

RESEARCH INTERESTS

Machine Learning, Information Retrieval, Natural Language Processing

EDUCATION

- 2008.9~Present,** PhD candidate, Language Technologies Institute,
Carnegie Mellon University
- 2006.9~2008.8,** Master's degree, Language Technologies Institute,
Carnegie Mellon University
- 2003.9~2006.7,** Master's degree, Dept. of Computer Science, Tsinghua University
- 1999.9~2003.7,** Bachelor's degree in Electronic Engineering, Tsinghua University

RESEARCH EXPERIENCE

- ◇ **Intern:** (6/2011~9/2011) Host, Amarnag Subramanya, asubram@google.com
Semantic Annotations From Text (SAFT) project, Google Inc.
- Use knowledge base to help information extraction from web pages.
- ◇ **Intern:** (5/2010~8/2010) Mentor, Kevyn Collins-Thompson, kevynct@microsoft.com
Context, Learning, and User Experience for Search (CLUES) group, Microsoft Research
- Learning to automatically modify users' search queries.
- ◇ **Research Assistant:** (6/2009~present) Advisor, William Cohen, wcohen@cs.cmu.edu
Machine Learning Department, Carnegie Mellon University
- Random walk inference on graphs (thesis topic)
 - Relational CRFs structure learning and hidden variable detection
- ◇ **Research Assistant:** (9/2007~5/2009) Advisor, Teruko Mitamura, teruko@cs.cmu.edu
Language Technologies Institute, Carnegie Mellon University
- Cross-lingual Question Answering System
 - Chinese Semantic Role Labeling
- ◇ **Research Assistant:** (7/2006~8/2007) Advisor, Yiming Yang, yiming@cs.cmu.edu
Language Technologies Institute, Carnegie Mellon University
- Evaluation metric for utility-based information distillation
 - Random walk on graph for multi-task learning and contextual retrieval
- ◇ **Intern:** (7/2003~6/2006) Mentor, Dr. Ji-Rong Wen, jrwen@microsoft.com
Web Search and Mining (WSM) group, Microsoft Research Asia
- Large scale clustering and classification of online products by visual and textual clues.
 - Compare personalized search strategies by simulating user experience from search log.
 - Compare learning methods to improve search engine relevance performance.
 - System that combines Window system configuration, system behavior log, and textual descriptions of user experiences to help automate PC troubleshooting.
-

◇ Student Research Training Program: (2/2001~6/2003)

Advisor, Prof. Zeng-Qi Sun, szq-dcs@mail.tsinghua.edu.cn

State Key Lab of Intelligent Technology and Systems (LITS), THU

- Bachelor thesis, Mining Spatial-Temporal Data Using Constructive Induction.
- I am one of the main developers of the Tsinghua Aeolus system, which placed first in the RoboCup Simulation League in 2001 and 2002.

PROGRAMMING SKILLS

Extended experience in JAVA, C++, Matlab, R, SQL, Perl

PUBLICATIONSNi Lao, Tom Mitchell, William W. Cohen, **Random Walk Inference and Learning in a Large Scale Knowledge Base**. EMNLP, 2011Jun Zhu, Ni Lao, Ning Chen, Eric P. Xing, **Conditional Topical Coding: an Efficient Topic Model Conditioned on Rich Features**. KDD, 2011Ni Lao, Jun Zhu, Liu Liu, Yandong Liu, William W. Cohen, **Efficient Relational Learning with Hidden Variable Detection**. NIPS, 2010Ni Lao, William W. Cohen, **Relational retrieval using a combination of path-constrained random walks**. *Machine Learning*, 2010, Volume 81, Number 1, Pages 53-67Ni Lao, William W. Cohen, **Fast Query Execution for Retrieval Models Based on Path Constrained Random Walks**. KDD, 2010Jun Zhu, Ni Lao, E. P. Xing, **Grafting-Light: Fast, Incremental Feature Selection and Structure Learning of Markov Random Fields**. KDD, 2010N. Lao, H. Shima, T. Mitamura, E. Nyberg. 2008. "**Query Expansion and Machine Translation for Robust Cross-Lingual Information Retrieval**", NTCIR-7, Japan.H. Shima, N. Lao, E. Nyberg, T. Mitamura. 2008. "**Complex Cross-lingual Question Answering as Sequential Classification and Multi-Document Summarization Task**", NTCIR-7, Japan.W. Zuo, N. Lao, Y. Geng, and K. Ma. 2008. **GeoSVM: an efficient and effective tool to predict species' potential distributions**. *Journal of Plant Ecology*, 1(2): 143-145.Y. Yang, A. Lad, N. Lao, A. Harpale, B. Kisiel, M. Rogati, **Utility-based information distillation over temporally sequenced documents**, SIGIR, pp. 31-38, 2007C. Yuan; N. Lao; J. Wen; J. Li; Z. Zhang; Y. Wang; W. Ma, **Automated Known Problem Diagnosis with Event Traces**, EuroSys, 2006.N. Lao, J. Wen, W. Ma, Y. Wang, **Combine High Level Symptom and Low Level State Information for Configuration Fault Diagnosis**, USENIX LISA, pp. 151-158, 2004.J. Wen, N. Lao, W. Ma, **Probabilistic Model for Contextual Retrieval**, SIGIR, 2004J. Yao, N. Lao, F. Yang, Y. Cai, Z. Sun: **Technical Solutions of TsinghuaAeolus for Robotic Soccer**. RoboCup, pp. 205-213, 2003
