

# Radford M. Neal

Curriculum Vitae, April 2008

## Biographical Information

### Personal Data

Date of Birth: 12 Sept 1956

Citizenship: Canadian

Home address:

21 Shaftesbury Avenue, #106  
Toronto, Ontario M4T 3B4  
(416) 964-1513

University address:

Department of Statistics  
University of Toronto M5S 3G3  
(416) 978-4970

Web page: <http://www.cs.utoronto.ca/~radford/>

E-mail address: [radford@stat.utoronto.ca](mailto:radford@stat.utoronto.ca)

### Degrees Awarded

BSc (Honours) in Computer Science, University of Calgary, 1977

MSc in Computer Science, University of Calgary, 1980

Thesis title: *An Editor for Trees*

Supervisor: David Hill

Ph.D in Computer Science, University of Toronto, 1995

Thesis title: *Bayesian Learning for Neural Networks*

Supervisor: Geoffrey Hinton

### Employment and Other Relevant Experience

Present position:

Professor, Department of Statistics and Department of Computer Science, University of Toronto, Canada Research Chair in Statistics and Machine Learning.

Appointed as Lecturer 1 July 1994.

Appointed as Associate Member (Limited Term) of the graduate faculty 13 January 1995.

Promoted to Assistant Professor 1 July 1995.

Promoted to Associate Professor and full member of the graduate faculty 1 July 1999.

Promoted to Full Professor 1 July 2001.

Awarded a Canada Research Chair in Statistics and Machine Learning 1 July 2003.

Cross-appointed to the Department of Public Health Sciences 14 February 2006.

Earlier teaching experience:

Sessional instructor in computer science at the University of Calgary for the following courses:

Summers 1979-1981 Taught half of the core second-year course on machine architecture and programming.

Winter 1981 Taught a fourth-year course on computer graphics.

1986/87 & 1987/88 Taught a fourth-year project course on system software design and development.

Teaching assistant in computer science at the University of Calgary (1977-1979) and the University of Toronto (1989).

Statistical consulting experience:

I have provided statistical consulting for a Canadian company regarding the analysis of mass spectroscopy data in 1990–1991, for a Canadian financial services company in 1997, for a US biomedical research company in 2004–2006, and for the Canadian Department of National Defence in 2006–2007.

Software engineering experience:

I did contract work as a software engineer on a variety of industrial projects, totaling approximately three years work, at various times between 1980 and 1989. These projects involved audio signal processing, computer graphics, data acquisition, oil and gas production accounting, data base design, and local-area networks.

From 1981 to 1985 I was a software engineer and research associate at the University of Calgary, first on a project to implement a code generator for a Simula compiler, and later as the system architect and team leader for software development in support of a research project in distributed computer systems.

I worked part-time as a programmer supporting researchers at the University of Calgary in computer science, management, and biology at various times between 1975 and 1980.

## Academic History

### Research Endeavours

I have active research interests in the following areas:

**Bayesian inference:** I am interested in fundamental issues concerning the Bayesian approach to statistical inference, in methods for solving the computational problems it gives rise to, and in the development of Bayesian models for practical applications.

**Markov chain Monte Carlo:** The only known way to solve many computational problems in Bayesian inference and statistical physics is to use Monte Carlo methods based on the simulation of a Markov chain. I am interested in improving these methods, and in applying them to problems of Bayesian inference, particularly for complex models such as those based on neural networks.

**Flexible models for regression and classification:** I am interested in methods for modeling regression and classification functions that allow a wide range of functional forms, including those based on neural networks and on Gaussian process priors.

**Latent variable models:** Probabilistic models with unobservable latent variables are widely used in some areas of science, and represent one possible approach to “concept learning” in artificial intelligence. I am interested in the development of these models, such as mixture models and models based on Dirichlet diffusion trees, and in their computational implementation, using Markov chain Monte Carlo and other methods.

**Empirical evaluation of statistical methods:** The merits of various regression and classification procedures, such as those based on neural networks, must ultimately be judged by their performance on realistic tasks. I am involved with the DELVE project, which aims to support such evaluations by providing a large collection of datasets and an environment in which they can be used to make valid comparisons.

**Data compression:** Efficient coding must be based on a probabilistic model for the data to be coded. I am interested in devising such models, and in efficiently using them for lossy and lossless compression of data.

**Error-correcting codes:** Information can also be encoded with a controlled amount of redundancy to allow correction of transmission errors. I am interested in developing such error-correcting codes that can be decoded using probabilistic methods.

**Current application areas:** Biological applications I am interested include haplotype inference, analysis of gene expression data, reconstruction of phylogenies, and analysis of spectroscopic data. I am also working on calibrating models for the history of glaciation, and on modeling the health effects of air pollution.

I have done work in the past on programming environments, compilers, operating systems, and computer graphics.

## Research Prize

In May 2004, I received the Lindley Prize, awarded every two years by the International Society for Bayesian Analysis for “innovative research in Bayesian Statistics” that was presented at the preceding ISBA or Valencia meeting. This award was for my paper on “Density modeling and clustering using Dirichlet diffusion trees”.

## Research Awards

- 2005 \$170,000 (\$34,000 per year) for the period 1 April 2005 to 31 March 2010, individual research grant from the Natural Sciences and Engineering Research Council (NSERC), for research on Bayesian analysis using flexible models.
- 2003 Awarded a Canada Research Chair, which comes with a \$10,000 per year research grant, commencing 1 July 2003.
- 2001 \$96,000 (\$24,000 per year) for the period 1 April 2001 to 31 March 2005, individual research grant from the Natural Sciences and Engineering Research Council (NSERC), for research on theory, computation, and applications of Bayesian inference.
- 1999 \$100,000 over the period from 1999 to 2003, Premier’s Research Excellence Award.
- 1999 \$30,000 (\$15,000 per year) for the period 1 April 1999 to 31 March 2001, individual research grant from the Natural Sciences and Engineering Research Council (NSERC), for research on theory and computation for realistic Bayesian models.
- 1998 \$58,914 (about \$15,000 per year) for the period 1 April 1998 to 31 March 2002, my portion of a grant from the Institute for Robotics and Intelligent Systems for

“Basic research and continuing education in learning algorithms”, with Y. Bengio (lead PI), R. Tibshirani, and F. Major.

1995 \$48,000 (\$12,000 per year) for the period 1 April 1995 to 31 March 1999, individual research grant from the Natural Sciences and Engineering Research Council (NSERC), for research on Markov chain Monte Carlo methods.

Became a Principal Investigator with G. Hinton (lead PI), R. Tibshirani, D. Lowe, and M. LeBlanc on a project to research applications of neural networks and other statistical methods for medical diagnosis and financial forecasting, financed by a grant of \$669,736 (over four years) from the Institute for Robotics and Intelligent Systems (IRIS).

1994 \$6,200 from the Connaught Fund, as a new faculty start-up grant.

## Scholarly and Professional Work

### Book

Neal, R. M. (1996) *Bayesian Learning for Neural Networks*, Lecture Notes in Statistics No. 118, New York: Springer-Verlag.

### Refereed Research Papers

Li, L., Zhang, J., and Neal, R. M. (2008) “A method for avoiding bias from feature selection with application to naive Bayes classification models”, *Bayesian Analysis*, posted online 2008-02-13, 25 pages. [Joint Author]

Sun, S., Greenwood, C. M. T., and Neal, R. M. (2007) “Haplotype inference using a Bayesian hidden Markov model”, *Genetic Epidemiology*, vol. 31, pp. 937-948. [Joint Author]

Jain, S. and Neal, R. M. (2007) “Splitting and merging components of a nonconjugate Dirichlet process mixture model” (with discussion), *Bayesian Analysis*, vol. 2, pp. 445-472. [Secondary Author]

Listgarten, J., Neal, R. M., Roweis, S. T., Wong, P., and Emili, A. (2007) “Difference detection in LC-MS data for protein biomarker discovery”, *Bioinformatics*, vol. 23, pp. e198-e204. [Joint Author]

Shahbaba, B. and Neal, R. M. (2007) “Improving classification when a class hierarchy is available using a hierarchy-based prior”, *Bayesian Analysis*, vol. 2, pp. 221-238, posted online 2006-11-16. [Joint Author]

Shahbaba, B. and Neal, R. M. (2006) “Gene function classification using Bayesian models with hierarchy-based priors”, *BMC Bioinformatics*, 7:448, 9 pages. [Joint Author]

Jain, S. and Neal, R. M. (2004) “A split-merge Markov chain Monte Carlo procedure for the Dirichlet process mixture model”, *Journal of Computational and Graphical Statistics*, vol. 13, pp. 158-182. [Joint Author]

Neal, R. M. (2003) “Slice sampling” (with discussion), *Annals of Statistics*, vol. 31, pp. 705-767.

- Neal, R. M. (2001) “Annealed importance sampling”, *Statistics and Computing*, vol. 11, pp. 125-139.
- Diaconis, P., Holmes, S., and Neal, R. M. (2000) “Analysis of a non-reversible Markov chain sampler”, *Annals of Applied Probability*, vol. 10, pp. 726-752. [Joint Author]
- Neal, R. M. (2000) “Markov chain sampling methods for Dirichlet process mixture models”, *Journal of Computational and Graphical Statistics*, vol. 9, pp. 249-265.
- Moffat, A., Neal, R. M., and Witten, I. H. (1998) “Arithmetic coding revisited”, *ACM Transactions on Information Systems*, vol. 16, pp. 256-294. [Joint author]
- Neal, R. M. and Dayan, P. (1997) “Factor analysis using delta-rule wake-sleep learning”, *Neural Computation*, vol. 9, pp. 1781-1803. [Primary author]
- MacKay, D. J. C. and Neal, R. M. (1996) “Near Shannon limit performance of low density parity check codes”, *Electronics Letters*, vol. 32, pp. 1645-1646. Reprinted with printing errors corrected in vol. 33, pp. 457-458. [Secondary author]
- Neal, R. M. (1996) “Sampling from multimodal distributions using tempered transitions”, *Statistics and Computing*, vol. 6, pp. 353-366.
- Dayan, P., Hinton, G. E., Neal, R. M., and Zemel, R. S. (1995) “The Helmholtz machine”, *Neural Computation*, vol. 7, pp. 1022-1037. [Secondary author]
- Hinton, G. E., Dayan, P., Frey, B. J., and Neal, R. M. (1995) “The wake-sleep algorithm for unsupervised neural networks”, *Science*, vol. 268, pp. 1158-1161. [Secondary author]
- Neal, R. M. (1994) “An improved acceptance procedure for the hybrid Monte Carlo algorithm”, *Journal of Computational Physics*, vol. 111, pp. 194-203.
- Neal, R. M. (1992) “Connectionist learning of belief networks”, *Artificial Intelligence*, vol. 56, pp. 71-113.
- Witten, I. H., Birtwistle, G. M., Cleary, J. G., Hill, D. R., Levinson, D., Lomow, G. A., Neal, R., Peterson, M., Unger, B. W., and Wyvill, B. L. M. (1983) “JADE: A distributed software prototyping environment”, *ACM Operating Systems Review*, vol. 7, no. 3. [Secondary author]
- Witten, I. H. and Neal, R. M. (1982) “Using Peano curves for bilevel display of continuous-tone images”, *IEEE Computer Graphics and Applications*, vol. 2, no. 3. [Joint author]

## Refereed Research Notes

- Neal, R. M. (2000) “On deducing conditional independence from  $d$ -separation in causal graphs with feedback:”, *Journal of Artificial Intelligence Research*, vol. 12, pp. 87-91.
- Gilks, W. R., Neal, R. M., Best, N. G., and Tan, K. K. C. (1997) “Corrigendum: Adaptive rejection Metropolis sampling”, *Applied Statistics*, vol. 46, pp. 541-542. [Joint author, not an author of the original paper]

Neal, R. M. (1993) “Comments on ‘A theoretical analysis of Monte Carlo algorithms for the simulation of Gibbs random field images’ ”, *IEEE Transactions on Information Theory*, vol. 39, p. 310.

Neal, R. M. (1992) “Asymmetric parallel Boltzmann machines are belief networks”, *Neural Computation*, vol. 4, pp. 832-834.

### Refereed Tutorial/Review Papers

Kass, R. E., Carlin, B. P., Gelman, A., and Neal, R. M. (1998) “Markov Chain Monte Carlo in Practice: A Roundtable Discussion”, *The American Statistician*, vol. 52, pp. 93-100. [Joint author]

Witten, I. H., Neal, R. M., and Cleary, J. G. (1987) “Arithmetic coding for data compression”, *Communications of the ACM*, vol. 30, pp. 520-540. [Joint author]

### Refereed Conference Papers

Neal, R. M. (2008) “Computing likelihood functions for high-energy physics experiments when distributions are defined by simulators with nuisance parameters”, in the proceedings of the PHYSTAT-LHC Workshop on Statistical Issues for LHC Physics, June 2007, CERN 2008-001, pp. 119-126.

Listgarten, J., Neal, R. M., Roweis, S. T., Puckrin, R., and Cutler, S. (2007) “Bayesian detection of infrequent differences in sets of time series with shared structure”, in B. Schölkopf, *et al.* (editors), *Advances in Neural Information Processing Systems 19* (aka NIPS\*2006), 8 pages. [Joint Author]

Meeds, E., Ghahramani, Z., Neal, R., and Roweis, S. (2007) “Modeling dyadic data with binary latent features”, in B. Schölkopf, *et al.* (editors), *Advances in Neural Information Processing Systems 19*, (aka NIPS\*2006), 8 pages. [Secondary Author]

Listgarten, J., Neal, R. M., Roweis, S. T., and Emili, A. (2005) “Multiple alignment of continuous time series”, in L. K. Saul, *et al.* (editors), *Advances in Neural Information Processing Systems 17*, 8 pages. [Joint Author]

Neal, R. M., Beal, M. J., and Roweis, S. T. (2004) “Inferring state sequences for non-linear systems with embedded hidden Markov models”, in S. Thrun, *et al.* (editors) *Advances in Neural Information Processing Systems 16*, 8 pages. [Joint author]

Neal, R. M. (2003) “Density modeling and clustering using Dirichlet diffusion trees”, in J. M. Bernardo, *et al.* (editors) *Bayesian Statistics 7*, Oxford University Press, pp. 619-629.

Harvey, M. and Neal, R. M. (2000) “Inference for belief networks using coupling from the past”, in C. Boutilier and M. Goldszmidt (editors), *Uncertainty in Artificial Intelligence: Proceedings of the Sixteenth Conference (2000)*, pp. 256-263. [Joint Author]

Neal, R. M. (1999) “Regression and classification using Gaussian process priors” (with discussion), in J. M. Bernardo, *et al.* (editors) *Bayesian Statistics 6*, Oxford University Press, pp. 475-501.

- MacKay, D. J. C. and Neal, R. M. (1995) “Good codes based on very sparse matrices”, in C. Boyd (editor) *Cryptography and Coding: 5th IAM Conference*, Lecture Notes in Computer Science No. 1025, pp. 100-111. Springer-Verlag. [Secondary author]
- Moffat, A., Neal, R., and Witten, I. H. (1995) “Arithmetic coding revisited”, in J. A. Storer and M. Cohn (editors) *Proceedings of the Fifth IEEE Data Compression Conference* (Snowbird, Utah, 1995), pp. 202-211, Los Alamitos, California: IEEE Computer Society Press. [Joint author]
- Neal, R. M. (1993) “Bayesian learning via stochastic dynamics”, in C. L. Giles, S. J. Hanson, and J. D. Cowan (editors) *Advances in Neural Information Processing Systems 5*, pp. 475-482, San Mateo, California: Morgan Kaufmann.

### Unrefereed Conference Papers

- Hinton, G. E., Dayan, P., To, A., and Neal, R. (1995) “The Helmholtz machine through time”, in F. Fogelman-Soulié and R. Gallinari (editors) *ICANN-95*, pp. 483-490. [Secondary author]
- Neal, R. M. (1992) “Bayesian mixture modeling”, in C. R. Smith, G. J. Erickson, and P. O. Neudorfer (editors) *Maximum Entropy and Bayesian Methods: Proceedings of the 11th International Workshop on Maximum Entropy and Bayesian Methods of Statistical Analysis* (Seattle, 1991), pp. 197-211, Dordrecht: Kluwer Academic Publishers.
- Neal, R. M., Lomow, G. A., Peterson, M., Unger, B. W., and Witten, I. H. (1984) “Experience with an inter-process communication protocol in a distributed programming environment”, *CIPS Session 84 Conference* (Calgary). [Primary author]
- Wyvill, B. L. M., Neal, R., Levinson, D., and Bramwell, B. (1984) “JAGGIES: A distributed hierarchical graphics system”, *CIPS Session 84 Conference* (Calgary). [Secondary author]

### Book Chapters

- Neal, R. M. and Zhang, J. (2006) “High dimensional classification with Bayesian neural networks and Dirichlet diffusion trees”, in I. Guyon, S. Gunn, M. Nikravesh, and L. A. Zadeh (editors) *Feature Extraction: Foundations and Applications*, Studies in Fuzziness and Soft Computing, Volume 207, Springer, pp. 265-2956. [Primary Author]
- Neal, R. M. (2006) “Classification with Bayesian neural networks”, in J. Quiñero-Candela, B. Magnini, I. Dagan, and F. D’Alche-Buc (editors) *Machine Learning Challenges. Evaluating Predictive Uncertainty, Visual Object Classification, and Recognising Textual Entailment*, Lecture Notes in Computer Science no. 3944, Springer-Verlag, pp.28-32.
- Neal, R. M. (1998) “Assessing relevance determination methods using DELVE”, in C. M. Bishop (editor), *Neural Networks and Machine Learning*, Berlin: Springer-Verlag, pp. 97-129.

Neal, R. M. (1998) “Suppressing random walks in Markov chain Monte Carlo using ordered overrelaxation”, in M. I. Jordan (editor), *Learning in Graphical Models*, pp. 205-228, Dordrecht: Kluwer Academic Publishers.

Neal, R. M. and Hinton, G. E. (1998) “A view of the EM algorithm that justifies incremental, sparse, and other variants”, in M. I. Jordan (editor), *Learning in Graphical Models*, Kluwer Academic Press, pp. 355-368, Dordrecht: Kluwer Academic Publishers. [Primary author]

## Other Publications

Neal, R. M. (2007) Book review: Pattern Recognition and Machine Learning, by Christopher M. Bishop, *Technometrics*, vol. 49, p. 366.

## Technical Reports

Technical Reports not duplicating other papers listed:

Shahbaba, B. and Neal, R. M. (2007) “Nonlinear models using Dirichlet process mixtures”, Technical Report No. 0707, Dept. of Statistics, 16 pages. [Secondary Author]

Neal, R. M. (2006) “Puzzles of anthropic reasoning resolved using full non-indexical conditioning”, Technical Report No. 0607, Dept. of Statistics, University of Toronto, 53 pages.

Neal, R. M. (2005) “Estimating ratios of normalizing constants using Linked Importance Sampling”, Technical Report No. 0511, Dept. of Statistics, University of Toronto, 37 pages.

Neal, R. M. (2005) “The short-cut Metropolis method”, Technical Report No. 0506, Dept. of Statistics, University of Toronto, 28 pages.

Neal, R. M. (2004) “Taking bigger Metropolis steps by dragging fast variables”, Technical Report No. 0411, Dept. of Statistics, University of Toronto, 9 pages.

Neal, R. M. (2004) “Improving asymptotic variance of MCMC estimators: Non-reversible chains are better”, Technical Report No. 0406, Dept. of Statistics, University of Toronto, 25 pages.

Neal, R. M. (2003) “Markov chain sampling for non-linear state space models using embedded hidden Markov models”, Technical Report No. 0304, Dept. of Statistics, University of Toronto, 9 pages.

Neal, R. M. (2002) “Circularly-coupled Markov chain sampling”, Technical Report No. 9910 (original version November 1999, revised version February 2002), Dept. of Statistics, University of Toronto, 49 pages.

Neal, R. M. (2001) “Transferring prior information between models using imaginary data”, Technical Report No. 0108, Dept. of Statistics, University of Toronto, 29 pages.

Neal, R. M. (2001) “Defining priors for distributions using Dirichlet diffusion trees”, Technical Report No. 0104, Dept. of Statistics, University of Toronto, 25 pages.

- Pinto, R. L. and Neal, R. M. (2001) “Improving Markov Chain Monte Carlo Estimators by Coupling to an Approximating Chain”, Technical Report No. 0101, Dept. of Statistics, University of Toronto, 13 pages. [Joint Author]
- Neal, R. M. (1997) “Monte Carlo implementation of Gaussian process models for Bayesian regression and classification”, Technical Report No. 9702, Dept. of Statistics, University of Toronto, 24 pages.
- Neal, R. M. (1993) *Probabilistic Inference Using Markov Chain Monte Carlo Methods*, Technical Report CRG-TR-93-1, Dept. of Computer Science, University of Toronto, 144 pages.
- Neal, R. M. (1992) “Bayesian training of backpropagation networks by the hybrid Monte Carlo method”, Technical Report CRG-TR-92-1, Dept. of Computer Science, University of Toronto, 21 pages.

Technical Reports largely similar to other papers listed:

- Li, L., Zhang, J., and Neal, R. M. (2007) “A method for avoiding bias from feature selection with application to naive Bayes classification models”, Technical Report No. 0705, Dept. of Statistics, University of Toronto, 21 pages. [Joint Author] [A revised version was published in *Bayesian Analysis*]
- Shahbaba, B. and Neal, R. M. (2006) “Gene function classification using Bayesian models with hierarchy-based priors”, Technical Report No. 0606, Dept. of Statistics, University of Toronto, 14 pages. [Joint Author] [A revised version was published in *BMC Bioinformatics*]
- Shahbaba, B. and Neal, R. M. (2005) “Improving classification when a class hierarchy is available using a hierarchy-based prior”, Technical Report No. 0510, Dept. of Statistics, University of Toronto, 11 pages. [Joint Author] [A revised version was published in *Bayesian Analysis*]
- Jain, S. and Neal, R. M. (2005) “Splitting and merging components of a nonconjugate Dirichlet process mixture model”, Technical Report No. 0507, Dept. of Statistics, University of Toronto, 37 pages. [Secondary Author] [A revised version was published in *Bayesian Analysis*]
- Neal, R. M. (2000) “Slice sampling”, Technical Report No. 2005, Dept. of Statistics, University of Toronto, 40 pages. [A revised version was published in *Annals of Statistics*]
- Jain, S. and Neal, R. M. (2000) “A split-merge Markov chain Monte Carlo procedure for the Dirichlet process mixture model”, Technical Report No. 2003, Dept. of Statistics, University of Toronto, 32 pages. [Joint Author] [A revised version was published in the *Journal of Computational and Graphical Statistics*]
- Neal, R. M. (1998) “Annealed importance sampling”, Technical Report No. 9805, Dept. of Statistics, University of Toronto, 25 pages (first version 18 February 1998, revised 1 September 1998). [A revised version was published in *Statistics and Computing*]
- Neal, R. M. (1998) “Markov chain sampling methods for Dirichlet process mixture models”, Technical Report No. 9815, Dept. of Statistics, University of Toronto, 17 pages. [A revised version was published in the *Journal of Computational and Graphical Statistics*]

- Neal, R. M. (1997) “Markov chain Monte Carlo methods based on ‘slicing’ the density function”, Technical Report No. 9722, Dept. of Statistics, University of Toronto, 27 pages. [Was revised and extended, and re-titled as “Slice sampling”]
- Diaconis, P., Holmes, S., and Neal, R. M. (1997) “Analysis of a non-reversible Markov chain sampler”, Technical Report BU-1385-M, Biometrics Unit, Cornell University, 26 pages. [Joint Author] [A revised version was published in *Annals of Applied Probability*]
- Neal, R. M. and Dayan, P. (1996) “Factor analysis using delta-rule wake-sleep learning”, Technical Report No. 9607, Dept. of Statistics, University of Toronto, 23 pages. [Primary Author] [A revised version was published in *Neural Computation*]
- Neal, R. M. (1995) “Suppressing random walks in Markov chain Monte Carlo using ordered overrelaxation”, Technical Report No. 9508, Dept. of Statistics, University of Toronto, 22 pages. [A revised version became a chapter in *Learning in Graphical Models*]
- Neal, R. M. (1994) “Sampling from multimodal distributions using tempered transitions”, Technical Report No. 9421, Dept. of Statistics, University of Toronto, 22 pages. [A revised version was published in *Statistics and Computing*]
- Neal, R. M. (1994) “Priors for infinite networks” Technical Report CRG-TR-94-1, Dept. of Computer Science, University of Toronto, 22 pages. [A revised version is part of *Bayesian Learning for Neural Networks*]
- Neal, R. M. (1991) “Bayesian mixture modeling by Monte Carlo simulation”, Technical Report CRG-TR-91-2, Dept. of Computer Science, University of Toronto, 23 pages. [A condensed version became “Bayesian mixture modeling”, in *Maximum Entropy and Bayesian Methods: Proceedings of the 11th International Workshop.*]
- Neal, R. M. (1990) “Learning stochastic feedforward networks”, Technical Report CRG-TR-90-7, Dept. of Computer Science, University of Toronto, 34 pages. [A revised version became “Connectionist learning of belief networks”, published in *Artificial Intelligence.*]

## Publicly-Released Research Software

**Software for flexible Bayesian modeling:** Implements the models and algorithms described in *Bayesian Learning for Neural Networks*, in “Monte Carlo implementation of Gaussian process models for Bayesian regression and classification”, in “Markov chain sampling for Dirichlet process mixture models”, and in “Density modeling and clustering using Dirichlet diffusion trees”. Also allows Markov chain Monte Carlo methods to be applied to distributions specified by simple formulas. Approximately 200 pages of documentation and 45,000 lines of C program code. First public release August 1995 (for neural network models only), most recent release November 2004.

**Software for Low Density Parity Check codes:** Supports research into Low Density Parity Check (LDPC) codes, algorithms for encoding them, and algorithms for decoding them using probability propagation. Also includes modules implementing operations on dense and sparse modulo-2 matrices. Approximately 50 pages of documentation and

7000 lines of C program code. First public release March 2000, most recent release February 2006.

**The DELVE software for assessing learning methods:** I contributed substantially to the design of the DELVE software, with C. E. Rasmussen, G. E. Hinton, R. Tibshirani, and others, and wrote about half of the 102 page user’s manual. First public release December 1996.

**Low-precision arithmetic coding software:** This consists of about 3 pages of documentation and 1300 lines of code. First public release July 1991; last updated May 1993. I also contributed to the design of the arithmetic coding software maintained by Alistair Moffat, Dept. of Computer Science, University of Melbourne, which implements the algorithms in “Arithmetic coding revisited”.

**R function for performing univariate slice sampling:** A small R function released March 2008.

### Talks at Conferences Without Proceedings

- Oct 2004 Sun, S., Greenwood, C. T. M., and Neal, R. M. “Haplotype inference using a Hidden Markov Model with efficient Markov Chain sampling”, poster presentation, American Society of Human Genetics annual meeting, Toronto. [Joint Author]
- Aug 2004 Neal, R. M. “Discussion of a talk by M. Jordan on “Variational inference in exponential families: Alternatives to MCMC”, Joint Statistical Meetings, Toronto.
- May 2004 Neal, R. M. “A new proof of Peskun’s theorem regarding the asymptotic variance of MCMC estimators”, ISBA 2004, Viña del Mar, Chile.
- Apr 2004 Tarasov, L., Neal, R., and Peltier, W. R. “A Bayesian calibration methodology applied to ice-sheet modeling”, European Geosciences Union general assembly, Nice. [Joint author, presented by L. Tarasov]
- Aug 2001 Neal, R. M. “Survival Analysis Using a Bayesian Neural Network”, Joint Statistical Meetings, Atlanta.
- Aug 1999 Neal, R. M. “Markov Chain Sampling Using Hamiltonian Dynamics”, Joint Statistical Meetings, Baltimore.
- June 1996 Neal, R. M. “Faster Markov chain Sampling Using Overrelaxation Methods”, Statistical Society of Canada annual meeting, Waterloo.
- May 1994 Neal, R. M. “Hierarchical Bayesian inference for Hidden Markov Models”, Statistical Society of Canada annual meeting, Banff.

### Invited Tutorial Presentation

- Dec 2004 Tutorial on “Bayesian Methods in Machine Learning”, NIPS 2004.

## Invited Workshop Participation

- June 2007 PHYSTAT-LHC Workshop on Statistical Issues for LHC Physics, Organizers: Louis Lyons, Albert De Roeck, and others.
- May 2007 Third Workshop on Monte Carlo Methods, Harvard University. Organizer: Jun Liu.
- July 2006 Workshop on “Statistical Inference Problems in High Energy Physics and Astronomy”, Banff International Research Station. Organizers: James Linnemann, Louis Lyons, and Nancy Reid.
- April 2006 Workshop on “Nonlinearity and Randomness in Complex Systems”, Dept. of Mathematics, State University of New York at Buffalo. Organizers: Gino Biondini, Bruce Pitman, and William Kath.
- June 2005 Workshop on “Mathematical Issues in Molecular Dynamics”, Banff International Research Station. Organizers: Robert Skeel and Paul Tupper.
- Aug 2004 Second Workshop on Monte Carlo Methods, Harvard University. Organizer: Jun Liu.
- Dec 2003 NIPS 2003 workshop on feature extraction, Whistler. Organizer: Isabelle Guyon.
- Sep 2002 First Cape Cod Workshop on Monte Carlo Methods, Hyannis, Massachusetts. Sponsored by the NSF, IMS, and ISBA.
- Dec 2001 Workshop on “New Trends in Optimization and Computational Algorithms”, Kyoto, Japan. Sponsored by the The Institute of Statistical Mathematics.
- May 2001 Workshop on “Statistical Physics and Capacity-Approaching Codes”, Trieste, Italy. Sponsored by the Abdus Salam International Centre for Theoretical Physics.
- Mar 2001 Workshop on “Statistical Mixtures and Latent Structure Models”, ICMS, Edinburgh, UK. Organizers: C. P. Robert and D. M. Titterton.
- Oct 1999 Workshop on “Relating Causal Structure to Conditional Independence Structure”, The Fields Institute, Toronto. Organizers: Thomas Richardson and Peter Spirtes.
- Aug 1999 Workshop on “Codes, Systems and Graphical Models”, Institute for Mathematics and its Applications, Minneapolis. Organizers: G. David Forney, *et al.*
- Oct 1998 Workshop on Monte Carlo Methods, part of the programme on “Probability and Its Applications” at the Fields Institute, Toronto. I was a co-organizer (with N. Madras and J. Rosenthal) of this workshop.
- Aug 1997 Programme on “Neural Networks and Machine Learning” at the Newton Institute, University of Cambridge. Organizers: Chris Bishop, *et al.*

- June 1996 AMS-IMS-SIAM Joint Workshop on “Stochastic Inference, Monte Carlo, and Empirical Methods”, Mt. Holyoke, Massachusetts, Organizers: Ian McKeague and Priscilla Greenwood.
- Apr 1995 Workshop on “Statistics and Neural Networks”, International Centre for Mathematical Sciences, Edinburgh. Organizers: Jim Kay and Mike Titterton.
- June 1994 AMS-IMS-SIAM Joint Workshop on “Markov Chain Monte Carlo Methods”, Mt. Holyoke, Massachusetts. Organizers: A. Gelfand, C. Geyer, P. Green, and A. F. M. Smith.
- Nov 1993 The “Statistical basis of learning and classification” section of the Computer Vision programme at the Newton Institute, University of Cambridge. Organizer: Brian Ripley.
- Apr 1992 Workshop on “Neural networks and nonparametric regression”, Math Sciences Research Institute, Berkeley. Organizers: Trevor Hastie and Rob Tibshirani.
- Dec 1991 Workshop on Bayesian neural network learning following the 1991 conference on Neural Information Processing Systems. Organizers: David MacKay and Steve Nowlan.

### Invited Lectures

- 17 Jan 2008 Dept. of Statistics, University of Toronto, “Constructing Efficient MCMC Methods Using Temporary Mapping and Caching”.
- 29 Nov 2007 Dept. of Statistics, University of Western Ontario, “Constructing Efficient MCMC Methods Using Temporary Mapping and Caching”.
- 27 Sep 2007 Canadian Institute for Theoretical Astrophysics, “A Critique of Anthropic Reasoning, with Applications to Cosmology and the Fermi Paradox”.
- 12 Feb 2007 Dept. of Computer Science, University of Toronto, “Constructing Efficient MCMC Methods Using Temporary Mapping and Caching”.
- 11 Dec 2006 Dept. of Statistics, Columbia University, “Constructing Efficient MCMC Methods Using Temporary Mapping and Caching”.
- 21 Feb 2003 Dept. of Mathematics and Statistics, York University, “Density Modeling and Clustering Using Dirichlet Diffusion Trees”.
- 12 Jun 2002 Cavendish Laboratory, Cambridge University, “Lossy Data Compression”.
- 10 Jun 2002 Dept. of Mathematics, University of Bristol, “Circularly-Coupled Markov Chain Sampling”.
- 22 May 2002 The Applied Biostatistics Association / Clinical Biostatistics Rounds, Toronto, “Survival analysis using a Bayesian neural network”.

- 12 Apr 2002 Dept. of Chemistry (Statistical Mechanics Group), University of Toronto: “Defining an Importance Sampling Distribution Using Hamiltonian Dynamics with Annealing”.
- 29 Nov 2001 Dept. of Statistics, University of Toronto: “Density Estimation and Clustering Using Dirichlet Diffusion Trees”.
- 1 Feb 2001 Dept. of Statistics, University of Toronto: “Circularly-Coupled Markov Chain Sampling”.
- 30 Jun 2000 Dept. of Statistics, Stanford University: “Circularly-Coupled Markov Chain Sampling”.
- 19 Apr 2000 Graduate School of Business, University of Chicago: “Circularly-Coupled Markov Chain Sampling”.
- 11 Mar 1999 Dept. of Physiology, University of Toronto: “Viewing Learning as Statistical Estimation for Models with Latent Variables”.
- 3 Dec 1998 Dept. of Statistics, University of Toronto: “Markov Chain Sampling Methods for Dirichlet Process Mixture Models”.
- 25 Nov 1998 Dept. of Mathematics and Statistics, York University: “Markov Chain Sampling Methods for Dirichlet Process Mixture Models”.
- 29 May 1998 Dept. of Engineering, University of Cambridge: “Annealed Importance Sampling”.
- 28 May 1998 Computer Science and Applied Mathematics, Aston University, Birmingham: “Annealed Importance Sampling”.
- 30 Nov 1997 Industrial Engineering and Management, Technion: “Bayesian Learning for Neural Networks”.
- 19 Nov 1997 Dept. of Statistics, Carnegie Mellon University: “Markov chain Monte Carlo methods based on ‘slicing’ the density function”.
- 20 Aug 1997 Dept. of Pure Mathematics and Mathematical Statistics, Cambridge University: “Markov chain Monte Carlo methods based on ‘slicing’ the density function”.
- 12 May 1997 Dept. of Statistics, University of Chicago: “Regression and classification using Gaussian process priors”.
- 4 April 1996 Dept. of Statistics, University of Western Ontario: “Markov chain Monte Carlo methods based on ‘slicing’ the density function”.
- 23 Feb 1995 Dept. of Statistics, University of Toronto: “Suppressing random walks in Markov chain Monte Carlo methods”.
- 14 Feb 1995 Dept. of Mathematics and Statistics, York University: “Sampling from multimodal distributions using tempered transitions”.
- 1 Dec 1994 Dept. of Statistics, University of Waterloo: “Sampling from multimodal distributions using tempered transitions”.

- 3 Feb 1994 Dept. of Statistics, University of Toronto: “Bayesian neural networks”.
- 11 Feb 1993 Dept. of Computer Science, University of Toronto: “Learning concepts with hierarchical latent variable models”.

## **Teaching Activities**

### **Undergraduate Courses Taught Recently (Fall 2003 – Spring 2008 )**

- Spring 2007 STA 414 — Statistical Methods for Machine Learning and Data Mining  
(cross listing of STA 2104).
- Fall 2006 CSC 411 — Machine Learning and Data Mining  
I was solely responsible for teaching this course.
- Fall 2006 STA 247 — Probability with Computer Applications  
I acted as coordinator for this course, and taught one of the two lecture sections.
- Spring 2006 STA 414 — Statistical Methods for Machine Learning and Data Mining  
(cross listing of STA 2104).
- Fall 2004 STA 247 — Probability with Computer Applications  
I acted as coordinator for this course, and taught one of the two lecture sections.
- Spring 2004 STA 410 — Statistical Computation (cross listing of STA 2102)
- Spring 2004 CSC 310 — Information Theory  
I was solely responsible for teaching this course.
- Fall 2003 STA 247 — Probability with Computer Applications  
This was a new course, for which I was solely responsible.

### **Graduate Courses Taught Recently (Fall 2003 – Spring 2008)**

- Spring 2007 STA 2104 — Statistical Methods for Machine Learning and Data Mining  
I was solely responsible for teaching this course.
- Spring 2006 STA 2104 — Statistical Methods for Machine Learning and Data Mining  
I was solely responsible for teaching this course.
- Fall 2004 CSC 2541 — Bayesian Methods for Machine Learning  
This was a new course, for which I was solely responsible.
- Spring 2004 STA 2102 — Computational Techniques in Statistics  
I was solely responsible for teaching this course.

Fall 2003      STA 3000 — Advanced Theory of Statistics

I taught the first half of this full-year course, in cooperation with Nancy Reid, who taught the second half.

## Graduate Students and Postdoctoral Fellows

I supervised the following graduate students, from the time listed to completion:

Longhai Li, PhD, Statistics, 2007 (from 2003)

Thesis title: *Bayesian Classification and Regression with High-Dimensional Features*

Peter Liu, MSc, Computer Science, 2007 (co-supervisor, from 2005)

Thesis title: *Using Gaussian Process Regression to Denoise Images and Remove Artefacts from Microarray Data*

Babak Shahbaba, PhD, Public Health Sciences (Biostatistics), 2007 (from 2003)

Thesis title: *Improving Classification Models When a Class Hierarchy is Available*

Jennifer Listgarten, PhD, Computer Science, 2006 (co-supervisor, from 2003)

Thesis title: *Analysis of Sibling Time Series Data: Alignment and Difference Detection*

Shuying Sun, PhD, Statistics, 2006 (co-supervisor, from 2003)

Thesis title: *Haplotype Inference Using a Hidden Markov Model with Efficient Markov Chain Sampling*

Ruxandra Pinto, PhD, Statistics, 2002 (from 1997)

Thesis title: *Improving Markov Chain Monte Carlo Estimators Using Overrelaxation and Coupling Techniques*

Sonia Jain, PhD, Statistics, 2002 (from 1998)

Thesis title: *Split-Merge Techniques for Bayesian Mixture Models*

Krunoslav Kovac, MSc, Computer Science, 2005 (from 2002)

Thesis title: *Multitask Learning for Bayesian Neural Networks*

Kiam Choo, MSc, Computer Science, 2000 (from 1997)

Thesis title: *Learning Hyperparameters for Neural Network Models Using Hamiltonian Dynamics*

Michael Harvey, MSc, Computer Science, 1999 (part-time, from 1996)

Thesis title: *Monte Carlo Inference for Belief Networks Using Coupling From the Past*

I supervised the following postdoctoral fellow:

Matthew Beal, Computer Science, 1 July 2003 – 31 July 2004.

I am currently supervising the following graduate student:

Jianguo Zhang, PhD, Statistics (from 2003, currently on leave)

I am currently co-supervising the following postdoctoral fellow:

Iain Murray, Computer Science, from August 2007

I am currently on the advisory committees of the following PhD students:

Jim Huang, ECE  
Zi Jin, Statistics  
Andriy Mnih, Computer Science  
Ilya Sutskever, Computer Science  
Tijmen Tieleman, Computer Science  
Philip Tong, Biochemistry

I played substantial roles on the advisory committees of the following PhD students (completion date given):

Edward Meeds, Computer Science, 2008  
Vladimir Jojic, Computer Science, 2007  
Marcus Lee, Management, 2003  
Yee Whye Teh, Computer Science, 2003  
Greg Kondrak, Computer Science, 2002  
Andrew Brown, Computer Science, 2002  
Brian Sallans, Computer Science, 2001  
Sageev Oore, Computer Science, 2001  
Alison Gibbs, Statistics, 2000  
Xiang Sun, Statistics, 1999  
Brendan Frey, Electrical and Computer Engineering, 1997  
Carl Edward Rasmussen, Computer Science, 1996

I served on the SGS examining committees for the following PhD students (in addition to serving on the examining committees of my own students):

Vladimir Jojic, Computer Science, 2007  
Sigfrido Iglesias-Gonzalez, Statistics, 2007  
Marcus Lee, Management, 2003  
Yee Whye Teh, Computer Science, 2003  
Andrew Brown, Computer Science, 2002  
Brian Sallans, Computer Science, 2001  
Sageev Oore, Computer Science, 2001  
Alison Gibbs, Statistics, 2000  
Xiang Sun, Statistics, 1999  
George Tomlinson, Biostatistics, 1998  
Tomoyuki Yamakami, Computer Science, 1997  
Brendan Frey, Electrical and Computer Engineering, 1997  
Carl Edward Rasmussen, Computer Science, 1996  
Margeurite Ennis, Statistics, 1996  
Kin Chung Lo, Economics, 1995

I served as an external examiner for the following PhD student:

Xiaoqiang Li, Statistics, University of Western Ontario, 2002.

I served as a reader/examiner for the following MSc students:

Ilya Sutskever, Computer Science, 2007  
Brian Sallans, Computer Science, 1998  
Rafal Kustra, Biostatistics, 1997

## **Administrative Positions**

### **Major Administrative Position**

I was Associate Chair for Graduate Studies in the Dept. of Statistics from January through August 2007.

### **Departmental Committee Work**

Listed by academic year, commencing in July.

Computing Committee, Dept. of Statistics: Member in 1995, 1996; Chair in 1997, 1998, 1999, 2000 (1st half); Member in 2000 (2nd half), 2001, Chair in 2002, 2003, 2004; Member in 2005 (2nd half), Chair in 2006.

Graduate Committee, Dept. of Statistics: Member in 1995, 1996, 1998 (2nd half), 2002, 2003, 2005 (2nd half), Chair in 2006.

Hiring Committee, Dept. of Statistics: Member in 1998 (1st half), 1999, 2000 (1st half), 2001, 2002, 2003.

Seminar Coordinator, Dept. of Statistics: 1999.

## **Other Professional Activity**

### **Organizational Activities**

In 1998, I co-organized (with N. Madras and J. Rosenthal) a workshop on Monte Carlo methods held at the Fields Institute in October 1998, as part of the programme on Probability and Its Applications.

In 1995/1996, I organized a series of local Toronto meetings involving researchers on Markov chain Monte Carlo methods from Statistics, Computer Science, and Chemistry. In 1997/1998, I was co-organizer for a similar series of meetings (with J. Rosenthal and S. Whittington).

I organized and chaired an invited session on “Bayesian inference for high dimensional problems: Models and computation” for the Statistical Computing Section of the ASA, as part of the August 1996 Joint Statistical Meetings in Chicago.

### **Refereeing**

I have refereed papers for the following journals:

Annals of Applied Probability (1)

Annals of Statistics (2)  
Annals of the Institute of Statistical Mathematics (1)  
Artificial Intelligence (2)  
Bayesian Analysis (3)  
Biometrics (1)  
Biometrika (2)  
Canadian Journal of Statistics (1)  
Communications in Statistics (1)  
Computational Statistics and Data Analysis (1)  
Electronic Journal of Probability (1)  
IEEE Communications Letters (1)  
IEEE Transactions on Computers (1)  
IEEE Transactions on Communications (1)  
IEEE Transactions on Information Theory (2)  
IEEE Transactions on Knowledge and Data Engineering (1)  
IEEE Transactions on Neural Networks (6)  
IEEE Transactions on Pattern Analysis and Machine Intelligence (3)  
IEEE Transactions on Signal Processing (1)  
Journal of Artificial Intelligence Research (1)  
Journal of Computational and Graphical Statistics (2)  
Journal of Lightwave Technology (1)  
Journal of Machine Learning Research (1)  
Journal of the American Statistical Association (7)  
Journal of the Royal Statistical Society B (5)  
Journal of Physics A (1)  
Machine Learning (2)  
Neural Computation (22)  
Neural Networks (2)  
Proceedings of the National Academy of Sciences (1)  
Random Structures and Algorithms (1)  
Statistics and Computing (4)  
Statistics and Probability Letters (2)

I have refereed submissions to the following conferences:

Neural Information Processing Systems (NIPS), 1993, 1997, 1999, 2003, 2004, 2005.  
International Workshop on Artificial Intelligence and Statistics, 1997.

I have also refereed 16 grant proposals for NSERC (Canada), 1 grant proposal for NERC (United Kingdom), and 1 grant proposal for FWF (Austria).