Francis Michael Ostrowski Ferraro

CONTACT INFORMATION

3400 N. Charles St.
Hackerman Hall 321
Baltimore, MD 21218

ferraro@cs.jhu.edu

RESEARCH INTERESTS

Language Understanding: event semantics, topic modeling, probabilistic modeling, latent syntactic and semantic induction, multi-modal modeling, cross-disciplinary models
Machine Learning: approximate inference, unsupervised learning, large-scale (parallel and distributed) inference

EDUCATION

The Johns Hopkins University, Baltimore, MD, USA

Ph.D. Student, Department of Computer Science, in progress

- Thesis Proposal: *Probabilistic Computational Event Semantics*
- Adviser: Benjamin Van Durme

M.S.E., Department of Computer Science, December 2013

- Qualifying project 1: "Toward tree substitution grammars with latent annotations" [T2]
- Qualifying project 2: "Evaluating progress in probabilistic programming through topic models" [P8]

The University of Rochester, Rochester, NY, USA

Summa cum Laude, with Highest Distinction

Honors B.S., Department of Computer Science, June 2011

- Honors Thesis: "Toward improving the automated classification of metonymy in text corpora" [T1]

B.S., Department of Mathematics, June 2011

Minor, Linguistics

REVIEWED PUBLICATIONS

P15. Francis Ferraro and Benjamin Van Durme. A Unified Bayesian Model of Scripts, Frames and Language. In *AAAI*, 2016

P14. Drew Reisinger, Rachel Rudinger, Francis Ferraro, Craig Harman, Kyle Rawlins, and Benjamin Van Durme. Semantic proto-roles. *Transactions of the Association for Computational Linguistics*, 3:475–488, 2015

P13. Rachel Rudinger, Pushpendre Rastogi, Francis Ferraro, and Benjamin Van Durme. Script induction as language modeling. In *EMNLP*, 2015

P12. Chandler May, Francis Ferraro, Alan McCree, Jonathan Wintrode, Daniel Garcia-Romero, and Benjamin Van Durme. Topic identification and discovery on text and speech. In *EMNLP*, 2015

P11. Francis Ferraro, Nasrin Mostafazadeh, Ting-Hao Kenneth Huang, Lucy Vanderwende, Jacob Devlin, Michel Galley, and Margaret Mitchell. A survey of current datasets for vision and language research. 2015

P10. 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. A Concrete Chinese NLP Pipeline. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations*, pages 86–90, Denver, Colorado, June 2015. Association for Computational Linguistics

P9. Francis Ferraro, Max Thomas, Matthew R. Gormley, Travis Wolfe, Craig Harman, and Benjamin Van Durme. Concretely Annotated Corpora. In *4th Workshop on Automated Knowledge Base Construction (AKBC)*, 2014

P8. Francis Ferraro, Benjamin Van Durme, and Yanif Ahmad. Evaluating progress in probabilistic programming through topic models. In *Proceedings of the NIPS Workshop on Topic Models*, 2013

P7. Francis Ferraro and Jason Eisner. A virtual manipulative for learning log-linear models. In *Proceedings of the ACL Workshop on Teaching NLP and CL*, 2013

P6. Francis Ferraro, Matt Post, and Benjamin Van Durme. Toward tree substitution grammars with latent annotations. In *The Proceedings of the Workshop on Inducing Linguistic Structure*, 2012

P5. Francis Ferraro, Matt Post, and Benjamin Van Durme. Judging grammaticality with count-induced tree substitution grammars. In *The Proceedings of the 7th Workshop on Innovative Use of NLP for Building Educational Applications*, 2012

P4. Francis Ferraro, Emily Kawaler, and Kenji Suzuki. A spinning tangent based cad system for detection of flat lesions in ct colongraphy. In *IEEE International Symposium on Biomedical Imaging (ISBI)*, 2011

P3. E. Tzoukermann, J. Neumann, J. Kosecka, C. Fermuller, I. Perera, F. Ferraro, B. Sapp, R. Chaudhry, and G. Singh. Language models for semantic extraction and filtering in video action recognition. In *AAAI Workshop on Language-Action Tools for Cognitive Artificial Agents*, 2011

P2. Benjamin Sapp, Rizwan Chaudhry, Xiaodong Yu, Gautam Singh, Ian Perera, Francis Ferraro, Evelyne Tzoukermann, Jana Kosecka, and Jan Neumann. Recognizing manipulation actions in arts and crafts shows using

domain specific visual and textual cues. In *The 3rd International Workshop on Video Event Categorization, Tagging and Retrieval for Real-World Applications (VECTaR2011)*, 2011

P1. K. Suzuki, I. Sheu, E. Kawaler, F. Ferraro, D.C. Rockey, and A.H. Dachman. Computer-aided detection (cade) of flat lesions in ct colonography (ctc) by means of a spinning-tangent technique. In *Radiological Society of North America (RSNA)*, 2010

Non-refereed Technical Reports

N2. David Etter, Francis Ferraro, Ryan Cotterell, Olivia Buzek, and Benjamin Van Durme. Nerit:named entity recognition for informal text. Technical Report 11, Human Language Technology Center of Excellence, Johns Hopkins University, 2013

N1. Frank Ferraro, Garrett Hall, and Andrew Wood. Refutation of aslam's proof that NP=P. Technical report, University of Rochester, 2009

Theses, Significant School Papers

T2. Francis Ferraro. Toward tree substitution grammars with latent annotations, 2013. With Matt Post and Benjamin Van Durme

T1. Francis Ferraro. Toward improving the automated classification of metonymy in text corpora. http://hdl.handle.net/1802/14985, 2011. Undergraduate Honors Thesis

Invited Talks

- "Document-level Modeling of Discourse." Applied Physics Laboratory, Johns Hopkins University. February 2nd, 2016.

Awards

- National Science Foundation Graduate Research Fellowship; Awarded 2011, On Tenure 2013 – present
- Outstanding Teaching Award, Department of Computer Science, Johns Hopkins University; 2013
- National Defense Science and Engineering Graduate Fellowship (NDSEG); 2011 (Declined)
- Finalist, Computing Research Association Outstanding Undergraduate Researchers; 2011
- Academic All-American, CoSIDA (College Sports Information Directors of America); 2011

- Robert L. Wells Prize, University of Rochester; 2011
- Provost's Circle Scholar, University of Rochester; 2010

Professional Experience

Microsoft Research Redmond, WA, USA

Research Intern **Summer 2015**

- Pursued multimodal research, establishing competitive baselines and curating a tiered dataset
- Advised by Margaret Mitchell, Ph.D.

Johns Hopkins University Baltimore, MD, USA

Graduate researcher, Frederick Jelinek Memorial Workshop **Summer 2014**

- Explored a computational approach to Dowty (1991)s proto-roles theory

Graduate researcher, SCALE Workshop **Summer 2013**

- Developed probabilistic models for narrative scenario induction
- Helped implement a user interface to facilitate downstream exploration of these models
- Wrote automated data extraction pipelines used as a primary processing step by nearly all participants ($\sim$ 35)
- Prepared and delivered multiple presentations to a large audience

Graduate research assistant, Deep Exploration and Filtering of Text (DARPA DEFT) **Spring 2013**

- Developed automatic way to evaluate efficiency of probabilistic programming languages/frameworks [P8]

Graduate researcher, SCALE Workshop **Summer 2012**

- Experimented with syntactic approaches to named entity recognition

Graduate research assistant, Human Language Technology Center of Excellence (HLTCOE) **Spring 2012**

- Developed a novel approach to automatically learn extended syntactic formalisms [P5, P6, T2]
- Delivered multiple presentations

Undergraduate researcher, CLSP Workshop **Summer 2010**

- Extracted action-related information from text for use in an automatic video-annotation system [P2, P3]
- Adapted a named-entity recognizer for use as a part-of-speech tagger and basic parser
- Obtained domain-specific action verbs and objects in a scalable and robust way
- Advised by Evelyne Tzoukermann, Ph.D.

University of Rochester Rochester, NY, USA

Research Collaborator **2010-2011**

- Studied ways to resolve underspecified referents in corpus-based text, with a focus on anaphora/pronoun co-reference and metonymy [T1]

University of Chicago Chicago, IL, USA

NSF REU Student Researcher, The University of Chicago Cancer Research Center **2010-2011**

- Was principal in developing a new detection scheme for colonic flat lesions based on mathematical morphology [P1, P4]
- Advised by Kenji Suzuki, Ph.D.

Advising

- William Povell, 2014 - 2016. Detecting and classifying attributes of answerable questions. (*High school*; Third place in regional Junior Science and Humanities Symposium)

Teaching Experience

Johns Hopkins University, Baltimore, MD, USA

Teaching Assistant, Department of Computer Science **Fall 2011, 2012**

Natural Language Processing (600.465: Graduate/undergraduate)

- Developed an interactive tutorial on log-linear models with Jason Eisner `http://cs.jhu.edu/~jason/tutorials/loglin/`
- Cowrote an introductory assignment/walkthrough tutorial with Jason Eisner that exposes students to (weighted) finite-state programming `http://cs.jhu.edu/~jason/465/hw-ofst/hw-ofst.pdf`
- Held weekly review sessions, graded assignments, and updated course materials

Theory of Computation (600.471: Graduate/undergraduate)

- Held office hours, provided assistance and graded assignments

University of Rochester, Rochester, NY, USA

Teaching Assistant, Department of Computer Science **Fall 2008 - Spring 2011**

Artificial Intelligence (CSC 242: Undergraduate)
Computation and Formal Systems (CSC 173: Undergraduate)
Computer Models and Limitations (CSC 280: Undergraduate)
Natural Language Understanding (CSC 247/447: Undergraduate/graduate)

- Led workshops, held office hours, provided assistance, graded assignments and updated course materials/software

Service

Reviewing

- Primary: NAACL 2016, ICML 2015, EMNLP 2015 (semantics)
- Secondary: EMNLP 2012 (semantics), ACL 2014 (EVENTS workshop), EMNLP 2014 (semantics), WWW 2015, EMNLP 2015

Departmental Service

- CLSP Admissions Committee (2013-2016)
- Undergraduate and Graduate Teaching/Curriculum Committee (2013-2014)

Conference/Meeting Organization

- Mid-Atlantic Student Colloquium on Speech, Language and Learning (2012)
  https://sites.google.com/site/studentcolloquiumsll/2012