

第二十二屆自然語言與語音處理研討會

ROCLING 2010 Program

September 1-2, 2009, National Chi Nan University, Nantou, Taiwan, ROC

September 1, 2010 (Wednesday) 9 : 10 ~ 20 : 00	
09:10-10:00	Registration
10:00-10:10	Opening Ceremony Chair: Prof. Jing-Shin Chang
10:10-11:20	Invited Talk Speaker: Prof. Hsin-Hsi Chen (陳信希), NTU
11:20-11:50	Coffee Break
11:50-12:50	Oral Session 1: Information retrieval and extraction
12:50-14:00	Lunch/NSC meeting for future directions
14:00-15:00	ACLCLP Meeting/Poster Session 1: NSC Projects
15:00-16:00	Oral Session 2: Speech recognition
16:00-16:30	Coffee Break
16:30-17:30	Oral Session 3: NLP applications/tools/resources
18:00-20:00	Banquet

September 2, 2010 (Thursday) 9 : 00 ~ 16 : 40	
09:00-10:10	Invited Talk Speaker: Dr. Jian SU Institute for Infocomm Research, Singapore
10:10-10:40	Coffee Break
10:40-12:00	Oral Session 4: Text/Opinion mining and segmentation
12:00-13:10	Lunch
13:10-14:30	Poster Session 2 : ROCLING Papers
14:30-15:00	Coffee Break
15:00-16:20	Oral Session 5: Speech analysis/synthesis
16:20-16:40	Closing Ceremony and Best Paper Award

Invited Talk I: Prof. Hsin-Hsi Chen (陳信希)

Time: Wednesday, September 1, 10:10-11:20

Title: Opinion Mining and Sentiment Analysis

Abstract:

Opinion mining attracts much attention in recent years due to that large scale of experiences are shared through different web platforms such as review sites, forums,

blogs, microblogs, and so on. Users' opinions are useful for many applications, e.g., buyer's best guide, marketing, public poll, advertisement, etc. In this talk, I will discuss the challenging issues of this research area, and give a brief survey from the issues of granularity, relevance, language and methodology. Evaluation benchmark is important for both training and testing. I will present the design of Chinese, English and Japanese opinion corpora for multilingual opinion analysis task (MOAT) in NTCIR series evaluation. Besides, linguistic resources are also indispensable. I will show NTU sentiment dictionary, the construction of Chinese opinion Treebank, and the uses of morphological and syntactic structures for Chinese opinion analysis. Finally, I will extend the discussion to other sentiment analysis. Both writer and reader emotions are my foci. I will employ the collaborative emotion tagging in Blogs and news materials to present several emotion analysis technologies.

Biography:

Professor Hsin-Hsi Chen received the B.S. and the M.S. degrees in computer science and information engineering in 1981 and 1983, respectively, and the Ph.D. degree in electrical engineering in 1988, all from National Taiwan University, Taipei, Taiwan.

Since August 1995, Professor Chen has been a professor in Department of Computer Science and Information Engineering, National Taiwan University. He served as chair of the department from 2002 to 2005. From 2006 to 2009, he was chief director of Computer and Information Networking Center. His research interests are computational linguistics, Chinese language processing, information retrieval and extraction, and web mining. He has published about 250 papers in these research areas.

Professor Chen is a member of ACL (1986-) and ACLCLP (1988-). He serves as the Board of Directors (1992-2009), deputy president (2005-2007), president (2007-2009) and advisory board chair (2010-2011) of Association for Computational Linguistics and Chinese Language Processing (ACLCLP), AFNLP Board Member (2006-2009) and ACL SIGHAN Vice-Chair/Chair-Elect (2009-2011). He is also Editorial Board member (2000-2008) of International Journal of Computational Linguistics and Chinese Language Processing, an Associate Editor (2007-2010) of ACM Transactions on Asian Language Information Processing, and an Editor of Journal of Information Science and Engineering (2010-2012). He is program co-chair of ACM SIGIR 2010, serves as steering committee members of CLEF, AIRS, and NTCIR, senior PC members of ACM SIGIR 2006, 2007, 2008 and 2009, area/track chairs of ACL-IJCNLP 2009 and ACM CIKM 2008, and PC members of many conferences (IJCAI, SIGIR, AIRS, ACL, COLING, EMNLP, NAACL, IJCNLP, WWW, and so on). He is one of organizers of CLIR, CLQA and MOAT tasks in NTCIR.

Homepage: <http://nlg.csie.ntu.edu.tw>

Invited Talk 2: Dr. Jian SU

Time: Thursday, September 2, 9:00-10:10

Title: Language Resource Assessment for Information Access

Abstract

In this talk, I'll share the experience in an EU frame work 6 Strategic Research project BOOTStrep when I was leading the final work package to access various language resources for information access purposes, which involves 4 EU partners including European Bioinformatics Institute, National Center of Text Mining, UK and two German Teams besides my team in Institute for Infocomm Research (I2R), Singapore. Bio-lexicon, Bio-ontology and Natural Language Processing tools are evaluated for their effectiveness for information retrieval and information extraction on gene regulation. The talk will cover the construction of document retrieval evaluation set and the evaluations on the incorporation of Named Entity Recognition and BioLexicon to Information Retrieval on both BOOTStrep evaluation set and Trec Genomics data besides the assessments on gene regulation extraction. I'll conclude my talk with the further extensions and future work.

Biography:

Dr. Su Jian (苏俭) is a senior scientist and group leader of information extraction and text mining in I2R where she established herself in the areas of Information Extraction, Co-reference Resolution, (Bio)Text Mining. Dr Su has published intensively in natural language processing (NLP) and bioinformatics conferences and journals, including 13 papers in ACL Annual Meetings, and one journal article in Computational Linguistics in recent years. Dr Su is active in professional services for the computational linguistics community. She has served as Editor / Member of Editorial Board for three international journals. She is program chair of ACL-IJCNLP09, publication chair of ACL 2007 and IJCNLP 2005, program chair of LBM 2007, Workshop Organizer of LREC2008 Workshop: Building and evaluating resources for biomedical text mining, and PC members of numerous NLP conferences including ACL, IJCNLP, COLING, and EMNLP.

Dr Su also led her team to achieve top performances in various information extraction / text mining benchmarking such as BioCreAtIve. She was responsible for the effort to establish the largest coreference annotation corpus on GENIA collection in the joint project with Tokyo U. She has been the Principal Investigator in multiple technology deployments including BioMedical Information Management, Intelligence Gathering, Legal / Standard Enforcement as well.

Homepage: <http://www.i2r.a-star.edu.sg/~sujian/>

Technical Program

Oral Session 1: Information retrieval and extraction

Time: Wednesday, September 1, 11:50-12:50

- 1 Tourism-Related Opinion Mining
Chuan-Jie Lin and Chao Pin-Hsien
 - 2 An Automatic Query Augmentation Model for Answering Well-Defined Questions in Top-1 Search Result
Jing-Shin Chang and Shu-Fan Shih
 - 3 鑑別式語言模型於語音辨識結果重新排序之研究
Chia-Wen Liu, Shih-Hsiang Lin and Berlin Chen
-

Oral Session 2: Speech recognition

Time: Wednesday, September 1, 15:00-16:00

- 1 發音事件驗證於多語辨識發音變異模型之產生
Pei-Shan Tsai, Han-Pin Shen and Chung-Hsien Wu
 - 2 整合邊際資訊於鑑別式聲學模型訓練方法之比較研究
Yueng-tien Lo and Berlin Chen
 - 3 A Study of Minimum Variance Modulation Filter for Robust Speech Recognition
Ren-hau Hsieh, Hao-tang Fan and Jehi-weih Hung
-

Oral Session 3: NLP applications/tools/resources

Time: Wednesday, September 1, 16:30-17:30

- 1 以語文特徵為基之中學閱讀測驗短文分級

Chao-Shainn Huang, Wei-Ti Kuo, Chia-Ling Li and Chao-Lin Liu

- 2 基於對照表以及語言模型之簡繁字體轉換
Min-Hsiang Li, Shih-Hung Wu, Ping-Che Yang and Tsun Ku
- 3 以共現資訊為基礎增進中學英漢翻譯試題與解答之詞彙對列
昭憲 黃, 昭麟 劉 and 元顯 曾

Oral Session 4: Text/Opinion mining and segmentation

Time: Thursday, September 2, 10:40-12:00

- 1 Term Contributed Boundary Feature using Conditional Random Fields for Chinese Word Segmentation Task
Mike Tian-Jian Jiang, Shih-Hung Liu, Cheng-Lung Sung and Wen-Lian Hsu
- 2 Japanese Opinion Word Translation Based on Unsupervised Word Sense Disambiguation in the Travel Domain
Jyun-Wei Huang and Richard Tzong-Han Tsai
- 3 Classifying mood in plurks
Mei-Yu Chen, Hsin-Ni Lin, Chang-An Shih, Yen-Ching Hsu, Pei-Yu Hsu and Shu-Kai Hsieh
- 4 中文短句之情緒分類
瑛澤 孫, 建良 陳, 峻杰 劉, 昭麟 劉 and 豐文 蘇

Oral Session 5: Speech analysis/synthesis

Time: Thursday, September 2, 15:00-16:20

- 1 以最佳化及機率分佈判斷漢字聲符之研究
Chia-Hui Chang, Shu-Ying Li, Sean Lin, Chia-Yi Huang and Jhih-ming Chen
- 2 Text-Independent Automatic Phone Segmentation Algorithm
Ching-Feng Lu and Hsiao-Chuan Wang
- 3 可變速中文文字轉語音系統
Qi-quan Huang, Chen-Yu Chiang, Yih-Ru Wang, Sin-Horng Chen and Hsiu-Min Yu
- 4 Advanced Modulation Spectrum Compensation Techniques for Robust Speech Recognition
Wei-Jeih Yeh, Wen-hsiang Tu and Jeh-wei Hung

Poster Session 2 : ROCLING Papers

Time: Thursday, September 2, 13:10-14:30

- 1 Identifying Correction Rules for Auto Editing
Anta Huang, Tsung-Ting Kuo, Ying-Chun Lai and Shou-De Lin
- 2 English article errors in Taiwanese students' EFL writing
Neil Edward Barrett and Li-mei Chen
- 3 結合HMM頻譜模型與ANN韻律模型之國語語音合成系統

Hung-yan Gu, Ming-Yen Lai and Sung-Fung Tsai

- 4 美國專利書「獨立項數」之搭配詞初探: 以LexisNexis資料庫為例
Darren Hsin-hung Lin and Shelley Ching-yu Hsieh
- 5 Dual-ALU Processor for Speech Signal Processing
Gin-Der Wu
- 6 基於網路社群之旅遊經驗及對應情境之情感意見分析研究
格致 齋 and 昆棋 賴
- 7 結合音長與發音特徵於GTB之腔調化語音辨識
Po-San Lai, Ming-chi Yen and Jui-Feng Yeh
- 8 應用直方圖均化於統計式未知詞萃取之研究
Yi-Cong Chen and Bor-Shen Lin
- 9 Qualia Modification in Noun-Noun Compounds: A Cross-Language Survey
Chih-yao Lee, Chia-hao Chang, Wei-chieh Hsu and Shu-Kai Hsieh
- 10 Discerning Emotions of Bloggers based on Topics – a Supervised Coreference Approach in Bengali
Dipankar Das and Sivaji Bandyopadhyay