

Katherine Ann Heller

Brain and Cognitive Science Department
Massachusetts Institute of Technology
Cambridge, MA 02139
kheller@mit.edu
<http://www.gatsby.ucl.ac.uk/~heller>
US Citizen

ACADEMIC BACKGROUND

NSF Postdoctoral Fellow
Massachusetts Institute of Technology
Jan. 2011-
Cambridge, MA USA

EPSRC Postdoctoral Fellow
Engineering Dept., University of Cambridge
Feb. 2008-Feb. 2011
Cambridge, UK

Ph.D., Machine Learning, 2008
“Efficient Bayesian Methods for Clustering”
Advisor: Prof. Zoubin Ghahramani
Gatsby Computational Neuroscience Unit, UCL, London, UK

M.S., Computer Science, 2003
Columbia University, NY, NY

B.S., Computer Science, 2000
B.S., Applied Mathematics and Statistics
State University of New York at Stony Brook, NY

AWARDS

NSF Postdoctoral Fellowship (\$240,000)
EPSRC Postdoctoral Fellowship in Theoretical Computer Science (£233,242)
Outstanding Paper Award, Rank Prize Symposium on Still and Moving Images (£250),
Rank Prize Fund
Graduate School Research Scholarship (£45,000), University College London
Overseas Research Scholarship (£30,000), Universities UK
NSF Graduate Research Fellowship (\$100,000)
Teaching Assistant Award(\$500), Numerical Algorithms and Complexity, Columbia
University
Paine Webber Technology and Engineering Careers Scholarship (\$5,000)
Reuters Information Technology, Inc. Scholarship (\$2,000)
NSF Scholar (\$2,000), Women in Science and Engineering (WISE) Program, SUNY
Stony Brook

REFEREED PUBLICATIONS

1. E. Airoldi, K.A. Heller, and R. Silva. Small sets of interacting proteins suggest functional linkage mechanisms via Bayesian analogical reasoning. *Bioinformatics (proceedings of the ISMB)*, 2011, to appear.

2. C. Blundell, Y.W. Teh, and K.A. Heller. Bayesian rose trees. In *Uncertainty in Artificial Intelligence (UAI)*, 2010.
3. S. Williamson, C. Wang, K.A. Heller, and D. Blei. The IBP compound Dirichlet process and its application to focused topic modeling. In *International Conference on Machine Learning (ICML)*, 2010.
4. H. Wallach, S. Jensen, L. Dicker, and K.A. Heller. An alternative prior process for nonparametric Bayesian clustering. In *Conference on AI and Statistics (AISTATS)*, 2010.
5. R. Silva, K.A. Heller, Z. Ghahramani, and E. Airoldi. Ranking relations using analogies. *Annals of Applied Statistics (AOAS)*, 4(2):615–644, 2010.
6. R.S. Savage, K.A. Heller, Y. Xu, Z. Ghahramani, W.M. Truman, M. Grant, K.J. Denby, and D.L. Wild. R/BHC:fast Bayesian hierarchical clustering for microarray data. *BMC Bioinformatics*, 10(242), 2009. **Highly Accessed.**
7. K.A. Heller, A. Sanborn, and N. Chater. Hierarchical learning of dimensional biases in human categorization. In *Neural Information Processing Systems (NIPS)*, 2009. **Spotlight Presentation.**
8. K.A. Heller, Y.W. Teh, and D. Gorur. Infinite hierarchical Hidden Markov Models. In *Conference on AI and Statistics (AISTATS)*, 2009.
9. Y. Xu, K.A. Heller, and Z. Ghahramani. Tree-based inference for Dirichlet process mixtures. In *Conference on AI and Statistics (AISTATS)*, 2009.
10. S. Mohamed, K.A. Heller, and Z. Ghahramani. Bayesian exponential family PCA. In *Neural Information Processing Systems (NIPS)*, 2008. **Spotlight Presentation.**
11. K.A. Heller, S. Williamson, and Z. Ghahramani. Statistical models for partial membership. In *International Conference on Machine Learning (ICML)*, 2008.
12. K.A. Heller and Z. Ghahramani. A nonparametric Bayesian approach to modeling overlapping clusters. In *Conference on Artificial Intelligence and Statistics (AISTATS)*, 2007.
13. R. Silva, K.A. Heller, and Z. Ghahramani. Analogical reasoning with relational Bayesian sets. In *Conference on Artificial Intelligence and Statistics (AISTATS)*, 2007.
14. K.A. Heller and Z. Ghahramani. A simple Bayesian framework for content-based image retrieval. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2006. **Outstanding Paper Award, Rank Prize Symposium.**
15. Z. Ghahramani and K.A. Heller. Bayesian sets. In *Advances in Neural Information Processing Systems (NIPS)*, 2005. **Oral Presentation.**
16. K.A. Heller and Z. Ghahramani. Randomized algorithms for faast Bayesian hierarchical clustering. In *PASCAL Workshop on Statistics and Optimization of Clustering*, 2005.
17. K.A. Heller and Z. Ghahramani. Bayesian hierarchical clustering. In *International Conference on Machine Learning (ICML)*, 2005.
18. S.J. Stolfo, F. Apap, E. Eskin, K.A. Heller, S. Hershkop, A. Honig, and K.M. Svore. A comparative evaluation of two algorithms for windows registry anomaly detection. *Journal of Computer Security*, 13(4):659–693, 2005.
19. K.A. Heller, K.M. Svore, A. Keromytis, and S.J. Stolfo. One class support vector machines for detecting anomalous windows registry accesses. In *International Conference on Data Mining (ICDM) Workshop on Data Mining for Computer Security*, 2003.

20. X.I. Zhang, K.A. Heller, I. Hefter, C.S. Leslie, and L.A. Chasin. Sequence information for the splicing of human pre-mRNA identified by support vector machine classification. *Genome Research*, 13:2637–2650, 2003.

BOOK CHAPTERS

21. S. Williamson, C. Wang, K.A. Heller, and D. Blei. Nonparametric mixed membership modelling using the ibp compound dirichlet process. In K. Mengerson, C. Robert, and D.M. Titterton, editors, *Mixture Estimation and Applications*, *In press*. J. Wiley, 2011.
22. C. Blundell, Y.W. Teh, and K.A. Heller. Discovering non-binary hierarchical structures with bayesian rose trees. In K. Mengerson, C. Robert, and D.M. Titterton, editors, *Mixture Estimation and Applications*, *In press*. J. Wiley, 2011.

WORKING PAPERS

23. S. Mohamed, K.A. Heller, and Z. Ghahramani. Generalised models for sparse latent factor discovery. In *submission*, 2011.

TECHNICAL REPORTS

1. K.A. Heller. Generalization using cross-category information. Technical report, Gatsby Unit, UCL, 2006.
2. K.A. Heller, K.M. Svore, and M. Kamvar. Q-HSK: A quantum simulation language. Technical report, Columbia University, 2002.

ABSTRACTS

1. K.A. Heller and J.M. Beck. Biologically plausible spike pattern source discovery using topic models. In *Computational and Systems Neuroscience (COSYNE)*, 2011.
2. K.A. Heller, Z. Ghahramani, B.J. de la Cruz, W.M. Truman, M. Grant, R.S. Savage, J.D. Stephenson, and D. Wild. Fast Bayesian clustering for microarray data. In *Mathematical and Statistical Aspects of Molecular Biology (MASAMB)*, 2008.
3. K. Borgwardt, K.A. Heller, Z. Ghahramani, H. Lightfoot, and D. Wild. Protein fold recognition using Bayesian information retrieval. In *Mathematical and Statistical Aspects of Molecular Biology (MASAMB)*, 2008.
4. K.A. Heller, Z. Ghahramani, and D. Wild. Efficient Bayesian hierarchical clustering for gene expression data. In *Valencia 8*, 2006.

PATENTS

Z. Ghahramani and K.A. Heller, W02007063328, *Information Retrieval*, 2007

TEACHING EXPERIENCE

Teaching Assistant **Gatsby Unit, UCL**
London, UK **2004**
Course title: “Unsupervised Learning”

Teaching Assistant **Columbia University**
NY, NY **2002**
Course title: “Computational Genomics”

Teaching Assistant Columbia University
NY, NY 2000
Course title: "Numerical Algorithms and Complexity"

Summer School Mathematics Teacher Long Island City High School
Queens, NY 2000
Sequential Mathematics 1 and 2, Math A

Educational Opportunity Program Tutor SUNY Stony Brook
Stony Brook, NY 1999
Course title: "Computer Science 2"

Teaching Assistant SUNY Stony Brook
Stony Brook, NY 1997
Course title: "Foundations of Computer Science 1"

ACADEMIC
SUPERVISION

Shakir Mohamed, PhD Student, University of Cambridge
Sinead Williamson, PhD Student, University of Cambridge
Yang Xu, 2007-2008, 4th Year Student, University of Cambridge (now a PhD student
at Carnegie Mellon University)

SEMINARS AND
TALKS

Intelligent Systems for Molecular Biology (ISMB), Vienna, Austria, July, 2011
Massachusetts Institute of Technology, Boston, MA, March, 2011
California Institute of Technology, Pasadena, CA, March, 2011
Columbia University, NY, NY, February, 2011
University of Chicago, Chicago, IL, February, 2011
Duke University, Durham, NC, February, 2011
Harvard University, Boston, MA, January, 2011
Rutgers, New Brunswick, NJ, January, 2011
University of Rochester, Rochester, NY, October, 2010
University of Pennsylvania, Philadelphia, PA, October, 2010
California Institute of Technology, Pasadena, CA, July, 2010
University of California, Irvine, CA, July, 2010
University of Edinburgh, Edinburgh, UK, April 2010
Workshop in Mixture Estimation and Applications, Edinburgh, UK, March 2010
New York University, NY, NY, February 2010
Columbia University, NY, NY, April 2009
Princeton University, Princeton, NJ, April 2009
University of Massachusetts, Amherst, MA, March 2009
University of California, Berkeley, CA, February 2009
University of Toronto, Toronto, Canada, November 2008
International Conference on Machine Learning (ICML), Helsinki, Finland, July 2008
Newton Institute, University of Cambridge, Cambridge, UK, February 2008.
University of Birmingham, Birmingham, UK, September 2007.

Rank Prize Fund Symposium: Interacting with Still and Moving Images, Windermere, UK, July 2007

Research Kitchen in Approximate Inference, Bath, UK, May 2007.

Conference on AI and Statistics (AISTATS), San Juan, Puerto Rico, March 2007.

Workshop for Women in Machine Learning (WiML), San Diego, USA, October 2006.

Microsoft Research, Cambridge, UK, May 2006.

Carnegie Mellon University, Pittsburgh, USA, April 2006.

Neural Information Processing Systems (NIPS), Vancouver, Canada, December 2005.

University of Sheffield, Sheffield, UK, October 2005.

International Conference on Machine Learning (ICML), Bonn, Germany, August 2005.

PASCAL Statistics and Optimization of Clustering Workshop, Windsor, UK, July 2005.

Carnegie Mellon University, Pittsburgh, USA, August, 2004.

University College London, UK, (many talks) 2003-.

REVIEWING

Grants

National Science Foundation (NSF)

Journals

Editorial Boards: Journal of Machine Learning Research (JMLR)

Reviewing: IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), Neural Networks, Journal of Machine Learning Research (JMLR), Neural Computation, Machine Learning Journal, Annals of Applied Statistics (AOAS), Journal of Computational and Graphical Statistics (JCGS), Journal of the American Statistical Association (JASA)

Conferences

Area Chair/SPC: Neural Information Processing Systems (NIPS) 2011, International Joint Conference on AI (IJCAI) 2011, Neural Information Processing Systems (NIPS) 2010, International Conference on Machine Learning (ICML) 2010

Reviewing/PC: International Conference on Machine Learning (ICML), Artificial Intelligence and Statistics (AISTATS), Neural Information Processing Systems (NIPS), Uncertainty in Artificial Intelligence (UAI), Conference on Computational Learning Theory (COLT)

ORGANIZATION AND SERVICE

Workshops Co-chair, ICML 2011, Bellevue, WA

Invited Panelist, Workshop for Women in Machine Learning (WiML), Vancouver, 2011

Organizer, NIPS Workshop on Applications of Topic Models: Text and Beyond, Vancouver, 2010

Student Organizer, AI and Statistics Conference, Barbados, 2005