2009 International Conference on Natural Language Processing and Knowledge Engineering

(NLP-KE 2009)

Dalian, China
24 – 27 September 2009

IEEE Catalog Number: CFP09811-PRT
ISBN: 978-1-4244-4539-4

TABLE OF CONTENTS

Study on Similarity of Simple Questions Based on the Catering Field 1
Q. Dong, S. Shi, H. Wang, X. Lv
A Conditional Mutual Information Based Selectional Association and Word Sense Disambiguation 6
X. Guo, D. Li, G. Clapworthy
Machine Translation Model Using Inductive Logic Programming 13
A. Hossny, K. Shaalan, A. Fahmy
Using Genetic Algorithm for Persian Grammar Induction 21
M. Arabsorkhi, H. Faili, M. Z. Jahroumi
PARSUMIST: A Persian Text Summarizer 27
M. Shamsfard, T. Akhavan, M. Erfani-Jourabchi
A Feature-Enhanced Smoothing Method for LDA Model Applied to Text Classification 34
D. X. Liu, W. R. Xu, J. N. Hu
Managing the Google Web 1T 5-Gram Data Set 41
A. Islam, D. Inkpen
Extending WordNet With Compound Nouns for Semi-Automatic Annotation in Data Integration Systems 46
D. Beneventano, S. Bergamaschi, S. Sorrentino
A Method For Stemming and Eliminating Common Words for Persian Text Summarization 54
M. Berenjkoob, R. Mehri, H. Khosravi, M. Ali
Integrating Divergent Models for Gene Mention Tagging 60
L. Li, R. Zhou, D. Huang, W. Liao
Recognizing Sentence Emotions Based on Polynomial Kernel Method Using Ren-CECps 67
C. Quan, F. Ren
The Picture Superiority Effect in Encoding and Retrieval Processes During Japanese Learning for Chinese Bilinguals 74
L. P. Mi, X. Liu, F. Ren
Simple Linguistic Processing Effect on Multi-Label Emotion Classification 82
Y. Wu
Hyponymy Acquisition from Chinese Text by SVM 87
F. Tian, F. Ren
Analysis on Degree Words for Chinese Emotion Expressions Based on Syntactic Parse and Rules 93
Y. Sun, C. Quan, F. Ren
Chinese Semantic Role Labeling Using CRFs and SVMs 99
Y. Tan, X. Wang, Y. Chen
Real-Word Spelling Correction Using Google Web 1T n-gram with Backoff 104
A. Islam, D. Inkpen
Unknown Place name Detection Base on YamCha for Japanese Guidance QA System 112
Y. Liu, F. Ren
Extended Super Function Based Chinese Japanese Machine Translation 117
X. Sun, F. Ren, D. Huang
Existential Negative Sentence Translation in Japanese-Chinese Machine Translation 125
H. Zhang, F. Ren
The Study on Text Emotional Orientation Based on a Three-Dimensional Emotion Space Model 130
H. Li, F. Ren
An Information Arrangement Technique for a Text Classification and Summarization Based on a Summarization Frame 136
S. Tsuchiya, E. Yoshimura, H. Watabe
Answer Generation for Chinese Cuisine QA System 141
L. Xia, Z. Teng, F. Ren
A Proposal for Creating a Chinese Emotion Thesaurus with Tag of Emotion Intesity 147
J. Li, F. Ren
A Study on Cross-Language Text Summarization Using Supervised Methods 155
L. Yu, F. Ren
Detect the Possible Spokesperson With an Omni-Directional Camera, in a Robot-Human Communication System 162
J. Ma, M. Suzuki, F. Ren

Towards the Relationship Between Semantic Web and NLP 167
R. Guo, F. Ren
Searching Semantically Similar Questions from a Large Community-Based Question Archive 175
M. Liu, Y. Liu, Q. Yang
Research on Knowledge Elements in Exponential Language Model 183
H. Jiang, X. Wang
Construction of English Composition Support System Based on Conveying Emotion 188
K. Matsumoto, J. Matsumoto, F. Ren
Chinese Text Orientation Analysis Based on Phrase 196
Y. Guo, Y. Zhou
Influence on Emotional Impression of Voice by Changing Prosodic Features 202
A. Sasaki, M. Adachi, M. Suzuki, R. Fuji
Designing a Japanese Idiom Education Support System for Overseas' Students 209
K. Yusuke, R. Li, F. Ren
Web Mining Based on VIPS in Intention-Based Information Retrieval 217
Q. Zhang, X. Jiang, J. Sun
Automatic Generation of Narrative Content for Digital Games 222
M. F. Caropreso, D. Inkpen, S. Khan, F. Keshtkar
Knowlege Element Analogy Relation Recognition Using Text and Graph Structure 230
W. Wang, Q. Zheng, Y. Chen
Chinese Coding Type Identification Based on Sub-Sentence Length Observation 238
G. Hi, P. Peng, X. Wu, L. Chen
Incorporating Multi-Task Learning in Conditional Random Fields for Chunking in Semantic Role Labeling 243
S. He, T. Zhang, X. Bai, X. Wang, Y. Dong
A New Evaluating Method for Chinese Text Summarization Not Requiring Model Summary 248
C. Wang, L. Li, Y. Zhong
Multi-Class Bootstrapping Learning Aspect-Related Terms for Aspect Identification 255
C. Zhang, J. Zhu
A Method of Building Chinese Field Association Knowledge from Wikipedia 261
L. Wang, S. Yata, E. S. Atlam, M. Fuketa, K. Morita, H. Bando, J. I. Aoe
Sentiment Classification for Chinese Reviews Based on Key Substring Features 266
Z. Zhai, H. Xu, J. Li, P. Jia
Research of the Method for Obtaining Semantic Knowledge in the System of Ci-Nlu Based Text Processing After Speech Recognition 274
J. Wang, L. Li, L. Long
Factors Influencing Dependency Parsing of Coordinating Structure 280
Y. Zhao, S. Gao, H. Liu
Building A Dictionary on Constituent Structure of Chinese Compounds 286
L. Qiu, X. Zhang, L. Mao
Analysis About Event Annotation and Information Structure in Suden Events Discourse 294
E. Yang, Q. Zeng, D. Zhu
Adaptation of a Named Entity Recognition System for the ESTER 2 Evaluation Campaign 299
C. Brun, M. Ehrmann
Japanese Dependency Analysis Using Fuzzy Support Vector Machines 307
H. Zhou, D. Huang, T. Yu
Opinion Analysis Based on a Fusion of Multiple Classifiers Approach 314
B. Xu, T. J. Zhao, D. Q. Zheng, Q. X. Chen
Chinese Sentence Similarity Based on Word Context and Semantic 319
T. Gu, F. Ren
Document Representation Combining Concepts and Words in Chinese (Text Categorization) 324
C. Che, H. Teng
A Fast Retrieval Algorithm for the Earth Mover's Distance Using EMD Lower Bounds and the Priority Queue 329
M. Shishibori, D. Koizumi, K. Kita
Multidimensional Concept Extraction and Validation from OLAP Requirements in NL 335
F. Bargui, H. Ben-Abdallah, J. Feki
An Ontology-Based Question Answering Method with the Use of Textual Entailment 343
S. Ou, D. Mekhaldi, C. Orasan
Comparison Between Two Arabic Tagsets 351
M. A. A. Rashwan, E. A. H. Khalil, A. Rafea

Chinese Maximal Noun Phrase Parsing Based on Cascaded Conditional Random Fields 359
D. Cai, X. Liu, Q. Zhou, N. Ye
Topic Control in A Free Conversation System 366
X. Song, K. Maeda, H. Kunimasa, H. Toyota, D. Han
Extracting Historical Terms Based on Aligned Chinese-English Parallel Corpora 372
X. Li, C. Che, L. Han, X. Liu
Term Relevance Estimation of Chinese Query Expansion 378
X. Tu, T. He, J. Luo
A Statistical Method for Uyghur Tokenization 385
B. Aisha, M. Sun
Classification of Persian Textual Documents Using Learning Vector Quanization 390
M. T. Pivevar, H. Feili, M. Soltani
Reader Emotion Classification of News Headlines 396
Y. Jia, Z. Chen, S. Yu
A Bi-Threshold Model for PP-Attachment Disambiguation Through Backing Off to 2 - Tuples Directly 402
B. Liao, H. Luo, L. Xiao
Translate Chinese Organization names Using Examples and Web 407
F. Ren, J. Zhu, H. Wang
Automatic Text Classification Using Modified Centroid Classifier 414
M. Elmarhumy
Entity Relation Extraction to Free Text 418
S. Zhang
Cassiopeia: A Model Based on Summarization and Clusterization Used for Knowledge Discovery in Textual Bases 423
M. V. C. Guelpeli, A. H. Branco, A. C. B. Garcia
A Weakly Supervised Optimize Method in Latent Semantic Indexing 431
D. Ji, D. Guo, D. Cai, Y. Bai
Research on Katakana Phrase Translation Based on Bi-Directional Integration 438
G. Zhang, Y. Gao, D. Ji, X. Ren
An Efficient Collaboration Filtering Algorithm Using SVD-Free Latent Semantic Indexing and Particle Swarm Optimization 444
A. Abdelwahab, H. Sekiya, M. Nishida
Automatic Evaluation of Translation Quality Using Expanded N-Gram Co-Occurrence 449
Y. Qin, Q. Wen, J. Wang
HMM - Based Phonemic Distance in Different Speaking Styles and Its Influence on Substitutions in Mandarin Speech Recongition 454
Z. Yang, W. Liu, Z. Lv
Studies on Automatic Recognition of Chinese Adverb CAI's Usages Based on Statistics 459
H. Zan, J. Zhang
Applying Tree Kernel to Chinese Syntactic Parsing Reranking 464
X. Zheng, L. Chen, B. Chang
An N-Gram Based Chinese Syllable Evaluation Approach for Speech Recognition Error Detection 470
X. Wang, L. Li
A New Language Model Adaptation Framework Using Modification of Structures of Background Corpus and Language Model 476
Z. Lv, W. Liu, Z. Yang
Contextual Binarization for Synatx-Based Machine Translation 480
Q. Chen, T. Yao
Generating and Mixing Feature Sets From Language Models for Sentiment Classification 485
Y. Jeong, Y. Kim, S. Kim, S. H. Myaeng, H. J. Oh
Perception of Syllables Pitch Contour in Sindhi Language 493
J. A. Mahar, G. Q. Memon, H. A. Shah
A Template-Based Approach to Extract Product Feartures and Sentiment Words 498
W. Zhao, Y. Zhou
Researches on Algorithm for Confidence Evaluation and Decision Modification of SVM 503
X. Zhao, Y. Zhou, H. He
Stochastic Arabic Hybrid Diacritizer 510
M. Rashwan, M. A. Badrashiny, M. Attia, S. Abdou, A. Rafea
A Novel Interpolated N-Gram Language Model Based on Class Hierarchy 518
Z. Lv, W. Liu, Z. Yang

A Method and Application of Automatic Term Extraction Using Conditional Random Fields 523
W. Fu, L. Li
An Unsupervised Method for Lexical Acquisition Based on Bootstrapping 528
Y. Zhang, Y. Zhou
Fuzzy Modeling to Conceptual Design of Mechanisms Based on Axiomatic Design 535
C. I. Cai, Y. W. Bai, Y. F. He, Z. M. Zhang
Protein-Protein Interaction Extraction From Biomedical Literatures Based on Modified SVM-KNN 542
L. Li, L. Jing, D. Huang
Data Extraction From Web Forums Based on Similarity of Page Layout 549
Y. Wang, B. Li, C. Lin
CCRFs: Cascaded Conditional Random Fields For Chinese POS Tagging 554
G. Shen, L. Qiu, C. Hu, K. Zhao
Using Sentiment Orientation Features for Mood Classification in Blogs 562
F. Keshtkar, D. Inkpen
A Computer-Assisted Dictionary-Making System for Chinese English Learner's Dictionary 568
W. Xiong, G. Chen
Ikhtasir - A User Selected Compression Ratio Arabic Text Summarization System 573
A. Azmi, S. Al-Thanyyan
Analyzing the Seventh Vowel of Classical Arabic 580
M. O. AlQahtany, Y. A. Alotaibi, S. A. Selouani
An Arabic Question - Answering System for Factoid Questions 587
W. Brini, M. Ellouze, S. Mesfar, L. Hadrich-Belguith, L. Miracl
Effect of Gaussian Densities and Amount of Training Data on Grapheme-Based Acoustic Modeling for Arabic 594
M. Elmahdy, R. Gruhn, W. Minker, S. Abdennadher
A Hybrid Approach for Arabic Multi-Word Term Extraction 599
I. Bounhas, Y. Slimani
Author Index