
Chenhai Xi

Department of Computer Science
5503 Sennott Square
University of Pittsburgh,
Pittsburgh, PA 15260

Phone: 412-726-8640
Email: chenhai@cs.pitt.edu

Objective

Looking for a summer intern position in the area of **machine learning** and/or **natural language processing**

Education

- **Ph.D. Candidate:** University of Pittsburgh, Department of Computer Science 2002 - present
- **M.S.:** Tsinghua University (China), Department of Computer Science December 2001
- **B.S.:** Tsinghua University (China), Department of Computer Science July 1999

Research Experience

- **Natural Language Processing Laboratory, University of Pittsburgh** 05/2004 - present
 - *Proposed a novel approach of combining a large amount of bilingual resources with a small amount of manually annotated data, and conducted the experiments with the language pair of English-Chinese to train a Chinese Part-of-Speech tagger. Experimental results showed that the proposed approach achieves a significant improvement over EM and self-training.
 - *Currently investigating how to apply the proposed approach to more complex tasks in which the output is not a class label but a structure, such as parsing
- **Department of Computer Science, University of Pittsburgh** 01/2003 - 08/2003
 - *Participated in the project PARTS: Power-Aware Real-Time Systems
- **Artificial Intelligence Laboratory, Tsinghua University** 09/1998 - 01/2002
 - *Designed and implemented a Chinese segmentation and part-of-speech tagging system, which includes a proper noun recognition subsystem

Teaching Experience

- Teaching Assistant of graduate course CS2750 (Machine Learning) 2005 and 2006 Spring
- Teaching Assistant of graduate course CS2150 (Design & analysis of algorithms) 2006 Spring

Publications

C. Xi and R. Hwa. A Backoff Model for Bootstrapping Resources for Non-English Languages. In *the Proceedings of HLT/EMNLP-05*, pages 851-858, Vancouver, Canada.

R. Xu, C. Xi, R. Melhem and D. Moss. Practical PACE for Embedded Systems. In *the Proceedings of 4th ACM International conference on Embedded Software(EMSOFT '04)*, Pisa, Italy, September, 2004.

C.Xi and M.Sun. Automatic Prediction of Chinese Phrase Boundary Location with Neural Networks. *Journal of Chinese Information Processing*, Vol 16, pages 20-26, 2002.

Programming Language Skills

C++, Java, Perl, Python, Matlab

References

References available on request.