Frank Wood

21 Claremont Ave. Apt. 93
New York, NY 10027, USA
Phone: +16467172202
fwood@stat.columbia.edu
http://www.stat.columbia.edu/∼fwood

Department of Statistics
Columbia University
1255 Amsterdam Ave.
New York, NY 10027, USA

Research Overview

My primary research area is applied statistical machine learning. My research interests range from the development of new probabilistic models and inference algorithms to real-world applications. My research contributions include new models and inference algorithms as well as novel applications of such models to problems in compression, neuroscience, robotics, and natural language processing.

Education

Ph.D., Computer Science, 2007
"Nonparametric Bayesian Models for Neural Data"
Advisor: Prof. Michael J. Black
Brown University, Providence, RI

M.S., Computer Science, 2004
Advisor: Prof. Michael J. Black
Brown University, Providence, RI

B.S., Computer Science, 1996
Cornell University, Ithaca, NY

Positions

Assistant Professor — **Columbia University**
New York, NY — **2009 – present**

Founder — **Betacular, Ltd.**
Zurich, CH — **2010 – present**

Postdoctoral Fellow — **Gatsby Unit, UCL**
London, UK — **2007 – 2009**

Consultant — **Stan James Limited**
Gibraltar — **2008 – 2009**
Stan James is a sports bookmaker. Initiated the development of mathematically consistent market pricing tools and introduced statistical market prediction methods.

Chief Executive Officer — **Interfolio, Inc.**
Washington, DC — **2002**
Interfolio is an online academic credential file management provider. Tripled revenue, doubled partner schools and user base, acquired ReferenceNow, LLC (a competitor), and negotiated contracts for debt reduction and strategic partnerships.

Principal Engineer — **AOL Time Warner**
Dulles, VA — **2000 – 2001**
Designed (then) the world's largest production image and mp3 search engines. Filed 5 patents for image, mp3, and text search innovations. Managed multimedia search engineering team.

Chief Executive Officer / Founder **ToFish!, Inc.**
Washington, DC **1998 – 2000**
ToFish! was a content-based image search technology company. Negotiated sale of company to AOL for a 400%+ return on investment. Managed marketing, sales, legal, finance, and technical teams. Arranged "friends and family" and early state venture financing. Negotiated contracts for sales, intellectual property, and employment.

Research Engineer **Lawrence Berkeley National Laboratory**
Berkeley, CA and Washington, DC **1997 – 1998**
Contributed to successful grant proposals for over $500,000. Communicated laboratory research to congressional staff. Designed and implemented virtual reality simulations.

Research Engineer **Cornell Theory Center**
Ithaca, NY **1996 – 1997**
Designed and implemented the first virtual reality "window manager." Implemented a military in-flight re-fueling simulator. Investigated novel super-computing computational steering algorithms.

Teaching Experience

Instructor **Columbia University**
New York, NY **Fall 2010 -**
Course title: "Data Mining/Statistical Machine Learning."

Instructor **Columbia University**
New York, NY **Fall 2009 -**
Course title: "Linear Regression Models."

Guest Lecturer **Princeton University**
Princeton, NJ **April 2009**
Course title: "Foundations of Probabilistic Modeling."

Guest Lecturer **Columbia University**
New York, NY **March 2009**
Course title: "Statistical analysis of neural data."

Guest Lecturer **University College London**
London, UK **January 2008**
Course title: "Advanced Topics in Machine Learning."

Teaching Assistant **Brown University**
Providence, RI **January 2005 – June 2005**
Course title: "Topics in brain computer interfaces."

Advising

J. Huggins, BS, Computer Science, Columbia University,
W. Neiswanger, BS, Applied Math, Columbia University,
N.Bartlett, MSc, Statistics, Columbia University,
J. Gasthaus, MSc, University College London, *Spike Sorting using Time-Varying Dirichlet Process Mixture Models, 2009*

Patents

M. J. Black, W. Wu, and F. Wood, 11,086,956, *Method and system for automatic decoding of motor cortical activity*, 2005
G. Pass and F. Wood, 6,671,402, *Representing an image with weighted joint histogram*, 2003

G. Pass and F. Wood, 6,522,782, *Image and text searching techniques*, 2003
G. Pass and F. Wood, 6,556,710, *Image searching techniques*, 2003
G. Pass and F. Wood, 6,622,780, *Indexing of images and/or text*, 2003
G. Pass and F. Wood, 6,522,779, *Representing an image with a posterized joint histogram*, 2003

Awards and Honors

IMSA Alumni Distinguished Leadership Award 2011
AISTATS Best Paper Award 2009
National Science Foundation Research Experience for Undergraduates, Cornell Theory Center, 1994
Honors College Scholar, University of Illinois at Chicago, 1992

Service

NIPS Area Chair, 2011
IJCAI Senior Program Committee, 2010
AISTATS Senior Program Committee, 2010
Columbia University Statistics Department Computing Committee, 2009
Gatsby external talks coordinator, 2008-2009

Support

Google research award $70,000, 2010
Xerox faculty research award $90,000, 2011

Frank Wood

Journal Publications

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

F. Wood and M. J. Black. A non-parametric Bayesian alternative to spike sorting. *Journal of Neuroscience Methods*, 173:1–12, 2008.

D. H. Grollman, O. C. Jenkins, and F. Wood. Discovering natural kinds of robot sensory experiences in unstructured environments. *Journal of Field Robotics*, 23:1077–1089, 2006.

F. Wood, M. Fellows, C. Vargas-Irwin, M. J. Black, and J. P. Donoghue. On the variability of manual spike sorting. *IEEE Transactions in Biomedical Engineering*, 51:912–918, 2004.

F. Wood, D. Brown, B. Amidon, J. Alferness, B. Joseph, R. E. Gillilan, and C. Faerman. Windowing and telecollaboration for virtual reality with applications to the study of a tropical disease. *IEEE Computer Graphics and Applications*, 16:72–78, 1996.

R. E. Gillilan and F. Wood. Visualization, virtual reality, and animation within the data flow model of computing. *Computer Graphics*, 29:55–58, 1995.

Refereed Conference Proceedings

A. Perotte, N. Bartlett Noıemie Elhadad, and F. Wood. Hierarchically supervised latent Dirichlet allocation. In *Advances in Neural Information Processing Systems*, to appear, 2012.

N. Bartlett and F. Wood. Deplump for streaming data. In *Data Compression Conference*, pages 363–372, 2011.

D. Pfau, N. Bartlett, and F. Wood. Probabilistic deterministic infinite automata. In *Advances in Neural Information Processing Systems*, pages 1930–1938, 2011.

N. Bartlett, D. Pfau, and F. Wood. Forgetting counts : Constant memory inference for a dependent hierarchical Pitman-Yor process. In *Proceedings of the 26th International Conference on Machine Learning*, pages 63–70, 2010.

J. Gasthaus, F. Wood, and Y.W. Teh. Lossless compression based on the Sequence Memoizer. In *Data Compression Conference*, pages 337–345, 2010.

F. Wood, C. Archambeau, J. Gasthaus, L. James, and Y.W. Teh. A stochastic memoizer for sequence data. In *Proceedings of the 26th International Conference on Machine Learning*, pages 1129–1136, 2009.

F. Wood and Y.W. Teh. A hierarchical nonparametric Bayesian approach to statistical language model domain adaptation. In *Artificial Intelligence and Statistics*, pages 607–614, 2009.

J. Gasthaus, F. Wood, D. Görür, and Y.W. Teh. Dependent Dirichlet process spike sorting. In *Advances in Neural Information Processing Systems*, pages 497–504, 2009.

P. Berkes, J.W. Pillow, and F. Wood. Characterizing neural dependencies with Poisson copula models. In *Advances in Neural Information Processing Systems*, pages 129 – 136, 2009.

F. Wood and T. L. Griffiths. Particle filtering for non-parametric Bayesian matrix factorization. In *Advances in Neural Information Processing Systems*, pages 1513–1520, 2006.

F. Wood, S. Goldwater, and M. J. Black. A non-parametric Bayesian approach to spike sorting. In *Proceedings of the 28th IEEE Conference on Engineering in Medicine and Biological Systems*, pages 1165–1169, 2006.

F. Wood, T. L. Griffiths, and Z. Ghahramani. A non-parametric Bayesian method for inferring hidden causes. In *Proceedings of the 22nd Conference on Uncertainty in Artificial Intelligence*, pages 536–543, 2006.

S. P. Kim, F. Wood, and M. J. Black. Statistical analysis of the non-stationarity of neural population codes. In *The First IEEE / RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics*, pages 259–299, 2006.

F. Wood, S. Roth, and M. J. Black. Modeling neural population spiking activity with Gibbs distributions. In *Advances in Neural Information Processing Systems*, pages 1537–1544, 2005.

F. Wood, Prabhat, J. P. Donoghue, and M. J. Black. Inferring attentional state and kinematics from motor cortical firing rates. In *Proceedings of the 27th IEEE Conference on Engineering in Medicine and Biological Systems*, pages 149–152, 2005.

F. Wood, M. Fellows, J. P. Donoghue, and M. J. Black. Automatic spike sorting for neural decoding. In *Proceedings of the 27th IEEE Conference on Engineering in Medicine and Biological Systems*, pages 4126–4129, 2004.

Technical Reports

F. Wood, D.H. Grollman, K.A. Heller, O.C. Jenkins, and M.J. Black. Incremental nonparametric Bayesian regression. Technical Report CS-08-07, Brown University, Department of Computer Science, 2008.

Workshop Publications

F. Wood and Y.W. Teh. A hierarchical, hierarchical Pitman Yor process language model. In *ICML/UAI Nonparametric Bayes Workshop*, 2008.

D. H. Grollman, O. C. Jenkins, and F. Wood. Discovering natural kinds of robot sensory experiences in unstructured environments. In *Advances in Neural Information Processing Systems Workshop on Machine Learning Based Robotics in Unstructured Environments*, 2005.

Abstracts

F. Wood and M. J. Black, Energy Based Models of Motor Cortical Population Activity, *Society for Neuroscience* , Washington, DC 2005

F. Wood, M. Fellows, J. P. Donoghue, and M. J. Black, Automatic Spike Sorting for Neural Decoding, *Statistical Analysis of Neural Data*, Pittsburg, PA 2004

F. Wood, M. Fellows, M. J. Black, and J. P. Donoghue, Accuracy of manual spike sorting: results for the Utah intracortical array, *Society for Neuroscience,* New Orleans, LA 2003

Reviewing	Neural Information Processing Systems Uncertainty in Artificial Intelligence Artificial Intelligence and Statistics International Conference on Machine Learning Journal of Machine Learning Research Association for the Advancement of Artificial Intelligence Journal of Neuroscience Methods IEEE Transactions on Biomedical Engineering IEEE Transactions on Pattern Analysis and Machine Intelligence International Joint Conferences on Artificial Intelligence Journal of Statistics and Computing
Invited Conference Presentations	*"Applied Virtual Reality"*, SigGraph, Course 14, Los Angeles, CA, 1997 *"The Sequence Memoizer"*, Information Theory and Applications Workshop, UCSD, 2011
Invited Talks	*"The Sequence Memoizer"* Columbia University, Brown University, University of Edinburgh, Oxford University, Australia National University 2009 *"Nonparametric Bayesian Natural Language Model Domain Adaptation"* Columbia University, Princeton University, University of Utah 2009 *"Nonparametric Bayesian Natural Language Model Domain Adaptation"* Radboud University, NL and Cambridge University, UK 2007 *"A Nonparametric Bayesian Alternative to Spike Sorting"* University College London, UK and Radboud University, NL 2007 *"Gentle Introduction to Infinite Gaussian Mixture Modeling"* Brown University, RI, 2006 *"Bayesian Decoding for Neural Prostheses"* Northwestern University, IL, 2005 *"Variability of Manual Spike Sorting for Multi-Electrode Arrays"* University of Chicago, IL, 2003