Katherine Ann Heller

Duke University
223B Old Chemistry
Durham, NC 27708
khelle r@stat.duke.edu
http://www.stat.duke.edu/∼kheller
US Citizen

Academic Background

Assistant Professor
Statistical Science, Center for Cognitive Neuroscience
Computer Science, Neurobiology
Duke University
Jan. 2012 -
Durham, NC USA

NSF Postdoctoral Fellow
Massachussetts Institute of Technology
Jan. 2011- 2012
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

Best Paper Award, Social Computing, Behavioral Modeling and Prediction (SBP)

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. J. Jiang, K.A. Heller, and T. Egner. Bayesian modeling of flexible cognitive control. *In press, Neuroscience and Biobehavioral Reviews*, 2014.
2. B. Letham, C. Rudin, and K.A. Heller. Growing a list. *Data Mining and Knowledge Discovery (ECML/PKDD)*, 2013. **ECML/PKDD Journal Track**.
3. C. Blundell, K.A. Heller, and J. Beck. Modelling reciprocating relationships with Hawkes processes. In *Neural Information Processing Systems (NIPS)*, 2012. **Spotlight Presentation**.
4. J. Beck, K.A. Heller, and A. Pouget. Complex inference in neural circuits with probabilistic population codes and topic models. In *Neural Information Processing Systems (NIPS)*, 2012. **Spotlight Presentation**.
5. W. Dong, K.A. Heller, and A. Pentland. Graph-coupled hmms for modeling the spread of infection. In *Uncertainty in Artificial Intelligence (UAI)*, 2012.
6. S. Mohamed, K.A. Heller, and Z. Ghahramani. Evaluating bayesian and l1 approaches for sparse unsupervised learning. In *International Conference on Machine Learning (ICML)*, 2012.
7. W. Dong, K.A. Heller, and A. Pentland. Modeling infection with multi-agent dynamics. In *Social Computing, Behavioral Modeling and Prediction (SBP)*, 2012. **Best Paper Award**.
8. J. T. Abbott, K. A. Heller, Z. Ghahramani, and T. L. Griffiths. Testing a Bayesian measure of representativeness using a large image database. In *Neural Information Processing Systems (NIPS)*, 2011.
9. 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.
10. C. Blundell, Y.W. Teh, and K.A. Heller. Bayesian rose trees. In *Uncertainty in Artificial Intelligence (UAI)*, 2010.
11. 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.
12. 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.
13. R. Silva, K.A. Heller, Z. Ghahramani, and E. Airoldi. Ranking relations using analogies. *Annals of Applied Statistics (AOAS)*, 4(2):615–644, 2010.
14. 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**.
15. 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**.
16. K.A. Heller, Y.W. Teh, and D. Gorur. Infinite hierarchical Hidden Markov Models. In *Conference on AI and Statistics (AISTATS)*, 2009.
17. Y. Xu, K.A. Heller, and Z. Ghahramani. Tree-based inference for Dirichlet process mixtures. In *Conference on AI and Statistics (AISTATS)*, 2009.

18. S. Mohamed, K.A. Heller, and Z. Ghahramani. Bayesian exponential family PCA. In *Neural Information Processing Systems (NIPS)*, 2008. **Spotlight Presentation**.
19. K.A. Heller, S. Williamson, and Z. Ghahramani. Statistical models for partial membership. In *International Conference on Machine Learning (ICML)*, 2008.
20. K.A. Heller and Z. Ghahramani. A nonparametric Bayesian approach to modeling overlapping clusters. In *Conference on Artificial Intelligence and Statistics (AISTATS)*, 2007.
21. R. Silva, K.A. Heller, and Z. Ghahramani. Analogical reasoning with relational Bayesian sets. In *Conference on Artificial Intelligence and Statistics (AISTATS)*, 2007.
22. 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**.
23. Z. Ghahramani and K.A. Heller. Bayesian sets. In *Advances in Neural Information Processing Systems (NIPS)*, 2005. **Oral Presentation**.
24. K.A. Heller and Z. Ghahramani. Randomized algorithms for faast Bayesian hierarchical clustering. In *PASCAL Workshop on Statistics and Optimization of Clustering*, 2005.
25. K.A. Heller and Z. Ghahramani. Bayesian hierarchical clustering. In *International Conference on Machine Learning (ICML)*, 2005.
26. 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.
27. 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.
28. 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

29. S. Mohamed, Z. Ghahramani, and K.A. Heller. A simple and general exponential family framework for partial membership and factor analysis. In E. Airoldi, D.M. Blei, E. Erosheva, and S.E. Fienberg, editors, *Handbook of Mixed Membership Models, In press*. Chapman and Hall, 2013.
30. S. Mohamed, K.A. Heller, and Z. Ghahramani. Bayesian approaches for sparse latent variable models: Reconsidering l1 sparsity. In I. Rish, A. Lozano, G. Cecchi, and A. Niculescu-Mizil, editors, *Practical Applications of Sparse Modeling: Biology, Signal Processing and Beyond, In press*. MIT Press, 2013.
31. 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. Titterington, editors, *Mixture Estimation and Applications*. J. Wiley, 2011.
32. 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. Titterington, editors, *Mixture Estimation and Applications*. J. Wiley, 2011.

Working Papers

33. R. Guo, C. Blundell, H. Wallach, and K.A. Heller. Bayesian echo chamber. In *submission*, 2014.

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

Grants

"BRAIN EAGER: Bayesian Models of Translational Neural Networks: Motivation and Reward" , National Science Foundation, August 2014 - 2016, $300,000.

"Workshop for Women in Machine Learning", National Science Foundation, August 2013 - 2015, $40,000.

"Bayesian Models of Social Behavior Using Online Resources", National Science Foundation, June 2013 - 2014, $78,500.00.

Bi-annual Interdepartmental Mini-Grant: "Machine Learning Seminar Series", Duke University, February 2013 - June 2014, $5,000.

Teaching Experience

Instructor — **Duke University**
Durham, NC — **2014**
Course title: "Advanced Machine Learning"

Instructor — **Duke University**
Durham, NC — **2013**
Course title: "Probability"

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"

Seminars and Talks

Johns Hopkins University, Baltimore, MD, November 2013

Information Theory and Applications (ITA), San Diego, CA, February 2012

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 Applicatons, 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 Board: 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: International Conference on Machine Learning (ICML) 2014, Artificial Intelligence and Statistics (AISTATS) 2013, Neural Information Processing Systems (NIPS) 2012, Uncertainty in Artificial Intelligence (UAI) 2012,Artificial Intelligence and Statistics (AISTATS) 2012, 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

Local Organizing Comittee, Conference on Bayesian Nonparametrics 2015

Organizer, Machine Learning Seminar Series, Duke University, 2013-

Board of Directors, Women in Machine Learning (WiML), 2012-

Panelist, National Science Foundation, CITraCS Postdoctoral Fellowships, 2012

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