12th International Conference on Natural Language Processing (ICON-2015)

Trivandrum, India
11 – 14 December 2015

ISBN: 978-1-5108-3787-4

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571

Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2015) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2017)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

Keynote Lecture 1: Scientific Paper Analysis
Yuji Matsumoto . . . . . .

Addressing Class Imbalance in Grammatical Error Detection with Evaluation Metric Optimization
Anoop Kunchukuttan and Pushpak Bhattacharyya . . . . . .

Words are not Equal: Graded Weighting Model for Building Composite Document Vectors
Pranjal Singh and Amitabha Mukerjee . . . . . .

Online Adspace Posts' Category Classification
Dhawal Joharapurkar, Vaishak Salin and Vishal Krishna . . . . . .

Noun Phrase Chunking for Marathi using Distant Supervision
Sachin Pawar, Nitin Ramrakhiyani, Girish K. Palshikar, Pushpak Bhattacharyya and Swapnil Hingmire . . . . . .

Self-Organizing Maps for Classification of a Multi-Labeled Corpus
Lars Bungum and Björn Gambäck . . . . . .

Word Sense Disambiguation in Hindi Language Using Hyperspace Analogue to Language and Fuzzy C-Means Clustering
Devendra K. Tayal, Leena Ahuja and Shreya Chhabra . . . . . .

Using Word Embeddings for Bilingual Unsupervised WSD
Sudha Bhingardive, Dhirendra Singh, Rudramurthy V and Pushpak Bhattacharyya . . . . . .

Compositionality in Bangla Compound Verbs and their Processing in the Mental Lexicon
Tirthankar Dasgupta, Manjira Sinha and Anupam Basu . . . . . .

IndoWordNet Dictionary: An Online Multilingual Dictionary using IndoWordNet
Hanumant Redkar, Sandhya Singh, Nilesh Joshi, Anupam Ghosh and Pushpak Bhattacharyya .

Let Sense Bags Do Talking: Cross Lingual Word Semantic Similarity for English and Hindi
Apurva Nagvenkar, Jyoti Pawar and Pushpak Bhattacharyya . . . . . .

A temporal expression recognition system for medical documents by
Naman Gupta, Aditya Joshi and Pushpak Bhattacharyya . . . . . .

An unsupervised EM method to infer time variation in sense probabilities
Martin Emms and Arun Jayapal . . . . . .

Solving Data Sparsity by Morphology Injection in Factored SMT
Sreelekha S, Piyush Dungarwal, Pushpak Bhattacharyya and Malathi D . . . . . .

Authorship Attribution in Bengali Language
Shanta Phani, Shibamouli Lahiri and Arindam Biswas . . . . . .

TransChat: Cross-Lingual Instant Messaging for Indian Languages
Diptesh Kanojia, Shehzaad Dhuliawala, Abhijit Mishra, Naman Gupta and Pushpak Bhattacharyya

A Database of Infant Cry Sounds to Study the Likely Cause of Cry
Shivam Sharma, Shubham Asthana and V. K. Mittal . . . . . .

Perplexed Bayes Classifier
Cohan Sujay Carlos . . . . . .

An Empirical Study of Diversity of Word Alignment and its Symmetrization Techniques for System Combination
Thoudam Doren Singh . . . . . .

Domain Sentiment Matters: A Two Stage Sentiment Analyzer
Raksha Sharma and Pushpak Bhattacharyya . . . . . .

Extracting Information from Indian First Names
Akshay Gulati . . . . . .

punct-An Alternative Verb Semantic Ontology Representation
Kavitha Rajan . . . . . .

SMT Errors Requiring Grammatical Knowledge for Prevention
Yukiko Sasaki Alam . . . . . .

Isolated Word Recognition System for Malayalam using Machine Learning
Maya Moneykumar, Elizabeth Sherly and Win Sam Varghese . . . . . .

Judge a Book by its Cover: Conservative Focused Crawling under Resource Constraints
Shehzaad Dhuliawala, Arjun Atreya V, Ravi Kumar Yadav and Pushpak Bhattacharyya . . . . . .

Text Normalization and Unit Selection for a Memory Based Non Uniform Unit Selection TTS in Malayalam
Gokul P., Neethu Thomas, Crisil Thomas and Dr. Deepa P. Gopinath . . . . . .

Morphological Analyzer for Gujarati using Paradigm based approach with Knowledge based and Statistical Methods
Jatayu Baxi, Pooja Patel and Brijesh Bhatt . . . . . .

Resolution of Pronominal Anaphora for Telugu Dialogues
Hemanth Reddy Jonnalagadda and Radhika Mamidi . . . . . .

A Study on Divergence in Malayalam and Tamil Language in Machine Translation Perceptive
Jisha P Jayan and Elizabeth Sherly . . . . . .

Automatic conversion of Indian Language Morphological Processors into Grammatical Framework (GF)
Harsha Vardhan Grandhi and Soma Paul . . . . . .

Logistic Regression for Automatic Lexical Level Morphological Paradigm Selection for Konkani Nouns
Shilpa Desai, Jyoti Pawar and Pushpak Bhattacharyya . . . . . .

Ruchi: Rating Individual Food Items in Restaurant Reviews
Burusothman Ahiladas, Paraneetharan Saravanaperumal, Sanjith Balachandran, Thamayanthy Sripalan and Surangika Ranathunga . . . . . .

Dependency Extraction for Knowledge-based Domain Classification
Lokesh Kumar Sharma and Namita Mittal . . . . . .

An Approach to Collective Entity Linking
Ashish Kulkarni, Kanika Agarwal, pararth Shah, Sunny Raj Rathod and Ganesh Ramakrishnan

Development of Speech corpora for different Speech Recognition tasks in Malayalam language
Cini Kurian . . . . . .

POS Tagging of Hindi-English Code Mixed Text from Social Media: Some Machine Learning Experiments
Royal Sequiera, Monojit Choudhury and Kalika Bali . . . . . .

Automated Analysis of Bangla Poetry for Classification and Poet Identification
Geetanjali Rakshit, Anupam Ghosh, Pushpak Bhattacharyya and Gholamreza Haffari . . . . . .

Sentence Boundary Detection for Social Media Text
Dwijen Rudrapal, Anupam Jamatia, Kunal Chakma, Amitava Das and Björn Gambäck . . . . . .

Mood Classification of Hindi Songs based on Lyrics
Braja Gopal Patra, Dipankar Das and Sivaji Bandyopadhyay . . . . . .

Using Skipgrams, Bigrams, and Part of Speech Features for Sentiment Classification of Twitter Messages
Badr Mohammed Badr and S. Sameen Fatima . . . . . .

A Hybrid Approach for Bracketing Noun Sequence
Arpita Batra and Soma Paul . . . . . .

Simultaneous Feature Selection and Parameter Optimization Using Multi-objective Optimization for Sentiment Analysis
Mohammed Arif Khan, Asif Ekbal and Eneldo Loza Mencía . . . . . .

Detection of Multiword Expressions for Hindi Language using Word Embeddings and WordNet-based Features
Dhirendra Singh, Sudha Bhingardive, Kevin Patel and Pushpak Bhattacharyya . . . . . .

Augmenting Pivot based SMT with word segmentation
Rohit More, Anoop Kunchukuttan, Pushpak Bhattacharyya and Raj Dabre . . . . . .

Using Multilingual Topic Models for Improved Alignment in English-Hindi MT
Diptesh Kanojia, Aditya Joshi, Pushpak Bhattacharyya and Mark James Carman . . . . . .

Triangulation of Reordering Tables: An Advancement Over Phrase Table Triangulation in Pivot-Based SMT
Deepak Patil, Harshad Chavan and Pushpak Bhattacharyya . . . . . .

Post-editing a chapter of a specialized textbook into 7 languages: importance of terminological proximity with English for productivity
Ritesh Shah, Christian Boitet, Pushpak Bhattacharyya, Mithun Padmakumar, Leonardo Zilio, Ruslan Kalitvianski, Mohammad Nasiruddin, Mutsuko Tomokiyo and Sandra Castellanos Páez . . . . . .

Generating Translation Corpora in Indic Languages:Cultivating Bilingual Texts for Cross Lingual Fertilization
Niladri Sekhar Dash, Arulmozi Selvraj and Mazhar Hussain . . . . . .

Translation Quality and Effort: Options versus Post-editing
Donald Sturgeon and John S. Y. Lee . . . . . .

Investigating the potential of post-ordering SMT output to improve translation quality
Pratik Mehta, Anoop Kunchukuttan and Pushpak Bhattacharyya . . . . . .

Applying Sanskrit Concepts for Reordering in MT
Akshar Bharati, , Prajna Jha, Soma Paul and Dipti M Sharma . . . . . .

Dialogue Act Recognition for Text-based Sinhala
Sudheera Palihakkara, Dammina Sahabandu, Ahsan Shamsudeen, Chamika Bandara and Surangika Ranathunga . . . . . .

A Semi Supervised Dialog Act Tagging for Telugu
Suman Dowlagar and Radhika Mamidi . . . . . .

Ranking Model with a Reduced Feature Set for an Automated Question Generation System
Manisha Satish Divate and Ambuja Salgaonkar . . . . . .

Natural Language Processing for Solving Simple Word Problems
Sowmya S Sundaram and Deepak Khemani . . . . . .

Analysis of Influence of L2 English Speakers' Fluency on Occurrence and Duration of Sentence-medial Pauses in English Readout Speech
Shambhu Nath Saha and Shyamal Kr. Das Mandal . . . . . .

Acoustic Correlates of Voicing and Gemination in Bangla
Aanusha Ghosh . . . . . .