

Department of Computer Science
University of Houston

DISTINGUISHED LECTURER SEMINAR 2012

WHEN: FRIDAY, SEPTEMBER 14, 2012

WHERE: PGH 232

TIME: 3:00 PM

SPEAKER: Dr. Pushpak Bhattacharyya, IIT Bombay

Host: Dr. Ioannis Kakadiaris



TITLE: Word Sense Disambiguation- a Multilingual Resource Conscious Perspective

Abstract:

Word Sense Disambiguation (WSD) is a fundamental problem in Natural Language Processing (NLP). Amongst various approaches to WSD, it is the supervised machine learning (ML) based approach that is the dominant paradigm today. However, ML based techniques need significant amount of resource in terms of sense annotated corpora which takes time, energy and manpower to create. Not all languages have this resource, and many of the languages cannot afford it.

We discuss ways of doing WSD under resource constraint. We describe a novel scoring function and an iterative algorithm based on this function to do WSD. This function separates the influence of the annotated corpus (corpus parameters) from the influence of wordnet (wordnet parameters), in deciding the sense. Next we describe how the corpus of one language can help WSD of another language, i.e., LANGUAGE ADAPTATION. This is presented in three setting of "complete", "some" and "no" annotation.

The talk is presented in a multilingual setting of Indian languages. The presentation is based on work done with PhD and Masters students and researchers: Rajat, Mitesh, Salil, Saurabh, Anup, Sapan and Piyush, and published in fora like ACL, COLING, EMNLP, GWC, ICON and so on.

BIO: Dr. Pushpak Bhattacharyya is a Professor of Computer Science and Engineering at IIT Bombay. He received his B.Tech from IIT Kharagpur, M.Tech from IIT Kanpur and PhD from IIT Bombay. He has held visiting positions at MIT, Cambridge, USA, Stanford University, USA and University Joseph Fourier, Grenoble, France. Dr. Bhattacharyya's research interests include Natural Language Processing, Machine Learning and Cross Lingual Search. He has had more than 130 publications in top conferences and journals and has served as program chair, area chair, workshop chair and PC member of top fora like ACL, COLING, LREC, SIGIR, CIKM, NAACL, GWC and others. He has guided 12 PhDs and over 150 masters and undergraduate students in their thesis work. Dr. Bhattacharyya plays a leading role in India's large scale projects on Machine Translation, Cross Lingual Search, and Wordnet and Dictionary Development. Dr. Bhattacharyya received a number of prestigious awards including IBM Innovation Award, Yahoo Faculty Award, United Nations Research Grant, Microsoft Research Grant, IIT Bombay's Patwardhan Award for Technology Development and Ministry of IT and Digital India Foundation's Manthan Award. Recently he has been appointed Associate Editor of the prestigious journal, ACM Transactions on Asian Language Information Processing.