Dennis N. Mehay (Curriculum Vitae)

5 Samuel J. Dr.
Merrimack, NH, USA 03054

Telephone: +1.614.256.1409
Email (professional): dmehay@bbn.com
Email (personal): dnmehay@gmail...
Citizenship: American

Research

I build software systems that do useful things with human language, ranging from information extraction (tagging, parsing and semantic event detection), to translation and paraphrase. Accordingly, my training and experience span many different disciplines related to natural language processing, most notably: machine learning (including "Big Data" variants thereof), algorithms, linguistic theory, psycholinguistics, information retrieval and software development.

Experience

- **Scientist**
  December 2012–present. Raytheon BBN Technolgies (Cambridge, MA)
  - Currently building translation systems for the BOLT and MADCAT-K projects.
  - Recent contributions:
    - Word-sense translation error detection with targeted user feedback (sole developer).
    - Predicting cross-lingual responses in bilingual conversations (co-developer).
    - Topic modeling for improved translation in conversations (co-developer).
    - Web application for Korean↔English document translation and post-editing (co-developer).
  - Previously worked on the FUSE project, which used NLP to uncover emerging trends in science that are buried in the prose and peer citations of scientific journal articles.

- **PhD Student and Researcher**
  June 2005–December 2012. The Ohio State University (Columbus, OH)
  - Developed novel uses of recursive syntax for phrase-based machine translation (dissertation).
  - Authored statistical tagging components for the OpenCCG natural language processing platform.
  - Used lexical and grammatical paraphrase (OpenCCG) to improve machine translation.
  - Researched unsupervised computational models of human sentence processing using neural networks.

- **Programming Consultant**
  Spring 2011. The Zerhouni Group [later i^3 Analytics, acquired by TEMIS] (Columbia, MD)
  - Harvested open-source biopharma data, automatically detecting names, affiliations and citations (Python).
  - Prototyped custom visualizations of interrelated players in the biopharma space (Java, Google Charts).

- **Programmer and Database Developer**
  June 2002–August 2003. Ajilon Professional, LLC (Louisville, KY)
  - Developed software to track enterprise-wide PC upgrades for a Fortune 100 client with 19,000 employees.

- **Computer Science Instructor**
  August 2001–June 2002. Sullivan University (Louisville, KY)
  - Developed and taught courses in computing basics, and in programming and data structures (Java).

Education

- PhD Computational Linguistics, The Ohio State University, 2012 (M.A., 2011).
- M.A.I. (Artificial Intelligence) Katholieke Universiteit Leuven, Belgium, 2004 (*magna cum laude*).
- M.A. Philosophy, Indiana University, Bloomington, 2001.
- B.A. French & Philosophy, Indiana University, Bloomington, 1999.

Skills and Portfolio

HIGHLIGHTS

- 15 years of software development experience in a variety of mainstream programming languages and across multiple programming paradigms (object-oriented, functional, logic and procedural).
- Experienced in machine learning, NLP, linguistics (from speech to semantics) and machine translation.
- Facility working with large data sets (large-scale processing in Hadoop/SGE, DBs, IR, etc.).
- Familiar with software design best practices (modular, expressive, documented code with tests), databases, major mark-up standards, writing and interfacing with RESTful APIs, etc.
- Proficiency in multiple natural languages.
- Open-source Software

COMPUTER LANGUAGES

- Have developed in: Java, Python, Haskell, C, C++, JavaScript, Scheme, Perl, among others.
- Know the basics of: Kivy (Python for Android/iPhone), R.
- Query and mark-up languages: SQL, XML, HTML, CSS, LaTeX, etc.

COMPUTING AND SOFTWARE DEVELOPMENT SKILLS

- Object-oriented, Functional and Logic (PROLOG) Programming
- Databases (sqlite3, MySQL, MongoDB, BerkeleyDB)
- Unit and Property-based (QuickCheck, JUnit "Theories") Testing
- *nix Command-line Text Processing (sed, grep, awk, etc.) and shell scripting
- Programming Language and Compiler Theory

NLP AND RELATED SKILLS

- Machine Learning
- Statistical Machine Translation
- Structured Prediction: Parsing, Tagging, Computational Semantics.
- Information Retrieval (Indri, Solr)
- Neural Networks for Language Processing

NATURAL LANGUAGES

- English (native)
- French (high proficiency, spoken and written)
- Dutch (basic conversation and reading)
- German (basic reading)
- Turkish (beginning conversation and reading)

OTHER PROFESSIONAL INTERESTS

- Decentralized, peer-to-peer ("blockchain") technologies
- Crowdsourcing and human computation

Peer-reviewed Publications

2015 Sankaranarayanan Ananthakrishnan, **Dennis Mehay**, Sanjika Hewavitharana, Rohit Kumar, Matt Roy, Enoch Kan. Lightly-Supervised Word Sense Translation Error Detection and Resolution in an Interactive Conversational Spoken Language Translation System. *Machine Translation*, Kluwer Academic Publishers, March 2015, 29(1), pp. 25–47.

2014 Sanjika Hewavitharana, **Dennis Mehay**, Sankaranarayanan Ananthakrishnan, Rohit Kumar and John Makhoul. Anticipatory Translation Model Adaptation for Bilingual Conversations. In Proceedings of the IWSLT-2014, Lake Tahoe, CA, USA.

Dennis N. Mehay, Sankaranarayanan Ananthakrishnan and Sanjika Hewavitharana. Lightly-Supervised Word Sense Translation Error Detection for an Interactive Conversational Spoken Language Translation System. In Proceedings of the EACL-2014, Gothenburg, Sweden.

2013 Sankaranarayanan Ananthakrishnan, Wei Chen, Rohit Kumar and **Dennis Mehay**. Source-Error Aware Phrase-Based Decoding for Robust Conversational Spoken Language Translation. In Proceedings of IWSLT-2013, Heidelberg, Germany.

Sanjika Hewavitharana, **Dennis Mehay**, Sankaranarayanan Ananthakrishnan and Prem Natarajan. Incremental Topic-Based Translation Model Adaptation for Conversational Spoken Language Translation. Proceedings of the ACL-2013, Sophia, Bulgaria.

2012 **Dennis N. Mehay** and Michael White. Shallow and Deep Paraphrasing for Improved Machine Translation Parameter Optimization. In Proceedings of AMTA-2012, San Diego, CA.

Dennis N. Mehay. Bean Soup Translation: Flexible, Linguistically-motivated Syntax for Machine Translation. PhD Dissertation. The Ohio State University, Columbus, OH.

Dennis N. Mehay and Chris Brew. CCG Syntactic Reordering Models for Phrase-based Machine Translation. In Proceedings of WMT-2012, Montreal, Canada

Dennis N. Mehay and Jon Dehdari. The Ohio State University Farsi-English System Submission to OpenMT 2012. In Proceedings of the NIST 2012 Open Machine Translation Workshop, Washington D.C.

2011 Stephen A. Boxwell, Chris Brew, Jason Baldridge, **Dennis Mehay** and Sujith Ravi. Semantic Role Labeling Without Treebanks? In Proceedings of the IJCNLP, Chiang Mai, Thailand.

2010 Stephen A. Boxwell, **Dennis N. Mehay**, and Chris Brew. What a Parser can Learn from a Semantic Role Labeler and Vice-Versa. In Proceedings of EMNLP-2010. Cambridge, MA.

2009 Judita Preiss, Jon Dehdari, Josh King, and **Dennis Mehay**. Refining the most frequent sense baseline. In Proceedings of SEW-2009 (at NAACL HLT), Boulder, Colorado.

Stephen A. Boxwell, **Dennis N. Mehay** and Chris Brew. Brutus: A Semantic Role Labeling System Incorporating CCG, CFG, and Dependency Features. In Proceedings of ACL-IJCNLP-09, Singapore.

Simon Dennis, **Dennis Mehay** and Srikar Yekollu. Predicting When Words May Appear: A Connectionist Model of Sentence Processing. In Proceedings of CogSci-09, Amsterdam, The Netherlands.

2008 Dominic Espinosa, Michael White and **Dennis Mehay**. Hypertagging: Supertagging for Surface Realization with CCG. In Proceedings of ACL-08:HLT. Columbus, OH.

2007   **Dennis N. Mehay** and Chris Brew. Bleuâtre: Flattening Syntactic Dependencies for MT Evaluation. In Proceedings of TMI-07, Skövde, Sweden.

2006   **Dennis Mehay**, Rik De Busser and Marie-Francine Moens. Labelling Generic Semantic Roles. In Proceedings of the IWCS-6, Tilburg, The Netherlands.

Invited Presentations

2009   Real-word Spelling Correction for CALL (with "DJ" Hovermale). The 6th Annual Midwest Computational Linguistics Colloquium (MCLC-6). Bloomington, IN.

2008   Tutorialist (with Dr. Philipp Koehn) at The Eighth Conference of the Association for Machine Translation in the Americas (AMTA-2008), Waikiki, Honolulu, Hawai'i.
Title: Statistical Machine Translation: Theory and Practice.

Honors & Awards

University Fellowship (one year), The Ohio State University, 2005.

Honorable Mention, NSF Graduate Research Fellowship, 2005.

Foreign Language Area Studies Fellowship in Turkish, The Ohio State University, 2006–2007.

Tuition Reimbursement Scholarship (from NAACL) for attending the Linguistic Society of America's Summer Institute, Stanford University, Summer 2007.

Targeted Investment Endowment (TIE) Fellowship, Department of Linguistics, The Ohio State University, Autumn 2007, Spring 2008.

Nominated for the OSU Presidential Fellowship, 2010 and 2011.