

Education

- 2007–present
(expected
8/2012) **PhD Candidate, Computer Science and Cognitive Science,**
University of Colorado Boulder, Boulder, Colorado.
Thesis Topic: Automatic Characterization of Tutorial Dialogues using Semantically-Informed Dialogue Act Representations
Committee: Wayne Ward (co-chair), Martha Palmer (co-chair), James Martin, Tamara Sumner, Sarel Van Vuuren, Barbara Fox
Areas of Focus: Natural Language Processing, Educational Technologies
- 2007–2010 **MS, Computer Science,**
University of Colorado Boulder, Boulder, Colorado.
- 1996–2000 **BS, Mathematical and Computer Sciences, Computer Science Specialty**
BS, Engineering, Electrical Specialty,
Colorado School of Mines, Golden, Colorado.

Professional Experience

- 2012–present **University of Colorado Boulder, Research Assistant,**
Center for Computational Language and Education Research, Boulder, Colorado.
SHARP: Strategic Health IT Advanced Research Projects - Extracting structured information from clinical notes using natural language processing - Supervised by Martha Palmer and Wayne Ward
- Summer 2011 **Microsoft Corporation, Research Intern,**
Microsoft Research - Machine Learning Group, Redmond, Washington.
Developed a machine learning and natural language processing based system for automatic generation and evaluation of educational questions from text. - Mentored by Sumit Basu and Lucy Vanderwende
- 2008–2012 **University of Colorado Boulder, Research Assistant,**
Center for Computational Language and Education Research, Boulder, Colorado.
Improving Science Learning through Dialogs with Virtual Tutors - Supervised by Wayne Ward and Sarel Van Vuuren
- 2006–2007 **Volunteers in Asia (VIA), Volunteer English Teacher,**
Pesantren Sunan Pandanaran, Yogyakarta, Indonesia.
Taught conversational English to middle and high school students
- 2005–2006 **Intel Corporation (acquired from HP), VLSI CAD Tools Software Engineer,**
Fort Collins Design Center, Fort Collins, CO.
Developed APIs, algorithms, and software tools to support VLSI design of high-end server microprocessors.
- 1999–2005 **Hewlett Packard Company, VLSI CAD Tools Software Engineer / Verification Engineer ,**
Microprocessor Technology Lab, Fort Collins, CO.
Developed software infrastructure to support VLSI design and architectural verification of high-end server microprocessors.

Research Interests

My research interests center around natural language processing, information retrieval, and machine learning, and I am motivated by the promise these areas hold for improving human computer interaction and for creating more personalized mechanisms for accessing and understanding information. Currently, I am investigating techniques for characterizing the quality of tutorial dialogue interactions and for automatic generation of questions.

Publications

Lee Becker, Sumit Basu, and Lucy Vanderwende. Mind the Gap: Learning to Choose Gaps for Question Generation. In *The 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Montreal, Canada, June 2012.

Lee Becker, Martha Palmer, Sarel van Vuuren, and Wayne Ward. Learning to Tutor Like a Tutor: Ranking Questions in Context. In *The 11th International Conference on Intelligent Tutoring Systems*, Crete, Greece, June 2012.

Lee Becker, Martha Palmer, Sarel van Vuuren, and Wayne Ward. Question Ranking and Selection in Tutorial Dialogues. In *The 7th Workshop on Innovative Use of NLP for Educational Applications*, Montreal, Canada, June 2012.

Caleb Phillips, Lee Becker, and Elizabeth Bradley. Strange Beta: An Assistance System for Indoor Rock Climbing Route Setting. *Chaos*, 22(1), March 2012.

Wayne Ward, Ron Cole, Daniel Bolanos, Cindy Buchenroth-Martin, Edward Svirsky, Sarel van Vuuren, Tim Weston, Jing Zheng, and Lee Becker. My Science Tutor: A conversational multimedia virtual tutor(submitted for review). *Journal of Educational Psychology*, 2012.

Lee Becker, Martha Palmer, Sarel van Vuuren, and Wayne Ward. Evaluating questions in context. In *AAAI 2011 Fall Symposium on Question Generation*, Arlington, VA USA, November 2011.

Caleb Phillips, Lee Becker, and Elizabeth Bradley. Strange Beta: An assistance system for indoor rock climbing route setting using chaotic variations and machine learning. Technical Report CU-CS-1088-11, University of Colorado Boulder, Boulder, CO USA, October 2011.

Wayne Ward, Ron Cole, Daniel Bolanos, Cindy Buchenroth-Martin, Edward Svirsky, Sarel van Vuuren, Tim Weston, Jing Zheng, and Lee Becker. My Science Tutor: A conversational multi-media virtual tutor for elementary school science. *ACM TSLP: Special Issue on Speech and Language Processing of Children's Speech for Child-machine Interaction Applications*, 7(1), August 2011.

Lee Becker, Wayne Ward, Sarel van Vuuren, and Martha Palmer. DISCUSS: A dialogue move taxonomy layered over semantic representations. In *IWCS 2011: The 9th International Conference on Computational Semantics*, Oxford, England, January 2011.

Lee Becker, Rodney D. Nielsen, Ifeyinwa Okoye, Tamara Sumner, and Wayne Ward. What's next? Target concept identification. In *Proceedings of the 3rd Workshop on Question Generation*, Pittsburgh, PA USA, June 2010.

Lee Becker, Wayne Ward, and Sarel van Vuuren. DISCUSS: Enabling detailed characterization of tutorial interactions through dialogue annotation. In *EDM 2011: Proceedings of the 3rd International Conference on Educational Data Mining.*, Pittsburgh, PA USA, June 2010.

Lee Becker, Rodney D. Nielsen, and Wayne H. Ward. What a pilot study says about running a question generation challenge. In *The 2nd Workshop on Question Generation*, Brighton, England, July 2009.

Rodney D. Nielsen, Lee Becker, and Wayne H. Ward. TAC 2008 CLEAR RTE system report: Facet-based entailment. In *Proceedings of the Text Analysis Conference*, Gaithersburg, Maryland, November 2008. National Institute of Standards and Technology.

Patents

Adaptively Presenting Content Based on User Knowledge, Sumit Basu, Lucy Vanderwende, and Lee Becker, Microsoft Corporation, Patent Pending, December 2012

Random code generation using genetic algorithms, Zachary Smith, Lee Becker, and David Heckman, Hewlett-Packard Company, US Patent 6918098, July 2005

Relevant Skills

Specialties: Natural Language Processing, Information Retrieval, Machine Learning, Data Mining

Natural Language Processing frameworks: ClearTK, UIMA, UIMAFit

Information Retrieval frameworks: Lucene

Programming Languages: Java, C#, C/C++, Python, Scala, Javascript, Matlab

Web programming frameworks: Django, Google Web Toolkit

Crowdsourcing: Amazon Mechanical Turk API

Open Source

Co-owner ClearTK - Toolkit for developing statistical natural language processing components
Contributor cTAKES - Clinical Text Analysis and Knowledge Extraction System

Guest Lectures

- 2010 Machine Learning - Computational Lexical Semantics, University of Colorado Boulder - Prof. Martha Palmer, 2010
- 2010 Semantic Role Labeling - Computational Lexical Semantics, University of Colorado Boulder - Prof. Martha Palmer, 2010
- 2010 Computational Linguistics - Colorado Math Circle
- 2002 VLSI Design - Microcontroller Architecture and Interfacing, Colorado School of Mines - Dr. William Hoff

Professional Activities

- Member Association for Computational Linguistics(ACL), 2009-present
- Member Association for the Advancement of Artificial Intelligence (AAAI), 2011-present
- Member International Artificial Intelligence in Education Society, 2009-present
- Reviewer FLAIRS 2012: The 25th Florida Artificial Intelligence Researchers Society Conference (Applied NLP Track), May 23-25, 2012, Marco Island, Florida.
- Reviewer Discourse and Dialogue: Special Issue on Question Generation, 2011
- Reviewer ACL 2011: The 49th Conference of the Association of Computational Linguistics in the area of Natural Language Processing Applications, June 19-24, 2011, Portland, OR.
- Reviewer FLAIRS 2011: The 24th Florida Artificial Intelligence Researchers Society Conference (NLP Track), May 18-20, 2011, Palm Beach, Florida.

775 37th Street • Boulder, CO 80303

☎ 720.382.8688 • ✉ lee.becker@colorado.edu • <http://www.leebecker.com>

- Program Committee Third Workshop on Question Generation, 2010, Pittsburgh, Pennsylvania.
- Program Committee The Fifth International Workshop on Question Generation, 2012, Intelligent Tutoring Systems, Chania, Crete, Greece.

Service: University of Colorado

- Student Representative Internal self-study committee, University of Colorado Institute of Cognitive Science, Fall 2011 - Responsible for interviewing stakeholders and writing reports detailing the current state and future plans for the institute.
- Student Representative Executive Committee, University of Colorado Institute of Cognitive Science, Fall 2010-Present
- Representative Student grants committee, University of Colorado Institute of Cognitive Science, Fall 2007 - present
- Social Chair Computer Science Graduate Student Social Committee, 2008-2009
- Website Design Meeting of the North American Association for Computational Linguistics Human Language Technologies (NAACL-HLT), 2009
- Website Design The Third Workshop on Issues in Teaching Computational Linguistics (TeachCL), 2008

Honors

- Honorable Mention National Science Foundation Graduate Research Fellowship Program
- Outstanding Graduate Award Colorado School of Mines Department of Mathematical and Computer Science, December 2000
- Scholarship Achievement Rewards for College Scientists (ARCS), Virginia Ramo Memorial Scholarship, 1999, 2000

References

- Dr. Martha Palmer Professor Linguistics, Computer Science, Institute of Cognitive Science, University of Colorado Boulder - martha.palmer@colorado.edu
- Dr. Wayne Ward Principal Research Scientist - Boulder Language Technologies and Research Professor Computer Science, University of Colorado Boulder - wward@bltek.com
- Dr. James Martin Professor Computer Science, Linguistics, Institute of Cognitive Science, University of Colorado Boulder - james.martin@colorado.edu