

Matthew R. Gormley

CONTACT INFORMATION	Johns Hopkins University 3400 N. Charles St., Hackerman Hall 321 Baltimore, MD 21218	<i>E-mail:</i> mrg@cs.jhu.edu <i>Phone:</i> 412-779-2018 www.cs.jhu.edu/~mrg/
RESEARCH INTERESTS	<i>Natural language processing:</i> grammar induction, dependency parsing, semantic parsing, knowledge base population, coreference resolution, topic modeling, computational semantics. <i>Machine learning:</i> approximate inference, unsupervised learning, decomposition techniques, nonparametrics, global optimization, low-resource learning, approximation-aware learning.	
EDUCATION	Johns Hopkins University <i>Center for Language and Speech Processing</i> <i>Human Language Technology Center of Excellence</i> - Ph.D. Candidate in Computer Science <ul style="list-style-type: none">◦ Advisors: Jason Eisner, Mark Dredze◦ Expected graduation date: Summer 2015 - M.S. in Computer Science, 2009 - 2011 <ul style="list-style-type: none">◦ GPA: 4.0 / 4.0 Carnegie Mellon University <i>School of Computer Science</i> - B.S. in Computer Science <ul style="list-style-type: none">◦ Double major in Cognitive Science◦ Cumulative QPA: 3.86 / 4.0◦ Advisor: Scott Fahlman	<i>2009 - present</i> <i>2003 - 2006</i>
FELLOWSHIPS	- Fred Jelinek Fellowship, <i>Johns Hopkins University</i> - Human Language Technologies Center of Excellence Fellowship, <i>Johns Hopkins University</i>	<i>2014 - 2015</i> <i>2009 - 2014</i>
WORK EXPERIENCE	Google - Semantic Parsing Group <i>Summer Intern (Mentor: Dan Bikel)</i> - Explored methods of confidence estimation for a semantic parser used in spoken language understanding. - Confidence estimation is useful both for downstream applications as well as future research in the area of semantic parsing itself. Johns Hopkins University <i>Ph.D. Candidate in Computer Science (Advisors: Mark Dredze, Jason Eisner)</i> - Introduced novel training methods for parsing: generative training method for globally normalized models; and backpropagation for approximation-aware training. - Studied low-resource approaches to joint semantic and syntactic dependency parsing. - Developed a global optimization method for unsupervised dependency parsing using branch-and-bound and the Dantzig-Wolfe decomposition. - Introduced the shared components topic model, in which topics are products of underlying component distributions. - Established four baselines, requiring various levels of parallel resources, for a new task in cross-lingual coreference resolution. Endeca Technologies - Information Transfer Layer Team <i>Software Engineer</i>	<i>2012</i> <i>2009 - present</i> <i>2007 - 2009</i>

- Developed first release of a file system crawler and expanded a platform for non-technical users to build dynamic web pages backed by the Endeca MDEX Engine.
- Investigated methods of statistical term extraction in order to improve search over a large corpus of documents.

Microsoft - Speech and Natural Language Group

2006

Summer Intern (Mentor: Kunal Mukerjee)

- Researched a new long-span language modeling technique based on Latent Semantic Analysis. Implemented several prototypes and performed experiments to test for word error rate improvements in LVCSR.
- Designed and implemented an API to expose the speech lattice, an acyclic directed graph of language and acoustic models.

Carnegie Mellon - Scone Knowledge Base Research

2005 - 2006

Research Assistant (Advisor: Scott Fahlman)

- Explored algorithms for learning constructions through collocation extraction techniques.
- Built an email classification system to select themes by aggregating salient Scone elements found using a Radical Construction Grammar parser.
- Developed new techniques for extracting symbolic knowledge from feature structures generated by a GLR parser.

TUTORIALS

Matthew R. Gormley and Jason Eisner. [Structured Belief Propagation for NLP](#).

To be presented at *ACL 2015*. Beijing.

Matthew R. Gormley and Jason Eisner. [Structured Belief Propagation for NLP](#).

Presented at *ACL 2014*. Baltimore.

IN PREPARATION

Matthew R. Gormley, Mark Dredze, and Jason Eisner. "Approximation-aware Dependency Parsing by Belief Propagation". In submission.

Matthew R. Gormley, Mo Yu, and Mark Dredze. "Enhancing Relation Extraction with Relation Representation Models". (The first two authors contributed equally). In submission.

Matthew R. Gormley, Mark Dredze, and Jason Eisner. "Parsing with PCFGs, MRFs, CRFs, and Everything in Between". In preparation.

PUBLICATIONS

Nanyun Peng, Francis Ferraro, Mo Yu, Nicholas Andrews, Jay DeYoung, Max Thomas, Matthew R. Gormley, Travis Wolfe, Craig Harman, Benjamin Van Durme, and Mark Dredze (2015). "A Concrete Chinese NLP Pipeline". In: *Proceedings of the NAACL Demonstration Session*.

Mo Yu, Matthew R. Gormley, and Mark Dredze (2015). "[Combining Word Embeddings and Feature Embeddings for Fine-grained Relation Extraction](#)". In: *Proceedings of NAACL*.

Matthew R. Gormley, Margaret Mitchell, Benjamin Van Durme, and Mark Dredze (2014). "[Low-Resource Semantic Role Labeling](#)". In: *Proceedings of ACL*.

Francis Ferraro, Max Thomas, Matthew R. Gormley, Travis Wolfe, Craig Harman, and Benjamin Van Durme (2014). "[Concretely Annotated Corpora](#)". In: *The NIPS 2014 AKBC Workshop*.

Mo Yu, Matthew R. Gormley, and Mark Dredze (2014). "[Factor-based Compositional Embedding Models](#)". In: *The NIPS 2014 Learning Semantics Workshop*.

- Matthew R. Gormley and Jason Eisner (2013). “Nonconvex Global Optimization for Latent-Variable Models”. In: *Proceedings of ACL*.
- Justin Snyder, Rebecca Knowles, Mark Dredze, Matthew Gormley, and Travis Wolfe (2013). “Topic Models and Metadata for Visualizing Text Corpora”. In: *Proceedings of the NAACL HLT Demonstration Session*.
- Matthew R. Gormley, Mark Dredze, Benjamin Van Durme, and Jason Eisner (2012). “Shared Components Topic Models”. In: *Proceedings of NAACL*.
- Courtney Napoles, Matthew Gormley, and Benjamin Van Durme (2012). “Annotated Gigaword”. In: *The NAACL 2012 AKBC-WEKEX Workshop*.
- Spence Green, Nicholas Andrews, Matthew R. Gormley, Mark Dredze, and Christopher D. Manning (2012). “Entity Clustering Across Languages”. In: *Proceedings of NAACL*.
- Matthew R. Gormley, Mark Dredze, Benjamin Van Durme, and Jason Eisner (2011). “Shared Components Topic Models with Application to Selectional Preference”. In: *The NIPS 2011 Learning Semantics Workshop*.
- Spence Green, Nicholas Andrews, Matthew R. Gormley, Mark Dredze, and Christopher D. Manning (2011). *Cross-lingual Coreference Resolution: A New Task for Multilingual Comparable Corpora*. Tech. rep. 6. HLTCOE, Johns Hopkins University.
- Matthew R. Gormley, Adam Gerber, Mary Harper, and Mark Dredze (2010). “Non-Expert Correction of Automatically Generated Relation Annotations”. In: *The NAACL Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk*.

TEACHING EXPERIENCE	JHU CS 476/676 Machine Learning in Complex Domains	<i>Spring 2013</i>
	<i>Teaching Assistant</i>	
	- Designed homework and exams for this new course	
	- Taught three background material lectures	
	JHU CS 103 Fundamentals of Practical Computing	<i>Spring 2012</i>
	<i>Co-Lecturer</i>	
	- Redesigned course material for the lectures, homework, and exams	
	- Taught half the course lectures	
	JHU CS 475 Machine Learning	
	<i>Guest Lecturer</i>	
	- Two guest lectures on Neural Networks and Deep Learning	<i>Fall 2014</i>
	- Two guest lectures on HMMs/CRFs/loopy belief propagation	<i>Spring 2014</i>
	- Guest lectures on Topic Modeling and Approximate Inference	<i>Fall 2011/2012</i>
ADVISEES	Undergraduate students at JHU:	
	- Sharon Li, 2013-2014. Grammar induction by classification.	
	- Justin Synder, 2011. Topic modeling visualization.	
HONORS AND AWARDS	- Certificate in Spanish Culinary Arts,	<i>2007</i>
	<i>El Centro Superior de Hosteleria y Turismo de Valencia</i>	
	- Dean’s List, <i>Carnegie Mellon University</i>	<i>All semesters</i>
	- International Baccalaureate Diploma	<i>2003</i>
PROFESSIONAL SERVICE	Reviewer: ACL, EMNLP, TACL, IJCNLP (Natural Language Processing)	
	Reviewer: ICML, AISTATS, TPAMI, NESCAI (Machine Learning)	

Entertainment Chair, ACL 2014.

Organizing committee member, JHU North American Computational Linguistics Olympiad (NACLO), 2011-2013.

PROFESSIONAL SKILLS	<i>Programming Languages:</i> Java, Python, C++, C, C#, Lisp, Perl, HTML, CSS, L ^A T _E X <i>Natural Languages:</i> English (native), Spanish (conversational), K'iche' (studied)
------------------------	---

LAST UPDATED	April 2015
--------------	------------