical Psychology* **44–1** 62–91.

BRAINERD, C. J. (1979). Markovian interpretations of conservation learning. *Psychological Review* **86-3** 181–213.

BRAINERD, C. J. (1980). Developmental invariance in a mathematical model of associative learning. *Child Development* **51(2)** 349–363.

BRAINERD, C. J. (1985). Three-state models of memory development: A review of advances in statistical methodology. *Journal of Experimental Child Psychology* **40** 375–394.

BULLINARIA, J. A. (1999). Free gifts from connectionist modelling. In Baddeley et al. (1999), chap. 12.

CHAUVIN, Y. and RUMMELHART, D. E. (eds.) (1995). *Backpropagation: Theory, architectures, and applications*. Hillsdale (NJ): Lawrence Erlbaum Associates.

CHIEN, J. T. and WANG, H. C. (1997). Telephone speech recognition based on bayesian adaptation of hidden Markov models. *Speech Communication* **22(4)** 369–384.

CHOMSKY, N. (1959a). On certain formal properties of grammars. *Information and Control* **2** 137–167.

CHOMSKY, N. (1959b). A review of B.F. Skinner's Verbal Behavior. *Language* **35–1** 26–58.

CHOMSKY, N. (1980). Rules and representations. *Bahavioral and Brain Sciences* **3** 1–61.

CHRISTIANSEN, M. H., CONWAY, C. M. and CURTIN, S. (2000). A connectionist single-mechanism account of rule-like behavior in infancy. In *Proceedings of the twenty-second annual conference of the cognitive science society* (L. R. Gleitman and A. K. Joshi, eds.). Lawrence Erlbaum Associates.

CLEEREMANS, A. and JIMENEZ, L. (1998). Implicit sequence learning: The truth is in the details. In *Handbook of Implicit Learning* (M. Stadler and P. Frensch, eds.). Thousand Oaks (Ca): Sage Publications, 323–364.

CLEEREMANS, A. and MCCLELLAND, J. L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology: General* **120** 235–253.

CLEEREMANS, A., SERVAN-SCHREIBER, D. and MCCLELLAND, J. L. (1989). Finite state automata and simple recurrent networks. *Neural Computation* **1** 372–381.

CLOGG, C. C. (1995). Latent class models. In Arminger et al. (1995).

DENNIS, J. E. and SCHNABEL, R. B. (1983). *Numerical methods for unconstrained optimization.* Englewood Cliffs, N.J.: Prentice-Hall.

DESTREBECQZ, A. and CLEEREMANS. A. (2001). Can sequence learning be implicit? new evidence with the process dissociation procedure. *Psychonomic Bulletin & Review* **8(2)** 343–350.

DIENES, Z., ALTMANN, G. T. and GAO, S.-J. (1999). Mapping across domains without feedback: A neural network model of transfer of implicit knowledge. *Cognitive Science* **23–1** 53–82.

DURBIN, M. A., EARWOOD, J. and GOLDEN, R. M. (2000). Hidden Markov models for coding story recall data. In *Proceedings of the twenty-second annual conference of the cognitive science society* (L. R. Gleitman and A. K. Joshi, eds.). Lawrence Erlbaum Associates.

EFRON, B. and TIBSHIRANI, R. J. (1986). Bootstrap methods for standard errors, confidence intervals and other measures of statistical accuracy. *Statistical Science* **1** 54–77.

EFRON, B. and TIBSHIRANI, R. J. (1993). *An introduction to the bootstrap.* Monographs on statistics and applied probability 57, New York: Chapman & Hall.

ELLIOTT, R. J., AGGOUN, L. and MOORE, J. B. (1995). *Hidden Markov models: Estimation and control.* New York: Springer Verlag.

ELMAN, J. L. (1990). Finding structure in time. *Cognitive Science* **14** 179–211.

ELMAN. J. L. (1993). Learning and development in neural networks: the importance of starting small. *Cognition* **48** 71–99.

FARAONE. S. V. (1986). A Statistical Analysis System (SAS) computer program for markov chain analysis. *Journal of Psychopathology and Behavioral Assessment* **8(4)** 367–379.

FODOR, J. A. (1981). Methodological solipsism considered as a research strategy in cognitive psychology. In *Representations.* Brighton: The Harvester Press, 225–253.

FODOR, J. A. and PYLYSHYN, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition* **28** 3–71.

FRENSCH, P. A., BUCHNER, A. and LIN, J. (1994). Implicit learning of unique and ambiguous serial transitions in the presence and absence of a distractor task. *Journal of Experimental Psychology: Learning, memory and cognition* **20–3** 567–584.

GERS, F. A. and SCHMIDHUBER, J. (2001). Lstm recurrent networks learn simple context free and context sensitive languages. *IEEE Transactions on Neural Networks* Accepted for publication.

GILES, C. L., MILLER, C. B., CHEN, D., CHEN, H. H., SUN, G. Z. and LEE, Y. C. (1992). Learning and extracting finite state automata with second-order recurrent neural networks. *Neural Computation* **4** 393–405.

GILL, P. E., MURRAY, W. and WRIGHT, M. H. (1981). *Practical optimization.* New York: Academic Press.

GRÜNWALD, P. (2001). Model selection based on minimum description length. *Journal of Mathematical Psychology* **44** 133–152.

HADLEY, R. F. (2000). Cognition and the computational power of connectionist networks. *Connection Science* **12–2** 95–110.

HARNAD, S. (1990). The symbol grounding problem. *Physica D* **42** 335–346.

HARTMAN, E. J., KEELER, J. D. and KOWALSKI, J. M. (1990). Layered neural networks with Gaussian hidden units as universal approximations. *Neural Computation* **2** 210–215.

HOPCROFT, J. and ULLMAN, J. (1979). *Introduction to automata theory, languages and computation.* Redwood City (CA): Addison-Wesley.

HOPCROFT, J. E., MOTWANI, R. and ULLMAN, J. D. (2001). *Introduction to Automata Theory, Languages, and Computation.* 2nd ed. Boston: Addison Wesley.

HORNIK, K., STINCHCOMBE, M. and WHITE, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks* **2** 359–366.

HUIZENGA, H. M., HESLENFELD, D. J. and MOLENAAR, P. C. M. (Under revision). Optimal design for eeg/meg source analysis. *Journal of the Royal Statistical Society, series C, Applied Statistics* .

INTRATOR, O. and INTRATOR, N. (2001). Interpreting neural-network results: a simulation study. *Computational Statistics & Data Analysis* **37** 373–393.

JANSEN, B. (2001). *Development of reasoning on the balance scale task: Psychometric assessment of cognitive strategies.* Ph.D. thesis, University of Amsterdam, Department of Psychology.

JIMENEZ, L. and MENDEZ, C. (1999). Which attention is needed for implicit sequence learning? *Journal of Experimental Psychology: Learning, Memory and Cognition* **25–1** 236–259.

JIMENEZ, L. and MENDEZ, C. (2001). Implicit sequence learning with competing implicit cues. *The Quarterly Journal of Experimental Psychology* **55–4** 345–369.

JIMENEZ, L., MENDEZ, C. and CLEEREMANS, A. (1996). Comparing direct and indirect measures of sequence learning. *Journal of Experimental Psychology: Learning, Memory and Cognition* **22–4** 948–969.

JÖRESKOG, K. and SÖRBOM, D. (1999). *LISREL 8 [Computer program].* Scientific Software International., Chicago.

KEMENY, J. G. and SNELL, J. (1960). *Finite Markov chains.* Princeton: Van Nostrand.

KINGMA, J. and REUVEKAMP, J. (1987a). Markov count: a program for computing the learning statistics of two-stage markov learning experiments. *Educational and Psychological measurement* **47(1)** 89–98.

KINGMA, J. and REUVEKAMP, J. (1987b). Markov-forget: A package for parameter estimation and hypothesis testing of 5, 6, 7, 8, 9 and 10-parameter two-state forgetting models. *Educational and Psychological measurement* **47(3)** 673–678.

KINTSCH, W. and MORRIS, C. J. (1965). Application of a Markov model to free recall and recognition. *Journal of Experimental Psychology* **69(2)** 200–206.

KROGH, A. (1998). An introduction to hidden Markov models for biological sequences. In *Computational methods in molecular biology* (S. L. Salzberg, D. B. Searls and S. Kasif, eds.), chap. 4. Elsevier, 45–63.

LANGEHEINE, R., PANNEKOEK, J. and VAN DE POL, F. (1995). Bootstrapping goodness-of-fit measures in categorical data analysis. Tech. rep., Statistics Netherlands, CBS.

LANGEHEINE, R., PANNEKOEK, J. and VANDEPOL, F. (1996). Bootstrapping goodness-of-fit measures in categorical data analysis. *Sociological methods and research* **24(4)** 492–516.

LEVELT, W. J. M. (1990). Are multilayer feedforward networks effectively turing machines? *Psychological Research* **52** 153–157.

LEWICKI, P., HILL, T. and BIZOT, E. (1988). Acquisition of procedural knowledge about a pattern of stimuli that cannot be articulated. *Cognitive Psychology* **20** 24–37.

LEWICKI, P., HOFFMAN, H. and CZYZEWSKA, M. (1987). Unconscious acquisition of complex procedural knowledge. *Journal of Experimental Psychology: Learning, memory and cognition* **13–4** 523–530.

LI, M. and VITÁNYI, P. (1997). *An introduction to Kolmogorov complexity and its applications*. 2nd ed. Graduate texts in computer science, New York etc.: Springer.

LIN, T. H. and DAYTON, C. M. (1997). Model selection information criteria for non-nested latent class models. *Journal of Educational and Behavioral Statistics* **22(3)** 249–264.

LIND, D. and MARCUS, B. (1995). *Symbolic dynamics and coding*. Cambridge University Press.

LUCE, R. D. (1986). *Reponse times: their role in inferring elementary mental organization*. Oxford: Oxford University press.

LYSTIG, T. C. and HUGHES, J. P. (2002/in press). Exact computation of the observed information matrix for hidden markov models. *Journal of Computational and Graphical Statistics* .

MARCUS, G. F., VIJAYAN, S., RAO, S. B. and VISHTON, P. M. (1999). Rule learning by seven-month-old infants. *Science* **283** 77–80.

MCCULLOCH, W. S. and PITTS, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* **5** 115–133.

MEEKER, W. Q. and ESCOBAR, L. A. (1995). Teaching about approximate confidence regions based on maximum likelihood estimation. *The American Statistician* **49(1)** 48–53.

MILLER, G. A. (1952). Finite Markov processes in psychology. *Psychometrika* **17** 149–167.

MILLER, G. A. and CHOMSKY, N. (1963). Finitary models of language users. In *Handbook of mathematical psychology* (R. Luce, R. R. Bush and E. Galanter, eds.), chap. 13. New York: Wiley.

MINSKY, M. and PAPERT, S. (1969). *Perceptrons*. Cambridge: Mit Press.

MOOIJAART, A. and VAN DER HEIJDEN, P. (1992). The EM algorithm for latent class analysis with equality constraints. *Psychometrika* **57(2)** 261–269.

NICOLSON, R. I. (1982). Shades of all-or-none learning: A stimulus sampling model. *British Journal of Mathematical and Statistical Psychology* **35(2)** 162–170.

NISSEN, M. J. and BULLEMER, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology* **19** 1–32.

OMLIN, C. and GILES, C. (1996). Extraction of rules from discrete-time recurrent neural networks. *Neural Networks* **9(1)** 41–51.

O'REILLY, R. C. and RUDY, J. W. (2000). Computational principles of learning in the neocortex and hippocampus. *Hippocampus* **10(4)** 389–397.

PERRUCHET, P. and AMORIM, M.-A. (1992). Conscious knowledge and changes in performance in sequence learning: Evidence against dissociation. *Journal of Experimental Psychology: Learning, Memory and Cognition* **18(4)** 785–800.

QUINLAN, P. (1991). *Connectionism and psychology: a psychological perspective on new connectionist research.* Hertfordshire: Harvester-Wheatsheaf.

RAAIJMAKERS, J. G. W. (1981). A general framework for the analysis of concept identification tasks. *Acta Psychologica* **49** 233–261.

RAAIJMAKERS, J. G. W. and SHIFFRIN, R. M. (In press). Models of memory. In *Handbook ofExperimental Psychology, Third Edition, Volume 2: Memory and Cognitive Processes* (H. Pashler and D. Medin, eds.). New York: John Wiley & Sons, Inc.

RABINER, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of IEEE* **77–2** 267–295.

RAMACHER, U. (1994). Hamiltonian dynamics of neural networks. In *Computational learning theory and natural learning systems* (S. Hanson, T. Pesche, M. Kearns and R. Rivest, eds.), vol. 2 of *Intersection between theory and experiment*, chap. 16. Boston: MIT Press, 251–277.

REBER, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior* **6** 317–327.

REBER, A. S. (1976). Implicit learning of synthetic languages: The role of the instructional set. *Journal of Experimental Psychology: Human Learning and Memory* **2–1** 88–94.

REBER, A. S. (1993). *Implicit learning and tacit knowledge : an essay on the cognitive unconscious.* New York etc.: Oxford University Press.

RESCORLA, R. A. and WAGNER, A. R. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In *Classical conditioning II: current research and theory* (A. H. Black and W. F. Prokasy, eds.), chap. 3. New York, N.Y. : Appleton-Century-Crofts.

RODRIQUEZ, P. (In press). Simple recurrent networks learn context-free and context-sensitive languages by counting. *Neural Computation* .

RODRIQUEZ, P., WILES, J. and ELMAN, J. L. (1999). A recurrent neural network that learns to count. *Connection Science* **11(1)** 5-40.

RUMMELHART, D. and MCCLELLAND, J. (eds.) (1986). *Parallel Distributed Processing.* Cambridge: MIT Press.

SCHMIDBAUER, O., CASACUBERTA, F., CASTRO, M. J. and HEGERL, G. (1993). Articulatory representation and speech technology. *Language and Speech* **36–2** 331–351.

SCHULTZ, T. R. and BALE, A. C. (In press). Neural network simulation of infant familiarization to artificial sentences: Rule-like behavior. *Infancy* .

SCHWARZ, G. (1978). Estimating the dimension of a model. *Annals of Statistics* **6** 461–464.

SEARLE, J. R. (1983). *Intentionality.* Cambridge: Cambridge University Press.

SEBER, G. A. F. and WILD, C. J. (1989). *Nonlinear regression.* Wiley series in probability and mathematical statistics, New York: John Wiley & Sons.

SEGER, C. A. (1997). Two forms of sequential implicit learning. *Consciousness and Cognition: An International Journal* **6** 108–131.

SHANKS, D. R. and JOHNSTONE, T. (1999). Evaluating the relationship between explicit and implicit knowledge in a sequential reaction time task. *Journal of Experimental Psychology: Learning, Memory and Cognition* **25–6** 1435–1451.

SHANKS, D. R. and PERRUCHET, P. (In press). Dissociation between priming and recognition in the expression of sequential knowledge. *Psychonomic Bulletin & Review* .

SHANKS, D. R. and ST-JOHN, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences* **17(3)** 367–447.

SHANNON, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal* **27** 623–656.

SHANON, B. (1993). *The Representational and the Presentational, an Essay on Cognition and the Study of Mind.* London: Harvester-Wheatsheaf.

SIEGELMANN, H. T. and SONTAG, E. D. (1995). On the computational power of neural nets. *Journal of Computer System Sciences* **50(1)** 132–150.

SMOLENSKY, P. (1988). On the proper treatment of connectionism. *Behavioral and Brain Sciences* **11** 1–74.

SONTAG, E. D. (1995). Automata and neural networks. In Arbib (1995), 119–124.

SPROTT, D. A. (1973). Normal likelihoods and their relation to large sample theory of estimation. *Biometrika* **60** 457–465.

STEVENS, J. (1996). *Applied Multivariate Statistics for the Social Sciences.* 3rd ed. Mahwah, NJ: Lawrence Erlbaum Associates.

SUN, R. (2000). Symbol grounding: a new look at an old idea. *Philosophical Psychology* **13–2** 149–172.

TABOR, W. (2000). Fractal encoding of context-free grammars in connectionist networks. *Expert Systems* **17(1)** 41–56.

TIMMER, J. and KLEIN, S. (1997). Testing the Markov condition in ion channel recordings. *Physical Review E* **55–3** 3306–3311.

TINO, P. and KÖTELES, M. (2000). Extracting finite state representations from recurrent neural networks trained on chaotic symbolic sequences. *IEEE Transactions on Neural Networks* In press.

TURING, A. M. (1950/1990). Computing machinery and intelligence. In *The Philosophy of Artificial Intelligence* (M. A. Boden, ed.). Cambridge: MIT Press.

VAN DER VELDE, F. (1993). Is the brain an effective turing machine or a finite-state machine? *Psychological Research* **55** 71–79.

VENZON, D. J. and MOOLGAVKAR, S. H. (1988). A method for computing profile-likelihood-based confidence intervals. *Applied Statistics* **37(1)** 87–94.

VERMUNT, J. (1997). *lEM: A general program for the analysis of categorical data [Program manual].*

VISSER, I. (1996). Mind rules: a philosophical essay on psychological rules and the the rules of psychology. Tech. Rep. X–96–01, Amsterdam: Institute for Logic, Language and Computation. Masters thesis under supervision of Prof. Dr. Renate Bartsch, reprints available from: ingmar@dds.nl.

VISSER, I. (2001). *Markovfit 1.6 [Computer program manual]. Program for fitting hidden Markov models for categorical data.* Copies of manual and source codes available from ingmar@dds.nl.

VISSER, I., RAIJMAKERS, M. E. J. and MOLENAAR, P. C. M. (2000). Confidence intervals for hidden Markov model parameters. *British journal of mathematical and statistical psychology* **53** 317–327.

VISSER, I., RAIJMAKERS, M. E. J. and MOLENAAR, P. C. M. (2001a). Hidden Markov model interpretations of neural networks. In *Connectionist models of Learning, Development and Evolution. The 6th Neural Computation and Psychology Workshop (NCPW6)* (R. M. French and J. P. Sougné, eds.). Springer-Verlag.

VISSER, I., RAIJMAKERS, M. E. J. and MOLENAAR, P. C. M. (2001b). Reaction times and predictions in implicit sequence learning: Testing predictions by the simple recurrent network. *Submitted for publication* Available from ingmar@dds.nl.

VISSER, I., RAIJMAKERS, M. E. J. and MOLENAAR, P. C. M. (2002). Associations and dissociations between direct and indirect measures of sequence knowledge. *Submitted for publication* Copies available from first author, ingmar@dds.nl.

WALD, A. (1943). Test of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society* **54** 426–482.

WHITTAKER, J. (1990). *Information Divergence*, chap. 4. Wiley series in probability and mathematical statistics, Chichester: John Wiley.

WICKENS, T. D. (1982). *Models for Behavior: Stochastic processes in psychology.* San Francisco: W. H. Freeman and Company.

WICKENS, T. D. (1989). *Multiway contingency tables analysis for the social sciences.* Hillsdale (NJ): Lawrence Erlbaum Associates.

WITTGENSTEIN, L. (1978). *Philosophical Investigations.* Oxford: Basil Blackwell. This is the standard English translation by G.E.M. Anscombe.

YANG, J., XU, Y. and CHEN, C. S. (1997). Human action learning via hidden Markov model. *IEEE Transactions on Systems, Man and Cybernetics* **27–1** 34–44.