

Daqing HE

BSc, MSc, PhD
Institute for Advanced Computer Studies
The University of Maryland, College Park, Maryland 20742, USA
Email: daqing@umiacs.umd.edu
Tel: +1 301-405-6766 *Fax:* +1 301-314-9658
Homepage: <http://www.umiacs.umd.edu/~daqingd>

Executive Summary

Career Goal:

To be a world famous researcher in information science, especially on topics related to applying language processing techniques to information access. At the same time, to be a well respected teacher in transferring knowledge of information science.

Career Highlights:

Research Scientist	Institute for Advanced Computer Studies, The University of Maryland, College Park, Maryland	3/02 - present
Research Fellow	School of Computing, Robert Gordon University, Aberdeen, United Kingdom	7/99 – 2/02
PhD Student	Division of Informatics, The University of Edinburgh, Edinburgh, United Kingdom	12/95 – 7/01
MSc Student	Dept. of Computer Science & Engineering, Beijing University of Aeronautics and Astronautics, Beijing, China	9/92 – 3/95
BSc Student	Dept. of Computer Science & Engineering, Beijing University of Aeronautics and Astronautics, Beijing, China	9/88 – 7/92

Summary of Expertise

Areas	Information retrieval (monolingual and multilingual), Web user context learning and user modeling, interactive retrieval interface design, computational linguistics (noun phrase analysis, statistical NLP, semantic analysis), and Web log mining and analysis.
Programming Language	Extensive experience in Java, Perl, C++, C, and Prolog. Good working knowledge of Unix, Windows and Linux
Language	Fluent in English, native speaker of Chinese

Education

PhD in Artificial Intelligence	Division of Informatics (former Dept of Artificial Intelligence), University of Edinburgh, Edinburgh, UK	12/95 – 7/01
Thesis Title	References to Graphical Objects in Intelligent Multimodal Interfaces	
Award	Overseas Research Studentship, Colin & Ethel Gordon Scholarship	
Supervisors	Dr Graeme Ritchie and Dr John Lee	
MSc in Computer Science	Dept. of Computer Science & Engineering, Beijing University of Aeronautics and Astronautics, Beijing, China	9/92 – 3/95
Dissertation Title	A Study of Three-value Logic and Its Application to the Detection of Functional Hazards in Combinational Circuits.	
Supervisors	Prof. Qiping Zhao and Prof. Huaimin Sun	
BSc in Computer Science	Dept. of Computer Science & Engineering, Beijing University of Aeronautics and Astronautics, Beijing, China	9/88 – 7/92
Dissertation Title	A Design of a Prolog Interpreter and its C Implementation	
Supervisor	Prof. Huaimin Sun	

Research Achievements

Research Scientist (3/02 - present)

Institute for Advanced Computer Studies, The University of Maryland (UMD), College Park, Maryland

- ❖ Team leader of UMD's participation in the High Accuracy Retrieval of Documents (HARD) track of Text REtrieval Conference (TREC) 2003. TREC is organized by the National Institute of Standards and Technology (NIST), and is the most prestigious gathering of the best scientists and teams in the world that are working on problems/techniques in information retrieval (IR).
 - Responsible for the design of the overall approach taken by the UMD HARD team.
 - Designed and implemented several key components of the system, including the automatic query expansion module, the document re-ranking module, and the ranked-list merging module.
 - Led effort in analyzing results, and writing report about the effort of the UMD HARD team.
 - Participated in the TREC conference to present the work of the team.

- ❖ Team leader of the Cross Language Information Retrieval (CLIR) team in UMD's participation in the surprise language experiment and dry-run of Translingual Information Detection, Extraction, and Summarization (TIDES) project, which is a research project funded by the Defense Advanced Research Projects Agency (DARPA).
 - Responsible for the overall design of a rapidly retargetable CLIR interactive system (called MIRACLE) for processing previous unseen languages.
 - Led the team to complete the implementation of the MIRACLE system for Hindi language within a month. MIRACLE was the first interactive CLIR system for Hindi in the world.
 - Led the team to complete the design and implementation of the batch CLIR systems for Hindi language within a month and for Cebuano language within 10 days. Initial test results showed that our batch CLIR systems achieved comparable results to monolingual Hindi retrieval.
 - Wrote a journal paper presenting the UMD CLIR team's effort in surprise language experiment. Listed as the first author of the paper.
- ❖ Team leader of the CLIR team in UMD's participation in the interactive track of the Cross Language Evaluation Forum (iCLEF) in the past two years. iCLEF is a TREC equivalent conference in Europe, and is a specialized gathering for CLIR researchers.
 - Responsible for the overall design of the experiments run by the UMD team, and conducted major parts of the experiments.
 - Responsible for the design and extension of the user-assisted query translation approach for interactive Cross Language Information Retrieval (CLIR), which was the main theme of UMD's participation in iCLEF in the past two years.
 - Led the effort in analyzing results, and wrote reports and presentations about UMD's iCLEF effort in the past two years. Listed as the first author of the reports.
- ❖ Team leader of the UMD team's participation in the Topic Detection and Tracking (TDT) track of TREC 2002.
 - Responsible for the overall design and implementation of a topic tracking system using language modeling and information retrieval techniques. The system achieved comparable tracking results to the state of art in TDT 2002 experiment.
 - Led effort in analyzing results, wrote reports, and presented at the TDT workshop of TREC conference. Listed as the first author of the report.

Research Fellow (7/99 - 2/02)

School of Computing, The Robert Gordon University, Aberdeen, UK.

- ❖ Responsible for the design of a Web user context learning model that uses both the user's personal search history and collaborative information from people who share the same interests to improve search effectiveness. Responsible for the implementation of the key context learning components in a Web search engine. Published several conference papers on this topic.
- ❖ Researched, surveyed, and compared theories and models about context, context information of users, and context models in communication, linguistic process, and information retrieval. Wrote a white paper about the outcome of the research, and based on the report, extended a context learning model in traditional information retrieval to a context learning model in Web information retrieval.
- ❖ Responsible for the design of a role-based context model for modeling the user in information retrieval. The role model extends traditional context model by viewing the user's actions from task-oriented view.

- ❖ Conducted text mining in several Web search log collections. Identified the problem of lacking sensible session boundaries in Web logs, which is a problem for context learning for Web information retrieval.
- ❖ Designed and implemented time interval based session identification method. Conducted experiments to test the method on several Web log collections. The results show that the method can provide reasonable boundaries for our context learning purpose. Wrote several conference papers about this work.
- ❖ Designed and implemented an evidence combination method, based on Dempster-Shafer theory, for Web search session identification. The method uses the outcome of time interval method and that of the search pattern method for Web search session identification as input, and generates statistical based inference about the possibility of having session delimiters at a given point. Conducted an experiment and demonstrated the effectiveness of the method. A journal paper, which I am the first author, was published on one of the most prestigious journals of IR research.

Graduate Research Assistant (12/95 - 7/01)

Division of Informatics (former Dept. of Artificial Intelligence), University of Edinburgh, UK

- ❖ Designed a computational model for interpreting English referring expressions in intelligent multimodal interfaces. The model is designed to handle reference ambiguities that could involve attributes of objects in the application domain and/or attributes of graphic icons on the screen. This kind of reference ambiguities, which I called “source ambiguities”, had been mentioned before in the literature, but it was my model that provides a computational means for resolving them for the first time.
- ❖ Designed a logical meaning representation language (MRL) to represent the meaning of natural language sentences. The advantages of the MRL is that it was designed to handle source ambiguity, and it provides a uniform meaning representation from the phase before the ambiguities are resolved all the way to the phase that the ambiguities have been solved. This again was the first MRL for handling source ambiguities.
- ❖ Developed an intelligent multimodal dialogue system to demonstrate the ability of the computational model for handling source ambiguities. I conducted experiment that involves human subjects to interact with the system. The results show that there was statistically significant difference concerning the user’s experience with the system with the ability of resolving source ambiguity and that without. The former is much preferred by the user.
- ❖ Researched, surveyed, and compared different theories and models in intelligent multimodal systems. Identified source ambiguities as an important and interesting problem that had not been fully studied in the literature. This work motivated me to work on resolving source ambiguities.

Graduate Research Assistant (9/92 – 3/95)

Department of Computer Science and Engineering, Beijing University of Aeronautics and Astronautics, Beijing, China

- ❖ Designed a resolving method for three-value logical equations, and constructed a computational model for detecting hazards in combinational circuits. The design was published in a journal paper, where I am the first author.

- ❖ Built a software system that identifies potential hazards existing in a combinational circuit based on the logical description of the circuit. The software was given the second place award in FengRu Cup student research competition in 1995.
- ❖ Researched, surveyed, and compared existing theories and methods in resolving three-value logical equations, and its application to detecting hazards.
- ❖ The dissertation that discusses the above designs and implementations received an excellent grade.

Undergraduate Research Assistant (9/91 – 7/92)

Department of Computer Science and Engineering, Beijing University of Aeronautics and Astronautics, Beijing, China

- ❖ Designed an interpreter for PROLOG programming language, and implemented the interpreter using C programming language.

Teaching Experience

Adjunct Faculty (9/03 – present)

College of Information Studies, University of Maryland, College Park, Maryland.

Course: Master Core Course 690: Information Technology

Duties:

- ❖ Instructed lecturing classes to a class of 20 Master students
- ❖ Organized discussion about course materials
- ❖ Organized lab sessions
- ❖ Constructed and graded examination papers
- ❖ Helped students in their team projects

Teaching Assistant (10/97 – 6/99)

Division of Informatics, University of Edinburgh, UK

Courses: Msc Prolog, Artificial Intelligence 1, Computer Science 1

Duties:

- ❖ Instructed and organized discussions about courses and course exercises;
- ❖ Marked examination papers

Teaching Assistant (9/93 – 1/94)

Dept of Computer Science & Engineering, Beijing University of Aeronautics & Astronautics, China.

Course: Machine Intelligence & Knowledge Engineering

Duties:

- ❖ Marked examination papers

Honors and Awards

Overseas Research Studentship (1995-1999)

Provided by UK government (highly competitive among all foreign students in UK).

Colin & Ethel Gordon Scholarship (1995-1999)

Provided by the University of Edinburgh (highly competitive among candidates applying for the university).

The second place award of FengRu Cup student research competition (1995)

Organized by Beijing University of Aeronautics and Astronautics (highly competitive among students in the university).

Elected president of Edinburgh Chinese Student & Scholar Association (1997-1998)

Edinburgh Chinese Student & Scholar Association is an association for all Chinese students and Scholar in the city of Edinburgh.

Professional Activities

- ❖ Invited Referee to review papers in
 - *Information Processing and Management*,
 - *ACM Transactions on Asian Language and Information Processing*, a special issue of the journal “Surprise Language Experiment”,
 - *User Modeling and User-Adapted Interaction*, a special issue of the journal “User Modeling for Web and Hypermedia Information Retrieval”.
- ❖ Member of the program committee for
 - “User Modeling, Machine Learning and Information Retrieval”, a workshop of User Modeling 2003.
 - The Student Section of COLING-ACL98.
- ❖ Member of
 - Association of Computational Linguistics,
 - Association of Computing Machinery,
 - American Society for Information Science and Technology,
 - British Computer Society.
- ❖ Invited talks at
 - Stanford University,
 - The University of Southern California,
 - The University of Maryland, and
 - The Robert Gordon University.
- ❖ Attended and presented papers at several conferences on Information Retrieval, Computational Linguistics, User Modeling, Intelligent User Interfaces and Artificial Intelligence, including TREC-2002, TDT-2002, ACL-2002, UM-2001, AH2000, ICMI2000, SIGIR2000, INLG2000, BCS-IRSG2000, COLING-ACL98, ECAI98, and ACL-EACL97.
- ❖ Web master: the TIDES project at UMD; Information Retrieval Group, School of Computer and Mathematical Sciences, The Robert Gordon University.
- ❖ Web master and mail-list organizer of the Chinese Student & Scholar Association in Edinburgh between 1996 and 1999.
- ❖ Certificate holder of the *Introduction Course of Teaching and Demonstrating*, the *Introduction Course of First Aid*, and the *BP Team Development Course*.

Publications

1. **Daqing He**, Dina Demner-Fushman. *HARD Experiment at Maryland: From Need Negotiation to Automated HARD Process*. In the proceedings of Text REtrieval Conference (TREC) 2003.
2. **Daqing He**, Douglas W. Oard, Jianqiang Wang, Jun Luo, Dina Demner-Fushman, Kareem Darwish, Philip Resnik, Sanjeev Khudanpur, Michael Nossal and Anton Leuski. *Making MIRACLES: Interactive Translingual Search for Cebuano and Hindi*. Accepted to *ACM Transactions of Asian Language Information Processing*, 2003.
3. **Daqing He**, Graeme Ritchie, John Lee. *Reference to Graphics Objects in Intelligent Multimodal Systems*. Journal paper in preparation.
4. Douglas W. Oard, David Doermann, Bonnie Dorr, **Daqing He**, Philip Resnik, Amy Weinberg, William Byrne, Sanjeev Khudanpur David Yarowsky, Anton Leuski, Philipp Koehn, and Kevin Knight, *Desparately Seeking Cebuano*. In the *Proceedings of the HLT-NAACL Conference*, Late Breaking Results, Edmonton, Canada, 2003.
5. **Daqing He**, Jianqiang Wang, Douglas W. Oard, Michael Nossal *User-Assisted Query Translation for Interactive CLIR..* In the Proceedings of SIGIR 2003, Toronto Canada 2003.
6. Ayse Goker, **Daqing He**. *Personalization via Collaboration in Web Retrieval System*. In 76th Annual Conference for American Society for Information Science and Technology ASIST 2003.
7. Bonnie J. Dorr, **Daqing He**, Jun Luo, Douglas W. Oard, Richard Schwartz, Jianqiang Wang, and David Zajic, *iCLEF 2003 at Maryland: Translation Selection and Document Selection*. In the *Proceedings of the Interactive track for the Cross-Language Evaluation Forum Workshop*, Trondheim, Norway, 2003.
8. **Daqing He**, Ayse Goker, David Harper. *Combining Evidence for Automatic Web Session Identification*. *Information Processing and Management*, 38(5): 727-742, 2002.
9. **Daqing He**. *Providing Context for Translation Alternative in interactive Cross Language Information Retrieval*. Project White paper. Institute for Advanced Computer Studies, University of Maryland, College Park, 2002.
10. **Daqing He**, Hyuk Ro Park, G. Craig Murray, Michael Subotin, and Douglas W. Oard. *TDT2002 Topic Tracking at University of Maryland*. In the *proceedings of the workshop of Topic Detection and Tracking TDT2002*, November, 2002.
11. **Daqing He**, Jianqiang Wang, Douglas W. Oard, Michael Nossal. *Comparing User-assisted and Automatic Query Translation*. In the *proceedings of the workshop of Cross Language Evaluation Forum CLEF2002*, September, 2002.
12. Shen Han, Ayse Goker, **Daqing He**. *Web user search pattern analysis for modeling query topic changes*. Presented at *User modeling for context-aware applications, a workshop of the 8th International Conference on User Modeling*, July, 2001.
13. **Daqing He**. *Towards a Role-based Context model*. Project White paper, School of Computing, Robert Gordon University, Aberdeen, Scotland, 2001.
14. **Daqing He**. *A review of Query Expansion Techniques in Information Retrieval*. Project White paper, School of Computing, Robert Gordon University, Aberdeen, Scotland, 2001.
15. Ayse Goker, **Daqing He**. *Analysing Web Search Logs to Determine Session Boundaries for User-Oriented Learning*. In *Adaptive Hypermedia and Adaptive Web-Based Systems, International Conference*, Trento Italy, Peter Brusilovsky, Oliviero Stock, Carlo Strapparava (Eds.) LNCS 1892, Pages 319-322, Springer, Berlin Germany, 2000.

16. **Daqing He**, Ayse Goker. *Detecting Session Boundaries from Web User Logs*. In *22nd Annual Colloquium on IR Research*, Pages 57-66, Cambridge, UK, 2000.
17. **Daqing He**, Graeme Ritchie, John Lee. *Resolving References to Graphical Objects in Multimodal Queries by Constraint Satisfaction*. In *Advances in Multimodal Interfaces, the 3rd International Conference*, Beijing China, Tieniu Tan, Yuanchun Shi, Wen Gao (Eds.) LNCS 1948, Pages 8-15, Springer, Berlin Germany, 2000.
18. **Daqing He**, *References to Graphical Objects in Intelligent Multimodal Interfaces*. PhD thesis, Division of Informatics, University of Edinburgh, Edinburgh, UK, 2000.
19. **Daqing He**, Graeme Ritchie, and John Lee. *Disambiguation between Visual Display and Represented Domain in Multimodal Interfaces*. Proceedings of Combining AI & Graphics for the Interface of the Future, a workshop of ECAI'98, Brighton, 1998.
20. **Daqing He**, Graeme Ritchie, and John Lee, *Referring to Displays in Multimodal Interfaces*. Proceedings of Referring Phenomena in a Multimedia Context and their Computational Treatment, a workshop of EAACL/ACL'97 Conference, Madrid, 1997.
21. Huacan He, Yonghuai Liu, **Daqing He**, and Hua Cheng, *Generalized Logic in Experience Thinking*, Science in China (series E), Vol.39, No.3, P226-234, 1996.
22. **Daqing He**, Qinqing Zhao, *Three-Value Logic Equations and the Detection of Functional Hazards in Combinational Switching Circuits*, *Journal of Beijing University of Aeronautics & Astronautics*, Vol.21, No.3, P38-44, 1995.
23. **Daqing He**, *A Study of Three-value Logic and its application to the Detection of Functional Hazards in Combinational Circuits*. Master of Science Dissertation. Department of Computer Science and Engineering, Beijing University of Aeronautics and Astronautics, Beijing, China, 1995.
24. Boyang Liu, **Daqing He**, Qinqing Zhao, *An Improvement on Self-Organising Feature Map Algorithm*, Proceedings of International Conference for Young Computer Scientist, Beijing, China, 1995.

Invited Talks and Presentations

1. “*HARD experiment at University of Maryland*”, Presented at HARD workshop, TREC 2003, Gaithersburg, Maryland, November 2003.
2. “*From Need Negotiation to Automated HARD process*”, Presented at TIDES project Maryland team site visit, October 2003.
3. “*Making MIRACLEs: Interactive Translingual Search for Cebuano and Hindi*”. Presented at CLIR minisummit, Center for the Study of Language and Information, Stanford University, October 2003.
4. “*Personalization via Collaboration in Web Retrieval System*”. Presented at 76th Annual Conference for American Society for Information Science and Technology ASIST 2003, Long Beach, October 2003.
5. “*Interactive Cross Language Information Retrieval*”. Presented at Research Review Day, University of Maryland, College Park, May 2003.
6. “*TDT2002 Topic Tracking at University of Maryland*.” Presented at the workshop of Topic Detection and Tracking TDT2002, Gaithersburg, Maryland, November 2002.
7. “*TDT2002 Dry-run at University of Maryland*”. Presented at the workshop of TDT 2002 dry-run, Los Angeles, California, April 2002.

8. *"Comparing User-assisted and Automatic Query Translation"*. Presented at Information Science Institute Seminar, University of Southern California, Santa Monica, California. April 2002.
9. *"Personalization in a Role-base Web Retrieval System"*. Presented at Linguistic Consortium of Cross Language and Information Processing Lab, University of Maryland, College Park, October 2001.
10. *"Web User Search Pattern Analysis for Modeling Query Topic Changes."* Presented at User Modeling for Context-aware Applications, a workshop of the 8th International Conference on User Modeling, July 2001.
11. *"Personalization in Web Retrieval"*. Presented at Burns Retreat, School of Computing, Robert Gordon University, Aberdeen, Scotland. June 2001
12. *"Analysing Web Search Logs to Determine Session Boundaries for User-Oriented Learning"*. Presented at Adaptive Hypermedia and Adaptive Web-Based Systems, International Conference, Trento Italy, July 2000.
13. *"Detecting Session Boundaries from Web User Logs."* Presented at 22nd Annual Colloquium on IR Research, Cambridge, UK. April 2000.
14. *"Resolving References to Graphical Objects in Multimodal Queries by Constraint Satisfaction."* Presented at the 3rd International Conference on Intelligent Multimodal Interfaces, Beijing China. October 2000.
15. *"References Ambiguities in Intelligent Multimodal Interfaces"*. Presented at School of Computing, Robert Gordon University, Aberdeen, UK. February 1999.
16. *"Disambiguation between Visual Display and Represented Domain in Multimodal Interfaces."* Presented at Combining AI & Graphics for the Interface of the Future, a workshop of ECAI'98, Brighton, August 1998.
17. *"Referring to Displays in Multimodal Interfaces"*. Presented at Referring Phenomena in a Multimedia Context and Their Computational Treatment, a workshop of EACL/ACL'97 Conference, Madrid, July 1997.

Programming and Natural Language Skills

- ❖ Programming Languages: extensive knowledge of Java, Perl, C/C++, HTML, Prolog and Tcl/Tk; good knowledge of Javascript and LISP.
- ❖ Operating Systems: good working knowledge of Unix, Windows and MS-DOS.
- ❖ Word Processing: extensive knowledge of MS-Office (Word, PowerPoint, Excel) and LaTeX.
- ❖ Natural Languages: Chinese and English.

References

1. Dr Douglas W. Oard,
Associate Professor,
College of Information Studies/Institute for Advanced Computer Studies,
The University of Maryland, College Park, MD 20742 USA
Tel: 301-4057590
Email: oard@glue.umd.edu

2. Dr Ayse Goker
Senior Lecturer (UK's equivalent position to Associate Professor),
School of Computing, The Robert Gordon University,
St Andrew Street, Aberdeen AB25 1HG UK
Tel: +44-1224-262713
Email: asga@comp.rgu.ac.uk

3. Dr Qiang Shen
Senior Lecturer,
Division of Informatics, The University of Edinburgh,
Room E6, 80 South Bridge, Edinburgh EH1 1HN, UK
Tel: +44-131-6502705
Email: qiangs@dai.ed.ac.uk

4. Dr Bonnie J. Dorr
Associate Professor,
Department of Computer Science/Institute for Advanced Computer Studies,
The University of Maryland, College Park, MD 20742, USA
Tel: 301-4056768
Email: bonnie@cs.umd.edu