

CURRICULUM VITAE

Advaith Siddharthan

Department of Computer Science, Columbia University
1214 Amsterdam Avenue, Mail Code 0401
New York, NY 10027, USA
Fax : +1 212 666 0140
Email: as372@cs.columbia.edu

Current position

- 2003- Post-Doctoral Research Scientist
Department of Computer Science, Columbia University
1214 Amsterdam Avenue, Mail Code 0401
New York, NY 10027, USA.
Dates: October 1 2003 to July 31 2005
Status: Full time
Main Duties and Responsibilities: To perform research and publish results, to assist in writing grant proposals, to oversee a software programmer for the purpose of transferring software from Columbia University to funding agencies, and to co-supervise Masters students taking up research projects for credit.

Education

- 2004 Ph.D. Computer Science, University of Cambridge, UK.
Dissertation Title: Syntactic Simplification and Text Cohesion
Research Group: Natural Language and Information Processing
Degree awarded: March 24, 2004 (Viva passed in September 2003)
Dates Attended: October 2000 to September 2003
- 2004 M.A. Computer Science, University of Cambridge, UK.
(Automatic upgrade of B.A. below)
Degree awarded: March 24, 2004
Dates Attended: October 1998 to June 2000
- 2000 B. A. Computer Science, University of Cambridge, UK.
Class: 1st
Degree awarded: July 1, 2000
Dates Attended: October 1998 to June 2000
- 1998 B. Sc. (Honours) Physics, University of Delhi, India.
Class: 1st
Degree awarded: March 14, 1999
Dates Attended: July 1995 to May 1998

Also participated in an Instructional Course in Quantum Computing, International Centre for Mathematical Sciences, Edinburgh (March, 2000) and the 13th European Summer School in Logic, Language and Information at Helsinki (August 2001).

Honours

- 2000-2003 Cambridge Commonwealth Trust Scholarship
- 2000-2003 British Overseas Research Student Award
- 2000-2003 Bateman Scholarship (Gonville & Caius College, Cambridge)
- 1998-2000 Cambridge Commonwealth Trust Scholarship
 - 1998 David Raja Ram Physics Prize (St Stephens's College, Delhi)
- 1997-1998 Kumari L. A. Meera Memorial Scholarship for Physics (St Stephens's College, Delhi)
 - 1997 Science Meritorious Award for being among the top 1% students in Delhi University

Work experience

- 2003- Post-doctoral research scientist at the Department of Computer Science, Columbia University, New York.
 - 1999 Summer internship in natural language processing for eight weeks at the SRI International Cambridge Computer Science Research Centre during the summer of 1999.
- 1997-1998 Research assistantship in a project on "Mathematics in the modern world" sponsored by the Department of Science and Technology, Government of India.
 - 1997 Summer research studentship in simulational problems in surface physics at the Indian Institute of Science (IISc), Bangalore, India for eight weeks in summer, 1997.

Teaching

Supervisor for the following undergraduate computer science courses at the University of Cambridge, UK. (totaling over 150 hours over three years, in groups of 1-4 students):

- 2002-2003 Information Retrieval
- 2002-2003 Information Theory and Coding
- 2000-2003 Natural Language Processing
- 2000-2003 Artificial Intelligence 1
- 2000-2003 Artificial Intelligence 2
- 2000-2003 Data Structures and Algorithms

Demonstrator for the following laboratory sessions in the Masters in Computer Speech, Text and Internet Technology at the University of Cambridge, UK. (totaling over 100 hours over three years, in groups of 15-30 students):

- 2001-2003 XML and the Internet
- 2001-2002 Question Answering Lab
- 2000-2002 Language Practicals 1 and 2 (Feature Structure Grammars)

Co-advisor (with Prof. Kathleen McKeown) for four Masters students at Columbia University on the research project:

- 2005 A multimodal interface for generating physician notes

Other professional activities

- 2004-2005 Reviewer for the Association for Computational Linguistics annual conference, the journals *Computational Linguistics* and *Natural Language Engineering* and miscellaneous workshops.
- 2002-2005 Member of the Association for Computational Linguistics
- 2002, 2005 Member of the Program Committee for the Student Research Workshop, the Association for Computational Linguistics annual conference (ACL 2002 and ACL 2005).
- 2001-2002 Book reviewer for the journal *Natural Language Engineering*.

Publications

- 2004 Advaith Siddharthan (Accepted). Syntactic Simplification and Text Cohesion. To appear in *Research on Language and Computation*, Kluwer Academic Publications, the Netherlands.
- 2004 Advaith Siddharthan & Ann Copestake. *Generating Referring Expressions in Open Domains*. In Proceedings of the 42th Meeting of the Association for Computational Linguistics Annual Conference (ACL 2004), Barcelona, Spain. Pages 407–414.
- 2004 Advaith Siddharthan, Ani Nenkova and Kathleen McKeown. *Syntactic Simplification for Improving Content Selection in Multi-Document Summarization*. In Proceedings of 20th International Conference on Computational Linguistics (COLING 2004), Geneva, Switzerland. Pages 896–902.
- 2004 David Farwell, Stephen Helmreich, Florence Reeder, Bonnie Dorr, Nizar Habash, Eduard Hovy, Lori Levin., Keith Miller, Teruko Mitamura, Owen Rambow and Advaith Siddharthan. *Interlingual Annotation of Multilingual Text Corpora*. In Proceedings of the Workshop on Frontiers in Corpus Annotation, NAACL/HLT' 04, Boston, MA. Pages 55–62.
- 2004 Bonnie Dorr, Lori Levin, Owen Rambow, David Farwell, Rebecca Green, Nizar Habash, Stephen Helmreich, Eduard Hovy, Keith Miller, Teruko Mitamura, Florence Reeder and Advaith Siddharthan. *Semantic Annotation and Lexico-Syntactic Paraphrase*. In Proceedings of the Workshop on Building Lexical Resources from Semantically Annotated Corpora, LREC'04, Lisbon, Portugal. Pages 47–52.
- 2004 Florence Reed, David Farwell, Teruko Mitamura, Stephen Helmreich, Bonnie Dorr, Nizar Habash, Eduard Hovy, Lori Levin., Keith Miller, Owen Rambow and Advaith Siddharthan. *Semantic Annotation for Interlingual Representation of Multilingual Texts*. In Proceedings of the Workshop on Beyond Named Entity Recognition Semantic labelling for NLP, LREC'04, Lisbon, Portugal. Pages 19–24.
- 2004 Florence Reeder, Bonnie Dorr, David Farwell, Nizar Habash, Stephen Helmreich, Eduard Hovy, Lori Levin, Teruko Mitamura, Keith Miller, Owen Rambow, Advaith Siddharthan, *Interlingual Annotation for MT Development*. In Proceedings of the 6th Conference of the Association for Machine Translation in the Americas (AMTA-2004), Georgetown University, Washington DC, 2004. Pages 236–245.
- 2004 Sasha Blair-Goldensohn, Dave Evans, Vassilios Hatzivassiloglou, Kathleen McKeown, Ani Nenkova, Rebecca Passonneau, Barry Schiffman, Andrew Schljajiker, Advaith Siddharthan and Segei Siegelman. *Columbia University at DUC 2004*. In Proceedings of the Document Understanding Workshop (DUC 2004) at HLT/NAACL 2004, Boston, MA. Pages 23-30.
- 2003 Advaith Siddharthan. *Preserving Discourse Structure when Simplifying Text*. In Proceedings of the European Natural Language Generation Workshop (EWNLG), in association with the 11th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2003), Budapest, Hungary. Pages 103–110.

- 2003 Advaith Siddharthan. *Resolving Pronouns Robustly: Plumbing the Depths of Shallowness*. In Proceedings of the Workshop on Computational Treatments of Anaphora, in association with the 11th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2003), Budapest, Hungary. Pages 7–14.
- 2003 Advaith Siddharthan. *Syntactic Simplification and Text Cohesion*. Ph.D. Thesis, University of Cambridge, U.K. (Also as Technical Report TR-597, University of Cambridge, August 2004.)
- 2002 Advaith Siddharthan. *Resolving Attachment and Clause Boundary Ambiguities for Simplifying Relative Clause Constructs*. In Proceedings of the Student Research Workshop, 40th Meeting of the Association for Computational Linguistics (ACL 2002), Philadelphia, USA. Pages 60–65.
- 2002 Advaith Siddharthan. *An Architecture for a Text Simplification System*. In Proceedings of the Language Engineering Conference 2002 (LEC 2002), Hyderabad, India. Pages 64–71.
- 2002 Advaith Siddharthan. *Resolving Relative Clause Attachment Ambiguities using Machine Learning Techniques and WordNet Hierarchies*. In Proceedings of the 5th National Colloquium for Computational Linguistics in the UK (CLUK 2002), Leeds, UK. Pages 45–49.
- 2002 Advaith Siddharthan & Ann Copestake. *Generating Anaphora for Simplifying Text*. In Proceedings of the 4th Discourse Anaphora and Anaphora Resolution Colloquium (DAARC 2002), Lisbon, Portugal. Pages 199–204.

Book reviews

- 2002 Advaith Siddharthan. *Review of Foundations of Statistical Natural Language Processing (Authors: Christopher D. Manning and Hinrich Schutze)*. In Natural Language Engineering, Volume 8, Issue 1, March 2002. Pages 91–92.
- 2001 Advaith Siddharthan. *Review of Advances in Automatic Text Summarization (Editors: Inderjeet Mani and Mark T. Maybury)*. In Natural Language Engineering, Volume 7, Issue 3, September 2001. Pages 271–272.
- 2001 Advaith Siddharthan. *Review of Building Natural Language Generation Systems (Authors: Ehud Reiter and Robert Dale)*. In Natural Language Engineering, Volume 7, Issue 3, September 2001. Pages 273–274.