

## SWAPNA SOMASUNDARAN

### EDUCATION

University of Pittsburgh (Pittsburgh, PA)	2010	Ph.D.	Computer Science
John Hopkins University (Baltimore, MD)	2002	M.S.E.	Computer Science
Mumbai University (Mumbai, India)	1999	B.E.	Electronics

### EMPLOYMENT HISTORY

2013-present	<p>Research Scientist, Research &amp; Development, Educational Testing Service, Princeton, NJ</p> <p><i>Research and Development:</i> Managed, lead or collaborated in research and development in areas of Natural Language processing for Automated Writing Evaluation, such as discourse coherence detection for essay scoring, scoring of picture-based items using word-level semantics, scoring of persuasive texts using opinion information, scoring of contrastive summaries, and narrative quality analysis.</p> <p><i>Leadership:</i> Co-lead the “Text and Speech Systems” sub initiative with a goal to advance research geared towards building resources and tools that support an array of NLP and Speech systems for aiding item generation, improving instruction, ensuring test security, in addition to creating foundational capabilities that can be applied to assessments.</p> <p><i>Mentoring:</i> Advised student intern on research projects, resulting in conference/workshop publications, and patent activity. Supervised Research Assistants and Associates on ETS projects.</p> <p><i>Miscellaneous:</i> Provided technical expertise and support for responding to various Request For Proposals, as well as internal initiatives. Assisted in the recruitment of full-time employees.</p>
2010-2012	<p>Research Scientist, Siemens Corporate Research, Siemens Corporation Corporate Research, Princeton, NJ, USA</p> <p><i>Research and Development:</i> Lead research in the areas of relationship mining and word spaces. Developed technologies for projects in domains such as fraud detection, personalized healthcare, patent invalidity search and fleet monitoring.</p>

*Project Acquisition and Management:* Played leading and supporting roles in the acquisition of customer projects via prototype development, customer workshop, presentations and discussions, Managed several customer projects ranging from small-scale prototype development to large-scale projects. Managed three software consultants.

*Proposal Writing:* Lead the writing of government proposals, involving multiple universities and departments within Siemens. Assisted in shaping proposals for internal funding.

*Mentoring:* Advised student interns on research projects, resulting in journal, conference and workshop publications, and patent applications. Co-advised interns from other groups.

*University Collaborations:* Lead an internally funded collaboration with UC Berkeley and ICSI FrameNet on developing frame semantics for encoding knowledge in the medical domain. Representing Siemens on the advisory board for SALTS (Rutgers Laboratory for the Study of Applied Language Technology and Society). Lead a DARPA proposal with CMU and University of Pittsburgh.

*Miscellaneous:* Initiated weekly reading club. Collaborated cross-department and cross-program on various projects via module development and discussions. Assisted in the recruitment of full-time employees, software consultants and interns.

2002-2004      Software Consultant, Parity Teltech, IBM TJ Watson Center, New York, USA

Tested and assisted the development of Unstructured Information Management Architecture (UIMA) architecture.

2000-2001      Senior Systems Engineer, Cognitive Systems Research Lab, Tata Infotech Limited, Mumbai, India

Developed a Question Answering system. Managed and coordinated a team of interns and outside experts.

1999            Associate Systems Engineer, Tata Infotech Limited, Mumbai, India

Implemented touch-screen info kiosks for the Central Railways.

1998-1999      Undergraduate Internship, Bhabha Atomic Research Center, Mumbai, India

Developed a system to automate the testing process for bus-based embedded system I/O cards.

## **PROFESSIONAL ACTIVITIES**

### **Research Interests**

Sentiment and Opinion Analysis, Discourse Analysis, Narrative Analysis, Relationship Mining, Natural Language Processing, Computational Linguistics

### **Awards**

SPOT award at ETS 2014, for outstanding team achievement in responding to a difficult RFP.

Best Reviewer Award, EMNLP 2010: Conference on Empirical Methods in NLP (2010)

Andrew Mellon Predoctoral Fellowship, University of Pittsburgh (2009-2010)

Best Graduate Student Research Award, Computer Science Department, University of Pittsburgh (2009)

Outstanding Paper Presentation Award, School of Arts and Sciences Grad Exposition, University of Pittsburgh (2009)

Popular Press, Research featured in Department of Homeland Security Network Newsletter (2007) (<http://www.dhsnetwork.org/Somasundaran.html> )

Incoming Doctoral Student Fellowship, Faculty of Arts and Sciences (FAS), University of Pittsburgh (2004)

Outstanding Employee Contribution Award, Tata Infotech Ltd, India (2001)

### **Organizer**

Textgraphs-5: ACL workshop on Graph-based Methods for Natural Language Processing (2010)

Vice President of Women in Computer Science (WiCS) (2009-2010)

### **Reviewer/Program Committee Member (2007-present)**

EMNLP 2014, COLING 2014, COLING-DEMO 2014, ACL 2014, EACL 2014, LREC 2014, Computational Linguistics Journal 2014, BEA 2014, TAFC (Transactions on Affective Computing) 2011, EMNLP 2011, ACL 2011, Text Graphs 2011, IJCNLP-SAAIP 2011, Emerging Applications of Natural Language Processing: Concepts and New Research (Book) 2010, IEEE Intelligent Systems 2010, ACL 2010, MLG 2010, COLING 2010, EMNLP 2010, NAACL 2010, ACL 2009, EMNLP 2009, ACL 2008, NESCAI 2008, EMNLP 2007, NESCAI 2007

## Computer Skills

Java, Perl, Shell scripts, UIMA (Unstructured Information Management Architecture), Eclipse, NetBeans, Linux, Mac OSX, Windows, Git, SVN, Gforge, SQL server management studio, Latex

## Languages

English, Hindi, Malayalam, Marathi

## PUBLICATIONS

### Ph.D. Thesis

Somasundaran, S. (2010). *Discourse-level Relations for Opinion Analysis*. University of Pittsburgh.

### Journals, Conference and Workshop Publications

(Publications where the first author is a student I advised are indicated by \*)

Swapna Somasundaran Jill Burstein and Martin Chodorow, (2014) *Lexical Chaining for Measuring Discourse Coherence Quality in Test-taker Essays* , COLING 2014, Dublin, Ireland

Swapna Somasundaran and Martin Chodorow, (2014) Automated Measures of Specific Vocabulary Knowledge from Constructed Responses (“Use These Words to Write a Sentence Based on this Picture”), In Proceedings of the Ninth Workshop on Innovative Use of NLP for Building Educational Applications, Baltimore, Maryland, ACL

Jill Burstein, Swapna Somasundaran and Martin Chodorow, (2014) *Finding your “inner-annotator”: An experiment in annotator independence for rating discourse coherence quality in essays*. 8th Linguistic Annotation Workshop (LAW VIII 2014) , Dublin, Ireland.

Beata Beigman Klebanov, Nitin Madnani, Jill Burstein and Swapna Somasundaran (2014) *Content Importance Models for Scoring Writing From Sources*. , In Proceedings of ACL 2014 (short papers) Baltimore, Maryland.

\* Tymoshenko, K., Somasundaran, S., Prabhakaran, V., & Damodar Shet, V. (2012). *Relation Mining in the Biomedical Domain using Entity-level Semantics*. ECAI 2012: 20th European Conference on Artificial Intelligence, Montpellier, France, August 27-31 (conditional accept, under review).

\* Li, D., Somasundaran, S., & Chakraborty, A. (2012). ERD-MedLDA: Entity relation detection using supervised topic models with maximum margin learning. *Journal of Natural Language Engineering (JNLE)*, 18, pp 263-289. doi:10.1017/S1351324912000058

\* Li, D., Somasundaran, S., & Chakraborty, A. (2011). *A Combination of Topic Models with Max-margin Learning for Relation Detection*. In Proceedings of TextGraphs-6: Graph-based

Methods for Natural Language Processing, Workshop at ACL-HLT 2011, Portland, Oregon.

- Somasundaran, S., & Wiebe, J. (2010). *Recognizing Stances in Ideological On-line Debates*. In Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text, pages 116-124, Los Angeles, CA. Association for Computational Linguistics, 2010.
- Somasundaran, S., & Wiebe, J. (2009). *Recognizing Stances in Online Debates*. ACL-IJCNLP: Joint conference of the 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing, Singapore, August 2-7, 2009.
- Somasundaran, S., Namata, G., Wiebe, J., & Getoor, L. (2009). *Supervised and Unsupervised Methods in Employing Discourse Relations for Improving Opinion Polarity Classification*. EMNLP 2009: Empirical Methods in Natural Language Processing, Singapore, August 6-7, 2009.
- Somasundaran, S., Namata, G., Getoor, L., & Wiebe, J. (2009). *Opinion Graphs for Polarity and Discourse Classification*. TextGraphs-4: Graph-based Methods for Natural Language Processing (ACL Workshop), Singapore, 7th August 2009.
- Somasundaran, S., Wiebe, J., & Ruppenhofer, J. (2008). *Discourse Level Opinion Interpretation*. Coling: International Conference on Computational Linguistics, Manchester, 18-22 August, 2008.
- Somasundaran, S., Ruppenhofer, J., & Wiebe, J. (2008). *Discourse Level Opinion Relations: An Annotation Study*. SIGdial Workshop on Discourse and Dialogue (ACL Workshop), Columbus, Ohio, June 2008.
- Ruppenhofer, J., Somasundaran, S., & Wiebe, J. (2008). *Finding the Sources and Targets of Subjective Expressions*. LREC: International Conference on Language Resources and Evaluation, Marrakech, Morocco, May 2008.
- Somasundaran, S., Ruppenhofer, J., & Wiebe, J. (2007). *Detecting Arguing and Sentiment in Meetings*. SIGdial Workshop on Discourse and Dialogue (ACL Workshop), Antwerp, Belgium, September 2007.
- Somasundaran, S., Wilson, T., Wiebe, J., & Stoyanov, V. (2007). *QA with Attitude: Exploiting Opinion Type Analysis for Improving Question Answering in On-line Discussions and the News*. ICWSM: International AAAI Conference on Weblogs and Social Media, Boulder, Colorado, March 26-28, 2007.
- Somasundaran, S., Wiebe, J., Hoffmann, P., & Litman, D. (2006). *Manual Annotation of Opinion Categories in Meetings*. ACL Workshop: Frontiers in Linguistically Annotated Corpora (Coling/ACL 2006), Sydney, Australia, 2006.

## Patent Activity

Somasundaran, S., Prabhakaran, V., Damodar Shet, V., & Tymoshenko, K. (2012). *A System and Method for Generating Knowledge Base for Medical Concepts*. (Patent application in progress: 2011E20301 US).

\* Tymoshenko, K., Somasundaran, S., & Damodar Shet, V. (2012). *Relation Mining in the Biomedical Domain using Entity-level Semantics*. (Patent application in progress: 2012E04345 US).

Somasundaran, S., Li, D., & Chakraborty, A. (2012). *Relationship Mining Using Latent Dirichlet Allocation*. (Patent application filed with USPTO: 2010P20118EP, 2010E19409 US).

## **Demo**

Wilson, T., Hoffmann, P., Somasundaran, S., Kessler, J., Wiebe, J., Choi, Y., Cardie, C., Riloff, E., & Patwardhan, S. (2005). *OpinionFinder: A system for subjectivity analysis*. HLT/EMNLP 2005: Human Language Technologies Conference/Conference on Empirical Methods in Natural Language Processing, Vancouver, Canada, 2005.

## **PRESENTATIONS**

Discourse-level Opinion Analysis (2009)

- Siemens Corporate Research, Princeton, NJ
- Indian Institute of Information Technology (IIIT), Hyderabad, India

Recognizing Stances in Online Debates (2009)

- AI forum, Intelligent Systems Program, University of Pittsburgh

Opinion is not only word deep: Discourse-level Relations for Opinion Analysis (2009)

- Arts and Sciences Graduate Student Expo, University of Pittsburgh

Discourse-level relations for opinion analysis (2009)

- Computer Science Research Competition, University of Pittsburgh

Discourse Level Opinion Interpretation (2008)

- AI forum, Intelligent Systems Program, University of Pittsburgh

Detecting Arguing and Sentiment in Meetings (2007)

- AI forum, Intelligent Systems Program, University of Pittsburgh

## **Teaching Experience**

2002 (Summer)      Instructor, Intermediate Programming, Johns Hopkins University

2005, 2006 (Fall)      Teaching Assistant, Introduction to NLP, University of Pittsburgh

2006 (Fall)	Teaching Assistant, Introduction to Artificial Intelligence, University of Pittsburgh
2005 (Fall)	Teaching Assistant, Discrete Structures for Computer Science, University of Pittsburgh
2001 (Fall)	Teaching Assistant, Database Systems, Johns Hopkins University