

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-2014 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 140 submissions, 54 papers were selected, 34 for full presentation and 20 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 Aravind K. Joshi, Lori Levin and Sobha L for giving the three keynote lectures at ICON. We would also like to thank the members of the Advisory Committee and Programme Committee for their support and co-operation in making ICON 2014 a success.

We thank Vishal Goyal, Chair, Student Paper Competition and Sandipan Dandapat, Chair, NLP Tools Contest for taking the responsibilities of the events.

We convey our thanks to P V S Ram Babu, G Srinivas Rao 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 Vasudeva Varma, Soma Paul, Radhika Mamidi, Manish Shrivastava, B Yegnanarayana, Kishore Prahallad and Suryakanth V Gangashetty and the team at Goa University lead by Ramdas Karmali and Ramrao Wagh along with large number of volunteers and many others for sharing the work and responsibilities of ICON. 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-2014, without whose overwhelming response the conference would not have been a success.

December 2014
Hyderabad

Dipti Misra Sharma
Rajeev Sangal
Jyoti D. Pawar