John DeNero

Contact Information

Google Research
1350 Charleston Road
Mountain View, CA 94043

Phone: (415) 203-1943
Email: denero@google.com
Web: http://www.denero.org

Research Interests

Statistical machine translation, natural language processing, machine learning, and computer science education

Current

Research scientist for Google Translate, Mountain View, CA — *2010-present*

Education

University of California, Berkeley
Ph.D., Computer Science — *2005-2010*
Advisor: Dan Klein

Stanford University
Master of Arts, Philosophy — *2001-2002*

Bachelor of Science, with distinction — *1998-2002*
Major: Mathematical and Computational Science
Secondary Major: Symbolic Systems

Honors and Awards

Google Global Intern Scholarship — *2009*

Teaching Effectiveness Award, UC Berkeley — *2008*

Outstanding Graduate Student Instructor, CS Division — *2007*

Outstanding Graduate Student Instructor, UC Berkeley — *2007*

Teaching Experience

Primary Course Instructor — *Fall 2011*
Instructor for CS 61A: Structure and Interpretation of Computer Programs, a 500+ student lower-division course. Delivered lectures and created all projects, assignments, lecture notes, and exams.

Primary Course Instructor — *Spring 2009*
Instructor for CS 188: Introduction to Artificial Intelligence, a 95 student upper-division course. Delivered 27 lectures, managed a 7-person course staff, and created all projects, assignments, and exams.

Graduate Student Instructor — *Fa '05, Sp '06 Fa '06, Fa '07*
Support staff for CS 188 in charge of discussion sections, labs, review sessions, and course projects.

Tutorial Instructor — *Sum '09 & '10*
Lecturer and project designer for a tutorial in machine translation during the Summer Workshop of the Center for Language and Speech Processing at Johns Hopkins University, co-taught with Adam Lopez.

Publications

Thesis

John DeNero. "Phrase Alignment Models for Statistical Machine Translation," in *Ph.D. Dissertation in Computer Science, UC Berkeley*, 2010.

Refereed Publications

David Golland, John DeNero, and Jakob Uszkoreit. "A Feature-Rich Constituent Context Model for Grammar Induction," in *Proceedings of the Association of Computational Linguistics*, 2012.

Spence Green and John DeNero. "A Class-Based Agreement Model for Generating Accurately Inflected Translations," in *Proceedings of the Association of Computational Linguistics*, 2012.

Mohit Bansal, John DeNero, and Dekang Lin. "Unsupervised Translation Sense Clustering," in *Proceedings of the North American Association of Computational Linguistics*, 2012.

Robert Moore and John DeNero. "L1 and L2 Regularization for Multiclass Hinge Loss Models," in *Proceedings of the Symposium on Machine Learning in Speech and Language Processing*, 2011.

John DeNero and Jakob Uszkoreit. "Inducing Sentence Structure from Parallel Corpora for Reordering," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 2011.

John DeNero and Klaus Machery. "Model-Based Aligner Combination Using Dual Decomposition," in *Proceedings of the Association of Computational Linguistics*, 2011.

John DeNero and Dan Klein. "Discriminative Modeling of Extraction Sets for Machine Translation," in *Proceedings of the Association of Computational Linguistics*, 2010.

John DeNero, Shankar Kumar, Ciprian Chelba, and Franz Och. "Model Combination for Machine Translation," in *Proceedings of the North American Association of Computational Linguistics*, 2010.

John DeNero and Dan Klein. "Teaching Introductory Artificial Intelligence with Pac-Man," in *Proceedings of the Symposium on Educational Advances in Artificial Intelligence*, 2010.

John DeNero and Dan Klein. "The Pac-Man Projects Software Package for Introductory Artificial Intelligence," in *Proceedings of the Symposium on Educational Advances in Artificial Intelligence, Model Assignments Track*, 2010.

Taylor Berg-Kirkpatrick, Alexandre Bouchard-Ct, John DeNero, and Dan Klein. "Painless Unsupervised Learning with Features," in *Proceedings of the North American Association of Computational Linguistics*, 2010.

John DeNero, David Chiang, and Kevin Knight. "Fast Consensus Decoding over Translation Forests," in *Proceedings of the Association of Computational Linguistics*, 2009.

Adam Pauls, John DeNero, and Dan Klein. "Consensus Training for Consensus Decoding in Machine Translation," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 2009.

John DeNero, Adam Pauls, and Dan Klein. "Asynchronous Binarization for Synchronous Grammars," in *Proceedings of the Association of Computational Linguistics, Short Paper Track*, 2009.

Aria Haghighi, John Blitzer, John DeNero, and Dan Klein. "Better Word Alignments with Supervised ITG Models," in *Proceedings of the Association of Computational Linguistics*, 2009.

John DeNero, Mohit Bansal, Adam Pauls, and Dan Klein. "Efficient Parsing for Transducer Grammars," in *Proceedings of the North American Association of Computational Linguistics*, 2009.

John DeNero, Alex Bouchard-Ct, and Dan Klein. "Sampling Alignment Structure under a Bayesian Translation Model," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 2008.

John DeNero and Dan Klein. "The Complexity of Phrase Alignment Models," in *Proceedings of the Association of Computational Linguistics, Short Paper Track*, 2008.

John DeNero and Alexandre Bouchard. "A Hierarchical Dirichlet Process Prior for a Conditional Model of Phrase Alignment," in *the Workshop on Unsupervised Models in NLP, Neural and Information Processing Systems*, 2008.

Aria Haghighi, John DeNero, and Dan Klein. "A* Search via Approximate Factoring," in *Proceedings of American Association of Artificial Intelligence, Nectar Track*, 2007.

John DeNero and Dan Klein. "Tailoring Word Alignments to Syntactic Machine Translation," in *Proceedings of the Association of Computational Linguistics*, 2007.

Aria Haghighi, John DeNero, and Dan Klein. "Approximate Factoring for A* Search," in *Proceedings of the North American Association of Computational Linguistics*, 2007.

John DeNero, Dan Gillick, James Zhang, and Dan Klein. "Why Generative Phrase Models Underperform Surface Heuristics," in *Proceedings of the Workshop on Statistical MT*, 2006.

Invited Talks

Invited Research Talk, Google Research *Host*: Jay Ponte	*2010*
Invited Research Talk, Yahoo! Research *Host*: Patrick Pantel	*2010*
Invited Research Talk, Microsoft Research *Host*: Robert Moore	*2010*
Invited Research Talk, Toyota Technological Institute, Chicago *Host*: Karen Livescu	*2010*
Invited Research Talk, New York University *Host*: Richard Cole	*2010*
Invited Research Talk, Johns Hopkins University *Host*: Mark Dredze	*2010*
Invited Research Talk, University of Maryland *Host*: Jimmy Lin	*2010*
Guest Lecture, UC Berkeley, Applied Natural Language Processing *Host*: Barbara Rosario	*2009*
Invited Research Talk, Yahoo! Research *Host*: Patrick Pantel	*2008*
Invited Research Talk, Information Sciences Institute *Host*: Kevin Knight	*2008*
Invited Research Talk, International Computer Science Institute *Host*: Dilek Hakkani-Tür	*2008*
Invited Research Talk, Microsoft Research Asia *Host*: Mu Li	*2008*
Invited Research Talk, SRI International *Host*: Jing Zheng	*2008*

Experience

Google Research, Mountain View, CA Research intern for Google Translate. Developed a method for combining heterogeneous machine translation models. *Hosts*: Jay Ponte and Shankar Kumar	*Summer 2009*
Information Sciences Institute, Marina del Rey, CA Research intern in machine translation. Developed a forest-based consensus decoding technique for hierarchical machine translation systems. *Hosts*: David Chiang and Kevin Knight	*Summer 2008*
McKinsey & Company, San Francisco, CA Business analyst for general management consulting projects in sourcing, human resources, risk-based pricing, and product strategy.	*2003-2004*

Professional Activities

Conference Reviewer: EMNLP '08, NAACL '09, ACL '09, EMNLP '09, MT Summit '09, NAACL '10, ACL '10, COLING '10, ACL '11, AAAI '11, EMNLP '11, NAACL '12, ACL '12, EMNLP '12, EAAI '12.

Software

The Pacman Projects
A series of five course projects that introduce students to core artificial intelligence concepts in an engaging and open-ended problem setting. Projects culminate in an intercollegiate agent-building competition. The projects have been shared with faculty at over 30 colleges and universities.
http://inst.eecs.berkeley.edu/ cs188/sp09/pacman.html

The Berkeley Aligner
A word alignment package for machine translation that implements agreement-based learning and a syntax-sensitive alignment model, authored and maintained with Percy Liang. The software has been downloaded over 4,000 times.
http://code.google.com/p/berkeleyaligner