

Table of Contents

Session 1

<i>TransType2 - An Innovative Computer-Assisted Translation System</i> José Esteban, José Lorenzo, Antonio S. Valderrábanos and Guy Lapalme	1
<i>Improving Domain-Specific Word Alignment for Computer Assisted Translation</i> Wu Hua and Wang Haifeng	5
<i>Constructing Transliteration Lexicons from Web Corpora</i> Jin-Shea Kuo and Ying-Kuei Yang	9
<i>Subsentential Translation Memory for Computer Assisted Writing and Translation</i> Jian-Cheng Wu, Thomas C. Chuang, Wen-Chi Shei and Jason S. Chang	13
<i>Customizing Parallel Corpora at the Document Level</i> Monica Rogati and Yiming Yang	17
<i>An Automatic Filter for Non-Parallel Texts</i> Chris Pike and I. Dan Melamed	21
<i>Exploiting Aggregate Properties of Bilingual Dictionaries For Distinguishing Senses of English Words and Inducing English Sense Clusters</i> Charles Schafer and David Yarowsky	25
<i>Interactive grammar development with WCDG</i> Kilian A. Foth, Michael Daum and Wolfgang Menzel	29
<i>Wide Coverage Symbolic Surface Realization</i> Charles Callaway	33
<i>Part-of-Speech Tagging Considering Surface Form for an Agglutinative Language</i> Do-Gil Lee and Hae-Chang Rim	37
<i>Is Conceptual Combination Influenced by Word Order?</i> Phil Maguire and Arthur Cater	41
<i>Corpus representativeness for syntactic information acquisition</i> Núria Bel	45
<i>Exploiting Unannotated Corpora for Tagging and Chunking</i> Rie Kubota Ando	49
<i>Improving Bitext Word Alignments via Syntax-based Reordering of English</i> Elliott Franco Drabek and David Yarowsky	53

Session 2

<i>Hierarchy Extraction based on Inclusion of Appearance</i> Eiko Yamamoto, Kyoko Kanzaki and Hitoshi Isahara	57
<i>Knowledge intensive e-mail summarization in CARPANTA</i> Laura Alonso, Irene Castellón, Bernardino Casas and Lluís Padró.....	61
<i>Finding Anchor Verbs for Biomedical IE Using Predicate-Argument Structures</i> Akane Yakushiji, Yuka Tateisi, Yusuke Miyao and Jun'ichi Tsujii	65
<i>Resource Analysis for Question Answering</i> Lucian Vlad Lita, Warren A. Hunt and Eric Nyberg	69
<i>TANGO: Bilingual Collocational Concordancer</i> Jia-Yan Jian, Yu-Chia Chang and Jason S. Chang	73
<i>Graph-based Ranking Algorithms for Sentence Extraction, Applied to Text Summarization</i> Rada Mihalcea	77
<i>Compiling Boostexter Rules into a Finite-state Transducer</i> Srinivas Bangalore	81
<i>Combining Lexical, Syntactic, and Semantic Features with Maximum Entropy Models for Information Extraction</i> Nanda Kambhatla	85
<i>On the Equivalence of Weighted Finite-state Transducers</i> Julien Quint.....	89
<i>A New Feature Selection Score for Multinomial Naive Bayes Text Classification Based on KL-Divergence</i> Karl-Michael Schneider	93

Session 3

<i>Incorporating topic information into semantic analysis models</i> Tony Mullen and Nigel Collier	97
<i>A Practical Solution to the Problem of Automatic Word Sense Induction</i> Reinhard Rapp.....	101
<i>Automatic clustering of collocation for detecting practical sense boundary</i> Saim Shin and Key-Sun Choi	105
<i>Co-training for Predicting Emotions with Spoken Dialogue Data</i> Beatriz Maeireizo, Diane Litman and Rebecca Hwa	109
<i>Multimodal Database Access on Handheld Devices</i> Elsa Pecourt and Norbert Reithinger	113
<i>Wysiwym with wider coverage</i> Richard Power and Roger Evans	117

<i>NLTK: The Natural Language Toolkit</i> Steven Bird and Edward Loper	121
<i>Dyna: A Language for Weighted Dynamic Programming</i> Jason Eisner, Eric Goldlust and Noah A. Smith	125
<i>MATCHkiosk: A Multimodal Interactive City Guide</i> Michael Johnston and Srinivas Bangalore	129
<i>Fragments and Text Categorization</i> Jan Blaták, Eva Mráková and Lubos Popelínsky	133