



**UvA-DARE (Digital Academic Repository)**

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

Visser, I.

[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

**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: <http://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

# Index

---

- AIC, *see* Akaike Information Criterion
- Akaike Information Criterion, 54
- all-or-none model, 60
- anticipation, 71, 80
- anticipations, 98
- automata extraction, 17, 24–25
  
- Bayesian Information Criterion, 54
- BIC, *see* Bayesian Information Criterion
- bootstrap, 36
  - non-naive, 36
  - parametric, 36
- Borel measurable function, 12
  
- cluster analysis, 17, 24
- cognitive architecture, 9
  - connectionist alternative, 10
- combinatorial syntax, 9
- computational power
  - and language recognition, 10
  - and the Turing machine, 14
  - of feedforward networks, 12
  - of recurrent neural network, 14
- concept identification, 58
- concept identification model, 60
- confidence interval, 57
- constrained maximum likelihood, 35
- covariance matrix, 33
- cross-entropy, 65
  
- deterministic functions, 13
- distributed representation, 15, 107
- distributed representations, 18
  - advantages, 20
  
- entropy, 28
  - Lempel-Ziv approximation, 29
- equality constraints, 50–52
- explicit knowledge, 72
  
- feedforward network, 12
  - language recognition, 13
- finite differences, 36
  
- formal languages
  - and computational power, 10
- formality condition, 2
- forward variables, 33
- fractal networks, 19
- free generation, 65
- free generation task
  - analysis, 89, 100
  - online, 86
- frequency effects, 20
- FSA, *see* finite state automaton
- functional equivalence, 19
  
- generation task, 72–73, 86
  - online, 73
- graceful degradation, 20
- grounding, 1
  
- hessian matrix, 33
- hidden Markov model
  - analysis of verbal reports, 98
  - and automata extraction, 17, 107
  - and rule-following, 4
  - applications, 25
  - comparison with Markov model, 31, 88, 110
  - definition, 25, 31, 48, 88, 109
  - degrees of freedom, 54
  - distance between, 65
  - distance to grammar, 95
  - dynamics, 26
  - equivalence with FSA, 63
  - examples, 31
  - for analyzing neural networks, 17
  - for implicit learning, 66
  - likelihood, 32, 49
- hidden unit activity, 17, 19, 24
- HMM, *see* hidden Markov model
  
- implicit learning
  - and explicit knowledge, 105
  - and rule-following, 4

- different measures, 69
- direct and indirect measures, 104
- explicit knowledge in, 72
- predictions by simple recurrent network, 71
- initial state probabilities, 50
- internal representations, 16
- k-means clustering, 24
- latent variable models, 4
- learning
  - and grounding, 1
  - importance of, 1, 11
- likelihood profile
  - cubic approximation, 39
  - linear approximation, 37
  - quadratic approximation, 39
- likelihood profiles, 35
- likelihood ratio distribution (figure), 37
- likelihood ratio statistic, 35, 52, 116
- localist representation, 107
- localist representations, 14
- Markov model, 31, 47, 110
  - parameter estimation, 58
- model selection criteria, 53, 114
- modification indices, 116
- multilayer network, 12
- neural network
  - interpretation, 16
- noise tolerance, 20
- partitioning, 24
- passage times, 51
- prediction error, 27, 56, 116
- prediction task, 72
- reaction times, 69
- Reber grammar
  - figure, 27
- recurrent neural network, 14
  - and Turing computability, 14
  - with distributed representations, 15
  - with localist representations, 14
- response function, 110
- RNN, *see* recurrent neural network
- rule-following, 106
  - according to Chomsky, 1
  - according to Wittgenstein, 1
- acquisition, 4
  - criteria for, 3, 5
  - examples, 1
  - in neural networks, 3
  - models of, 2
- rules, 94
  - in neural networks, 24
- secant method, 37
- second-order RNN, 17
- sequence learning, 64, 69
  - different measures, 99
- simple recurrent network, 25, 70
  - and context-free languages, 18
  - and finite state automata, 16
  - and regular languages, 16
  - and subsymbols, 19
  - description, 15
  - figure, 15
  - for sequence learning, 70
  - in sequence learning, 87
  - memory, 25
  - prediction, 27
  - predictions for sequence learning, 71
- simple recurrent network (figure), 25, 70
- SRN, *see* simple recurrent network
- standard error, 57
- starting values, 49, 119
- state space model, 17
- subsymbolic processing, 15, 19
- subsymbols, 19-20
  - and hidden unit activity, 19
- symbol manipulation, 10
- universal function approximation, 12
  - limitations, 13
- verbal reports
  - analysis with HMMs, 98
- win-stay/lose-shift model, 60