Jason Smith
2905 N. Charles St. Apt. 212
Baltimore, MD 21218
(504)432-6520
jrs026@gmail.com

Research Interests

Machine translation, large-scale mining of annotated data, dynamic programming, finite-state automata, unsupervised and semi-supervised machine learning.

Education

Louisiana Tech University

- *Master of Science*, Computer Science, GPA: 4.0, August 2006
- *Bachelor of Science*, Computer Science, GPA: 4.0, May 2005
- Louisiana Tech President's Scholarship
- National Merit Scholarship

Johns Hopkins University

- Ph.D. student, September 2006-Present

Professional Experience

Louisiana Tech University

- Teaching Assistant (Fall 2005)
- Research Assistant (Winter 2006 - Summer 2006)

Worked as a teaching assistant for Advanced Data Structures and Algorithms
Developed and implemented algorithms for specializations of finite automata.

Johns Hopkins University

- Teaching Assistant (Spring 2007 - Fall 2008, Spring 2010, Spring 2011)
- CLSP Workshop Participant (Summer 2007)
- Research Assistant on GALE Program (Summer 2008 - Spring 2009)

Worked as a teaching assistant for Natural Language Processing and Declarative Methods
Assisted in developing a toolkit for coreference resolution and worked on feature extraction from Wikipedia
Worked on system combination for machine translation for the GALE program (with Damianos Karakos and Sanjeev Khudanpur)

Microsoft Research

- Intern (Summer 2009)

Worked with Chris Quirk and Kristina Toutanova on a system for extracting parallel data from Wikipedia

Google
- Intern (Summer 2011)

Worked with Peng Xu and Jakob Uszkoreit on a system for extracting parallel data from mined Web documents

Teaching

Natural Language Processing (with Jason Eisner), Fall 2007, 2008
Teaching Assistant

Declarative Methods (with Jason Eisner), Spring 2007, 2008, 2010, 2011
Teaching Assistant

References

Adam Lopez - Advisor (alopez@cs.jhu.edu)
Chris Callison-Burch - Advisor (ccb@cs.jhu.edu)
Jason Eisner - Research Supervisor (jason@cs.jhu.edu)

Publications

Conference Publications

J. Smith, J. Eisner, *Unsupervised Learning on an Approximate Corpus.* Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies

A. V. Rosti, X. He, D. Karakos, G. Leusch, Y. Cao, M. Freitag, S. Matsoukas, H. Ney, J. Smith, B. Zhang, *Review of Hypothesis Alignment Algorithms for MT System Combination via Confusion Network Decoding* Proceedings of the Seventh Workshop on Statistical Machine Translation 2012

J. Smith, C. Quirk, K. Toutanova, *Extracting Parallel Sentences from Comparable Corpora using Document Level Alignment.* Proc. of the 11th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 2010.

D. Karakos, J. Smith, S. Khudanpur, *Hypothesis Ranking and Two-Pass Approaches for Machine Translation System Combination.* Proc. of the 35th International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2010.

M. Dreyer, J. Smith, J. Eisner, *Latent-Variable Modeling of String Transductions With Finite-State Methods.* Proc. of the Conference on Empirical Methods in Natural Language Processing (EMNLP), October 2008, Honolulu, Hawaii.

Y. Versley, S. Ponzetto, M. Poesio, V. Eidelman, A. Jern, J. Smith, X. Yang, A.Moschitti, *BART: A Modular Toolkit for Coreference Resolution.* Companion Volume of the Proc. of the 46th Annual Meeting of the Association for Computational Linguistics, June 16-18, 2008.

Y. Versley, S. Ponzetto, M. Poesio, V. Eidelman, A. Jern, J. Smith, X. Yang, A.Moschitti, *BART: A Modular Toolkit for Coreference Resolution.* Proc. of the 6th International Conference on Language Resources and Evaluation, May 28-30, 2008, Marrakech, Morocco.

C. Campeanu, A. Paun, J. R. Smith, *Tight Bounds for the State Complexity of Deterministic Cover Automata*, Proc. of Descriptional Complexity of Formal Systems, 8th Workshop, June 21-23, 2006, Las Cruces, New Mexico, USA.

C. Campeanu, A. Paun, J.R. Smith, *An Incremental Algorithm for Minimal Deterministic Finite Cover Automata*, Proc. of Tenth International Conference on Implementation and Application of Automata (CIAA05), June 27-29, 2005, Sophia Antipolis, France.

Journal Publications

C. Campeanu, A. Paun, J.R. Smith, *Incremental construction of minimal deterministic finite cover automata.* Theoretical Computer Science, 363(2), 2006, 135-148.

Book Chapters

A. V. Rosti, E. Matusov, J. Smith, N. Ayan, J. Eisner, D. Karakos, S. Khudanpur, G. Leusch, Z. Li, S. Matsoukas, H. Ney, R. Schwartz, B. Zhang, J. Zheng, *Confusion Network Decoding for MT System Combination.* Handbook of Natural Language Processing and Machine Translation 2011

Service

Conference reviewing: ACL 2010, COLING 2010, EMNLP 2011, EMNLP 2012, NAACL 2013