Conference on Computational Linguistics and Speech Processing

第二十一屆自然語言與語音處理研討會

September 1-2, 2009, National Chung Hsing University, Taichung, Taiwan, ROC

http://rocling2009.cs.nchu.edu.tw/

September 1, 2009 (Tuesday)　9：20～20：00		
09:20-09:30	**Opening Ceremony**	Chair: Prof. Ming-Shing Yu
09:30-10:30	**Invited Talk**	Speaker: Dr. **Chin-Hui Lee** (李錦輝), Georgia Tech
10:30-10:50	Coffee Break	
10:50-12:10	**Oral Session 1：speech analysis and synthesis**	
12:10-13:00	Lunch	
13:00-14:10	**ACLCLP Meeting/Poster Session 1: NSC Projects**	
14:10-15:30	**Oral Session 2 : information retrieval and extraction**	
15:30-15:50	Coffee Break	
15:50-17:10	**Oral Session 3 : phonetics/phonology analysis and application**	
17:40-20:00	Banquet	

September 2, 2009 (Wednesday)　9：10～16：30		
09:10-10:10	**Invited Talk**	Speaker: Dr. **Keh-Jiann Chen** (陳克健), Academia Sinica
10:10-10:30	Coffee Break	
10:30-11:50	**Oral Session 4 : NLP applications, tools and resources**	
11:50-12:50	Lunch	
12:50-14:30	**Poster Session 2：　ROCLING Papers**	
14:30-14:50	Coffee Break	
14:50-16:10	**Oral Session 5 : speech recognition and understanding**	
16:10-16:30	**Closing Ceremony and Best Paper Award**	Chair: Prof. Ming-Shing Yu

Invited Speaker : Chin-Hui Lee

Topic

Universal Phone Modeling for Multilingual Automatic Speech Recognition

Abstract

Building on recent successes in automatic speech recognition (ASR), the next big research challenge will be multilingual ASR (MASR) capable of exceeding human performance. It is always believed that the MASR problem is too big to address for researchers knowing only a few languages or any research groups with limited resources. Language-specific acoustic modeling has always been a practical approach to designing high performance ASR systems for a particular language. However for data-limited languages the system accuracy is usually poor. Extending to MASR a popular technique is to group together all training speech data from all the available languages, find a set of fundamental phone units that cover all the languages, and train a set of universal phone models (UPMs) that can be used to characterize all the phones and triphones for all the languages being considered. Language-adaptive models have recently been shown to improve over language-specific models in some special situations. This common set of phones is usually derived from the collection of International Phonetic Alphabet (IPA) which was mainly defined phonetically, and was shown in previous studies to give non-satisfactory MASR performance because of the inconsistency and a lack of full knowledge in defining the IPA. Due to our recent success in modeling and detecting speech attributes across multiple languages it seem reasonable to explore these fundamental units as shared structures spanning over all spoken languages that can be used for large vocabulary MASR of all the languages seen or unseen during training.

In this talk we first review our attribute-based system for continuous phone recognition with little or no language-specific speech training data by integrating three levels of information from: (1) frame based speech attribute detectors, (2) artificial neural network based phone event mergers, and (3) decoding based evidence verifiers. We report on experimental results on Japanese phone recognition with the OGI Multilingual Speech Corpus. It is interesting to note that a good performance can be achieved without using any Japanese speech training data. If the set of common units and the shared acoustic-phonetic properties among different languages can be properly utilized, we believe a good multilingual phone recognizer can be designed, and a high performance MASR system based on automatic speech attribute transcription (ASAT) can eventually be realized. We demonstrate advantages in designing attribute detectors for place and manner of articulation, and share preliminary result on how they can be used to provide diagnostic information for correcting errors in state-of-the-art large vocabulary ASR systems. We believe MASR is an excellent platform for the acoustic and phonetic communities to pull together a collaborative effort to address key research issues, such as defining a common set of units based on acoustic phonetic attributes that can be directly and reliably detected from the speech signal.

Autobiography

Dr. **Chin-Hui Lee** (李錦輝) is a professor at School of Electrical and Computer Engineering, Georgia Institute of Technology. Dr. Lee received the B.S. degree in Electrical Engineering from National Taiwan University, Taipei, in 1973, the M.S. degree in Engineering and Applied Science from Yale University, New Haven, in 1977, and the Ph.D. degree in Electrical Engineering with a minor in Statistics from University of Washington, Seattle, in 1981.Dr. Lee started his professional career at Verbex Corporation, Bedford, MA, and was involved in research on connected word recognition. In 1984, he became affiliated with Digital Sound Corporation, Santa Barbara, where he engaged in research and product development in speech coding, speech synthesis, speech recognition and signal processing for the development of the DSC-2000 Voice Server. Between 1986 and 2001, he was with Bell Laboratories, Murray Hill, New Jersey, where he became a Distinguished Member of Technical Staff and Director of the Dialogue Systems Research Department. His research interests include multimedia communication, multimedia signal and information processing, speech and speaker recognition, speech and language modeling, spoken dialogue processing, adaptive and discriminative learning, biometric authentication, and information retrieval. From August 2001 to August 2002 he was a visiting professor at School of Computing, The National University of Singapore. In September 2002, he joined the Faculty Georgia Institute of Technology. Prof. Lee has participated actively in professional societies. He is a member of the IEEE Signal Processing Society (SPS), Communication Society, and the International Speech Communication Association (ISCA). In 1991-1995, he was an associate editor for the IEEE Transactions on Signal Processing and Transactions on Speech and Audio Processing. During the same period, he served as a member of the ARPA Spoken Language Coordination Committee. In 1995-1998 he was a member of the Speech Processing Technical Committee and later became the chairman from 1997 to 1998. In 1996, he helped promote the SPS Multimedia Signal Processing Technical Committee in which he is a founding member.

Dr. Lee is a Fellow of the IEEE, and has published more than 300 papers and 25 patents on the subject of automatic speech and speaker recognition. He received the SPS Senior Award in 1994 and the SPS Best Paper Award in 1997 and 1999, respectively. In 1997, he was awarded the prestigious Bell Labs President's Gold Award for his contributions to the Lucent Speech Processing Solutions product. Dr. Lee often gives seminal lectures to a wide international audience. In 2000, he was named one of the six Distinguished Lecturers by the IEEE Signal Processing Society. He was also named one of the two ISCA's inaugural Distinguished Lecturers in 2007-2008. Recently he won the SPS's 2006 Technical Achievement Award for "Exceptional Contributions to the Field of Automatic Speech Recognition".

Home page: http://users.ece.gatech.edu/~chl/

Invited Speaker : Keh-Jiann Chen

Topic

E-HowNet- a Lexical Semantic Representation System and its Relation to Morphology, Syntax and Semantics

Abstract

Natural language understanding is a long-term goal for NLP. Compositional or generative approaches for semantic processing need a framework for representing lexical knowledge and carrying semantic composition capability. E-HowNet is a frame-based entity-relation model extended from HowNet to define concepts. It intends to achieve following goals:

- Representing word senses (concepts) by primitives or well-defined concepts.
- Performing semantic composition and decomposition.
- Achieving near canonical sense representations.
- Universal and language independent representation.

In this talk we will present the designing methodology of E-HowNet and discuss the sense representational issues for various semantic types of words including both function words and content words. We will talk about the research issues of how compositional processing can extend lexical sense representations to form phrasal/sentential senses. We will focus our attention on the issues of interactions among morphology, syntax and semantics and their logical consequences. We will also point out some important semantic features expressed in E-HowNet, in particular the coarse-grained semantic features of object, attribute, and value, and show differences in their syntactic behaviors and how they applied to NLP.

Autobiography

Dr. **Keh-Jiann Chen** (陳克健) obtained a B.S. in mathematics from National Cheng Kung University in 1972. He received a Ph.D. in computer science from the State University of New York at Buffalo in 1981. Since then he joined the Institute of Information Science as an associate research fellow and became a research fellow in 1989. He was the deputy director of the institute from August 1991 to July 1994.

His research interests include Chinese language processing, lexical semantics, lexical knowledge representation, and corpus linguistics. He had been and continued in developing the research environments for Chinese natural language processing including Chinese lexical databases, corpora, Treebank, lexical analyzer and parsers.

Dr. Chen is one of the founding members of the Association of Computational Linguistic and Chinese Language Processing (also known as ROCLING). He had served as 2nd term president of the association from 1991 to 1993. Currently he is the board member of the Chinese Language Computer Society, the advisory board member of the International Journal of Computational Linguistics and Chinese Language Processing, and the editor of journal of Computer Processing of Oriental Language.

Home page: http://www.iis.sinica.edu.tw/pages/kchen/index_zh.html

Oral session 1 (10:50-12:10): speech analysis and synthesis

10:50-11:10 普遍化相似度比率鑑別分析

11:10-11:30 Wavelet Energy-Based Support Vector Machine for Noisy Word Boundary Detection With Speech Recognition Application

11:30-11:50 Noise-Robust Speech Features Based on Cepstral Time Coefficients

11:50-12:10 A Study of Sub-band Feature Statistics Compensation Techniques Based on a Discrete Wavelet Transform for Robust Speech Recognition

Oral session 2 (14:10-15:30): information retrieval and extraction

14:10-14:30 Sampling the Web as Training Data for Text Classification

14:30-14:50 Query Formulation by Selecting Good Terms

14:50-15:10 中英文專利文書之文句對列

15:10-15:30 意見持有者辨識之研究

Oral session 3 (15:50-17:10) : phonetics/phonology analysis and application

15:50-16:10 Tonal effects on voice onset time: Stops in Mandarin and Hakka

16:10-16:30 Latent Prosody Model-Assisted Mandarin Accent Identification

16:30-16:50 Sample-based Phone-like Unit Automatic Labeling in Mandarin Speech

16:50-17:10 基於離散倒頻譜之頻譜包絡估計架構及其於語音轉換之應用

Oral session 4 (10:30-11:50) : NLP applications, tools and resources

10:30-10:50 電腦輔助句子重組試題編製

10:50-11:10 On the Use of Topic Models for Large-Vocabulary Continuous Speech Recognition

11:10-11:30 Improving Translation Fluency with Search-Based Decoding and a Monolingual Statistical Machine Translation Model for Automatic Post-Editing

11:30-11:50 Unsupervised Question Classification and Answering based on WordNet and Wikipedia

Oral session 5 (14:50-16:10): speech recognition and understanding

14:50-15:10 應用句型結構與部份樣本樹於對話行為偵測之研究

15:10-15:30 資源受限運算環境下華英混雜語音辨識系統

15:30-15:50 A Study of Sub-band Modulation Spectrum Compensation for Robust Speech Recognition
15:50-16:10 A Study of Hybrid-based Cepstral Statistics Normalization Techniques for Robust Speech Recognition

Poster session 2 : ROCLING papers

1. Automatic Term Pair Extraction from Bilingual Patent Corpus
2. On The Learning of Chinese Aspect Marker le through Interactive Multimedia Program
3. A Framework for Machine Translation Output Combination
4. A Voice Conversion System based on Formant and LSF Mapping without Using Parallel Corpus
5. Speech Enhancement Technique Based on Blind Source Separation for Far-Field Noisy Speech Recognition
6. Hierarchical Web Document Classification Based on Hierarchically Trained Domain Specific Words
7. 中文混淆字集應用於錯別字模板自動產生
8. Consolidation of Robust Speaker and Speech Recognition for Intelligent Doorway Application
9. Voice Activity Detection Using Spectral Entropy in Bark-Scale Wavelet Domai
10. 讓格書寫 以及 台華互譯 初探