

ICON-2015

**12th International  
Conference on Natural  
Language Processing**

**Proceedings of the Conference**

11-14 December 2015  
IIITM-Kerala, Trivandrum, India

© 2015 NLP Association of India (NLPAI)

## Preface

Research in Natural Language Processing (NLP) has taken a noticeable leap in the recent years. Tremendous growth of information on the web and its easy access has stimulated large interest in the field. India with multiple languages and continuous growth of Indian language content on the web makes a fertile ground for NLP research. Moreover, industry is keenly interested in obtaining NLP technology for mass use. The internet search companies are increasingly aware of the large market for processing languages other than English. For example, search capability is needed for content in Indian and other languages. There is also a need for searching content in multiple languages, and making the retrieved documents available in the language of the user. As a result, a strong need is being felt for machine translation to handle this large instantaneous use. Information Extraction, Question Answering Systems and Sentiment Analysis are also showing up as other business opportunities.

These needs have resulted in two welcome trends. First, there is much wider student interest in getting into NLP at both postgraduate and undergraduate levels. Many students interested in computing technology are getting interested in natural language technology, and those interested in pursuing computing research are joining NLP research. Second, the research community in academic institutions and the government funding agencies in India have joined hands to launch consortia projects to develop NLP products. Each consortium project is a multi-institutional endeavour working with a common software framework, common language standards, and common technology engines for all the different languages covered in the consortium. As a result, it has already led to development of basic tools for multiple languages which are inter-operable for machine translation, cross lingual search, hand writing recognition and OCR.

In this backdrop of increased student interest, greater funding and most importantly, common standards and interoperable tools, there has been a spurt in research in NLP on Indian languages whose effects we have just begun to see. A great number of submissions reflecting good research is a heartening matter. There is an increasing realization to take advantage of features common to Indian languages in machine learning. It is a delight to see that such features are not just specific to Indian languages but to a large number of languages of the world, hitherto ignored. The insights so gained are furthering our linguistic understanding and will help in technology development for hopefully all languages of the world.

For machine learning and other purposes, linguistically annotated corpora using the common standards have become available for multiple Indian languages. They have been used for the development of basic technologies for several languages. Larger set of corpora are expected to be prepared in near future.

This volume contains papers selected for presentation in technical sessions of ICON-2015 and short communications selected for poster presentation. We are thankful to our excellent team of reviewers from all over the globe who deserve full credit for the hard work of reviewing the high quality submissions with rich technical content. From 134 submissions, 56 papers were selected, 31 for full presentation and 25 for poster presentation, representing a variety of new and interesting developments, covering a wide spectrum of NLP areas and core linguistics.

We are deeply grateful to *Yuji Matsumoto, Nara Institute of Science and Technology (NAIST), Japan* for giving the keynote lecture at ICON. We would also like to thank the members of the Advisory Committee and Programme Committee for their support and co-operation.

We thank Sudip Kumar Naskar, Chair, Student Paper Competition and Manish Shrivastava and Amitav Das, Chairs, NLP Tools Contest for taking the responsibilities of the events.

We convey our thanks to P V S Ram Babu, G Srinivas Rao, G Namratha and A Lakshmi Narayana, International Institute of Information Technology (IIIT), Hyderabad for their dedicated efforts in successfully handling the ICON Secretariat. We also thank IIIT Hyderabad team of Peri Bhaskararao, Vasudeva Varma, Soma Paul, Radhika Mamidi, Manish Shrivastava, B Yegnanarayana, Suryakanth V Gangashetty and Anil Kumar Vuppala. We heart-fully express our gratitude to Rajeev R R, Maya Moneykumar, VRCLC team members, Research Scholars and student volunteers for their timely help with sincere dedication to make this conference a success.

We also thank all those who came forward to help us in this task.

Finally, we thank all the researchers who responded to our call for papers and all the participants of ICON-2015, without whose overwhelming response the conference would not have been a success.

December 2015  
Trivandrum

Dipti Misra Sharma  
Rajeev Sangal  
Elizabeth Sherly

**Advisory Committee:**

Aravind K Joshi, University of Pennsylvania, USA (Chair)

**Conference General Chair:**

Rajeev Sangal, IIT (BHU), Varanasi, India

**Programme Committee:**

Elizabeth Sherly, IIITM-Kerala, Trivandrum, India (Chair)

Dipti Misra Sharma, IIIT Hyderabad, India (Co-Chair)

**Tools Contest Chairs:**

Manish Shrivastava, IIIT Hyderabad, India

Amitav Das, NIIT University, Rajasthan, India

**Organizing Committee:**

Rajeev R R, IIITM-K, Trivandrum, India (Chair)



## Organized by



International Institute of Information  
Technology, Hyderabad



Natural Language Processing  
Association, India



IIITM-Kerala, Trivandrum



LDC-IL, CIIL Mysore

## Sponsors



Microsoft Research, India



Kerala State Council for Science,  
Technology & Environment





# Referees

We gratefully acknowledge the excellent quality of refereeing we received from the reviewers. We thank them all for being precise and fair in their assessment and for reviewing the papers in time.

A Kumaran	Jim Maddock	Royal Sequeira
A R Balamurali	Joakim Nivre	Sachin Pawar
Abhijit Mishra	Jyoti Pareek	Samar Husain
Aditi Sharan	Jyoti Pawar	Sandipan Dandapat
Aditya Joshi	K V Subbarao	Sanjukta Ghosh
Ajit Kumar	Kalika Bali	Santanu Pal
Alok Parlikar	Kamal Garg	Satarupa Guha
Amba Kulkarni	Keh-Yih Su	Shashi Narayan
Amitava Das	Kishorjit	Shruti Rijhwani
Anandaswarup	Nongmeikapam	Silpa Kanneganti
Vadapalli	Kunal Chakma	Sivaji Bandyopadhyay
Anil Kumar Singh	Lars Bungum	Sivanand Achanta
Anil Kumar Vuppala	Litton Kurisinkel	Sobha L
Anil Thakur	Maaz Anwar	Soma Paul
Aniruddha Tammewar	Maite Giménez	Somnath Banerjee
Anoop Kunchukuttan	Malhar Kulkarni	Sopan Kolte
Anupam Jamatia	Manish Shrivastava	Srinivas Bangalore
Anupam Mondal	Matthias Huck	Sriram Venkatapathy
Aravind Ganapathiraju	Monojit Choudhury	Subhash Chandra
Ashwini Vaidya	Mounika K V	Sudip Kumar Naskar
Asif Ekbal	N Vasudevan	Sunayana Sitaram
Ayushi Dalmia	Neha Prabhugaonkar	Suryakanth V
Ayushi Pandey	Nicoletta Calzolari	Gangashetty
B Bajibabu	Nikhil Pattisapu	Sutanu Chakraborti
Balaji Jagan	Nikhilesh Bhatnagar	Swapnil Chaudhari
Bharat Ram Ambati	Niladri Chatterjee	Tapabrata Mondal
Bharathi Raja Asoka	Niladri Sekhar Dash	Tejas Godambe
Chakravarthi	Owen Rambow	Thamar Solorio
Bhaskararao Peri	Paolo Rosso	Thoudam Doren Singh
Bhuvana Narasimhan	Parminder Singh	Umamaheswari E
Bira Chandra Singh	Parth Gupta	Vandan Mujadia
Bjorn Gambäck	Partha Talukdar	Vasudeva Varma
Bonnie Webber	Pattabhi Rao	Vigneshwaran
Braja Gopal Patra	Pawan Goyal	Muralidaran
Brijesh Bhatt	Pranaw Kumar	Vijaysundar Ram
C V Jawahar	Prateek Bhatia	Vinay Kumar Mittal
Debasis Ganguly	Preethi Raghavan	Vineet Chaitanya
Deepak Padmanabhan	Priya Radhakrishnan	Vishal Goyal
Dhananjaya Gowda	Pruthwik Mishra	
Dipankar Das	Pushpak	
Dipti Misra Sharma	Bhattacharyya	
Dwijen Rudrapal	Radhika Mamidi	
Elizabeth Sherly	Rafiya Begum	
Enrique Flores	Rajeev R R	
Fei Xia	Rajeev Sangal	
Ganesh Katrapati	Rajesh Bhatt	
Gautam Mantena	Rakesh Balabantaray	
Geethanjali Rakshit	Raksha Sharma	
Girish Palshikar	Ranjani Parthasarathi	
Gurpreet Singh Lehal	Ratish Surendran	
Harikrishna K V	Raveesh Motlani	
Hema A Murthy	Riyaz Ahmad Bhat	



## Table of Contents

<i>Keynote Lecture 1: Scientific Paper Analysis</i> Yuji Matsumoto .....	1
<i>Addressing Class Imbalance in Grammatical Error Detection with Evaluation Metric Optimization</i> Anoop Kunchukuttan and Pushpak Bhattacharyya .....	2
<i>Words are not Equal: Graded Weighting Model for Building Composite Document Vectors</i> Pranjal Singh and Amitabha Mukerjee .....	11
<i>Online Adspace Posts' Category Classification</i> Dhawal Joharapurkar, Vaishak Salin and Vishal Krishna .....	20
<i>Noun Phrase Chunking for Marathi using Distant Supervision</i> Sachin Pawar, Nitin Ramrakhiani, Girish K. Palshikar, Pushpak Bhattacharyya and Swapnil Hingmire .....	29
<i>Self-Organizing Maps for Classification of a Multi-Labeled Corpus</i> Lars Bungum and Björn Gambäck .....	39
<i>Word Sense Disambiguation in Hindi Language Using Hyperspace Analogue to Language and Fuzzy C-Means Clustering</i> Devendra K. Tayal, Leena Ahuja and Shreya Chhabra .....	49
<i>Using Word Embeddings for Bilingual Unsupervised WSD</i> Sudha Bhingardive, Dharendra Singh, Rudramurthy V and Pushpak Bhattacharyya .....	59
<i>Compositionality in Bangla Compound Verbs and their Processing in the Mental Lexicon</i> Tirthankar Dasgupta, Manjira Sinha and Anupam Basu .....	65
<i>IndoWordNet Dictionary: An Online Multilingual Dictionary using IndoWordNet</i> Hanumant Redkar, Sandhya Singh, Nilesh Joshi, Anupam Ghosh and Pushpak Bhattacharyya ..	71
<i>Let Sense Bags Do Talking: Cross Lingual Word Semantic Similarity for English and Hindi</i> Apurva Nagvenkar, Jyoti Pawar and Pushpak Bhattacharyya .....	79
<i>A temporal expression recognition system for medical documents by</i> Naman Gupta, Aditya Joshi and Pushpak Bhattacharyya .....	84
<i>An unsupervised EM method to infer time variation in sense probabilities</i> Martin Emms and Arun Jayapal .....	89
<i>Solving Data Sparsity by Morphology Injection in Factored SMT</i> Sreelekha S, Piyush Dungarwal, Pushpak Bhattacharyya and Malathi D .....	95
<i>Authorship Attribution in Bengali Language</i> Shanta Phani, Shibamouli Lahiri and Arindam Biswas .....	100

<i>TransChat: Cross-Lingual Instant Messaging for Indian Languages</i>	
Diptesh Kanojia, Shehzaad Dhuliawala, Abhijit Mishra, Naman Gupta and Pushpak Bhattacharyya	106
<i>A Database of Infant Cry Sounds to Study the Likely Cause of Cry</i>	
Shivam Sharma, Shubham Asthana and V. K. Mittal	112
<i>Perplexed Bayes Classifier</i>	
Cohan Sujay Carlos	118
<i>An Empirical Study of Diversity of Word Alignment and its Symmetrization Techniques for System Combination</i>	
Thoudam Doren Singh	124
<i>Domain Sentiment Matters: A Two Stage Sentiment Analyzer</i>	
Raksha Sharma and Pushpak Bhattacharyya	130
<i>Extracting Information from Indian First Names</i>	
Akshay Gulati	138
<i>punct-An Alternative Verb Semantic Ontology Representation</i>	
Kavitha Rajan	144
<i>SMT Errors Requiring Grammatical Knowledge for Prevention</i>	
Yukiko Sasaki Alam	152
<i>Isolated Word Recognition System for Malayalam using Machine Learning</i>	
Maya Moneykumar, Elizabeth Sherly and Win Sam Varghese	158
<i>Judge a Book by its Cover: Conservative Focused Crawling under Resource Constraints</i>	
Shehzaad Dhuliawala, Arjun Atreya V, Ravi Kumar Yadav and Pushpak Bhattacharyya	166
<i>Text Normalization and Unit Selection for a Memory Based Non Uniform Unit Selection TTS in Malayalam</i>	
Gokul P., Neethu Thomas, Crisil Thomas and Dr. Deepa P. Gopinath	172
<i>Morphological Analyzer for Gujarati using Paradigm based approach with Knowledge based and Statistical Methods</i>	
Jatayu Baxi, Pooja Patel and Brijesh Bhatt	178
<i>Resolution of Pronominal Anaphora for Telugu Dialogues</i>	
Hemanth Reddy Jonnalagadda and Radhika Mamidi	183
<i>A Study on Divergence in Malayalam and Tamil Language in Machine Translation Perceptive</i>	
Jisha P Jayan and Elizabeth Sherly	189
<i>Automatic conversion of Indian Language Morphological Processors into Grammatical Framework (GF)</i>	
Harsha Vardhan Grandhi and Soma Paul	197

<i>Logistic Regression for Automatic Lexical Level Morphological Paradigm Selection for Konkani Nouns</i> Shilpa Desai, Jyoti Pawar and Pushpak Bhattacharyya .....	203
<i>Ruchi: Rating Individual Food Items in Restaurant Reviews</i> Burusothman Ahiladas, Paraneetharan Saravanaperumal, Sanjith Balachandran, Thamayanthy Sri- palan and Surangika Ranathunga .....	209
<i>Dependency Extraction for Knowledge-based Domain Classification</i> Lokesh Kumar Sharma and Namita Mittal .....	215
<i>An Approach to Collective Entity Linking</i> Ashish Kulkarni, Kanika Agarwal, pararth Shah, Sunny Raj Rathod and Ganesh Ramakrishnan	219
<i>Development of Speech corpora for different Speech Recognition tasks in Malayalam language</i> Cini Kurian .....	229
<i>POS Tagging of Hindi-English Code Mixed Text from Social Media: Some Machine Learning Experi- ments</i> Royal Sequiera, Monojit Choudhury and Kalika Bali .....	237
<i>Automated Analysis of Bangla Poetry for Classification and Poet Identification</i> Geetanjali Rakshit, Anupam Ghosh, Pushpak Bhattacharyya and Gholamreza Haffari .....	247
<i>Sentence Boundary Detection for Social Media Text</i> Dwijen Rudrapal, Anupam Jamatia, Kunal Chakma, Amitava Das and Björn Gambäck .....	254
<i>Mood Classification of Hindi Songs based on Lyrics</i> Braja Gopal Patra, Dipankar Das and Sivaji Bandyopadhyay .....	261
<i>Using Skipgrams, Bigrams, and Part of Speech Features for Sentiment Classification of Twitter Mes- sages</i> Badr Mohammed Badr and S. Sameen Fatima .....	268
<i>A Hybrid Approach for Bracketing Noun Sequence</i> Arpita Batra and Soma Paul .....	276
<i>Simultaneous Feature Selection and Parameter Optimization Using Multi-objective Optimization for Sentiment Analysis</i> Mohammed Arif Khan, Asif Ekbal and Eneldo Loza Mencía .....	285
<i>Detection of Multiword Expressions for Hindi Language using Word Embeddings and WordNet-based Features</i> Dhirendra Singh, Sudha Bhingardive, Kevin Patel and Pushpak Bhattacharyya .....	295
<i>Augmenting Pivot based SMT with word segmentation</i> Rohit More, Anoop Kunchukuttan, Pushpak Bhattacharyya and Raj Dabre .....	303
<i>Using Multilingual Topic Models for Improved Alignment in English-Hindi MT</i> Diptesh Kanojia, Aditya Joshi, Pushpak Bhattacharyya and Mark James Carman .....	308

<i>Triangulation of Reordering Tables: An Advancement Over Phrase Table Triangulation in Pivot-Based SMT</i>	
Deepak Patil, Harshad Chavan and Pushpak Bhattacharyya .....	316
<i>Post-editing a chapter of a specialized textbook into 7 languages: importance of terminological proximity with English for productivity</i>	
Ritesh Shah, Christian Boitet, Pushpak Bhattacharyya, Mithun Padmakumar, Leonardo Zilio, Ruslan Kalitvianski, Mohammad Nasiruddin, Mutsuko Tomokiyo and Sandra Castellanos Páez .....	325
<i>Generating Translation Corpora in Indic Languages: Cultivating Bilingual Texts for Cross Lingual Fertilization</i>	
Niladri Sekhar Dash, Arulmozi Selvraj and Mazhar Hussain .....	333
<i>Translation Quality and Effort: Options versus Post-editing</i>	
Donald Sturgeon and John S. Y. Lee .....	343
<i>Investigating the potential of post-ordering SMT output to improve translation quality</i>	
Pratik Mehta, Anoop Kunchukuttan and Pushpak Bhattacharyya .....	351
<i>Applying Sanskrit Concepts for Reordering in MT</i>	
Akshar Bharati, , Prajna Jha, Soma Paul and Dipti M Sharma .....	357
<i>Dialogue Act Recognition for Text-based Sinhala</i>	
Sudheera Palihakkara, Dammina Sahabandu, Ahsan Shamsudeen, Chamika Bandara and Surangika Ranathunga .....	367
<i>A Semi Supervised Dialog Act Tagging for Telugu</i>	
Suman Dowlagar and Radhika Mamidi .....	376
<i>Ranking Model with a Reduced Feature Set for an Automated Question Generation System</i>	
Manisha Satish Divate and Ambuja Salgaonkar .....	384
<i>Natural Language Processing for Solving Simple Word Problems</i>	
Sowmya S Sundaram and Deepak Khemani .....	394
<i>Analysis of Influence of L2 English Speakers' Fluency on Occurrence and Duration of Sentence-medial Pauses in English Readout Speech</i>	
Shambhu Nath Saha and Shyamal Kr. Das Mandal .....	403
<i>Acoustic Correlates of Voicing and Gemination in Bangla</i>	
Aanusha Ghosh .....	413

# Conference Program

**Saturday, December 12, 2015**

**+ 9:00-9:35 Inaugural Ceremony**

**+ 9:35-10:30 Keynote Lecture by Yuji Matsumoto**

*Keynote Lecture 1: Scientific Paper Analysis*

Yuji Matsumoto

**+ 10:30-11:00 Tea Break**

**+ 11:00-13:05 Technical Session I: Statistical Methods**

*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 Ramrakhiani, Girish K. Palshikar, Pushpak Bhattacharyya and Swapnil Hingmire

*Self-Organizing Maps for Classification of a Multi-Labeled Corpus*

Lars Bungum and Björn Gambäck

**Saturday, December 12, 2015 (continued)**

**+ 11:00-13:05 Technical Session II: WSD and Lexicon**

*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, Dharendra 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

**+ 13:05-14:00 Lunch**

**+ 14:00-15:30 Poster and Demo Session:**

*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



**Saturday, December 12, 2015 (continued)**

*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

**Saturday, December 12, 2015 (continued)**

*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

**+ 15:30-16:00 Tea Break**

**+ 16:00-17:40 Technical Session III: Emerging Areas**

*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

**Saturday, December 12, 2015 (continued)**

**+ 16:00-17:40 Technical Session IV : Sentiment Analysis**

*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

**+ 17:40-18:40 NLP AI Meeting**

**+ 19:00-20:00 Cultural Program**

**+ 20:00-20:30 Dinner**

**Sunday, December 13, 2015**

**+ 9:30-10:30 Panel Discussion**

**+ 10:30-11:00 Tea Break**

**Sunday, December 13, 2015 (continued)**

**+ 11:00-13:05 Technical Session V: Statistical Machine Translation**

*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 Pérez

*Generating Translation Corpora in Indic Languages: Cultivating Bilingual Texts for Cross Lingual Fertilization*

Niladri Sekhar Dash, Arulmozi Selvraj and Mazhar Hussain

**+ 11:00-13:05 Technical Session VI: NLP Tools Contest**

**+ 13:20-14:20 Lunch**

**+ 14:00-15:30 Technical Session VII: Machine Translation**

*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

**Sunday, December 13, 2015 (continued)**

**+ 14:00-15:30 Technical Session VIII: Dialog System and Question**

*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

**+ 15:30-16:00 Tea Break**

**+ 16:00-17:30 Technical Session IX: Speech Processing**

*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

**+ 17:30-18:00 Valedictory Function**

