

CONFERENCE PROGRAM

FIRST DAY: 22 March 2004 (Monday)

9:00 – 9:15	Opening Ceremony		
9:15 – 9:55	Opening Address: Prof. Makoto Nagao		
9:55 – 10:00	Break		
10:00 – 10:50	Invited Speech: <i>Probabilistic Models for Rich Linguistic Information</i> (Prof. Mark Johnson)		
10:50 – 11:10	Refreshment Break		
	<i>Taggers, Chunkers, Shallow Parsers – I</i>	<i>Text and Sentence Generation</i>	<i>Machine Translation and Multilinguality – I</i>
11:10 – 11:35	<i>High Speed Unknown Word Prediction Using Support Vector Machine For Chinese Text-to-Speech Systems</i> Juhong Ha, Yu Zheng and Gary Geunbae Lee	<i>Detection of Incorrect Case Assignments in Automatically Generated Paraphrases of Japanese Sentences</i> Atsushi Fujita, Kentaro Inui and Yuji Matsumoto	<i>Automatic Learning of Parallel Dependency Treelet Pairs</i> Yuan Ding and Martha Palmer
11:35 – 12:00	<i>Chinese Chunk Identification Using SVMs plus Sigmoid</i> Yongmei Tan, Tianshun Yao, Qing Chen and Jingbo Zhu	<i>Feature Selection and Machine Learning for Pronominalization</i> Ji-Eun Roh and Jong-Hyeok Lee	<i>Example-based Machine Translation without Saying Inferable Predicate</i> Eiji Aramaki, Sadao Kurohashi, Hideki Kashioka and Hideki Tanaka
12:00 – 13:35	Lunch Break		
	<i>Statistical Models and Machine Learning for NLP – I</i>	<i>NLP Software and Application – I</i>	<i>Panel Discussion</i>
13:35 – 14:00	<i>A Three Level Cache-based Adaptive Chinese Language Model</i> Junlin Zhang, Weimin Qu, Le Sun, Lin Du and Yufang Sun	<i>Automatic Genre Detection of Web Documents</i> Chul Su Lim, Kong Joo Lee and Gil Chang Kim	<i>Panel on Emerging Asian Language Processing Efforts</i>
14:00 – 14:25	<i>Capturing Long Distance Dependency in Language Modeling: An Empirical Study</i> Jianfeng Gao and Hisami Suzuki	<i>Statistical Substring Reduction in Linear Time</i> Xueqiang Lü, Le Zhang and Junfeng Hu	
14:25 – 14:50	<i>Word Folding: Taking the Snapshot of Words Instead of the Whole</i> Jin-Dong Kim and Jun'ichi Tsujii	<i>You don't have to think twice if you carefully tokenize</i> Stefan Klatt and Bernd Bohnet	

14:50 – 15:00	Break		
	<i>Information Retrieval – I</i>	<i>Theories and Formalisms for Morphology, Syntax and Semantics – I</i>	<i>Poster Presentation – I</i>
15:00 – 15:25	<i>Information Flow Analysis with Chinese Text</i> Paulo Cheong, Dawei Song, Peter Bruza and Kam-Fai Wong	<i>The Automatic Acquisition of Verb Subcategorisations and their Impact on the Performance of an HPSG Parser</i> John Carroll and Alex C. Fang	<i>Using a Smoothing Maximum Entropy Model for Chinese Nominal Entity Tagging</i> – Jinying Chen, Nianwen Xue and Martha Palmer
15:25 – 15:50	<i>BBS Based Hot Topic Retrieval Using Back-Propagation Neural Network</i> Lan You, Yongping Du, Jiayin Ge, Xuanjing Huang and Lide Wu	<i>FML-Based SCF Predefinition Learning for Chinese Verbs</i> Xiwu Han, Tiejun Zhao and Muyun Yang	<i>Robust Speaker Identification System Based on Wavelet Transform and Gaussian Mixture Model</i> – Wan-Chen Chen, Ching-Tang Hsieh and Eugene Lai <i>Deterministic dependency structure analyzer for Chinese</i> – Yuchang Cheng, Masayuki Asahara and Yuji Matsumoto
15:50 – 16:20	Refreshment Break		
	<i>Semantic Disambiguation – I</i>	<i>Text Mining in Biomedicine – I</i>	<i>Building a parallel bilingual syntactically annotated corpus</i> – Jan Čuřín, Martin Čmejrek, Jiří Havelka and Vladislav Kuboň
16:20 – 16:45	<i>Influence of WSD on Cross-Language Information Retrieval</i> In-Su Kang, Seung-Hoon Na and Jong-Hyeok Lee	<i>SVM-based Biological Named Entity Recognition using Minimum Edit-Distance Feature Boosted by Virtual Examples</i> Eunji Yi, Gary Geunbae Lee and Soo-Jun Park	<i>Fast Reinforcement Learning of Dialogue Policies using Stable Function Approximation</i> – Matthias Denecke, Kohji Dohsaka and Mikio Nakano
16:45 – 17:10	<i>Improving Word Sense Disambiguation by Pseudo Samples</i> Xiaojie Wang and Yuji Matsumoto	<i>Mining Biomedical Abstracts: What's in a Term?</i> Goran Nenadić, Irena Spasić and Sophia Ananiadou	<i>Selecting Prosody Parameters for Unit Selection Based Chinese TTS</i> – Minghui Dong, Kim-Teng Lua and Jun Xu <i>Making Use of Furigana</i> – Gary Kacmarcik <i>Stochastic Word-Spacing System with Dynamic Increase of Word List</i> – Mi-young Kang, Sung-ja Choi, Ae-sun Yoon and Hyuk-chul Kwon <i>Resolution of Modifier-Head Relation Gaps using Automatically Extracted Metonymic Expressions</i> – Yoji Kiyota, Sadao Kurohashi and Fuyuko Kido <i>Headword Percolation in a Multi-Parser Architecture for Natural Language Understanding</i> – Po Chui Luk, Kui Xu and Helen Meng

			<p><i>Recognition of HTML Table Structure</i> – Hidetaka Masuda, Shuichi Tsukamoto and Hiroshi Nakagawa</p> <p><i>Improving Relevance Feedback in the Language Modeling Approach: Maximum a Posteriori Probability Criterion and Three-component Mixture Model</i> – Seung-Hoon Na, In-Su Kang and Jong-Hyeok Lee</p> <p><i>A Persistent Feature-Object Database for Intelligent Text Archive Systems</i> – Takashi Ninomiya, Jun'ichi Tsujii and Yusuke Miyao</p> <p><i>Improving Quality of the Web Corpus</i> – Youichi Sekiguchi and Kazuhide Yamamoto</p> <p><i>Detecting sentence boundaries in Japanese speech transcriptions using a morphological analyzer</i> – Sachie Tajima, Hidetsugu Nanba and Manabu Okumura</p> <p><i>Improving PinYin to Chinese Conversion with a Whole Sentence Maximum Entropy Model</i> – Le Zhang and Tianshun Yao</p> <p><i>How Effective is Query Expansion for Finding Novel Information?</i> – Min Zhang and Shaoping Ma</p>
17:10 – 17:20	Break		
	<i>Word Segmentation – I</i>	<i>Lexical Semantics, Ontology and Linguistic Resource – I</i>	<i>Information Extraction, Q/A – I</i>
17:20 – 17:45	<p><i>Unsupervised Segmentation of Chinese Corpus Using Accessor Variety</i> Haodi Feng, Kang Chen, Chunyu Kit and Xiaotie Deng</p>	<p><i>Acquiring Bilingual Named Entity Translations from Content-aligned Corpora</i> Tadashi Kumano, Hideki Kashioka, Hideki Tanaka and Takahiro Fukusima</p>	<p><i>A Novel Pattern Learning Method for Open Domain Question Answering</i> Yongping Du, Xuanjing Huang, Xin Li and Lide Wu</p>
17:45 – 18:10	<p><i>Chinese Unknown Word Identification Using Class-based LM</i> Guohong Fu and Kang-Kwong Luke</p>	<p><i>Visual Semantics and Ontology of Eventive Verbs</i> Minhua Ma and Paul Mc Kevitt</p>	<p><i>Chinese Named Entity Recognition Based on Multilevel Linguistic Features</i> Honglei Guo, Jianmin Jiang, Gang Hu and Tong Zhang</p>
19:00	Banquet		

SECOND DAY: 23 March 2004 (Tuesday)

9:00 – 9:50	Invited Speech: <i>The Impact of Information Technology on Communication and Linguistics</i> (Prof. Ching-Chun Hsieh)		
9:50 – 10:10	Refreshment Break		
	<i>Text Mining</i>	<i>Dialogue and Discourse</i>	<i>Natural Language Technology in Mobile IR and Text Processing User Interfaces</i>
10:10 – 10:35	<i>A Study of Semi-Discrete Matrix Decomposition for LSI in Automated Text Categorization</i> Qiang Wang, XiaoLong Wang and Yi Guan	<i>Improving Noun Phrase Coreference Resolution by Matching Strings</i> Xiaofeng Yang, Guodong Zhou, Jian Su and Chew Lim Tan	<i>Dit4dah: Predictive Pruning For Morse Code Text Entry: Towards Entry Systems For the Seriously Impaired</i> Kumiko Tanaka-Ishii and Ian Frank
10:35 – 11:00	<i>Systematic Construction of Hierarchical Classifier in SVM-based Text Categorization</i> Yongwook Yoon, Changki Lee and Gary Geunbae Lee	<i>Zero Pronoun Resolution based on Automatically Constructed Case Frames and Structural Preference of Antecedents</i> Daisuke Kawahara and Sadao Kurohashi	<i>Spoken versus Written Queries for Mobile Information Access: an Experiment on Mandarin Chinese</i> Heather Du and Fabio Crestani
11:00 – 11:25	<i>Categorizing Unknown Text Patterns for Information Extraction Using a Search Result Mining Approach</i> Chien-Chung Huang, Shui-Lung Chuang and Lee-Feng Chien	<i>Combining Labeled and Unlabeled Data for Learning Cross-document Structural Relationships</i> Zhu Zhang and Dragomir Radev	<i>An Interactive Proofreading System for Inappropriately Selected Words on Using Predictive Text Entry</i> Hideya Iwasaki and Kumiko Tanaka-Ishii
11:25 – 11:35	Break		
	<i>Information Retrieval – II</i>	<i>Theories and Formalisms for Morphology, Syntax and Semantics – II</i>	<i>FSA, Parsing Algorithms</i>
11:35 – 12:00	<i>Phoneme-based Transliteration of Foreign Names for OOV Problem</i> Wei Gao, Kam-Fai Wong and Wai Lam	<i>Corpus-oriented Grammar Development for Acquiring a Head-driven Phrase Structure Grammar from the Penn Treebank</i> Yusuke Miyao, Takashi Ninomiya and Jun'ichi Tsujii	<i>Data-Oriented Parsing and the Penn Chinese Treebank</i> Mary Hearne and Andy Way
12:00 – 12:25	<i>Window-based Method for Information Retrieval</i> Qianli Jin, Jun Zhao and Bo Xu	<i>Implementing the Syntax of Japanese Numeral Classifiers</i> Emily M. Bender and Melanie Siegel	<i>Iterative CKY parsing for Probabilistic Context-Free Grammars</i> Yoshimasa Tsuruoka and Jun'ichi Tsujii
12:25 – 13:30	Lunch Break		
13:30	Excursion		

THIRD DAY: 24 March 2004 (Wednesday)

9:00 – 9:50	Invited Speech: <i>Language Technology for E-Memory Applications</i> (Prof. Hans Uszkoreit)		
9:50 – 10:10	Refreshment Break		
	<i>Taggers, Chunkers, Shallow Parsers – II</i>	<i>Information Extraction, Q/A – II</i>	<i>Interactive Poster / Demo Session</i>
10:10 – 10:35	<i>Syntactic Analysis of Long Sentences Based on S-clauses</i> Mi-Young Kim and Jong-Hyeok Lee	<i>Causal Relation Extraction Using Cue Phrase and Lexical Pair Probabilities</i> Du-Seong Chang and Key-Sun Choi	
10:35 – 11:00	<i>A Nearest-Neighbor Method for Resolving PP-Attachment Ambiguity</i> Shaojun Zhao and Dekang Lin	<i>A re-examination of IR techniques in QA system</i> Yi Chang, Hongbo Xu and Shuo Bai	
11:00 – 11:10	Break		
	<i>Semantic Disambiguation – II</i>		
11:10 – 11:35	<i>The Role of Semantic Information in Learning Question Classifiers</i> Xin Li, Dan Roth and Kevin Small		
11:35 – 12:00	<i>Concept-Based Sense Disambiguation for Korean Nouns</i> You-Jin Chung, Kyonghi Moon and Jong-Hyeok Lee		
12:00 – 13:35	Lunch Break		
	<i>Statistical Models and Machine Learning for NLP - II</i>	<i>Word Segmentation – II</i>	<i>Panel Discussion</i>
13:35 – 14:00	<i>Flexible Margin Selection for Reranking with Full Pairwise Samples</i> Libin Shen and Aravind K. Joshi	<i>Chinese New Word Identification Based on Character Parsing Model</i> Yao Meng, Hao Yu and Fumihito Nishino	<i>Panel on Multilingual NLP for Public Information Services (2008 Digital Olympics)</i>
14:00 – 14:25	<i>Comparing Entropies within the Chinese Language</i> Benjamin K Tsou, Tom B Y Lai and Ka-po Chow	<i>The Use of SVM for Chinese New Word Identification</i> Hongqiao Li, Chang-Ning Huang, Jianfeng Gao and Xiaozhong Fan	
14:25 – 14:50	<i>Bilingual Chunk Alignment Based on Interactional Matching and Probabilistic Latent Semantic Indexing</i> Feifan Liu, Qianli Jin, Jun Zhao and Bo Xu	<i>An Example-based Study on Chinese Word Segmentation Using Critical Fragments</i> Qinan Hu, Haihua Pan and Chunyu Kit	
14:50 – 15:00	Break		

	<i>NLP Software and Application – II</i>	<i>Text Mining in Biomedicine – II</i>	<i>Poster Presentation – II</i>
15:00 – 15:25	<i>Natural Language Database Access using Semi-Automatically Constructed Translation Knowledge</i> In-Su Kang, Jae-Hak J. Bae and Jong-Hyeok Lee	<i>Annotation of Gene Products in the Literature with Gene Ontology Terms using Syntactic Dependencies</i> Jung-jae Kim and Jong C. Park	<i>Improving Back-Transliteration by Combining Information Sources</i> – Slaven Bilac and Hozumi Tanaka <i>A Graph Grammar Approach to Map between Dependency Trees and Topological Models</i> – Bernd Bohnet
15:25 – 15:50	<i>Specification Retrieval – How to Find Attribute-Value Information on the Web</i> Minoru Yoshida and Hiroshi Nakagawa	<i>Unsupervised Event Extraction from Biomedical Literature using Co-occurrence Information and Basic Patterns</i> Hong-woo Chun, Young-sook Hwang and Hae-chang Rim	<i>The Hinoki Treebank: A Treebank for Text Understanding</i> – Francis Bond, Sanae Fujita, Chikara Hashimoto, Kaname Kasahara, Shigeko Nariyama, Eric Nichols, Akira Ohtani, Takaaki Tanaka and Shigeaki Amano <i>Chinese Treebanks and Grammar Extraction</i> – Keh-Jiann Chen and Yu-Ming Hsieh
15:50 – 16:20	<i>Refreshment Break</i>		
	<i>Machine Translation and Multilinguality – II</i>	<i>Lexical Semantics, Ontology and Linguistic Resource – II</i>	<i>Using a Paraphraser to Improve Machine Translation Evaluation</i> – Andrew Finch, Yasuhiro Akiba and Eiichiro Sumita <i>Mining Table Information on the Internet</i> – Sung-won Jung, Gi-deuk Han and Hyuk-chul Kwon <i>Parsing Mixed Constructions in a Type Feature Structure Grammar</i> – Jong-Bok Kim and Jaehyung Yang <i>Collecting Evaluative Expressions for Opinion Extraction</i> – Nozomi Kobayashi, Kentaro Inui, Yuji Matsumoto, Kenji Tateishi and Toshikazu Fukushima <i>Deep Analysis of Modern Greek</i> – Valia Kordoni and Julia Neu <i>User Adaptation in MT-mediated Communication</i> – Kentaro Ogura, Yoshihiko Hayashi, Saeko Nomura and Toru Ishida <i>Learning to Filter Junk E-Mail from Positive and Unlabeled Examples</i> – Karl-Michael Schneider
16:20 – 16:45	<i>Bilingual Sentence Alignment Based on Punctuation Statistics and Lexicon</i> Thomas C. Chuang, Jian-Cheng Wu, Tracy Lin, Wen-Chie Shei and Jason S. Chang	<i>A Novel Approach to Improve Word Translations Extraction from Non-Parallel, Comparable Corpora</i> Yun-Chuang Chiao, Jean-David Sta and Pierre Zweigenbaum	
16:45 – 17:10	<i>Practical Translation Pattern Acquisition from Combined Language Resources</i> Mihoko Kitamura and Yuji Matsumoto	<i>Acquiring Selectional Preferences in a Thai Lexical Database</i> Canasai Kruengkrai, Thatsanee Charoenporn, Virach Sornlertlamvanich and Hitoshi Isahara	

			<p><i>A Collaborative Ability Measurement for Co-Training – Dan Shen, Jie Zhang, Jian Su, Guodong Zhou and Chew Lim Tan</i></p> <p><i>Word Sense Disambiguation using Heterogeneous Language Resources – Kiyooki Shirai and Takayuki Tamagaki</i></p> <p><i>A Comparative Study on the Use of Labeled and Unlabeled Data for Large Margin Classifiers – Hiroya Takamura and Manabu Okumura</i></p> <p><i>An English-Hindi Statistical Machine Translation System – Raghavendra Udapa U and Tanveer A Faruque</i></p> <p><i>N-fold Templated Piped Correction – Dekai Wu, Grace Ngai and Marine Carpuat</i></p> <p><i>Tagging Complex NEs with Maxent Models: Layered Structures versus Extended Tagset – Deyi Xiong, Hongkui Yu and Qun Liu</i></p>
17:10 – 17:20	Break		
17:20 – 17:40	Best Paper Award and Closing Session		