Andrew McCallum

Department of Computer Science
University of Massachusetts Amherst
140 Governors Drive
Amherst, MA 01003

Email: mccallum@cs.umass.edu
Phone: (413) 545-1323
Fax: (413) 545-1789
Web: http://www.cs.umass.edu/∼mccallum

November 2008

1 Curriculum Vitae

1.1 Primary Research Interests

- Information extraction from text, Web mining.
- Data mining. Information integration. Knowledge Management.
- Bibliometrics. Digital libraries of research literature.
- Social network analysis.
- Natural language processing.
- Machine learning, graphical models, probabilistic reasoning.
- Probabilistic programming languages.
- Artificial intelligence.

1.2 Education

- Post-doctoral Research Fellow, School of Computer Science, Carnegie Mellon University, 1996. Advisers: Sebastian Thrun and Tom Mitchell. Machine Learning for Text, Language Modeling. Reinforcement Learning, Intelligent Control for Office Space.

- Ph.D. in Computer Science, University of Rochester, Rochester New York, 1995. Thesis title: "Reinforcement Learning with Selective Perception and Hidden State". Adviser: Dana Ballard; committee: Leslie Kaelbling (MIT), Chris Brown (Rochester), Randal Nelson (Rochester). Best Doctoral Dissertation of 1995-96, University of Rochester. Nominated for ACM Best Dissertation Award. (300 citations in Google Scholar)

- M.S. in Computer Science, University of Rochester, Rochester, New York, 1992.

- B.A. in Computer Science (*summa cum laude*, Phi Beta Kappa), Dartmouth College, Hanover, New Hampshire, 1989.

- High School: The North Carolina School of Science and Mathematics. 1985.

1.3 Professional Positions

- Associate Professor, Department of Computer Science, University of Massachusetts, Amherst, MA. Fall 2003–present. Director of Information Extraction and Synthesis Laboratory. Fall 2004–present.

  Currently advising ten Ph.D. students, two postdocs, and several undergraduates. Teaching graduate and undergraduate classes.

Research projects include (a) Unified inference to manage uncertainty across NLP, information extraction, information integration and data mining. (b) Extraction and mining of bibliometric data; building Rexa.info, an enhanced alternative to Google Scholar. (c) Efficient inference and training in conditional random fields, including relational models combining logic and probability, and new probabilistic programming languages going beyond logic. (d) Structured topic models: Bayesian latent variable models of text and other modalities, including authors, citations, time-stamps, geo-spatial locations. (e) Learning by reading. (f) Managing uncertainty in NLP dialog systems.

- Research Associate Professor, Department of Computer Science, University of Massachusetts, Amherst, MA. September 2002–September 2003.

- Vice President of Research & Development, Director of Pittsburgh office of WhizBang Labs, Inc. January 2000–May 2002.

  Set research direction for this 170-person company, identified key research problems and technical approaches, led two research-related software projects, managed group of research scientists, directed 30-person Pittsburgh office, joined sales team on visits with key customers.

- Adjunct Faculty, Carnegie Mellon University, Center for Automated Learning and Discovery, and Language Technology Institute. 1998–2002.

  Co-taught a graduate-level class, gave guest lectures at graduate and undergraduate levels, advised three PhD graduate students, advised graduate and undergraduate research projects, gave invited academic research presentations.

- Research Scientist and Research Coordinator, Just Research (Justsystem Pittsburgh Research Center). January 1997–December 1999.

  Spearheaded research in technology for statistical text processing.

  Leader of the project that created "Cora," a domain-specific search engine for computer science research papers—an early contemporary of CiteSeer. (With Kamal Nigam, Kristie Seymore and Jason Rennie.) This project resulted in eight refereed publications.

  Created *libbow*, an extensive software library for document classification, document clustering, information retrieval and information extraction written in C, released as open source. This is one of the most widely used software packages for document classification, with many hundreds of users from all over the world.

  Recruited and advised undergraduate and graduate summer interns, including Dayne Freitag (CMU), Kamal Nigam (CMU), Kristie Seymore (CMU), Thomas Minka (MIT), and many undergraduates, including Jason Rennie (now at MIT).

- Research Fellow, Computer Science Department, University of Rochester. January–June 1996. With Dana Ballard and Mary Hahoe.

  Psychophysics experiments with human eye-movements during the performance of highway driving tasks in virtual reality: testing models of short-term memory.

- Consultant, Net Community Inc., Portland, Oregon, 1996.

- Research Assistant, Computer Science Department, University of Rochester. Spring 1991, Spring 1992, Spring, Summer & Fall 1993, Summer & Fall 1994, Spring & Summer 1995.

  Accelerating reinforcement learning with techniques from Kohonen Maps. Learning Hidden Markov Models for tasks with incomplete perception. Memory-based techniques for fast learning of context-dependent tasks. Visual routines and deictic strategies on images generated from realistic images of a 3-D world. Human eye-tracking while chasing cars in virtual reality.

- Research Staff, Biomedical Information Communications Center, Oregon Health Sciences University. Summer 1992, Fall 1989–Summer 1990.

  Using Genetic Algorithms to categorize diseases from their symptoms. Designing graphical user interfaces and novel graphical representations for complex medical data.

- Programmer, IBM, Research Triangle Park, North Carolina. Summer & Winter 1988.

1.4 Grant and Contract Proposals and Activities

- PI, with Co-PIs Fernando Pereira and Ben Taskar. "Dynamically-Structured Conditional Random Fields for Complex, Natural Domains." NSF CISE Robust Intelligence, Medium. 2008-2011. (Acceptance rate for this program was about 5%.)

- PI, DARPA DIESEL program seedling (David Gunning, PM). Summer 2007–Summer 2008.

- IBM Faculty Partnership Award. 2007.

- PI with Padhraic Smyth (UC Irvine) (Co-PI) on NSF CRI, "Improving Experimental Computer Science with a Searchable Web Portal for Datasets." Summer 2006–Summer 2009.

- PI, IC Postdoctoral Fellowship. (Funding for a postdoc of my choice in my lab). "Resource-Bounded Information Gathering for Efficient Entity Resolution and Link Analysis using Probabilistic Reasoning and Decision Analysis." Fall 2006–Fall 2008, with option for one additional year.

- PI, "Aiding Collaboration through Probabilistic Methods on Text and Semi-Structured Data." IARPA. January 2006–August 2010.

- PI, Monster.com Inc, Research gift. 2006-2007.

- Co-PI with James Allan (PI) and Bruce Croft (Co-PI), DARPA GALE Program, "Information Distillation in Nightingale." September 2005–September 2010.

- PI with David Jensen (Co-PI) on NSF Medium ITR, "Unified Graphical Models of Information Extraction and Data Mining with Application to Social Network Analysis." (Acceptance rate for this program was below 10%) September 2003–September 2009, with option for extention.

- PI with Fernando Pereira (collaborative PI) and John Lafferty (collaborative PI), NSF Medium ITR, "Machine Learning for Sequences and Structured Data: Tools for Non-Experts." (Given highest rating by all reviewers. Acceptance rate for this program was below 10%.) September 2004–August 2007.

- PI, DARPA/IBM UIMA. "UIMA Integration with a Machine Learning for Language Toolkit." May 2005–May 2006.
- PI, Intelligence Technology Innovation Center (ITIC) grant. "Confidence Measures for Information Extraction of Entities, Relations and Object Correspondence." January 2005–January 2006.
- PI, DARPA IPTO, "Enduring Personal Cognitive Assistant" (CALO). May 2003–January 2009.
- IBM Faculty Partnership Award. 2004.
- PI, with Lance Ramshaw (prime contractor PI), BBN. U.S. Department of the Interior, ARDA, "Statistical Models for Information Extraction for REFLEX." September 2004–August 2007.
- Co-PI with Jared Freeman of Aptima, Inc. "Automated Diagnosis of Usability Problems Using Statistical Computational Methods" STTR. September 2003–May 2006.
- PI, Microsoft Research faculty gift. 2003-2004.
- Co-PI with Tom Dietterich (PI). NSF IIS "Student Travel Scholarships for ICML 2003"

1.5 Honors

- AAAI Fellow, 2009.
- UMass NSM "Distinguished Research Award" (given to 2/250 faculty per year), 2007.
- Kavli Fellow, National Academy of Sciences, 2007.
- UMass Lilly Teaching Fellow, 2005–2006.
- "Best Paper Honorable Mention," Proceedings of AAAI, 2004.
- National Science Foundation, Information Technology Research (ITR) Awards, 2003, 2004.
- IBM Faculty Partnership Award, 2004–2005, 2007–2008.
- Best Doctoral Dissertation of 1995–1996, University of Rochester.
- Nominated for ACM Best Ph.D. Dissertation Award, 1995.
- Phi Beta Kappa, 1989.
- Rufus Choates Scholar, and graduation *summa cum laude*. Dartmouth College. 1989.
- IBM T. J. Watson Scholarship. 1985.

2 Publications and Presentations

In November 2008 *Google Scholar* reported over 13,000 citations to 103 of the 155 research articles listed below. (Recent and other low-citation articles do not have their citation counts listed below.)

3 of the articles below had more than 1000 citations. 30 of the articles below had more than 100 citations.

The *h-index* for this collection of articles was 47. (See http://en.wikipedia.org/wiki/Hirsch_number for a description of this measure, and http://www.brics.dk/∼mis/hnumber.html for a script that calculates it.)

2.1 Submitted and Pending Journal Papers

[1] Wei Li and Andrew McCallum. "Pachinko Allocation: Scalable Mixture Models of Topic Correlations" Submitted to the Journal of Machine Learning Research, (JMLR).

[2] Ron Bekkerman, Ran El-Yaniv, Andrew McCallum "Multi-Way Distributional Clustering via Pairwise Interactions" Submitted to Journal of Machine Learning Research (JMLR).

2.2 Journal Publications

[1] Charles Sutton and Andrew McCallum. "Piecewise Training for Structured Prediction" Machine Learning Journal, (MLJ) 2009.

[2] Andrew McCallum, Xuerui Wang and Andres Corrada-Emmanuel. "Topic and Role Discovery in Social Networks with Experiments on Enron and Academic Email." Journal of Artificial Intelligence Research (JAIR), Volume 30, October 2007.

[3] Charles Sutton, Andrew McCallum and Khashayar Rohanimanesh. "Dynamic Conditional Random Fields." *Journal of Machine Learning Research (JMLR)*, Vol. 7, 2006. (100 citations in Google Scholar)

[4] Aron Culotta, Trausti Kristjannson, Andrew McCallum and Paul Viola. "Interactive Information Extraction and Active Learning with Constrained Conditional Random Fields." Artificial Intelligence Journal, Vol. 170 (No. 12), 2006.

[5] Xing Wei, Bruce Croft and Andrew McCallum. "Table Extraction for Answer Retrieval" Information Retrieval Journal, Vol. 9, (No. 5), November, 2006.

[6] Fuchun Peng and Andrew McCallum. "Information Extraction from Scientific Papers with Conditional Models," Information Processing and Management, Volume 42 (Number 4), pages 963-979, 2006.

[7] Andrew McCallum. "Information Extraction: Distilling Structured Data from Unstructured Text." ACM Queue, Vol. 3, (No 9), November 2005. (Invited article)

[8] Wei Li and Andrew McCallum. "Rapid Development of Hindi Named Entity Recognition Using Conditional Random Fields and Feature Induction." ACM Transactions on Asian Language Information Processing. Volume 2, Issue 3, September 2003. (18 citations in Google Scholar)

[9] Andrew McCallum, Kamal Nigam, Jason Rennie, Kristie Seymore. "Automating the Construction of Internet Portals with Machine Learning." Information Retrieval Journal, Volume 3, pages 127-163. Kluwer. 2000. (214 citations in Google Scholar)

[10] Kamal Nigam, Andrew McCallum, Sebastian Thrun and Tom Mitchell. "Text Classification from Labeled and Unlabeled Documents using EM." Machine Learning Journal, 2000. (1027 citations in Google Scholar)

[11] Mark Craven, Dan DiPasquo, Dayne Freitag, Andrew McCallum, Tom Mitchell, Kamal Nigam, Sean Slattery. "Learning to Construct Knowledge Bases from the World Wide Web". Artificial Intelligence, 118(1-2). pp 69-114. 2000. (317 citations in Google Scholar)

[12] William Cohen, Andrew McCallum, Dallan Quass. "Learning to Understand the Web." IEEE Data Engineering Bulletin. September 2000, Vol. 23, No. 3. Pages 17-24. (Invited article) (21 citations in Google Scholar)

[13] Andrew McCallum. "Hidden State and Reinforcement Learning with Instance-Based State Identification". In *IEEE Transactions on Systems, Man and Cybernetics*, (Special Issue on Robot Learning), 26 (3), June 1996 (54 citations in Google Scholar)

2.3 Book Chapters and Sections

[1] Andrew McCallum and Charles Sutton. Contributed a 1.5 page case study to Daphne Koller's book "Structured Probabilistic Models", 2008.

[2] David Cohn, Rich Caruana and Andrew McCalllum. "Semi-supervised Clustering with User Feedback." In *Constrained Clustering: Advances in Algorithms, Theory, and Applications.* Editors: Sugato Basu, Ian Davidson, and Kiri Wagstaff. Chapman & Hall/CRC Press, Data Mining and Knowledge Discovery Series. 2008.

[3] Andrew McCallum, Xuerui Wang and Natasha Mohanty. "Joint Group and Topic Discovery from Relations and Text" In *Statistical Network Analysis*, Lecture Notes in Computer Science (LNCS), Springer-Verlag. 2007.

[4] Charles Sutton and Andrew McCallum. "An Introduction to Conditional Random Fields for Relational Learning." In Introduction to Statistical Relational Learning. Edited by Lise Getoor and Ben Taskar. MIT Press. 2006.

[5] Kamal Nigam, Andrew McCallum and Tom Mitchell. "Semi-Supervised Text Classification Using EM." *Semi-supervised Learning.* Olivier Chapelle, Alexander Zien, and Bernhard Schlkopf, eds. 2006.

2.4 Refereed Conference Publications

[1] Michael Wick, Khashayar Rohanimanesh, Karl Schlutz and Andrew McCallum. "A Unified Approach for Schema matching, Coreference and Canonicalization. " Proceedings of the Conference on Knowledge Discovery and Data Mining, (KDD), 2008. (10% acceptance rate for long oral presentation)

[2] Robert Hall, Charles Sutton, Andrew McCallum. "Unsupervised Deduplication using Cross-field Dependencies." In Conference on Knowledge Discovery and Data Mining (KDD), 2008. (10% acceptance rate for long oral presentation)

[3] Gregory Druck, Gideon Mann and Andrew McCallum. "Learning from Labeled Features using Generalized Expectation Criteria." Proceedings of ACM Special Interest Group on Information Retreival, (SIGIR), 2008. (17% acceptance rate)

[4] David Mimno and Andrew McCallum. "Topic Models Conditioned on Arbitrary Features with Dirichlet-multinomial Regression," Conference on Uncertainty in Artificial Intelligence, (UAI), 2008. (Plenary presentation.) (13% acceptance rate for oral presentation)

[5] Gideon Mann and Andrew McCallum. "Generalized Expectation Criteria for Semi-Supervised Learning of Conditional Random Fields," Proceedings of Association of Computational Linguistics, (ACL), 2008. (20% acceptance rate)

[6] Vidit Jain, Erik Learned-Miller, and Andrew McCallum. "People-LDA: Anchoring Topics to People Using Face Recognition," International Conference on Computer Vision (ICCV), 2007. (23% acceptance rate)

[7] Xuerui Wang, Andrew McCallum and Xing Wei. "Topical N-grams: Phrase and Topic Discovery, with an Application to Information Retrieval," Proceedings of the 7th IEEE International Conference on Data Mining (ICDM), 2007. (19% acceptance rate)

[8] Aron Culotta, Michael Wick, Robert Hall, Matthew Marzilli and Andrew McCallum. "Canonicalization of Database Records using Adaptive Similarity Measures." Conference on Knowledge Discovery and Data Mining (KDD), 2007. (18% acceptance rate)

[9] Xuerui Wang, Chris Pal and Andrew McCallum. "Generalized Component Analysis for Text with Heterogeneous Attributes." Conference on Knowledge Discovery and Data Mining (KDD), 2007. (18% acceptance rate)

[10] Greg Druck, Chris Pal, Xiaojin Zhu and Andrew McCallum. "Semi-Supervised Classification with Hybrid Generative/Discriminative Methods." Conference on Knowledge Discovery and Data Mining (KDD), 2007. (18% acceptance rate)

[11] David Mimno and Andrew McCallum. "Expertise Modeling for Matching Papers with Reviewers." Conference on Knowledge Discovery and Data Mining (KDD), 2007. (18% acceptance rate)

[12] Wei Li, David Blei and Andrew McCallum. "Nonparametric Bayes Pachinko Allocation." Conference on Uncertainty in Artificial Intelligence (UAI), 2007. ($\sim$30% acceptance rate) (7 citations in Google Scholar)

[13] Charles Sutton and Andrew McCallum. "Improved Dynamic Schedules for Belief Propagation." Conference on Uncertainty in Artificial Intelligence (UAI), 2007. ($\sim$30% acceptance rate)

[14] Gideon Mann and Andrew McCallum. "Simple, Robust, Scalable Semi-supervised Learning via Expectation Regularization." International Conference on Machine Learning (ICML), 2007. (29% acceptance rate) (15 citations in Google Scholar)

[15] Charles Sutton and Andrew McCallum. "Piecewise Pseudolikelihood for Efficient Training of Conditional Random Fields." International Conference on Machine Learning (ICML), 2007. (29% acceptance rate) (12 citations in Google Scholar)

[16] David Mimno, Wei Li and Andrew McCallum. "Mixtures of Hierarchical Topics with Pachinko Allocation." International Conference on Machine Learning (ICML), 2007. (29% acceptance rate)

[17] Aron Culotta, Michael Wick and Andrew McCallum. "First-Order Probabilistic Models for Coreference Resolution." North American Association of Computational Linguistics / Human Language Technology (NAACL/HLT), 2007. (24% acceptance rate) (13 citations in Google Scholar)

[18] Gideon Mann and Andrew McCallum. "Efficient Computation of Entropy Gradient for Semi-Supervised Conditional Random Fields." North American Association of Computational Linguistics / Human Language Technology (NAACL/HLT), 2007. (short paper) (36% acceptance rate) (5 citations in Google Scholar)

[19] David Mimno and Andrew McCallum. "Mining a digital library for influential authors." Joint Conference on Digital Libraries (JCDL), 2007. (∼30% acceptance rate)

[20] David Mimno and Andrew McCallum "Organizing the OCA: Learning Faceted Subjects from a Library of Digital Books." Joint Conference on Digital Libraries (JCDL), 2007. (∼30% acceptance rate)

[21] Pallika Kanani, Andrew McCallum and Chris Pal. "Improving Author Coreference by Resource-bounded Information Gathering from the Web." Twentieth International Joint Conference on Artificial Intelligence (IJCAI), 2007. (15% acceptance rate) (10 citations in Google Scholar)

[22] Pallika Kanai and Andrew McCallum. "Resource-bounded Information Gathering for Correlation Clustering." Conference on Computational Learning Theory (COLT) Open Problems Track, 2007. (∼40% acceptance rate)

[23] Gary Huang, Erik Learned Miller and Andrew McCallum. "Cryptogram Decoding for Optical Character Recognition." International Conference on Document Analysis and Recognition (ICDAR), 2007. (66% acceptance rate)

[24] Xuerui Wang and Andrew McCallum. "Topics over Time: A Non-Markov Continuous-Time Model of Topical Trends." Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), pp. 424-433, 2006. (10% accepted) (37 citations in Google Scholar)

[25] Michael Wick, Aron Culotta, Andrew McCallum. "Learning field compatibilities to extract database records from unstructured text." Empirical Methods in Natural Language Processing (EMNLP), 2006. (18% accepted) (8 citations in Google Scholar)

[26] Gideon Mann, David Mimno and Andrew McCallum. "Bibliometric Impact Measures Leveraging Topic Analysis." Joint Conference on Digital Libraries (JCDL) 2006. (Full paper.) (11% accepted) (15 citations in Google Scholar)

[27] Wei Li and Andrew McCallum. "Pachinko Allocation: DAG-structured Mixture Models of Topic Correlations." Proceedings of the International Conference on Machine Learning (ICML), 2006. (∼20% accepted) (39 citations in Google Scholar)

[28] Chris Pal, Charles Sutton and Andrew McCallum. "Sparse Forward-Backward using Minimum Divergence Beams for Fast Training of Conditional Random Fields." In the proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), vol. 5, pp. 581-584, 2006. (∼25% accepted) (11 citations in Google Scholar)

[29] Andrew McCallum, Chris Pal, Greg Druck and Xuerui Wang. "Multi-Conditional Learning: Generative/Discriminative Training for Clustering and Classification." Proceedings of the 21st National Conference on Artificial Intelligence (AAAI), pp. 433-439, 2006. (30% accepted) (20 citations in Google Scholar)

[30] Aron Culotta, Andrew McCallum, Jonathan Betz. "Integrating probabilistic extraction models and data mining to discover relations and patterns in text." Proceedings of Human Language Technology / North American Association of Computational Linguistics (HLT-NAACL), 2006. (25% accepted) (24 citations in Google Scholar)

[31] Charles Sutton, Michael Sindelar, and Andrew McCallum. "Reducing Weight Undertraining in Structured Discriminative Learning." Proceedings of Human Language Technology / North American Association of Computational Linguistics (HLT-NAACL), 2006. (25% accepted) (5 citations in Google Scholar)

[32] Chris Pal and Andrew McCallum. "CC Prediction with Graphical Models." The Third Conference on Email and Anti-Spam, (CEAS), 2006. (36% accepted) (6 citations in Google Scholar)

[33] Xuerui Wang, Natasha Mohanty and Andrew McCallum. "Group and Topic Discovery from Relations and Text." Neural Information Processing Systems (NIPS), 2005. (One of 7% accepted for an oral or spotlight presentation.) (30 citations in Google Scholar)

[34] Michael Kelm, Chris Pal and Andrew McCallum. "Combining Generative and Discriminative Methods for Pixel Classification with Multi-Conditional Learning." The Proceedings of International Conference of Pattern Recognition (ICPR), 2006. ($\sim$50% accepted) (5 citations in Google Scholar)

[35] Shaolei L. Feng, R. Manmatha and Andrew McCallum. "Exploring the Use of Conditional Random Field Models and HMMs for Historical Handwritten Document Recognition." IEEE International Conference on Document Image Analysis for Libraries (DIAL 06), pp. 30-37. 2006. ($\sim$50% accepted) (5 citations in Google Scholar)

[36] Charles Sutton and Andrew McCallum "Composition of Conditional Random Fields for Transfer Learning." Conference on Human Language Technologies and Empirical Methods in Natural Language Processing (HLT/EMNLP), 2005. (32% accepted) (18 citations in Google Scholar)

[37] Nadia Ghamrawi and Andrew McCallum. "Collective Multi-Label Classification." Conference on Information and Knowledge Management (CIKM), (full paper), 2005. (17% accepted) (24 citations in Google Scholar)

[38] Aron Culotta and Andrew McCallum "Joint Deduplication of Multiple Record Types in Relational Data." Conference on Information and Knowledge Management (CIKM), (short paper), 2005. ($\sim$25% accepted) (10 citations in Google Scholar)

[39] Wei Li and Andrew McCallum. "Semi-Supervised Sequence Modeling with Syntactic Topic Models." Proceedings of the National Conference on Artificial Intelligence (AAAI), 2005. (27% accepted) (17 citations in Google Scholar)

[40] Aron Culotta and Andrew McCallum. "Reducing Labeling Effort for Structured Prediction Tasks." Proceedings of the National Conference on Artificial Intelligence (AAAI), 2005. (27% accepted) (10 citations in Google Scholar)

[41] Yu Gu, Andrew McCallum and Don Towsley. "Detecting Anomalies in Network Traffic Using Maximum Entropy Estimation." Internet Measurement Conference, 2005. (24% accepted) (33 citations in Google Scholar)

[42] Andrew McCallum, Kedar Bellare and Fernando Pereira. "A Conditional Random Field for Discriminatively-trained Finite-state String Edit Distance." Conference on Uncertainty in AI (UAI), 2005. (34% accepted) (30 citations in Google Scholar)

[43] Charles Sutton and Andrew McCallum. "Piecewise Training for Undirected Models." Conference on Uncertainty and Artificial Intelligence (UAI), 2005. (34% accepted) (40 citations in Google Scholar)

[44] Ron Bekkerman, Ran El Yaniv and Andrew McCallum. "Multi-way Distributional Clustering via Pairwise Interactions." International Conference on Machine Learning (ICML), 2005. (30% accepted) (35 citations in Google Scholar)

[45] Andrew McCallum, Andres Corrada, Xuerui Wang. "Topic and Role Discovery in Social Networks." International Joint Conference on Artificial Intelligence (IJCAI), 2005. (18% accepted, to appear) (72 citations in Google Scholar)

[46] Ron Bekkerman and Andrew McCallum. "Disambiguating Web Appearances of People in a Social Network." The World Wide Web Conference (WWW), 2005. (14% accepted) (86 citations in Google Scholar)

[47] Andrew McCallum, Andres Corrada, Xuerui Wang. "A Probabilistic Model for Topic and Role Discovery in Social Networks and Message Text." International Conference on Intelligence Analysis (IA), 2005. (16% acceptance rate)

[48] Andrew McCallum and Ben Wellner. "Conditional Models of Identity Uncertainty with Application to Noun Coreference." Neural Information Processing Systems (NIPS), 2004. (25% accepted) (78 citations in Google Scholar)

[49] Ben Wellner, Andrew McCallum, Fuchun Peng and Michael Hay. "An Integrated, Conditional Model of Information Extraction and Coreference with Application to Citation Matching." Uncertainty in Artificial Intelligence (UAI), 2004. (30% accepted) (63 citations in Google Scholar)

[50] Aron Culotta, Ron Bekkerman and Andrew McCallum. "Extracting Social Networks and Contact Information from Email and the Web." Conference on Email and Spam, 2004. (35% accepted) (71 citations in Google Scholar)

[51] Charles Sutton, Khashayar Rohanimanesh and Andrew McCallum. "Dynamic Conditional Random Fields: Factorized Probabilistic Models for Labeling and Segmenting Sequence Data." International Conference on Machine Learning (ICML), 2004. (32% accepted) (40 citations in Google Scholar)

[52] Fuchun Peng, Fang-fang Feng and Andrew McCallum. "Chinese Segmentation and New Word Detection using Conditional Random Fields." International Conference on Computational Linguistics (COLING), Geneva, Switzerland, 2004. ($\sim$20% accepted) (79 citations in Google Scholar)

[53] Trausti Kristjannson, Aron Culotta, Paul Viola and Andrew McCallum. "Interactive Information Extraction with Constrained Conditional Random Fields." Nineteenth National Conference on Artificial Intelligence (AAAI 2004). San Jose, CA. (Winner of Honorable Mention Award.) (26% accepted) (48 citations in Google Scholar)

[54] Fuchun Peng and Andrew McCallum. "Accurate Information Extraction from Research Papers using Conditional Random Fields." Proceedings of Human Language Technology Conference and North American Chapter of the Association for Computational Linguistics (HLT-NAACL), 2004. (26% accepted) (101 citations in Google Scholar)

[55] Aron Culotta and Andrew McCallum. "Confidence Estimation for Information Extraction." Proceedings of Human Language Technology Conference and North American Chapter of the Association for Computational Linguistics (HLT-NAACL), 2004, short paper. (26% accepted) (36 citations in Google Scholar)

[56] Rajat Raina, Yirong Shen, Andrew Y. Ng, Andrew McCallum. "Classification with Hybrid Generative/Conditional Models." Proceedings of Neural Information Processing Systems (NIPS), 2003. (18% accepted) (61 citations in Google Scholar)

[57] Andrew McCallum. "Efficiently Inducing Features of Conditional Random Fields." Conference on Uncertainty in Artificial Intelligence (UAI), 2003. (33% accepted) (169 citations in Google Scholar)

[58] Andrew McCallum, Wei Li. "Early Results for Named Entity Recognition with Conditional Random Fields, Feature Induction and Web-Enhanced Lexicons." Conference on Natural Language Learning (CoNLL), 2003. (144 citations in Google Scholar)

[59] David Pinto, Andrew McCallum, Xing Wei and Bruce Croft. "Table Extraction Using Conditional Random Fields." SIGIR 2003. (17% accepted) (140 citations in Google Scholar)

[60] David Blei, Drew Bagnell and Andrew McCallum. "Learning with Scope, with Application to Information Extraction and Classification." Uncertainty in Artificial Intelligence (UAI), 2002. (18 citations in Google Scholar)

[61] John Lafferty, Andrew McCallum and Fernando Pereira. "Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data." In *The Proceedings of the Eighteenth International Machine Learning Conference*, (ICML-2001). (1504 citations in Google Scholar)

[62] Nick Roy and Andrew McCallum. "Toward Optimal Active Learning through Sampling Estimation of Error Reduction." In *The Proceedings of the Eighteenth International Machine Learning Conference*, (ICML-2001). (182 citations in Google Scholar)

[63] Andrew McCallum, Dayne Freitag and Fernando Pereira. "Maximum Entropy Markov Models for Information Extraction and Segmentation." In *The Proceedings of the Seventeenth International Machine Learning Conference*, (ICML-2000). (441 citations in Google Scholar)

[64] Andrew McCallum, Kamal Nigam and Lyle Ungar. "Efficient Clustering of High-Dimensional Data Sets with Application to Reference Matching" *The Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, (KDD-2000). (235 citations in Google Scholar)

[65] Dayne Freitag and Andrew McCallum. "Information Extraction with HMM Structures Learned by Stochastic Optimization." *Proceedings of the Seventeenth National Conference on Artificial Intelligence*, (AAAI-2000). (152 citations in Google Scholar)

[66] Huan Chang, David Cohn and Andrew McCallum. "Learning to Create Customized Authority Lists." In *The Proceedings of the Seventeenth International Machine Learning Conference*, (ICML-2000). (56 citations in Google Scholar)

[67] Andrew McCallum, Kamal Nigam, Jason Rennie and Kristie Seymore "A Machine Learning Approach to Building Domain-Specific Search Engines," *Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence*, (IJCAI-99). (100 citations in Google Scholar)

[68] Jason Rennie and Andrew McCallum. "Using Reinforcement Learning to Spider the Web Efficiently." In *The Proceedings of the Sixteenth International Machine Learning Conference*, (ICML-99). (197 citations in Google Scholar)

[69] Andrew McCallum, Ronald Rosenfeld, Tom Mitchell and Andrew Ng. "Improving Text Classification by Shrinkage in a Hierarchy of Classes." In *The Proceedings of the Fifteenth International Machine Learning Conference*, (ICML-98). (332 citations in Google Scholar)

[70] Andrew McCallum and Kamal Nigam. "Employing EM in Pool-Based Active Learning for Text Classification." In *The Proceedings of the Fifteenth International Machine Learning Conference*, (ICML-98). (240 citations in Google Scholar)

[71] Doug Baker, Andrew McCallum "Distributional Clustering of Words for Text Classification." *Proceedings of the 21th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, (SIGIR-98). (385 citations in Google Scholar)

[72] Kamal Nigam, Andrew McCallum, Sebastian Thrun and Tom Mitchell. "Learning to Classify Text from Labeled and Unlabeled Documents." AAAI-98. (225 citations in Google Scholar)

[73] Mark Craven, Dan DiPasquo, Dayne Freitag, Andrew McCallum, Tom Mitchell, Kamal Nigam, Sean Slattery. "Learning to Extract Knowledge from the World Wide Web." AAAI-98. (437 citations in Google Scholar)

[74] Andrew McCallum. "Learning to Use Selective Attention and Short-Term Memory in Sequential Tasks". *From Animals to Animats 5: Proceedings of the Fifth International Conference on Simulation of Adaptive Behavior*, (SAB-96). (98 citations in Google Scholar)

[75] Andrew McCallum. "Instance-Based Utile Distinctions for Reinforcement Learning". In *The Proceedings of the Twelfth International Machine Learning Conference*, (ICML-95). Morgan Kaufmann Publishers, Inc. 1995. (132 citations in Google Scholar)

[76] Andrew McCallum. "Instance-Based State Identification for Reinforcement Learning". In *Advances of Neural Information Processing Systems*, (NIPS 7). 1995. (47 citations in Google Scholar)

[77] Andrew McCallum. "Overcoming Incomplete Perception with Utile Distinctions Memory". In *The Proceedings of the Tenth International Machine Learning Conference*, Amherst, Massachusetts. (ICML-93). (123 citations in Google Scholar)

[78] William Garrett, Ricardo Bianchini, Leonidas Kontothanassis, Andrew McCallum, Jeff Thomas, Robert Wisniewski and Michael L. Scott. "Linking Shared Segments". In *Proceedings of the USENIX Winter '93 Technical Conference*, pages 13–27, San Diego, CA, January 1993.

[79] Andrew McCallum. "Using Transitional Proximity for Faster Reinforcement Learning". In *The Proceedings of the Ninth International Machine Learning Conference*. (ICML-92). (19 citations in Google Scholar)

[80] Andrew McCallum and Kent Spackman. "Using Genetic Algorithms to Learning Disjunctive Rules from Examples". In *The Proceedings of the Seventh International Machine Learning Conference*. (ICML-90). (22 citations in Google Scholar)

2.5 Patents

[1] Dallan W. Quass, Tom M. Mitchell, Andrew K. McCallum, and William Cohen. "Method for Learning and Combining Global and Local Regularities for Information Extraction and Classification." U.S. Patent No.: 6,892,189. Application No.: 09/771,008; Filed: Jan 26, 2001. Issued May 10, 2005.

2.6 Invited Articles

[1] Dallan Quass, Andrew McCallum, William Cohen. "Unlocking the Information in Text." The Future of Software, Winter 2000/2001.

2.7 Theses

[1] Andrew McCallum. "Reinforcement Learning with Selective Perception and Hidden State", (Ph.D. Thesis). Department of Computer Science, University of Rochester. December 1995. (300 citations in Google Scholar)

2.8 Refereed Workshop Publications

[1] Andrew McCallum, Khashayar Rohanemanesh, Michael Wick, Karl Schultz, Sameer Singh. "FACTORIE: Efficient Probabilistic Programming for Relational Factor Graphs via Imperative Declarations of Structure, Inference and Learning." In Proceedings of the NIPS 2008 Workshop on Probabilistic Programming Languages.

[2] David Mimno, Hanna Wallach and Andrew McCallum. "Gibbs sampling for Logistic Normal Topic Models with Graph Based Priors." In Proceedings of the NIPS 2008 Workshop on Analyzing Graphs: Theory and Applications.

[3] Michael Wick, Khashayar Rohanimanesh, Andrew McCallum, and AnHai Doan. "A Discriminative Approach to Ontology Alignment." In the International Workshop on New Trends in Information Integration (NTII) at the conference for Very Large Databases (VLDB Workshop), Auckland, New Zealand, 2008.

[4] Hanna Wallach, Charles Sutton, Andrew McCallum. "Bayesian Modeling of Dependency Trees Using Hierarchical Pitman-Yor Priors." In International Conference on Machine Learning, Workshop on Prior Knowledge for Text and Language Processing. (ICML Workshop), 2008.

[5] David Mimno, Hanna M. Wallach and Andrew McCallum. "Community-based Link Prediction with Text." In Proceedings of the NIPS 2007 Workshop on Statistical Network Modeling (NIPS Workshop), 2007.

[6] Gregory Druck, Gerome Miklau and Andrew McCallum. "Learning to Predict the Quality of Contributions to Wikipedia", AAAI Workshop on Wikipedia and AI, (AAAI Workshop), 2008.

[7] Gregory Druck, Gideon Mann and Andrew McCallum. "Leveraging Existing Resources using Generalized Expectation Criteria." NIPS Workshop on Learning Problem Design, 2007.

[8] Kedar Bellare, Partha Pratim Talukdar, Giridhar Kumaran, Fernando Pereira, Mark Liberman, Andrew McCallum and Mark Dredze. "Lightly-Supervised Attribute Extraction for Web Search. NIPS Workshop on Machine Learning for Web Search," 2007.

[9] Kedar Bellare and Andrew McCallum. "Learning Extractors from Unlabeled Text using Relevant Databases." Sixth International Workshop on Information Integration on the Web (IIWeb), collocated with AAAI, 2007.

[10] Pallika Kanani and Andrew McCallum. "Efficient Strategies for Improving Partitioning-Based Author Coreference by Incorporating Web Pages as Graph Nodes." Sixth International Workshop on Information Integration on the Web (IIWeb), collocated with AAAI, 2007.

[11] David Mimno and Andrew McCallum. "Probabilistic Representations for Integrating Unreliable Data Sources." Sixth International Workshop on Information Integration on the Web (IIWeb), collocated with AAAI, 2007.

[12] Aron Culotta, Andrew McCallum. "Tractable Learning and Inference with High-Order Representations" ICML Workshop on Open Problems in Statistical Relational Learning, 2006.

[13] Aron Culotta, Andrew McCallum. "Practical Markov logic containing first-order quantifiers with application to identity uncertainty." HLT Workshop on Computationally Hard Problems and Joint Inference in Speech and Language Processing, 2006. (5 citations in Google Scholar)

[14] Wei Li, Xuerui Wang and Andrew McCallum. "A Continuous-Time Model of Topic Co-occurrence Trends." Proceedings of the 21st National Conference on Artificial Intelligence Workshop on Event Extraction and Synthesis, pp. 48-53, 2006.

[15] Chris Pal, Xuerui Wang, Michael Kelm and Andrew McCallum. "Multi-Conditional Learning for Joint Probability Models with Latent Variables." NIPS Workshop on Advances in Structured Learning for Text and Speech Processing, 2005.

[16] Wei Li and Andrew McCallum. "Pachinko allocation: A Directed Acyclic Graph for Topic Correlations." NIPS Workshop on Nonparametric Bayesian Methods, 2005.

[17] Aron Culotta, Andrew McCallum. "Learning clusterwise similarity with first-order features." NIPS Workshop on the Theoretical Foundations of Clustering, 2005.

[18] Xuerui Wang, Natasha Mohanty and Andrew McCallum. "Group and Topic Discovery from Relations and Text." KDD Workshop on Link Discovery: Issues, Approaches and Applications (LinkKDD) 2005.

[19] Charles Sutton and Andrew McCallum. "Joint Parsing and Semantic Role Labeling." Conference on Natural Language Learning (CoNLL) Shared Task, 2005. (9 citations in Google Scholar)

[20] Andrew McCallum, Andres Corrada-Emmanuel, Xuerui Wang. "The Author-Recipient-Topic Model for Topic and Role Discovery in Social Networks: Experiments with Enron and Academic Email." NIPS'04 Workshop on "Structured Data and Representations in Probabilistic Models for Categorization" (Also Technical Report UM-CS-2004-096), 2004. (19 citations in Google Scholar)

[21] Charles Sutton and Andrew McCallum. "Collective Segmentation and Labeling of Distant Entities in Information Extraction." ICML workshop on Statistical Relational Learning, 2004. (38 citations in Google Scholar)

[22] Jerod Weinman, Al Hansen and Andrew McCallum. "Sign Detection in Natural Images with Conditional Random Fields." IEEE International Workshop on Machine Learning for Signal Processing, 2004. (16 citations in Google Scholar)

[23] Hema Ragavan, James Allan and Andrew McCallum, "An Exploration of Entity Models, Collective Classification and Relation Description." KDD Workshop on Link Analysis and Group Detection, August 2004. (13 citations in Google Scholar)

[24] Andrew McCallum, Khashayar Rohanimanesh and Charles Sutton. "Dynamic Conditional Random Fields for Jointly Labeling Multiple Sequences." NIPS*2003 Workshop on Syntax, Semantics, Statistics, 2003. (18 citations in Google Scholar)

[25] Andrew McCallum and Ben Wellner. "Toward conditional models of identity uncertainty with application to proper noun coreference." IJCAI Workshop on Information Integration on the Web, 2003. (58 citations in Google Scholar)

[26] Andrew McCallum and David Jensen. "A Note on the Unification of Information Extraction and Data Mining using Conditional-Probability, Relational Models." IJCAI'03 Workshop on Learning Statistical Models from Relational Data, 2003. (35 citations in Google Scholar)

[27] Andrew McCallum. "Multi-Label Text Classification with a Mixture Model Trained by EM." AAAI'99 Workshop on Text Learning. (114 citations in Google Scholar)

[28] Andrew McCallum, Kamal Nigam, Jason Rennie and Kristie Seymore "Building Domain-Specific Search Engines with Machine Learning Techniques." AAAI-99 Spring Symposium on Intelligent Agents in Cyberspace. (124 citations in Google Scholar)

[29] Kamal Nigam, John Lafferty, Andrew McCallum. "Using Maximum Entropy for Text Classification." IJCAI'99 Workshop on Information Filtering. (296 citations in Google Scholar)

[30] Andrew McCallum and Kamal Nigam. "Text classification by bootstrapping with keywords, EM and shrinkage." Working Notes of ACL 1999 Workshop for the Unsupervised Learning in Natural Language Processing, pp. 52-58, 1999. (74 citations in Google Scholar)

[31] Rosie Jones, Andrew McCallum, Kamal Nigam and Ellen Riloff. "Bootstrapping for Text Learning Tasks." IJCAI-99 Workshop on Text Mining: Foundations, Techniques and Applications. 1999. (62 citations in Google Scholar)

[32] Dayne Freitag and Andrew McCallum "Information Extraction with HMMs and Shrinkage. " AAAI'99 Workshop on Machine Learning for Information Extraction. (236 citations in Google Scholar)

[33] Kristie Seymore, Andrew McCallum, Roni Rosenfeld "Learning Hidden Markov Model Structure for Information Extraction." AAAI'99 Workshop on Machine Learning for Information Extraction. (219 citations in Google Scholar)

[34] Andrew McCallum and Kamal Nigam. "A Comparison of Event Models for Naive Bayes Text Classification." AAAI-98 Workshop on "Learning for Text Categorization". (1246 citations in Google Scholar)

[35] Andrew McCallum. "Reduced Training Time for Reinforcement Learning with Hidden State". In *The Proceedings of the Eleventh International Machine Learning Workshop, "Reinforcement Learning"*, Rutgers University. 1994.

[36] Andrew McCallum. "Short-Term Memory in Visual Routines". In *The Working Notes of the AAAI Spring Symposium Series, "Toward Physical Interactions and Manipulation"*, Stanford University. 1994.

[37] Andrew McCallum. "Short-Term Memory for Visual Routines". In *The Proceedings of the Intelligent Robotic Systems Workshop, (IRS'94)*. Grenoble, France. 1994.

2.9 Unrefereed Papers

[1] Chris Pal, Xuerui Wang and Andrew McCallum. "Transfer Learning for Enhancing Information Flow in Organizations and Social Networks." Technical Note, 2007.

[2] Gary Huang, Erik Learned Miller and Andrew McCallum. "Cryptogram Decoding for Optical Character Recognition." UMass Technical Report UM-CS-2006-045, 2006.

[3] Xuerui Wang and Andrew McCallum. "A Note on Topical N-grams." UMass Technical Report UM-CS-2005-071, 2005 (8 citations in Google Scholar)

[4] Aron Culotta, Andrew McCallum. "A conditional model of deduplication for multi-type relational data." University of Massachusetts IR-443, 2005.

[5] Andrew McCallum, Xuerui Wang and Chris Pal. "Predictive Random Fields: Latent Variable Models Fit by Multiway Conditional Probability with Applications to Document Analysis." UMass Technical Report UM-CS-2005-053, 2005.

[6] Don Metzler, W. Bruce Croft and Andrew McCallum. "Direct Maximization of Rank-Based Metrics for Information Retrieval." CIIR Technical Report IR-429, 2005.

[7] Charles Sutton, Michael Sindelar, and Andrew McCallum. "Feature Bagging: Preventing Weight Undertraining in Structured Discriminative Learning." Center for Intelligent Information Retrieval, University of Massachusetts Technical Report IR-402. 2005.

[8] Ron Bekkerman, Andrew McCallum and Gary Huang. "Automatic Categorization of Email into Folders: Benchmark Experiments on Enron and SRI Corpora." Center for Intelligent Information Retrieval, Technical Report IR-418. 2005. (42 citations in Google Scholar)

[9] Andrew McCallum, Xuerui Wang and Chris Pal. "Predictive Random Fields: Latent Variable Models Fit by Multiway Conditional Probability with Applications to Document Analysis." UMass Technical Report UM-CS-2005-053, 2005.

[10] Charles Sutton and Andrew McCallum. "Fast, Piecewise Training for Discriminative Finite-state and Parsing Models." Center for Intelligent Information Retrieval Technical Report IR-403. 2005

[11] Aron Culotta, David Kulp, and Andrew McCallum. "Gene Prediction with Conditional Random Fields." UMass Technical Report UM-CS-2005-028, University of Massachusetts, Amherst, April 2005. (14 citations in Google Scholar)

[12] Andrew McCallum and Nadia Ghamrawi. "Collective Multilabel Text Classification." Technical Report UM-CS-2004.

[13] Wei Li and Andrew McCallum. "A Note on Semi-supervised Learning using Markov Random Fields." Technical Note, February 3, 2004. (4 citations in Google Scholar)

[14] James Allan, Jay Aslam, Nicholas Belkin, Chris Buckley, Jamie Callan, Bruce Croft, Sue Dumais, Norbert Fuhr, Donna Harman, David J. Harper, Djoerd Hiemstra, Thomas Hofmann, Eduard Hovy, Wessel Kraaij, John Lafferty, Victor Lavrenko, David Lewis, Liz Liddy, R. Manmatha, Andrew McCallum, Jay Ponte, John Prager, Dragomir Radev, Philip Resnik, Stephen Robertson, Roni Rosenfeld, Salim Roukos, Mark Sanderson, Rich Schwartz, Amit Singhal, Alan Smeaton, Howard Turtle, Ellen Voorhees, Ralph Weischedel, Jinxi Xu and ChengXiang Zhai "Challenges in information retrieval and language modeling: report of a workshop held at the center for intelligent information retrieval, University of Massachusetts Amherst." Journal of SIGIR Forum, Volume 37 (number 1), pp 31-47, 2003. (87 citations in Google Scholar)

[15] Andrew McCallum and Ben Wellner. "Object Consolidation by Graph Partitioning with a Conditionally Trained Distance Metric." Proceedings of the KDD Workshop on Data Cleaning. 2003. (16 citations in Google Scholar)

[16] David Cohn, Rich Caruana and Andrew McCallum. "Semi-supervised Clustering with User Feedback." Technical note. 2000. (84 citations in Google Scholar)

[17] Doug Baker, Thomas Hofmann, Andrew McCallum and Yiming Yang. "A Hierarchical Probabilistic Model for Novelty Detection in Text." 1999.

[18] Andrew McCallum. "Efficient Exploration in Reinforcement Learning with Hidden State." AAAI Fall Symposium on Model directed Autonomous Systems, 1997. (15 citations in Google Scholar)

[19] Andrew McCallum. "Persia User's Manual—A Scheme Interface to SGI's Performer 3D Graphics Library". Computer Science Department, University of Rochester. 1995.

[20] Andrew McCallum. "Utile Suffix Memory for Reinforcement Learning with Hidden State". TR 549, Department of Computer Science, University of Rochester. December 1994.

[21] Andrew McCallum. "First Results with Instance-Based State Identification". TR 502, Department of Computer Science, University of Rochester. 1994. (7 citations in Google Scholar)

[22] Andrew McCallum. "Learning with Incomplete Selective Perception", (Thesis proposal). TR 453, Department of Computer Science, University of Rochester. March 1993.

[23] Andrew McCallum. "First Results with Utile Distinctions Memory for Reinforcement Learning". TR 446, Department of Computer Science, University of Rochester. December 1992.

[24] William Garrett, Ricardo Bianchini, Leonidas Kontothanassis, Andrew McCallum, Jeff Thomas, Robert Wisniewski and Michael L. Scott. "Dynamic Sharing and Backward Compatibility on 64-Bit Machines". TR 418, Department of Computer Science, University of Rochester. April 1992. (13 citations in Google Scholar)

[25] Andrew McCallum. "Using Transitional Proximity for Faster Reinforcement Learning". Department of Computer Science, University of Rochester. 1992.

[26] Andrew McCallum. "Faster Reinforcement Learning with On-Line Value Iteration". In *The Proceedings of the Buffalo Graduate Conference on Computer Science*, SUNY Buffalo. 1992.

2.10 Invited Talks and Presentations

[1] Duke University, (and UNC, NSCU), **Triangle Computer Science Distinguished Lecturer Series**. November 2008.

[2] Columbia University. Seminar Series. October 1, 2008.

[3] Google Research, New York. October 2, 2008.

[4] Universite Pierre et Marie Curie (Paris, France), Department of Computer Science. July 11, 2008.

[5] INRIA (Sophia Antipolis, France). July 16, 2008.

[6] EURECOM (Sophia Antipolis, France). July 17, 2008.

[7] Xerox Research Center Europe (Grenoble, France). July 21, 2008.

[8] Google Research, Mountainview, CA. April 17, 2008.

[9] Carnegie Mellon University, CMU Machine Learning Department-Google Seminar. (Contact Kamal Nigam). March 2008.

[10] University of California Irvine, **Distinguished Speaker Series**, Center for Machine Learning and Intelligent Systems. February 2008.

[11] Kavli Frontiers of Science Symposium, National Academy of Sciences, Fall 2007.

[12] Oak Ridge National Lab, Fall 2007.

[13] Isenberg School of Management, University of Massachusetts. Operations Research / Management Science, Fall Speaker Series. November 3, 2007.

[14] Tsinghua University (China), Department of Computer Science. July 16, 2007.

[15] Peking University (China), Department of Computer Science. July 13, 2007.

[16] Microsoft Research Asia (Beijing, China). July 12, 2007. Two-day invited tutorial. As part of a series with Trevor Hastie, Stanford University.

[17] Kennedy School of Government, Harvard University. March 5, 2007.

[18] Second Annual Information Theory and Applications Workshop, UC San Dieago. Invited talk. (Over 500 researchers in attendence.) February 4, 2007.

[19] BBN. January 9, 2007.

[20] UMass National Center for Digital Government. December 2006.

[21] National Science Foundation. Invited talk. CISE IIS. October 5, 2006.

[22] The Twelfth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), Invited talk. August 2006.

[23] Carnegie Mellon University, AI Seminar Series. (Contact Tom Mitchell). March 2006.

[24] DARPA ISAT workshop study on "Adaptive and Interactive Representations." MIT. (Contact: Kendra Moore, DARPA PM.) July 2006.

[25] 23rd International Conference on Machine Learning (ICML), Workshop on Statistical Network Analysis: Models, Issues and New Directions. Invited talk. (Contact: Stephen Feinberg.) July 2006.

[26] 23rd International Conference on Machine Learning (ICML), Workshop on Open Problems in Statistical Relational Learning. Invited talk. (Contact: Lise Getoor.) July 2006.

[27] Neural Information Processing Systems (NIPS), Workshop on Open Problems and Challenges for Nonparametric Bayesian Methods in Machine Learning. (Contact Ye Whye Teh.) December 2005.

[28] Neural Information Processing Systems (NIPS), Workshop on Bayesian Methods for Natural Language Processing. (Contact Hal Daume.) December 2005.

[29] SRI International, Menlo Park, CA. CALO year End Status. (Contact Bill Marks.) October 2005.

[30] Stanford University, Computer Science Department, Broad Area Colloquium. (Contact: Andrew Ng.) October 2006.

[31] Yahoo, Inc. (Contact: Byron Dom) October 2006.

[32] University of Edinburgh. (Contact Miles Osborne and Chris Williams.) August 2005.

[33] ACL Workshop on Feature Engineering. Invited keynote presentation. "Recent Advances in Machine Learning Methods for Feature Engineering." (Contact Eric Ringger.) June 2005.

[34] Microsoft Research. (Contact Sue Dumais.) June 2005.

[35] University of Washington. (Contact Pedro Domingos.) June 2005.

[36] National Library of Medicine. (Contact John Wilbur.) May 2005.

[37] Department of Homeland Security, Workshop on Text Analysis, Invited Presentation. (Contact Tom Potok.) May 2005.

[38] Google NYC. (Contact Craig Nevill-Manning.) April 2005.

[39] UPenn, Department of Computer Science. (Contact: Fernando Pereira) April 2005.

[40] MIT, CSAIL. (Contact Leslie Kaelbling.) March 2005.

[41] Dagstuhl Workshop, Germany. Keynote address. (Contact Nicholas Kushmerick). February 2005.

[42] Alphatech Research. (Contact: Tom Stephenson.) January 2005.

[43] UC Berkeley. (Contact: Michael I. Jordan). November 2004.

[44] Google Research. (Contact Peter Norvig). November 2004.

[45] Lawrence Livermore National Laboratory. (Contact: Tina Eliassi-Rad). November 2004.

[46] Invited tutorial, Johns Hopkins, Center for Language and Speech Processing, NSF Summer School on Human Language Technology, 2004.

[47] Cornell University, Department Seminar Series. (Contact: Claire Cardie). April 2004.

[48] University of Washington, AI Seminar Series. (Contact: Oren Etzioni). December 2003.

[49] Microsoft Research. (Contact: Sue Dumais). December 2003.

[50] Carnegie Mellon University, AI Seminar Series. (Contact: Tuomas Sandholm). Fall 2003.

[51] BBN, Information Extraction Seminar Series. (Contact: Scott Miller). Fall 2003.

[52] Xerox Research Center Europe. (Contact: Nicola Cancedda). September 2003.

[53] University of Pennsylvania, Seminar Series. (Summer 2003).

[54] University of Wisconsin. (Contact: Jude Shavlik). May 2003.

[55] Invited by Ted Senator to a DARPA workshop assessing the feasibility of a new DARPA program on "Misinformation Detection." Selected to be one of three workshop report co-authors. February 2003.

[56] MIT AI Lab. (Contact: Leslie Kaelbling). "Information Extraction from the World Wide Web: Discriminative Finite State Models, Feature Induction and Scoped Learning". AI Lab Seminar Series. February 2003.

[57] University of Massachusetts Amherst, Math and Statistics Department. (Contact: Paola Sebastiani). "Conditional Random Fields for Finite State Sequence Analysis with Application to Information Extraction from the World Wide Web." November 2002.

[58] Brown University. (Contact: Thomas Hofmann). "Turning the Web into a Knowledge Base: Information Extraction with Finite State Models." Industrial Partners Program symposium. November 2002.

[59] University of Pennsylvania. (Contact: Fernando Pereira) "Learning from Scoped Regularities." October 2002.

[60] Google. (Contact: Peter Norvig) "Turning the Web into a Knowledge Base: Information Extraction with Finite State Models." May 2002.

[61] University of Texas at Austin. (Contact: Ray Mooney) "Information Extraction with Finite State Models." February 2002.

[62] IJCAI-2001 Workshop on Adaptive Text Extraction and Mining. Invited talk. (Contact: Nicholas Kushmerick) "Information Extraction with Machine Learning". Ralph Grisham gave the other invited talk. August 2001.

[63] AT&T Shannon Labs. (Contact: Fernando Pereira and Michael Kearns) "Automatically Building Internet Portals using Machine Learning". September 1999.

[64] IBM Almaden Research Lab. (Contact: Shivakumar Vaithaynathan) "Automatically Building Domain-Specific Search Engines using Machine Learning". March 1999.

[65] Stanford University, Computer Science Department. (Contact: Pat Langley and Daphne Koller) "Automatically Building Domain-Specific Search Engines using Machine Learning". March 1999.

[66] SRI. (Contact: Moises Goldszmidt and Andreas Stolcke) "Automatically Building Domain-Specific Search Engines using Machine Learning". March 1999.

[67] MIT AI Lab. (Contact: Paul Viola) "Text Classification with Limited Labeled Data". November 1998.

[68] MIT Media Lab. (Contact: Pattie Maes and Tony Jebara) "Text Classification with Limited Labeled Data". November 1998.

[69] Carnegie Mellon University, AI Seminar. (Contact: Tai Sing Lee) "Two Methods for Improving Text Classification when there is Sparse Training Data". April 1998.

[70] Media Lab, MIT. Invited talk. "Learning Visual Routines for Highway Driving—Goldilocks meets Reinforcement Learning". November 1996.

[71] AAAI Fall Symposium, (Symposium on 'Learning Complex Behaviors in Adaptive Intelligent Systems'). "Learning Task-Relevant State Spaces with a Utile Distinction Test". November 1996.

[72] Simulation of Adaptive Behavior. "Learning to Use Selective Attention and Short-Term Memory in Sequential Tasks". September 1996.

[73] Center for Visual Science, Department of Psychology, University of Rochester. Invited talk. "A Model for Learning Visual Routines with Short Term Memory". April 1996.

[74] Carnegie Mellon University, Reinforcement Learning Group. Invited talk. "Addressing Selective Attention and Hidden State with Utile Distinctions in Feature-Space and History-Space". April 1996.

[75] SUNY Geneseo, Department of Computer Science. Invited talk. "Learning Where to Look When Driving on a Crowded Highway". March 1996.

[76] Carnegie Mellon University, Reinforcement Learning Group. Invited talk. "First Results with Instance-Based State Identification for Reinforcement Learning". April 1994.

[77] AAAI Spring Symposium, ('Toward Physical Interaction and Manipulation'), Stanford University. "Short-Term Memory in Visual Routines". March 1994.

[78] Xerox PARC. Invited talk. "A Nearest-Neighbor Approach to Short-Term Context While Performing Tasks with Visual Routines". March 1994.

[79] Brown University. Invited talk. "Learning Hidden Markov Models for Short-Term Memory in Reinforcement Learning". June 1993.

[80] University of Buffalo Graduate Conference on Computer Science. "Speeding Reinforcement Learning with On-line Value Iteration". Fall 1992.

3 Service Activities

3.1 Professional Service

2 program co-chairmanships (ICML, CEAS), 3 program area chairmanships (ACL, IJCAI, NIPS), 2 journal editorial board memberships (MLJ, then JMLR, FnT-ML), 4 board memberships (IMLS, CRA-CCC, CoRR, DMIFG), 5 years officer-level service to international conferences (NIPS and ICML), 50+ international conference program committee memberships, 5 grant proposal review panels, co-organizer of 7 workshops at international conferences, 2 tutorials at international conferences (NIPS, KDD).

- Program Co-chair, International Conference on Machine Learning (ICML), with Sam Roweis, 2008.
- Board, CRA Computating Community Consortium. http://www.cra.org/ccc. 2007–2009.
- Program Committee Area Chair, Association of Computational Linguistics (ACL), 2008.
- Program Committee Area Chair, International Joint Conference on Artificial Intelligence (IJCAI), 2005.
- Program Co-chair, Conference on Email and Spam, 2005.
- Program Committee Area Chair, North American Association of Computational Linguistics / Human Language Technologies (NAACL/HLT), 2004.
- Program Committee Area Chair, Neural Information Processing Systems (NIPS), 2003.
- Action Editor, Journal of Machine Learning Research, 2003–. Founding Member of Editorial Board, 2001–.
- Editoral Board Member, Foundations and Trends in Machine Learning, (Michael Jordan, editor), 2007–.
- Board member, International Machine Learning Society (IMLS), 2006–.
- Board member, ACM Computing Research Repository (CoRR), 2004–.

- Board member, Data Mining and Information Fusion Group, Computer Science and Telecommunications Board, The National Academies. 2005–.
- Co-organizer IPAM 2007 Workshop on "Social Network Analysis".
- Online Proceedings Chair, Neural Information Processing Systems (NIPS), 2003-2008.
- Fundraising Chair for International Conference on Machine Learning (ICML), 2003. Raised over $40k for student travel scholarships.
- Program Committee, Empirical Methods in Natural Language Processing (EMNLP), 2004–2007.
- Program Committee, World Wide Web Conference, 2004–2005.
- Program Committee, Neural Information Processing (NIPS), 1998–2007.
- Program Committee, International Conference on Machine Learning (ICML), 1998–2007, 2009.
- Program Committee, Conference on Email and Spam (CEAS), 2004.
- Program Committee, International Joint Conference on Artificial Intelligence (IJCAI), 2007, 2005, 2003, 2001, 1999, 1997, 1995.
- Program Committee, ACM Special Interest Group on Information Retrieval (SIGIR), 2003.
- Program Committee, Conference of the American Association for Artificial Intelligence (AAAI), 1997, 1998, 2000, 2004, 2005, 2006.
- Program Committee, Uncertainty in Artificial Intelligence (UAI), 2004–2006.
- Program Committee, Association for Computational Linguistics (ACL), 2004–2007, 2009.
- Program Committee, Empirical Methods in Natural Language Processing, 2004–2006.
- Reviewer, ACM-SIAM Symposium on Discrete Algorithms (SODA), 2005.
- Panel Session member, Conference on Human Language Technology, 2004.
- Invited tutorial, Johns Hopkins, Center for Language and Speech Processing, NSF Summer School on Human Language Technology, 2003.
- Grant proposal review panel. NSF Robust Intelligence. May 2008.
- Grant proposal review panel. NSF ITR. June 2002.
- Grant proposal review panel. NSF SBIR. 2004.
- Grant proposal review panel. U. S. Department of Energy. 2005.
- Invited tutorial at Knowledge Discovery and Data Mining Conference (KDD), "Information Extraction from the World Wide Web," with William Cohen. Summer 2003.

- Program area chair, "Algorithms and Architectures" section co-chair, Neural Information Processing Systems Conference (NIPS), with Geoffrey Hinton, Chris Burges, Yoram Singer, Martin Wainwright and Alex Smola. 2003.
- Gave invited tutorial at Neural Information Processing Systems Conference (NIPS), "Information Extraction from the World Wide Web," with William Cohen. Fall 2002.
- Co-organizer NAACL/HLT 2006 Workshop on "Computationally hard problems and joint inference in speech and language processing", with Charles Sutton and Jeff Bilmes.
- Co-organizer KDD 2003 Workshop on "Record Linkage", with Sheila Tejada.
- Co-Organizer of IJCAI workshop "Text Learning: Beyond Supervision,", with Lillian Lee, Tony Jebara and Kamal Nigam). 2001.
- Co-Organizer of IJCAI Workshop "Machine Learning for Information Filtering," with Thorsten Joachims, Mehran Sahami and Lyle Ungar. 1999.
- Co-Organizer of NIPS*98 Workshop "Integrating Supervised and Unsupervised Learning," with Rich Caruana, Virginia de Sa and Michael Kearns. 1998.
- Co-Organizer of ICML/AAAI Workshop "Learning for Text Categorization," with Mehran Sahami, Mark Craven, and Thorsten Joachims. 1998.
- Reviewer for many journals, including: Machine Learning Journal (MLJ), Journal of Machine Learning Research (JMLR), Journal of Artificial Intelligence Research (JAIR), BMC Bioinformatics, IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), IEEE Systems, Man and Cybernetics (SMC), IEEE Computer Vision and Pattern Recognition (CVPR), Journal for the American Society of Information Science and Technology (JASIST), IEEE Transactions on Speech and Language Processing (TSLP).
- Chief Maintainer of GNUstep Project, appointed by Richard Stallman, January 1996-August 1998. GNUstep is the Free Software Foundation's effort to implement NeXT Computer Inc.'s OpenStep object-oriented GUI standard. Also made small contributions to gcc compiler and emacs editor.

3.2 Departmental Service Activities

- Graduate Program Committee, Computer Science Department, University of Massachusetts Amherst. 2008–2009.
- Strategic Planning Committee, Computer Science Department, University of Massachusetts Amherst. 2008–2009.
- Outreach Committee (Chair), Computer Science Department, University of Massachusetts Amherst. 2006–2008.
- Executive Committee, Computer Science Department, University of Massachusetts Amherst. 2006–2007.

- Curriculum Committee, Computer Science Department, University of Massachusetts Amherst. 2005–2006, 2007-2008.
- Diversity Committee, Computer Science Department, University of Massachusetts Amherst. 2005–2006.
- Special Events/Colloquia Committee (chair), Computer Science Department, University of Massachusetts Amherst. 2004–2005.
- Personnel Committee, Computer Science Department, University of Massachusetts Amherst. 2004–2005.
- Search Committee for Department Chair, Computer Science Department, University of Massachusetts Amherst. 2003–2004.
- Outreach Committee, Computer Science Department, University of Massachusetts Amherst. 2003–2004.
- Space Committee, Computer Science Department, University of Massachusetts Amherst. 2003–2004.
- Faculty Recruiting Committee, Computer Science Department, University of Massachusetts Amherst. 2002–2003.
- Creator and organizer of new UMass Computer Science weekly discussion series, "Research Discussion Breakfast" in which a faculty members present and discuss their research with other faculty. 2008–2009.
- Creator and organizer of new UMass Computer Science monthly discussion series, "Big Picture Advice Breakfast" in which a a senior faculty member gives advice and participates in discussion with junior faculty. 2004–2005.
- Creator and organizer of new UMass Computer Science weekly seminar series, "Machine Learning and Friends Lunch". Average attendance: 50 students, faculty and staff. 2002–.
- Graduate Student Representative, Computer Science Department, University of Rochester. Attended faculty meetings, representing the graduate students. 1993–1994.
- Curriculum Committee, Computer Science Department, University of Rochester. 1992–1993.
- Admissions Committee, Computer Science Department, University of Rochester. 1991–1992.

3.3 Software Packages and Services

- FACTORIE. A toolkit for probabilistic programming. Implements factor graphs, MCMC inference, discriminative training. Innovative approach based on combining declarative and procedural knowledge. Extremely scalable and efficient.
- Rexa. A Web portal for computer science research, supporting search and analysis for research papers, people, conferences, journals, universities, conferences, grants, and the links between them. Launched as http://www.rexa.info in April 2006.

- MALLET. Machine Learning for Language Toolkit created at University of Massachusetts Amherst. Released as Open Source. Personally wrote more than 30k lines of Java; now includes about 50k lines. In active use at MIT, UPenn, Berkeley, Stanford, CMU, and other places, including China, France, England. Fall 2002–present. (168 citations in Google Scholar)
- C4. Co-architected and helped develop first stages of the foundation classes for WhizBang's machine learning software libraries. System eventually grew to over 1 million lines of Java code. 2000-2002.
- Cora. Research paper search engine, an early contemporary of CiteSeer.org, 1998–2002.
- Libbow. Software package for statistical text classification, clustering and retrieval. Has thousands of users from all over the world. Subject of an article by John Udell in InfoWorld magazine. Nearly 100k lines of C. Released under GPL. 1996-2000. (462 citations in Google Scholar)
- RLkit. Software library that makes it easy to test various reinforcement learning algorithms in different environments with different sensory-motor systems. Objective-C and Guile. 1992-1995.
- Persia. Toolkit for building virtual reality environments on SGI Onyx RealityEngine. The kit is based on SGI's Performer library and an embeddable Scheme interpreter. 1994-1995.

4 Advising and Teaching

4.1 Doctoral Committees

[1] Aron Culotta, University of Massachusetts Amherst. Chair. Graduated 2008.

[2] Charles Sutton, University of Massachusetts Amherst. Chair. Graduated 2008.

[3] Wei Li, University of Massachusetts Amherst. Chair. Graduated 2007.

[4] Hoifung Poon, **University of Washington**. Adviser, Pedro Domingos.

[5] Nick Matsakis, **MIT**. Adviser, David Karger.

[6] Daniel Lowd, **University of Washington**. Adviser, Pedro Domingos.

[7] Razvan Bunescu, **University of Texas at Austin**. Adviser, Ray Mooney.

[8] Hal Daume, **University of Southern California**. Adviser, Daniel Marcu. Graduated 2006.

[9] Kathryn Flack, **Linguistics Department, University of Massachusetts Amherst**. Adviser, John McCarthy. Graduated 2008.

[10] Jerod Weinman, University of Massachusetts Amherst. Advisers, Erik Learned-Miller and Allen Hansen. Graduated 2007.

[11] Brendan Burns, University of Massachusetts Amherst. Adviser, Oliver Brock. Graduated 2007.

[12] Donald Metzler, University of Massachusetts Amherst. Adviser, Bruce Croft. Graduated 2007.

[13] Xing Wei, University of Massachusetts Amherst. Adviser, Bruce Croft. Graduated 2007.

[14] Jennifer Neville, University of Massachusetts Amherst. Adviser, David Jensen. Graduated 2006.

[15] Xiaoyan Li, University of Massachusetts Amherst. Adviser, Bruce Croft. Graduated 2006.

[16] David Stracuzzi, University of Massachusetts Amherst. Adviser, Paul Utgoff. Graduated 2006.

[17] Andy Arnt, University of Massachusetts Amherst. Adviser, Shlomo Zilberstein. Graduated 2005.

[18] Andrew Schein, **University of Pennsylvania**. Adviser, Lyle Ungar. Graduated 2005.

[19] Kamal Nigam, Carnegie Mellon University. Adviser, Tom Mitchell. Graduated 2001.

[20] Kristie Seymore, Carnegie Mellon University. Adviser, Roni Rosenfeld.

4.2 Current Graduate and Undergraduate Advisees

[1] Xuerui Wang (UMass Ph.D. student). Group, role and topic discovery in social network analysis. 2004–.

[2] Kedar Bellare (UMass Ph.D. student). Discriminatively-trained alignment models. 2004–.

[3] Pallika Kanani (UMass Ph.D. student). Resource-bounded information gathering for coreference and information extraction. 2005–.

[4] Greg Druck (UMass Ph.D. student). Semi-supervised learning for information extraction. 2005–.

[5] David Mimno (UMass Ph.D. student). Social network analysis and corpus discovery using structured topic models. 2005–.

[6] Michael Wick (UMass Masters/Ph.D. student). Efficient inference in learning in relational factor graphs, especially for entity resolution and information integration 2005–.

[7] Karl Schultz (UMass Masters/Ph.D student). Joint inference for information extraction, natural language processing, dialog.

[8] Jason Naradowsky (UMass Ph.D. student). Latent variable models of natural language.

[9] Limin Yao (UMass Ph.D. student). Information extraction and probabilistic databases.

[10] Sameer Singh (UMass Masters/Ph.D. student). Accurate and scalable entity resolution.

[11] Megan Day (UMass undergraduate). Bibliometrics and topic analysis of bio-medical research literature.

4.3 Past Graduate and Undergraduate Advisees and Research Supervised

[1] Robert Hall (UMass Masters student). Markov-chain Monte Carlo methods for coreference and parsing. 2006-2008. (Now PhD student at Carnegie Mellon University.)

[2] Matthew Marzilli (UMass undergraduate student). Statistical models of canonicalization from multiple extracted mentions. (Now PhD student at UMass in Software Engineering.)

[3] Gary Huang (UMass Ph.D. student). Interactive information extraction and active learning. 2003–. (Now jointly advised with Erik Learned-Miller. Fall 2006 on internship at Google.)

[4] Michael Sindelar (UMass undergraduate student). Feature bagging for regularization in training conditional random fields. (Michael was then first-author on an accepted paper to EMNLP Conference, 2006.)

[5] Ron Bekkerman (UMass Ph.D. student). Co-reference of entities across email and the Web. 2003–2005. (Now working with James Allan.)

[6] Benjamin Wellner (UMass Ph.D. student). Markov random fields for co-reference resolution. 2002–2004. (On long-term parental leave.)

[7] Michael Mammarella (UMass Masters student, now Ph.D. student at UCLA). Active and targeted spidering of the Web. 2003-2004.

[8] Ben Lambert (UMass undergraduate, now in Ph.D. program at CMU). Extracting a database of people relations from the World Wide Web. (Senior honors thesis.) 2002–2003.

[9] Nadia Ghamrawi (UMass Ph.D. student). Conditional probability models for multi-labeled data. 2003–2004.

[10] Hema Raghavan (UMass Ph.D. student). Synthesis project. 2003. (Joint with James Allan)

[11] David Blei (UC Berkeley Ph.D. student, now Assistant Professor at Princeton) and Drew Bagnell (CMU Ph.D. student, now Research Assistant Professor at CMU). Scoped learning for integrated local/global learning with language and layout. 2002.

[12] Kamal Nigam (CMU Ph.D. student, now Assistant Director, Google Labs, Pittsburgh). Document classification and clustering with unlabeled data and EM, (with Tom Mitchell). 1998–2001.

[13] Kary Meyers (CMU Ph.D. student). Language modeling with the Dirichlet-Multinomial (with Larry Wasserman). 2000.

[14] Nicholas Roy (CMU Ph.D. student, now Assistant Professor, MIT). Active learning with statistical sampling. 2000.

[15] Alexandrin "Sasha" Popescul (U Penn Ph.D. student, now at AskJeeves.com). Distributional clustering for integrating site-level information into document classification. 2000.

[16] Andrew Ng (UC Berkeley Ph.D. student, now Assistant Professor, Stanford). More accurate posterior probabilities for Bayesian classifiers. 2000.

[17] Thomas Minka (MIT Ph.D. student, now Research Scientist, Microsoft). Learning distance metrics classifiers trained with labeled and unlabeled data. 1999.

[18] Kristie Seymore (CMU Ph.D. student, now COO software start-up company). Information Extraction with hidden Markov models. 1998-2000.

[19] Jason Rennie (CMU undergraduate, now Ph.D. student at MIT) Efficient spidering with Reinforcement Learning. 1998-1999. Document classification with large negative classes. 1997.

[20] Doug Baker (CMU Ph.D. student, now running a lab at Nokia). Word merging for document classification. 1998.

[21] Jessica Bayliss (Rochester Ph.D. student). Reinforcement learning with short-term memory. 1996.

[22] Josh Richardson (Rochester undergraduate). Building graphical simulations of highway environments. 1995.

[23] Matt Hohlfeld (Rochester undergraduate). A Scheme interface to SGI's 'Performer' Graphics 3D Library. 1995.

[24] Jennifer Alexander (Rochester undergrad). Reinforcement learning for a cheetah learning to stalk and pounce. 1994–1996.

[25] Chandler Chao (Rochester undergraduate). Learning to perform tasks with deictic primitives on bitmap-vision perception. 1994.

[26] Mia Stern (Rochester undergraduate, went on to get a Ph.D. from UMass). Reinforcement learning exploration strategies. 1992.

[27] Jason Anderson (Rochester undergraduate). A machine learning simulator implemented in Objective C. 1992.

4.4 Teaching Experience

- Instructor, "Our Networked World," UMass CMPSCI 191c, Newly-created UMass undergraduate-level three-credit course in economic/political/social networks, graph properties, generation, clustering, navigation, privacy and behavior. Co-instructor with David Jensen. Fall 2008.
- Instructor, "Introduction to Natural Language Processing," UMass CMPSCI 585, UMass graduate and undergraduate-level three-credit course with lectures, quizzes, readings, class project, programming and writing assignments. Fall 2007, Spring 2004, Fall 2004.
- Instructor, "Introduction to Problem Solving with Computers," UMass CMPSCI 121, UMass undergraduate-level three-credit introduction to computer programming. 150 students. Fall 2006.
- Instuctor, "H15 Honors Section," Honors section to CMPSCI 121, teaching the "big ideas" of computer science, including computational complexity, finite-state automata, parsing, search, cryptography, robotics, uncertainty.
- Instructor, "Computational Linguistics," UMass CMPSCI 591N. Newly-designed interdisciplinary course for linguistics and computer science undergraduates and graduate students, covering computational approaches to parsing, anaphora resolution, verb sub-categorization, lexicon discovery, semantics and pragmatics. Course designed during Lilly Fellowship. 22 students, from both departments. Spring 2006.
- Instructor, "Introduction to Problem Solving with Computers," UMass CMPSCI 121, UMass undergraduate-level three-credit introduction to computer programming. $\sim$120 students. Spring 2005.

- Instructor, "Computational Social Network Analysis," UMass CMPSCI CS791S, graduate-level three-credit course with readings, discussion, written assignments and class project. Co-taught with David Jensen. Fall 2003.
- Instructor, "Statistical Information Extraction," UMass CMPSCI 791E graduate-level three-credit course with lectures, quizzes, readings and class project. Spring 2003.
- Guest Lecturer, "Information Retrieval," UMass CMPSCI graduate-level course. Fall 2002.
- Guest Lecturer, "Artificial Intelligence," UMass CMPSCI graduate-level course. Fall 2002.
- Guest Lecturer, "Advanced Topics in Information Retrieval," UMass CMPSCI graduate-level seminar. Lectures on maximum entropy and finite state models. Fall 2002.
- Project-customer guest, "Software Engineering", UMass CMPSCI undergraduate-level course. Fall 2002.
- Instructor, "Machine Learning for Text Mining", CMU CS graduate-level course. Semester-long course with lectures and class project. Co-taught with Tom Mitchell, Fernando Pereira and William Cohen. 2000.
- Guest Lecturer, "Machine Learning", CMU CS graduate-level course. (Main instructor was Roni Rosenfeld). 2001.
- Guest Lecturer, "Machine Learning", CMU CS undergraduate-level course. (Main instructor was Sebastian Thrun). 1999, 2000.
- Guest Lecturer, "Information Retrieval", CMU CS graduate-level course. 1998, 1999. (Main instructor was Yiming Yang.)
- Instructor, "Introduction to Computer Systems", University of Rochester, (CSC104). Fall 1995, Fall 1991. Introductory programming class for non-majors. Taught two lectures per week for entire semester. Redesigned course lecture format for greater interactivity and student involvement; new design was then used by other instructors.
- Teaching Assistant and Lecturer, "Artificial Intelligence", University of Rochester. Spring 1992. Lectured one fourth of the classes. Assistant for Professor Chris Brown.
- Teaching and Project Assistant, Rochester First-Year Graduate Student Boot-Camp Course (CSC400), 1992. Organized machine learning component.
- Undergraduate teaching assistant, "Artificial Intelligence", "Operating Systems", Dartmouth College. 1988–1989. For Professors Peter Sandon and Matt Bishop.

4.5 Public Tutorials, Oral and Written

Collated information, duplicated from above.

- Invited tutorial at Neural Information Processing Systems Conference (NIPS), "Information Extraction from the World Wide Web," with William Cohen. Fall 2002.

- Invited tutorial at Knowledge Discovery and Data Mining Conference (KDD), "Information Extraction from the World Wide Web," with William Cohen. Summer 2003.
- Invited tutorial, Johns Hopkins, Center for Language and Speech Processing, NSF Summer School on Human Language Technology, 2003.
- Invited tutorial, Johns Hopkins, Center for Language and Speech Processing, NSF Summer School on Human Language Technology, 2004.
- Charles Sutton and Andrew McCallum. "An Introduction to Conditional Random Fields for Relational Learning." In Introduction to Statistical Relational Learning. Edited by Lise Getoor and Ben Taskar. MIT Press. 2006. (87 citations in Google Scholar)
- Andrew McCallum. "Information Extraction: Distilling Structured Data from Unstructured Text." ACM Queue, volume 3, number 9, November 2005. (An overview of information extraction by machine learning methods, written for people not familiar with machine learning, especially CTOs and other people in business.) (30 citations in Google Scholar)

5 Personal Information

- Born: 1967
- Nationality: United States of America
- Marital Status: Married, two children
- Languages: English (Native language), French (9 years of experience; attended one semester in University of Toulouse, France), Greek (Limited knowledge).