Contents

Syntax and Syntactic Analysis

Pointwise Prediction and Sequence-Based Reranking for Adaptable Part-of-Speech Tagging .......................... 3
Shinsuke Mori, Yosuke Nakata, Graham Neubig, and Tetsuro Sasada

Semantics and Semantic Analysis

Japanese Semantic Role Labeling with Hierarchical Tag Context Trees .... 21
Yasuhiro Ishihara and Koichi Takeuchi

Spoken Language and Dialogue

Arabic Dialect Identification Using a Parallel Multidialectal Corpus ......... 35
Shervin Malmasi, Eshrag Refaee, and Mark Dras

Filled Pause Detection in Indonesian Spontaneous Speech ................. 54
Auliya Sani, Dessi Puji Lestari, and Ayu Purwarianti

Automatic Extraction Phonetically Rich and Balanced Verses for Speaker-Dependent Quranic Speech Recognition System .......... 65
Rahmi Yuwan and Dessi Puji Lestari

Corpora and Corpus-Based Language Processing

3-Step Parallel Corpus Cleaning Using Monolingual Crowd Workers ....... 79
Toshiaki Nakazawa, Sadao Kurohashi, Hayato Kobayashi, Hiroki Ishikawa, and Manabu Sassano

Text and Message Understanding

Active Learning to Remove Source Instances for Domain Adaptation for Word Sense Disambiguation .................................. 97
Hiroyuki Shinnou, Yoshiyuki Onodera, Minoru Sasaki, and Kanako Komiya

Context Representation with Word Embeddings for WSD ............... 108
Hiromu Sugawara, Hiroya Takamura, Ryohei Sasano, and Manabu Okumura
Contents

Information Extraction and Text Mining

Location Mention Detection in Tweets and Microblogs
Shervin Malmasi and Mark Dras

123

Recognizing and Normalizing Temporal Expressions in Indonesian Texts
Paramita Mirza

135

Named Entity Recognizer Trainable from Partially Annotated Data
Tetsuro Sasada, Shinsuke Mori, Tatsuya Kawahara, and Yoko Yamakata

148

Relation Extraction Using Semantic Information
Jian Xu, Qin Lu, and Minglei Li

161

Incorporating Tweet Relationships into Topic Derivation
Robertus Nugroho, Diego Molla-Aliod, Jian Yang, Youliang Zhong,
Cecile Paris, and Surya Nepal

177

Information Retrieval and Question-Answering

Detecting Vital Documents Using Negative Relevance Feedback in
Distributed Realtime Computation Framework
Shun Kawahara, Kazuhiro Seki, and Kuniaki Uehara

193

Language Learning

Measuring Readability for Learners of English as a Foreign Language by
Linguistic and Learner Features
Katsunori Kotani and Takehiko Yoshimi

211

Machine Translation

Machine Translation Method Based on Non-compositional Semantics
(Word-Level Sentence-Pattern-Based MT)
Jun Sakata, Jin’ichi Murakami, Masato Tokuhisa, and Masaki Murata

225

The Application of Phrase Based Statistical Machine Translation
Techniques to Myanmar Grapheme to Phoneme Conversion
Ye Kyaw Thu, Win Pa Pa, Andrew Finch, Jinfu Ni, Eiichiro Sumita,
and Chiori Hori

238

Collecting Bilingual Technical Terms from Japanese-Chinese Patent
Families by SVM
Lijuan Dong, Zi Long, Takehito Utsuro, Tomoharu Mitsuhashi,
and Mikio Yamamoto

251

Author Index

263
Computational Linguistics
14th International Conference of the Pacific Association for Computational Linguistics, PACLING 2015, Bali, Indonesia, May 19-21, 2015, Revised Selected Papers
Hasida, K.; Purwarianti, A. (Eds.)
2016, X, 263 p. 82 illus., Softcover