

SHILPA ARORA

PHONE +1 917 496 3662 • EMAIL : SHILPAA@CS.CMU.EDU
HTTP://WWW.CS.CMU.EDU/~SHILPAA/

OBJECTIVE

To pursue a career in the field of research in Natural Language Processing (NLP) with a focus on Information Extraction, Question Answering and Active Learning

EDUCATION

Masters in Language Technologies (Fall 2006-Present), School of Computer Science, Carnegie Mellon University, USA (www.lti.cs.cmu.edu) (GPA - 3.78/4.0)

Masters of Science in Computer Science (2004-05): Singapore-MIT Alliance (CGPA - 4.43/5.0) - An alliance between Massachusetts Institute of Technology (MIT), National University of Singapore (NUS) and Nanyang Technological University (NTU) (<http://web.mit.edu/sma/>)

Bachelor of Engineering, Computer Engineering (2000-04): Nanyang Technological University, Singapore (www.ntu.edu.sg). First Class honors (among top 5% of 440 students). (Major GPA - 3.77/4.0)

ACHIEVEMENTS

- **Siebel Scholarship** award for academic year 2007-08 at CMU (<http://www.siebelscholars.com>)
- **Research Fellowship** for Masters Degree at Carnegie Mellon University, 2006-08
- **Graduate Fellowship** to pursue Masters Degree with Singapore-MIT Alliance, 2004-05
- **Nanyang Scholarship** for Bachelor's Degree at Nanyang Tech Univ (NTU), 2000-04
- Finalist of **Microsoft's .NET development competition** - Imagine Cup 2004 (Singapore)
- **First Prize Winner, GRID JAVA Developers Challenge**: Buddy-Finder application, organized by Motorola & ST Mobile Data, national level competition open to educational institutions and companies in Singapore, 2003
- **Dean's Letter of Academic Excellence** (awarded to top 20 students), B.E., NTU, 2000-04

PUBLICATIONS

- Arora, S., Lin F., Shima H., Wang M., Nyberg, E. and Mitamura T., *Tree Conditional Random Fields for Japanese Semantic Role Labeling*. (to be submitted)
- Badaskar S., Agarwal S. and Arora S., *Identifying Real or Fake Articles: Towards better Language Modeling*, Third International Joint Conference on Natural Language Processing, Hyderabad, India, Jan 2008 (to appear)
- Arora S. and Pathak M., *Learning More with Less: Reducing Annotation Effort with Active and Interactive Learning*, LTI Student Research Symposium, CMU, September 2007
- Agarwal S. and Arora S., *Context Based Word Prediction for Texting Language*, 8th RIAO Conference on Large Scale Semantic Access to Content (Text, Image, Video and Sound), Pittsburgh, USA, May 2007
- Arora S., *Using Paraphrases for Information Extraction*, Masters Thesis, Singapore-MIT Alliance, NUS, June, 2005. Supervisors: Dr. Ng Hwee Tou (NUS) and Dr. Leslie Kaelbling (MIT)
- Arora S. and Emmanuel S., *Adaptive Spread Spectrum Based Watermarking of Speech*, 9th NUROP Congress, NTU, September 2003
- Arora S. and Emmanuel S., *Real-Time Adaptive Speech Watermarking for Scheme for Mobile Applications*, 4th ICICS & 4th IEEE PCM Conference, Singapore, December 2003
- **Literature Surveys**:
 - Arora S., Agarwal S., *Active Learning for NLP*, December 2007
 - Arora S., Nyberg, E., *Semantic Role Labeling*, December 2006
 - Arora S., Ng., H.T., *Acquiring and Using Paraphrases - Various approaches of paraphrase acquisition and applications in NLP*, December 2004

WORK EXPERIENCE

Graduate Research Assistant, Language Technology Institute, SCS, CMU (Aug'06-Present)

Projects:

- ROSETTA - part of DARPA sponsored GALE program (<http://www.darpa.mil/ipto/programs/gale/>)

- JAVELIN - sponsored by DTO's AQUAINT program (http://durazno.lti.cs.cmu.edu/wiki/moin.cgi/Javelin_Project)
- Interactive Annotation Learning (IAL)

Roles & Responsibilities:

- Managed UIMA Component Repository (UCR)- Joint work with IBM Research
- Support for UIMA components for the interoperability demo for Gale (MEMT Component)
- Developed Interactive Annotation Learning (IAL) framework for named entities - aim is to develop a strategy to reduce the user annotation effort required to learn a concept. Framework is being used in a LTI course as a class project
- Semantic Role Labeling (SRL) for improving the performance of the JAVELIN multilingual question answering system. Developed in-house Tree-CRFs based SRL for Japanese (one of the first work in SRL for Japanese and SRL using Tree-CRFs)

Analyst , Consulting workforce - Accenture, Singapore (Jun '05-Jun '06)

Project: DBS Core Banking India & Indonesia Project (Corporate, Enterprise and Consumer banking)

Roles & Responsibilities:

- Application failover planning and testing. Banking interface components - designing, deployment and testing. Proposed, implemented and tested the approach for Disaster Recovery system. Support for live production system
- Process optimization: Improved and managed firewall application and tracking process -bottleneck & error prone system. Automated the application deployment procedure
- Interacted closely with DBS (client), IBM (client's technology provider) and Infosys (sub-contractor)

Industrial Attachment, Research Scientist - Singapore Institute of Manufacturing Technology (SIMTech) (Jan- Jun '03)

Project: Remote operation of an All Terrain Robotic Vehicle (ATRV-2, iRobot)

Roles & Responsibilities:

- *Graphical User Interface* for the remote operation of a teleoperated vehicle (iRobot ATRV-2) using wireless LAN, and RF modem for serial port communication
- *Video Streaming* over wireless LAN for sending video feedback from the vehicle to the operator for analysis and image processing

SKILLS AND ABILITIES

Programming Skills: C, C++, C#, Java, J2ME, VB, Microsoft Visual Studio, Eclipse, .NET, Measurement Studio, NI Instruments, VHDL, LEX & YACC, Flex & Bison, Perl, latex, MySQL, Postgres, WebSphere, IBM Tivoli Scheduler, SVN, CVS

Machine learning & NLP tools: Minorthird, Mallet Toolkit, UIMA, OpenFST, WordNet 2.0, Wnjn 1.6 (JNI interface to WordNet)

Languages: Well versed in English and Hindi. Introductory knowledge of French

EXTRA-CURRICULAR ACTIVITIES

Organizer, Opportunities for Undergraduate Research in Computer Science Conference, CMU, 2007 (<http://www.cs.cmu.edu/ourcs/>) Sponsored by Microsoft Research (MSR, USA)

- Research conference for Undergraduate Women in CS, attended by 60 students internationally
- Organized by CMU for the first time and first of its kind in the US

Social Director, Organizing Committee for World Universities Debating Championship, NTU, 2003-04

- World's largest non-sporting student event - 900 participants from 100 universities in 50 countries

Vice President, Student's Initiative For Technopreneurship (SIFT), NTU, 2002-03

REFERENCES: Available upon request