
Yee Whye Teh

Curriculum Vitae

Gatsby Computational Neuroscience Unit Webpage: <http://www.gatsby.ucl.ac.uk/~ywteh>
University College London Email: ywteh@gatsby.ucl.ac.uk
17 Queen Square Phone: +44-207-679-1199
London WC1N 3AR, United Kingdom. Fax: +44-207-679-1193
Citizenship: Malaysian Mobile: +44-755-731-9260

Academic History

Mar 2011 - June 2011 Visitor.

Wellcome Trust Centre for Human Genetics, University of Oxford.
Host: Gilean McVean and Chris Holmes.
Research areas: statistical genetics.

Jan 2007 - Present Lecturer.

Gatsby Computational Neuroscience Unit, University College London.
Research areas: Bayesian methods for machine learning and applications.

Aug 2005 - Dec 2006 Lee Kuan Yew Postdoctoral Fellow.

Department of Computer Science, National University of Singapore.
Host: Wee Sun Lee.
Research areas: Bayesian methods for machine learning and applications.

Feb 2003 - Dec 2004 Postdoctoral Fellow.

Division of Computer Science, University of California at Berkeley.
Supervisors: Michael I. Jordan, David A. Forsyth.
Research areas: probabilistic models for machine learning and computer vision.

Jan 2000 - Jan 2003 Doctor of Philosophy.

Department of Computer Science, University of Toronto.
Supervisor: Geoffrey E. Hinton.
Thesis: Bethe Free Energy and Contrastive Divergence Approximations for Undirected Graphical Models.

Jul 1999 - Aug 2001 Affiliate research student.

Gatsby Computational Neuroscience Unit, University College London.
Supervisor: Geoffrey E. Hinton

Sep 1997 - Dec 2000 Master of Science.

Department of Computer Science, University of Toronto.
Supervisor: Geoffrey E. Hinton.
Thesis: Learning to Parse Images.

Sep 1994 - Aug 1997 Bachelor of Mathematics.
Faculty of Mathematics, University of Waterloo.
Double honours in Computer Science and Pure Mathematics.
Graduated on the Dean's Honours List.

Apr 1997 - Aug 1997 Research assistant.
Department of Computer Science, University of Waterloo.
Supervisor: Fahiem Bacchus.

Supervision

Ph.D. students

2010 - Present Dominik Beste (second supervisor)
2009 - Present Lloyd Elliott.
2008 - Present Jan Gasthaus.
2008 - Present Charles Blundell.
2008 - Present Dorota Glowacka (second supervisor)
2007 - Present Vinayak Rao.

Postdoctoral trainees and fellows

2010 - Present Andriy Mnih.
2007 - 2010 Dilan Görür.
2007 - 2009 Frank Wood.

Other students

2011 Sam Patterson (MSc)
2011 Marius Cobzarenco (MSc)
2010 Krikamol Muandet (MSc)
2008 Jan Gasthaus (MSc; co-supervised with Frank Wood).
2007 Daniel Roy (summer intern)
2006 Cai Junfu (BSc; co-supervised with Wee Sun Lee).

Ph.D. examiner/committee member

2011 Carl Scheffler, University of Cambridge.
2011 Daniel Roy, Massachusetts Institute of Technology.
2010 Shahid Iqbal, University College London.
2008 Hanna Wallach, University of Cambridge.
2007 David Stern, University of Cambridge.

First and second year PhD viva examiner

2010 Cristina Garcia Cifuentes, Computer Science, UCL

2008 Zhuoran Wang, Computer Science, UCL

Scholarships and Awards

2009 Best Paper Award (with Frank Wood).
International Conference on Artificial Intelligence and Statistics.

1998, 2002 Ontario Graduate Scholarship (CDN\$11859, CDN\$15000).
Ministry of Training, Colleges and Universities, Ontario, Canada.

1999 - 2002 University of Toronto Fellowship (Tuition and living expenses).
University of Toronto, Canada.

1997 University of Toronto Open Masters Fellowship (CDN\$7200).
University of Toronto, Canada.

Grants

2011 Composing Learning for Artificial Cognitive Systems.
European Union STREP. EUR4.5 million over 4 years. co-PI (along with 9 other co-PIs).

2010 Renewal of funding for the Gatsby Computational Neuroscience Unit.
Gatsby Charitable Foundation.
GBP6.9 million over 5 years. co-PI (along with 4 other co-PIs).

2005 Lee Kuan Yew Postdoctoral Fellowship.
Lee Kuan Yew Endowment Fund, Singapore.
Salary and SGD177990 grant over three years.

Professional Activities

Conference organizer

2010 International Conference on Artificial Intelligence and Statistics, Sardinia, Italy.
Program co-chair (with Mike Titterton).

Workshop organizer

2011 Gatsby Unit/CIFAR Workshop on Hierarchical Distributed Representations.
Windsor, United Kingdom.
Co-organizers: Geoffrey E. Hinton and Peter Dayan.

- 2009** Nonparametric Bayes.
NIPS, Whistler, Canada.
Co-organizers: Dilan Görür, Francois Caron, David Dunson, Zoubin Ghahramani, Michael I. Jordan.
- 2009** Grammar Induction, Representation of Language and Language Learning.
NIPS, Whistler, Canada.
Co-organizers: Alex Clark, Dorota Glowacka, John Shawe-Taylor, Chris Watkins.
- 2008** Nonparametric Bayes.
ICML/UAI/COLT, Helsinki, Finland.
Co-organizers: Romain Thibaux, Athanasios Kottas, Zoubin Ghahramani, Michael I. Jordan.
- 2008** Prior knowledge for text and language processing.
ICML/UAI/COLT, Helsinki, Finland.
Co-organizers: Marc Dymetman, Guillaume Bouchard, Hal Daume III.
- 2005** Bayesian methods for natural language processing.
NIPS, Whistler, Canada.
Co-organizer: Hal Daume III.
- 2005** Open problems and challenges for nonparametric Bayesian methods in machine learning.
NIPS, Whistler, Canada.
Co-organizer: Matthew J. Beal.
- 2003** Nonparametric Bayesian methods and infinite models.
NIPS, Whistler, Canada.
Co-organizer: Matthew J. Beal.

Conference advisory board

2009, 2011 Bayesian Nonparametrics Workshop.

Conference programme committee member

2009 International Conference on Artificial Intelligence and Statistics.

2009 International Joint Conference on Artificial Intelligence.

2007, 2009, 2011(declined) International Conference on Machine Learning.

2007, 2008 Neural Information Processing Systems.

Conference reviewer

Artificial Intelligence and Statistics
Computer Vision and Pattern Recognition
Empirical Methods on Natural Language Processing
International Conference on Machine Learning
International Joint Conference on Artificial Intelligence

Neural Information Processing Systems
Uncertainty in Artificial Intelligence

Associate editor

2009 - Present Bayesian Analysis

Editorial board member

2009 - Present Journal of Machine Learning Research **2010 - Present** Machine Learning Journal

Journal reviewer

Artificial Intelligence Journal
Annals of the Institute of Statistical Mathematics
Bayesian Analysis
IEEE Signal Processing Magazine
IEEE Transactions on Image Processing
IEEE Transactions on Knowledge and Data Engineering
IEEE Transactions on Neural Networks
IEEE Transactions on Pattern Analysis and Machine Intelligence
IEEE Transactions on Speech, Audio, and Language Processing
International Journal of Neural Systems
Journal of Computational and Graphical Statistics
Journal of Machine Learning Research
Journal of the American Statistical Association
Neural Computation
Statistics and Computing

Grant reviewer

Biotechnology and Biological Sciences Research Council, United Kingdom
Engineering and Physical Sciences Research Council, United Kingdom
Leverhulme Trust, United Kingdom
National Research Foundation, Singapore
National Science Foundation, USA
Netherlands Organisation for Scientific Research, The Netherlands

Teaching

Courses

2011 Bayesian nonparametrics.
PhD Course at Section for Cognitive Systems,
Denmark Technical University, Informatics, Denmark.

- 2010** Probabilistic and Bayesian Modelling.
Signal Theory and Communications Department, University Carlos III in Madrid.
Week-long graduate-level intensive course.
- 2008 - 2010** Machine Learning II (co-taught with Maneesh Sahani).
Gatsby Computational Neuroscience Unit, University College London.
Ph.D. level course.
Course on a number of advanced topics in machine learning.
- 2007 - 2010** Unsupervised and Probabilistic Learning (co-taught with Maneesh Sahani).
Gatsby Computational Neuroscience Unit, University College London.
M.Sc. and Ph.D. level course.
Introduction to probabilistic and graphical models for unsupervised learning.

Tutorials

- 2011** Tutorial on combinatorial stochastic processes and Bayesian nonparametrics.
Neural Information Processing Systems, Granada, Spain.
- 2011** Tutorial on Bayesian nonparametrics.
Machine Learning Summer Schools,
Singapore (July) and Bordeaux, France (September).
- 2008 - 2010** Clustering: tutorial and practical course.
Adaptive Modelling of Complex Data. University College London.
- 2010** Tutorial on Bayesian nonparametrics and applications.
Korean Advanced Institute of Science and Technology, South Korea.
- 2010** Tutorial on Bayesian nonparametrics and applications.
Workshop on Probabilistic Modeling for Computer Vision Applications.
Center for Mathematical Investigations, Mexico.
- 2009** Tutorial on Bayesian nonparametrics.
Machine Learning Summer School, University of Cambridge, UK.
Microsoft Research Cambridge, UK.
University of Toronto, Canada.
- 2008** Tutorial on hierarchical clustering.
EPSRC Winter School on Mathematics for Data Modelling, University of Sheffield, UK.
- 2007** Tutorial on Dirichlet processes and hierarchical Dirichlet processes.
Machine Learning Advanced Tutorial Series, University of Cambridge, UK.
- 2007** Bayesian nonparametric models: applications in machine learning and beyond.
Issac Newton Institute Workshop on Bayesian Nonparametric Regression, UK.
- 2007** Dirichlet processes: tutorial and practical course.
Machine Learning Summer School.
Max Planck Institute for Biological Cybernetics, Tübingen, Germany.

2006 Dirichlet processes.
Max Planck Institute for Biological Cybernetics Retreat, Tübingen, Germany.

Invited Conference and Workshop Talks

2012 Institute of Mathematical Sciences Programme on Meeting the Challenges of High Dimension—
Statistical Methodology, Theory and Applications, Singapore.

2012 Nordita Programme on Biology and Physics of Information Processing, Sweden.

2011 Conference on Natural Language Learning, Oregon, USA. (**keynote**)

2011 Bayesian Nonparametrics Workshop, Veracruz, Mexico.

2010 MCMC for Bayesian Nonparametric Models: A Bag of Tricks.
NIPS Workshop on Monte Carlo Methods, Canada.

2010 Hierarchical Bayesian Nonparametric Models for Language and Text (**keynote**).
ACM-SIGIR Workshop on Feature Generation and Selection for Information Retrieval, Switzerland.

2010 Hierarchical Bayesian Nonparametric Models for Language and Text.
International Society for Bayesian Analysis World Meeting, Valencia, Spain.

2010 Bayesian Rose Trees.
CRiSM Workshop on Model Uncertainty, Warwick, UK.

2010 Bayesian Rose Trees.
Workshop in Mixture Estimation and Applications, ICMS, Edinburgh, UK.

2009 Nonparametric Surrogate Priors.
Gatsby Deep Learning Workshop, London, UK.

2009 Bayesian Nonparametrics in Document and Language Modelling.
European Workshop on Challenges in Modern Massive Data Sets, Copenhagen, Denmark.

2009 The Mondrian Process: Bayesian Nonparametrics for Relational Data.
Bayesian Nonparametrics Workshop, Turin, Italy.

2008 Collapsed Variational Inference for Infinite State Bayesian Networks.
London Mathematical Society Durham Symposium on Mathematical Aspects of Graphical Models, UK.

2008 Bayesian Language Models.
CRiSM/INI Workshop on Bayesian Analysis of High Dimensional Data, Warwick, UK.

2007 Setting the Stage: Complementary Priors and Variational Bounds.
NIPS Deep Learning Symposium, Vancouver, Canada.

2007 Bayesian Agglomerative Clustering with Coalescents.
PASCAL Entente-Cordiale Workshop, Paris, France.

- 2007** A Hierarchical Bayesian Language Model based on Pitman-Yor Processes.
PASCAL Entente-Cordiale Workshop, London, UK.
- 2006** Some Applications of Hierarchical Dirichlet Processes (**keynote**).
Bayesian Nonparametrics Workshop, Jeju, Korea.
- 2006** Beam Sampling for the Infinite Hidden Markov Model (**keynote**).
ICML Workshop on Bayesian Nonparametric Models, Pittsburgh, USA.
- 2004** Hierarchical Dirichlet Processes.
Snowbird Learning Workshop, Utah, USA.
- 2002** Automatic Alignment of Local Representations.
NIPS Workshop on Spectral Methods, Vancouver, Canada.
- 2001** Passing and Bouncing Messages for Generalized Inference.
Gatsby Approximate Inference Workshop, London, UK.

Invited Seminars

- 2011** Computer Science and Information Systems, Birkbeck College, London, UK.
Decision Sciences, Bocconi University, Italy.
Wellcome Trust Centre for Human Genetics, University of Oxford, UK.
Engineering Science, University of Oxford, UK.
Computing, University of Oxford, UK.
CSAIL, Massachusetts Institute of Technology, USA.
Engineering, University of Cambridge, UK.
Institute of Statistical Mathematics, Japan.
Tokyo Institute of Technology, Japan.
NTT Research Labs, Japan.
- 2010** Defence Science Organization (DSO), Singapore.
Informatics, University of Edinburgh, UK.
Google Research and Computer Science, University of Waterloo, Canada.
Computer Science, University of Southern California, USA.
Salk Institute, University of Southern California, USA.
Information and Computer Science, University of California at Irvine, USA.
Computer Science, University of Glasgow, UK.
Statistics, University of Leeds, UK.
Statistics, UCL, UK.
Computer Science, University of Sheffield, UK.
- 2009** Hong Kong University of Science and Technology, Hong Kong.
Computer Science, University of Sheffield, UK.
Engineering, University of Cambridge, UK.
Computer Science, University of Birmingham, UK.

- 2008** Atomic Energy and Alternative Energies Commission, Saclay, France.
 Yahoo! Research Silicon Valley, USA.
 Google Research, USA.
 NEC Laboratories America, USA.
 Information and Computer Science, University of California at Irvine, USA.
 Computer Science, University of California at Berkeley, USA.
 Engineering, University of Cambridge, UK.
 Computer Science, Royal Holloway University of London, UK.
 Computer Science, University of Bristol, UK.
 School of Computing, National University of Singapore, Singapore.
 Nanyang Technological University, Singapore.
- 2007** Radboud University Nijmegen, The Netherlands.
 Computer Science, University of Toronto, Canada.
 Computer Science, University of Manchester, UK.
 Mathematics, Statistics and Actuarial Science, University of Kent, UK.
 Informatics, University of Edinburgh, UK.
- 2006** Information and Computer Science, University of California at Irvine, USA.
 Computer Science and Engineering, State University of New York at Buffalo, USA.
- 2005** Computer Science, University of Toronto, Canada.
- 2004** Microsoft Research Redmond, USA.
 Intel Research, USA.
 Computer Science, University of Pennsylvania, USA.
 Computer Science, University of Massachusetts Amherst, USA.
 CSAIL, Massachusetts Institute of Technology, USA.
 Computer Science, New York University, USA.
 Computer Science, Columbia University, USA.
 Computer Science, University of Toronto, Canada.
 School of Computing, National University of Singapore, Singapore.
- 2003** Honda Research Institute, USA. Computer Science, University of Toronto, Canada.
 Information and Computer Science, University of California at Irvine, USA.
- 2002** Redwood Neuroscience Institute, Menlo Park, USA.
 Computer Science, University of Toronto, Canada.
 Gatsby Computational Neuroscience Unit, UCL, UK.
 Computer Science, Carnegie Mellon University, USA.
- 2001** Computer Science, University of Waterloo, Canada.

Publications

Invited Journal Papers

1. F. Wood, J. Gasthaus, C. Archambeau, L. James, and Y. W. Teh. The sequence memoizer. *Communications of the Association for Computing Machines*, 54(2):91–98, 2011.

Refereed Journal Papers

1. D. Görür and Y. W. Teh. Concave-convex adaptive rejection sampling. *Journal of Computational and Graphical Statistics*, 2011. doi: 10.1198/jcgs.2011.09058.
2. Y. W. Teh, M. I. Jordan, M. J. Beal, and D. M. Blei. Hierarchical Dirichlet processes. *Journal of the American Statistical Association*, 101(476):1566–1581, 2006.
3. G. E. Hinton, S. Osindero, and Y. W. Teh. A fast learning algorithm for deep belief networks. *Neural Computation*, 18(7):1527–1554, 2006a.
4. G. E. Hinton, S. Osindero, M. Welling, and Y. W. Teh. Unsupervised discovery of non-linear structure using contrastive backpropagation. *Cognitive Science*, 30(4):725–731, 2006b.
5. M. Welling and Y. W. Teh. Linear response algorithms for approximate inference in graphical models. *Neural Computation*, 16:197–221, 2004.
6. Y. W. Teh, M. Welling, S. Osindero, and G. E. Hinton. Energy-based models for sparse overcomplete representations. *Journal of Machine Learning Research*, 4:1235–1260, Dec 2003.
7. M. Welling and Y. W. Teh. Approximate inference in Boltzmann machines. *Artificial Intelligence*, 143(1):19–50, 2003.

Review Articles

1. P. Orbanz and Y. W. Teh. Bayesian nonparametric models. In *Encyclopedia of Machine Learning*. Springer, 2010.
2. Y. W. Teh. Dirichlet processes. In *Encyclopedia of Machine Learning*. Springer, 2010.
3. Y. W. Teh and M. I. Jordan. Hierarchical Bayesian nonparametric models with applications. In N. Hjort, C. Holmes, P. Müller, and S. Walker, editors, *Bayesian Nonparametrics*. Cambridge University Press, 2010.

Book Chapters

1. C. Blundell, Y. W. Teh, and K. A. Heller. Bayesian rose trees. In C. P. Robert, K. Mengersen, and M. Titterton, editors, *Mixture Estimation and Applications*. John Wiley & Sons, 2011.

Invited Conference Papers

1. G. E. Hinton, M. Welling, Y. W. Teh, and S. Osindero. A new view of ICA. In *Proceedings of the International Conference on Independent Component Analysis and Blind Signal Separation*, volume 3, 2001.
2. G. E. Hinton and Y. W. Teh. Discovering multiple constraints that are frequently approximately satisfied. In *Proceedings of the Conference on Uncertainty in Artificial Intelligence*, volume 17, pages 227–234, 2001.

Peer-Reviewed Conference Papers

1. M. Welling and Y. W. Teh. Bayesian learning via stochastic gradient Langevin dynamics. In *Proceedings of the International Conference on Machine Learning*, 2011.
2. R. Silva, C. Blundell, and Y. W. Teh. Mixed cumulative distribution networks. In *Proceedings of the International Conference on Artificial Intelligence and Statistics*, 2011.
3. J. Gasthaus and Y. W. Teh. Improvements to the sequence memoizer. In *Advances in Neural Information Processing Systems*, 2010.
4. C. Blundell, Y. W. Teh, and K. A. Heller. Bayesian rose trees. In *Proceedings of the International Conference on Uncertainty in Artificial Intelligence*, 2010.
5. J. Gasthaus, F. Wood, and Y. W. Teh. Lossless compression based on the sequence memoizer. In *Data Compression Conference*, 2010.
6. Y. W. Teh and D. Görür. Indian buffet processes with power-law behavior. In *Advances in Neural Information Processing Systems*, volume 22, pages 1838–1846, 2009.
7. V. Rao and Y. W. Teh. Spatial normalized gamma processes. In *Advances in Neural Information Processing Systems*, volume 22, pages 1554–1562, 2009.
8. A. Asuncion, M. Welling, P. Smyth, and Y. W. Teh. On smoothing and inference for topic models. In *Proceedings of the International Conference on Uncertainty in Artificial Intelligence*, volume 25, 2009.
9. F. Wood, C. Archambeau, J. Gasthaus, L. F. James, and Y. W. Teh. A stochastic memoizer for sequence data. In *Proceedings of the International Conference on Machine Learning*, volume 26, pages 1129–1136, 2009.
10. F. Doshi, K. T. Miller, J. Van Gael, and Y. W. Teh. Variational inference for the Indian buffet process. In *JMLR Workshop and Conference Proceedings: AISTATS 2009*, volume 5, pages 137–144, 2009.
11. K. A. Heller, Y. W. Teh, and D. Görür. Infinite hierarchical hidden Markov models. In *JMLR Workshop and Conference Proceedings: AISTATS 2009*, volume 5, pages 224–231, 2009.
12. F. Wood and Y. W. Teh. A hierarchical nonparametric Bayesian approach to statistical language model domain adaptation. In *JMLR Workshop and Conference Proceedings: AISTATS 2009*, volume 5, pages 607–614, 2009.
13. G. R. Haffari and Y. W. Teh. Hierarchical Dirichlet trees for information retrieval. In *Proceedings of the Annual Meeting of the North American Association for Computational Linguistics and the Human Language Technology Conference*, 2009.
14. D. M. Roy and Y. W. Teh. The Mondrian process. In *Advances in Neural Information Processing Systems*, volume 21, pages 1377–1384, 2009.
15. D. Görür and Y. W. Teh. An efficient sequential Monte-Carlo algorithm for coalescent clustering. In *Advances in Neural Information Processing Systems*, volume 21, pages 521–528, 2009.

16. J. Van Gael, Y. W. Teh, and Z. Ghahramani. The infinite factorial hidden Markov model. In *Advances in Neural Information Processing Systems*, volume 21, pages 1697–1704, 2009.
17. J. Gasthaus, F. Wood, D. Görür, and Y. W. Teh. Dependent Dirichlet process spike sorting. In *Advances in Neural Information Processing Systems*, volume 21, pages 497–504, 2009.
18. G. Quon, Y. W. Teh, E. Chan, T. Hughes, M. Brudno, and Q. Morris. A mixture model for the evolution of gene expression in non-homogeneous datasets. In *Advances in Neural Information Processing Systems*, volume 21, pages 1297–1304, 2009.
19. M. Welling, Y. W. Teh, and H. J. Kappen. Hybrid Variational/Gibbs collapsed inference in topic models. In *Proceedings of the International Conference on Uncertainty in Artificial Intelligence*, volume 24, 2008.
20. J. Van Gael, Y. Saatici, Y. W. Teh, and Z. Ghahramani. Beam sampling for the infinite hidden Markov model. In *Proceedings of the International Conference on Machine Learning*, volume 25, pages 1088–1095, 2008.
21. Y. W. Teh, H. Daume III, and D. M. Roy. Bayesian agglomerative clustering with coalescents. In *Advances in Neural Information Processing Systems*, volume 20, pages 1473–1480, 2008a.
22. Y. W. Teh, K. Kurihara, and M. Welling. Collapsed variational inference for HDP. In *Advances in Neural Information Processing Systems*, volume 20, pages 1481–1488, 2008b.
23. H. L. Chieu, W. S. Lee, and Y. W. Teh. Cooled and relaxed survey propagation for MRFs. In *Advances in Neural Information Processing Systems*, volume 20, pages 297–304, 2008.
24. Y. J. Lim and Y. W. Teh. Variational Bayesian approach to movie rating prediction. In *Proceedings of KDD Cup and Workshop*, 2007.
25. J. F. Cai, W. S. Lee, and Y. W. Teh. NUS-ML: Improving word sense disambiguation using topic features. In *Proceedings of the International Workshop on Semantic Evaluations*, volume 4, 2007a.
26. J. F. Cai, W. S. Lee, and Y. W. Teh. Improving word sense disambiguation using topic features. In *Proceedings of the Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-coNLL)*, pages 1015–1023, 2007b. URL <http://www.aclweb.org/anthology/D/D07/D07-1108>.
27. Y. W. Teh, D. Görür, and Z. Ghahramani. Stick-breaking construction for the Indian buffet process. In *Proceedings of the International Conference on Artificial Intelligence and Statistics*, volume 11, 2007a.
28. K. Kurihara, M. Welling, and Y. W. Teh. Collapsed variational Dirichlet process mixture models. In *Proceedings of the International Joint Conference on Artificial Intelligence*, volume 20, 2007.
29. Y. W. Teh, D. Newman, and M. Welling. A collapsed variational Bayesian inference algorithm for latent Dirichlet allocation. In *Advances in Neural Information Processing Systems*, volume 19, pages 1353–1360, 2007b.

30. Y. W. Teh. A hierarchical Bayesian language model based on Pitman-Yor processes. In *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*, pages 985–992, 2006b. URL <http://www.aclweb.org/anthology/P/P06/P06-1124>.
31. E. P. Xing, K. Sohn, M. I. Jordan, and Y. W. Teh. Bayesian multi-population haplotype inference via a hierarchical Dirichlet process mixture. In *Proceedings of the International Conference on Machine Learning*, volume 23, pages 1049–1056, 2006.
32. M. Welling, T. Minka, and Y. W. Teh. Structured region graphs: Morphing EP into GBP. In *Proceedings of the International Conference on Uncertainty in Artificial Intelligence*, volume 21, 2005.
33. Y. W. Teh, M. Seeger, and M. I. Jordan. Semiparametric latent factor models. In *Proceedings of the International Workshop on Artificial Intelligence and Statistics*, volume 10, 2005b.
34. Y. W. Teh, M. I. Jordan, M. J. Beal, and D. M. Blei. Sharing clusters among related groups: Hierarchical Dirichlet processes. In *Advances in Neural Information Processing Systems*, volume 17, 2005a.
35. J. Edwards, Y. W. Teh, D. A. Forsyth, R. Bock, M. Maire, and G. Vesom. Making latin manuscripts searchable using gHMM's. In *Advances in Neural Information Processing Systems*, volume 17, 2005.
36. T. Miller, A. C. Berg, J. Edwards, M. Maire, R. White, Y. W. Teh, E. Learned-Miller, and D. A. Forsyth. Faces and names in the news. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, 2004.
37. M. Welling, M. Rosen-Zvi, and Y. W. Teh. Approximate inference by markov chains on union spaces. In *Proceedings of the International Conference on Machine Learning*, volume 21, 2004.
38. M. Welling and Y. W. Teh. Linear response algorithms for approximate inference in graphical models. *Neural Computation*, 16:197–221, 2004.
39. Y. W. Teh and M. Welling. On improving the efficiency of the iterative proportional fitting procedure. In *Proceedings of the International Workshop on Artificial Intelligence and Statistics*, volume 9, 2003.
40. Y. W. Teh and S. Roweis. Automatic alignment of local representations. In *Advances in Neural Information Processing Systems*, volume 15, 2003.
41. S. Kakade, Y. W. Teh, and S. Roweis. An alternate objective function for Markovian fields. In *Proceedings of the International Conference on Machine Learning*, volume 19, 2002.
42. Y. W. Teh and M. Welling. The unified propagation and scaling algorithm. In *Advances in Neural Information Processing Systems*, volume 14, 2002.
43. M. Welling and Y. W. Teh. Belief optimization for binary networks : A stable alternative to loopy belief propagation. In *Proceedings of the Conference on Uncertainty in Artificial Intelligence*, volume 17, 2001.

44. Y. W. Teh and G. E. Hinton. Rate-coded restricted Boltzmann machines for face recognition. In *Advances in Neural Information Processing Systems*, volume 13, 2001.
45. G. E. Hinton, Z. Ghahramani, and Y. W. Teh. Learning to parse images. In *Advances in Neural Information Processing Systems*, volume 12, 2000.
46. F. Bacchus and Y. W. Teh. Making forward chaining relevant. In *Proceedings of the International Conference on Artificial Intelligence Planning Systems*, 1998.

Conference Abstracts

1. Y. W. Teh, F. Wood, J. Gasthaus, C. Archambeau, and L. F. James. Hierarchical Bayesian nonparametric models for language and text. In *International Society for Bayesian Analysis World Meeting*, 2010.
2. Y. W. Teh, M. I. Jordan, M. J. Beal, and D. M. Blei. Hierarchical Dirichlet processes. In *Snowbird Learning Workshop*, 2004b.
3. M. J. Beal, Y. W. Teh, and M. I. Jordan. Infinite hidden Markov models via the hierarchical Dirichlet process. In *Snowbird Learning Workshop*, 2004.

Technical Reports

1. W. S. Lee, X. Zhang, and Y. W. Teh. Semi-supervised learning in reproducing kernel Hilbert spaces using local invariances. Technical Report TRB3/06, School of Computing, National University of Singapore, 2006.
2. Y. W. Teh. A Bayesian interpretation of interpolated Kneser-Ney. Technical Report TRA2/06, School of Computing, National University of Singapore, 2006a.
3. M. Seeger, Y. W. Teh, and M. I. Jordan. Semiparametric latent factor models. Technical report, Division of Computer Science, University of California at Berkeley, 2005.
4. Y. W. Teh, M. I. Jordan, M. J. Beal, and D. M. Blei. Hierarchical Dirichlet processes. Technical Report 653, Department of Statistics, University of California at Berkeley, 2004a.
5. Y. W. Teh and M. Welling. Passing and bouncing messages for generalized inference. Technical Report GCNU TR 2001-01, Gatsby Computational Neuroscience Unit, University College London, 2001.

In Press and Under Review

1. T. L. Berg, A. C. Berg, J. Edwards, M. Maire, R. White, Y. W. Teh, E. Learned-Miller, and D. A. Forsyth. Names and faces. *International Journal of Computer Vision*, under review. *International Journal of Computer Vision*.