Rocling 2007 Program		
September 6, 2007 (Thursday) 9：30～18：00		
International Conference Room in NTU Library		
9:20~9:30	Opening Session	
9:30~10:30	Invited Speaker Session Speaker: Janyce Wiebe	
10:30~10:50	Coffee Break	
10:50~12:10	Session 1: Information Retrieval	
	以英語寫作輔助為目的之語料庫語句檢索方法	Jyi-Shane Liu, Pei-Chun Horng, and Ching-Ying Lee
	Bayesian Topic Mixture Model for Information Retrieval	Meng-Sung Wu, Hsuan-Jui Hsu, and Jen-Tzung Chien
	Korean-Chinese Cross-Language Information Retrieval Based on Extension of Dictionaries and Transliteration	Yu-Chun Wang, Tzong-Han Richard Tsai, Hsu-Chun Yen, and Wen-Lian Hsu
Forte Oliver Garden in Ming Da Hall		
12:10~13:00	Lunch	
International Conference Room in NTU Library		
13:00~14:00	ACLCLP Meeting/Poster Session: NSC Projects	
14:10~15:50	Session 2: Speech Recognition	
	Feature Statistics Compensation for Robust Speech Recognition in Additive Noise Environments	Tsung-hsueh Hsieh and Jeih-weih Hung
	Cyberon Voice Commander 多國語言語音命令系統	何泰軒 and 劉進榮

	改善以最小化音素錯誤為基礎的鑑別式聲學模型訓練於中文連續語音辨識之研究	劉士弘, 朱芳輝, and 陳柏琳
	Study of the Voice Activity Detection Techniques for Robust Speech Feature Extraction	Wen-Hsiang Tu and Jeih-weih Hung
15:50~16:20	Coffee Break	
16:20~18:00	Session 3: Classification & Disambiguation	
	從不同韻律格式驗證階層式韻律架構並兼論對語音科技的應用	鄭秋豫 and 蘇昭宇
	多語聲學單位分類之最佳化研究	呂道誠, 呂仁園, 江永進, and 許鈞南
	詞義辨識:機器學習演算法特徵的選取與組合	高紹航 and 高照明
	Word Translation Disambiguation via Dependency	Meng-Chin Hsiao, Kun-Ju Yang, and Jason S. Chang
Forte Oliver Garden in Ming Da Hall		
18:30~20:00	Banquet	
September 7, 2007（Friday）9：00～16：40		
International Conference Room in NTU Library		
9:00~10:00	Invited Speaker Session Speaker: Hwee Tou Ng	
10:00~10:20	Coffee Break	
10:20~12:00	Session 4: Information Extraction	
	Knowledge Representation for Interrogatives in E-HowNet	Shu-Ling Huang, You-Shan Chung, Yueh-Yin Shih, and Keh-Jiann Chen

	Question Analysis and Answer Passage Retrieval for Opinion Question Answering Systems	Lun-Wei Ku, Yu-Ting Liang, and Hsin-Hsi Chen
	基於統計方法之中文搭配詞自動擷取	張翠芸 and 柯淑津
	以部落格語料進行情緒趨勢分析	楊昌樺, 高虹安, and 陳信希

Forte Oliver Garden in Ming Da Hall	
12:00~13:00	Lunch

International Conference Room in NTU Library		
13:00~14:40	Poster Session	
	中文詞彙語意資料的整合及擷取	高照明
	A Study on Prosodic Modeling for Isolated Mandarin Words	Chi-feng Chen, Chen-yu Chiang, Yih-ru Wang, and Sin-Horng Chen
	以中文十億詞語料庫為基礎之兩岸詞彙對比研究	洪嘉翡, 黃居仁, and 許銘維
	VOT productions of word-initial stops in Mandarin and English: A cross-linguistic study	Li-mei Chen, Kuan-Yi Chao, and Jui-Feng Peng
	台灣共通語言	Ming-Shing Yu
	中文詞義全文標記語料庫之設計與雛形製作	柯淑津, 黃居仁, 洪嘉翡, 劉詩音, 簡卉伶, and 蘇依莉
	Predicting Trends of Stock Prices with Text Classification Techniques	Jiun-Da Chen, Tai-Ping Wang, and Chao-Lin Liu
	基於階層架構資訊及關鍵詞語義擴展的階層式目錄整合研究	Cheng-Tse Hung, Ing-Xiang Chen, Ping-Jung Wu, and Cheng-Zen Yang
	Speech recognition of mandarin syllables using both linear predict coding cepstra and Mel frequency cepstra	Tze Fen Li and Shui-Ching Chang
	Improving Retrieval Effectiveness by Document Reranking and Local Expansion	Wang Wen-Chi and Lin Bor-Shen
	針對數學與科學教育領域之電腦輔助英中試題翻譯系統	呂明欣, 高照明, 劉昭麟, and 張俊彥

	Word sense induction using independent component analysis	Petr Simon and Jia-Fei Hong
14:40~15:00	Coffee Break	
15:00~16:20	Session 5: Speech Processing	
	Language Identification on Code-Switching Speech	Chyng-Leei Chu, Dau-cheng Lyu, and Ren-yuan Lyu
	An HNM Based Method for Synthesizing Mandarin Syllable Signal	Hung-Yan Gu and Yen-zuo Zhou
	The Role of Sound Changes in the Speech Recognition System: A Phonetic Analysis of the Final Nasal Shift in Mandarin	James H. Yang
16:20~16:40	Best Paper Award and Closing Session	

Invited Speaker: Janyce Wiebe

Topic

Subjectivity and Sentiment Analysis

Abstract

A growing area of research, "subjectivity analysis" is the computational study of affect, opinions, and sentiments expressed in text. Blogs, editorials, reviews (of products, movies, books, etc.), and even "objective" newspaper articles (which include many opinions and sentiments) are just some of the genres for which accurate identification and interpretation of opinions is critical for full text understanding. Subjectivity analysis will support developing tools for information analysts in governmental, commercial, and political domains who want to automatically track attitudes and feelings in the news and on-line forums. How do people feel about the latest iPod? Is there a change in the support for the new Medicare bill? A system able to automatically identify and extract opinions and sentiments from text would be an enormous help to someone sifting through the vast amounts of news and web data, trying to answer these kinds of questions. In this talk, I will first describe a corpus annotated with rich information about opinions and sentiments, and then present experiments using that data to develop and evaluate automatic systems. In particular, I will describe experiments in recognizing the "contextual polarity" of expressions, i.e., whether a phrase is being used to express a positive or negative sentiment, considering the context in which it appears. I will also describe experiments exploring interactions between subjectivity and word sense, showing that subjectivity is a property that can be associated with word meanings and that subjectivity classification can be beneficial for word sense disambiguation.

Autobiography

Jan Wiebe is a professor of computer science and Director of the Intelligent Systems Program at the University of Pittsburgh. Her research with students and colleagues has been in discourse processing, pragmatics, word-sense disambiguation, and probabilistic classification in NLP. Her most recent work investigates automatically recognizing and interpretating expressions of opinions and sentiments in text, to support NLP applications such as question answering, information extraction, text categorization, and summarization. Her current and past professional roles include NAACL Program Committee Chair, NAACL Executive Board member, Computational Linguistics and Language Resources and Evaluation Editorial Board member, AAAI Workshop Co-Chair, ACM Special Interest Group on Artificial Intelligence (SIGART) Vice-Chair, and ACM-SIGART/AAAI Doctoral Consortium Chair.

Invited Speaker: Hwee Tou NG

Topic

Recent Advances in Word Sense Disambiguation:
Scaling Up, Sense Prior Estimation, and Integration into Statistical Machine Translation

Abstract

In this talk, I will introduce several research issues associated with word sense disambiguation (WSD), which is the task of determining the correct meaning, or sense, of a word in context. I will present recent work completed in my research group to address these issues. The first issue concerns the scaling up of WSD. Although supervised WSD gives good accuracy, the lack of sense-tagged training data has hampered the progress of WSD. I will present our approach of using parallel texts to scale up WSD. Using this approach, our WSD system participated in SemEval-2007, where our system achieved the highest and second highest accuracy in the coarse-grained and fine-grained English all-words task, among 16 and 14 participating systems respectively. The second issue concerns the accuracy drop of a WSD system, when it is applied to texts drawn from a different domain with different sense priors. I will present results showing improved WSD accuracy after applying class prior estimation algorithms and using well calibrated probabilities. The third issue concerns the perceived lack of applications utilizing WSD. We integrated our state-of-the-art WSD system into Hiero, a state-of-the-art hierarchical phrase-based statistical machine translation system. We found that the use of WSD improves translation quality, and the improvement is statistically significant.

Autobiography

Dr. Hwee Tou NG (黃偉道博士) is an associate professor of computer science at the National University of Singapore, program co-chair (Computer Science Program) of the Singapore-MIT Alliance, and a senior faculty member at the NUS Graduate School for Integrative Sciences and Engineering. He is the Editor-in-Chief of the ACM Transactions on Asian Language Information Processing (TALIP), and an editorial board member of Natural Language Engineering journal. He has also served as an editorial board member of Computational Linguistics journal (2004 - 2006). He was program co-chair of ACL-2005 conference, and has served on the program committees of many past conferences including ACL, SIGIR, AAAI, and IJCAI. He is the secretary of ACL SIGNLL, and a SENSEVAL committee member of ACL SIGLEX. His research focuses on natural language processing and information retrieval.