

# Curriculum Vitae

Bridget T. McInnes  
Department of Computer Science and Engineering  
University of Minnesota Twin Cities  
Minneapolis, MN, 55455

bthomson@cs.umn.edu

## Current Status

I am completing a PhD in Computer Science at the University of Minnesota Twin Cities under the advisement of Dr. Ted Pedersen and Dr. John Carlis. I am currently conducting my research at the National Library of Medicine under the advisement of Dr. Lan Aronson.

## Research Interests

My researching interests are in Natural Language Processing and Computational Linguistics in the biomedical and clinical domain. My interests lie in the automatic disambiguation of ambiguous words in biomedical and clinical text as well as the automatic medical coding and indexing of these documents. The long term goal of my research is to improve the automatic coding and indexing of medical text and clinical records in such a way that ambiguity is no longer an issue and coding and indexing can become an automatic/semi-automatic procedure.

## Education

- **University of Minnesota Duluth**  
M.S., Computer Science, December 2004  
Thesis Title: *Extending the Log-Likelihood Ratio to Improve Collocation Identification*  
Advisor: Ted Pedersen
- **University of Minnesota Duluth**  
B.S., Computer Science

## Professional Experience

- Student Research Scientist  
Lister Hill National Center for Biomedical Communications, Bethesda, Maryland  
Mar. 08 - present
  - Conduct bio-NLP related research focusing on automatically determining the sense of a word that has multiple sense in biomedical text.

- Research Assistant  
Computer Science and Engineering Department, University of Minnesota Twin Cities  
Sep. 06 - Mar. 08
  - Conduct bio-NLP related research including hypothesis definition, experimental design and implementation, data analysis and publication writing.
- Teaching Assistant  
Computer Science Department, University of Minnesota Duluth  
Computer Science and Engineering Department, University of Minnesota Twin Cities  
Sep. 02 - May. 06
  - Conducted programming labs and recitations for computer science classes including Software Development, Software Analysis and Design, Natural Language Processing, Data Structures and Algorithms, Data Modeling and Discrete Math.
- Research and Development Intern  
Thomson Legal and Regulatory, Eagan, MN.  
May 05 - Aug 05  
May 06 - Aug 06
  - Research information extraction and semantic parsing of legal news briefs.
- Student Research Scientist  
Mayo Clinic, Rochester, MN  
Jun. 03 - Sep. 03  
Jun. 04 - Sep. 04
  - Researched and implemented a spelling correction tool to automatically correct misspelled words in clinical notes
  - Researched term identification in clinical notes

## **Honors and Awards**

- National Library of Medicine's Student Research Participation Program Fellowship, 2008  
Awarded by the National Library of Medicine, Bethesda, Maryland
- Graduate Assistance in Areas of National Need (GAANN) Fellowship, 2006-2009  
Awarded by the Computer Science and Engineering Department, University of Minnesota Twin Cities
- Most Outstanding Teaching Assistant , 2004  
Awarded by the Graduate School, University of Minnesota Duluth
- National Science Foundation scholarship, 2001-2002  
Awarded by the Computer Science Department, University of Minnesota Duluth

## Refereed Conference Publications

- McInnes, B. (2008) An Unsupervised Vector Approach to Biomedical Term Disambiguation: Integrating UMLS and Medline. To appear in the Association for Computational Linguistics Student Research Workshop.
- McInnes, B. & Pedersen, T. & Carlis, J. (2007) Using UMLS Concept Unique Identifiers (CUIs) for Word Sense Disambiguation in the Biomedical Domain. In *Proceedings of the Annual Symposium of the American Medical Informatics Association*. pages 533-37, Nov. 2007, Chicago, IL. [acceptance rate 45%]

## Refereed Workshops Publications

- McInnes, B. & Pedersen, T. & Pakhomov, S. (2007) Determining the Syntactic Structure of Medical Terms in Clinical Notes. In *Proceedings of the ACL Workshop BioNLP 2007: Biological, translational and clinical language processing*. pp. 9-16, Prague, Czech Republic. [acceptance rate 29%]

## Workshop Publications

- Schilder F. & McInnes B. (2006) TLR at DUC 2006: Approximate tree similarity and a new evaluation regime. In *Proceedings of the Document Understanding Conference (DUC)*, New York, NY, USA.
- Schilder F. & McInnes B. (2006) Word and tree-based similarities for textual entailment. In *Proceedings of the Second PASCAL Challenges Workshop on Recognizing Textual Entailment (RTE-2)*, Venice, Italy.
- Schilder, F. & McCullom, A. & Zhou, A. & McInnes, B. (2005) TLR at DUC: Tree Similarity. In *Proceedings of the Document Understanding Conference (DUC)*, Vancouver, Canada.
- McInnes, B. & Pedersen, T. (2003) The Duluth Word Alignment System. In *Proceedings of the NAACL Workshop on Building and Using Parallel Texts: Data Driven Machine Translation and Beyond*. pp. 40-43. Edmonton, Canada.

## Published Abstracts

- McInnes, B. & Pedersen, T. & Carlis, J. (2007) Using Domain Specific Information for Word Sense Disambiguation. In *The Grace Hopper Conference for Women in Computing*, October 2007, Orlando, Florida. (poster presentation)
- Pakhomov S. & McInnes, B. (2005) Resolving Structural Ambiguity of Medical Terms with Statistical Model Fitting. In *Proceedings of the Linguistic Society of America (LSA)*. Oakland, CA. (panel presentation)

- McInnes, B. & Pakhomov, S. & Pedersen T. & Chute, C. (2004) Incorporating Bigram Statistics to Spelling Correction Tools. In *Medinfo 2004: Proceedings of the 11th World Congress on Medical Informatics*. pp. 182. San Francisco, CA. IOS Press. (poster presentation) [acceptance rate: 73%]

### **Workshop Reviewing**

- ACL 2005 Workshop on Building and Using Parallel Texts: Data Driven MT and Beyond, June 2005, Ann Arbor, MI. [3 papers]

### **Participation in Shared Tasks and Comparative Evaluations**

- Medical NLP Challenge - Classifying Clinical Free Text Using Natural Language Processing (2007): Participated with Ted Pedersen and Serguei Pakhomov. Organized by the Computational Medicine Center.

### **Professional Society Memberships**

- Association for Computational Linguistics
- American Medical and Informatic Association
- Association for Computational Machinery