



Word Sense Induction for Machine Translation

Prof. Min ZHANG

Distinguished Professor

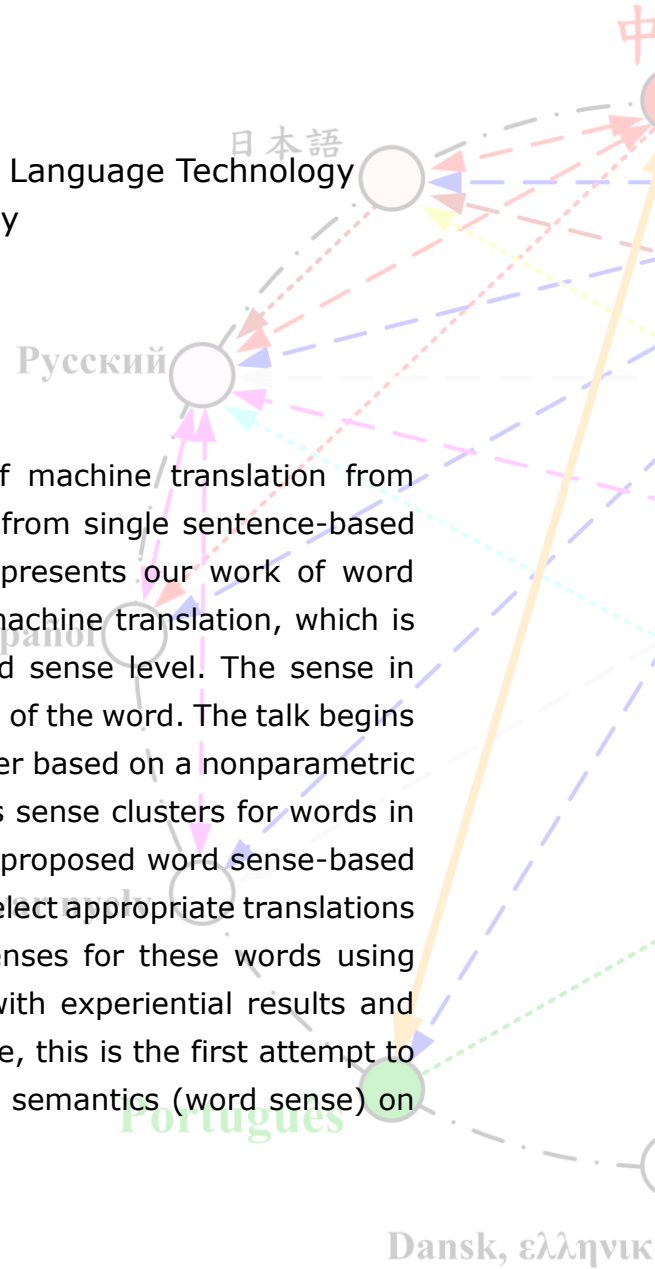
Director of the Research Center for Human Language Technology

School of Computer Science and Technology

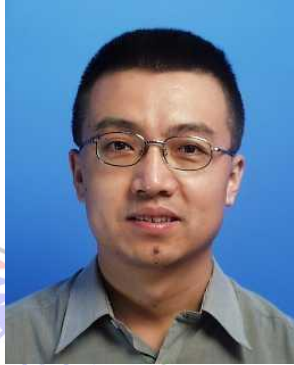
Soochow University, China

Abstract

We have witnessed the research progress of machine translation from phrase/syntax-based to semantics-based and from single sentence-based to discourse and document-based. This talk presents our work of word sense-based translation model for statistical machine translation, which is one of semantics-based SMT research at word sense level. The sense in which a word is used determines the translation of the word. The talk begins with how to build a broad-coverage sense tagger based on a nonparametric Bayesian topic model that automatically learns sense clusters for words in the source language, and then focuses on the proposed word sense-based translation model that enables the decoder to select appropriate translations for source words according to the inferred senses for these words using maximum entropy classifiers. The talk ends with experiential results and some conclusions. To the best of our knowledge, this is the first attempt to empirically verify the positive impact of lexical semantics (word sense) on translation quality.



Biography



Dr. Min Zhang, a distinguished professor, Director of the Research Center of Human Language Technology and Vice Dean of the school of computer science at Soochow University (China), received his Bachelor degree and Ph.D. degree in computer science from Harbin Institute of Technology (China) in 1991 and 1997, respectively. He worked in Academy and Industry from 1997-2013 at Singapore and South Korea. His current research interests include machine translation, natural language processing and big search. He has co-authored more than 150 papers in leading journals and conferences, and co-authored/co-edited 15 books published by Springer, IEEE CPS and COLIPS. He is an associated editor of IEEE T-ASLP and Acta Automatic Sinica. He was the recipient of several awards in China and oversea. He is the PI of three NSFC projects, including 1 for Distinguished Young Scholars, 1 Key Program and 1 General Program.

Venue: E11-3033

Date: 28 January, Thursday

Time: 11:00a.m.~12:00p.m.

Contacts: Dr. Derek F. Wong / Dr. Lidia S. Chao
Telephone: (853) 8822-8051
Website: <http://nlp2ct.cis.umac.mo/>

Address:
E11-2023, University of Macau,
Avenida da Universidade,
Taipa, Macau, China