



Trishna

LTRC NEWSLETTER

Vol: 4 Issue :1

February 2008

January was an action packed month for the students and professors in LTRC. First the 15-day Winter School, followed by IJCNLP-2008 (International Joint Conference in Natural Language Processing) held between 7th and 12th January. Quite a hectic way to usher in the New Year! Both events were very fruitful for IIIT's LTRC community as we were able to interact with international students and researchers and meet some of the most reputed scholars in the field. We also had the opportunity to listen to the latest NLP research and participate in the tutorials, workshops and demos. Several students from the department also presented their own work.

For many of the students taking part in the TCS Winter School, the interactive nature of the lectures and the group work for projects was immensely beneficial. Ideas were shared among the students and researchers over a longer duration as compared to a conference or seminar. Some well known NLP scholars were present to deliver lectures on Machine Translation, Morphological Processing and Natural Language Semantics. The format provided ample time for questions and the discussion sessions towards the end were often the most interesting. Everyone enjoyed the social dinner at Walden, Whisper Valley organized towards the end of the School. Many have expressed an interest in continuing or extending the project work carried out. Most of the participants that I interacted with were very happy with the research oriented environment at IIIT and the format of the School.

Another important event that took place was the series of meetings on NLP in India, where the sharing and creation of resources for NLP was discussed.

Right after, it was time for IJCNLP- 08. Students and the Local Organizing Committee had been working well in advance for this international conference. The first leg of the conference comprised the tutorial session that was held in IIIT-H. The main conference began on 8th of January in ISB (Indian School of Business), Gachibowli. There were 238 delegates with 195 from foreign countries who attended the conference and nearly half of them were students.

The airy and spacious surroundings of ISB provided a pleasant atmosphere for the conference. Some research students were sponsored by the Asia Fund and were able to attend the sessions of their choice. Students were exposed to current research in the areas of Machine Translation, Information Extraction and Statistical NLP.

The conference ended with the release of NLP tools for Indian languages. The hard work put in by the Local Organizing Committee headed by Prof. Sangal and Dr. R Bagga as well as the student volunteers helped to ensure the smooth functioning of all the events. For most of the students, the event resulted in some great exposure to the latest research areas and created a lot of motivation to make their own contribution.

-Ashwini Vaidya



Winter school participants at the social dinner

Winter School Report

Vishal Goyal

Punjabi University, Patiala

TCS NLP Winter School 2008 had been organized at IIIT Hyderabad from 24th December to 7th January, 2008. Aravind K. Joshi, University of Pennsylvania, Srinivas Bangalore, AT&T Research Labs, USA, Rajeev Sangal, Director, International Institute for Information Technology, Hyderabad, Sriram Venkatapathy, Lecturer and Ph.D. participant at IIIT, Hyderabad were the organizers. There were 54 participants mainly from Punjab, Noida, Mumbai, Hyderabad, Chennai, Nepal, Sri Lanka, Pune, New Delhi and Jaipur.

On the first day, Dr. Rajeev Sangal and Dr. Dipti Misra Sharma introduced the participants with the Computational Paninian Model. There were a number of experts throughout the world who delivered talks on various areas of NLP. Sivaji Bandyopadhyay, Sudeshna Sarkar, Sobha. L., Srinivas Bangalore, Pushpak Bhattacharya, Gerard Huet, Gina Levow, Ganesh Ramakrishnan, Philipp Koehn, Sadao Kurohashi, Fei Xia, William Lewis, Aravind Joshi, Michael Gamon, Yuji Matsumoto, Junichi Tsujii were among the key speakers.

The winter school was basically a platform to meet budding researchers in the field of NLP. All the participants had gained a lot of knowledge about the basics as well as the

advances in the NLP fields. The research environment in IIIT campus was excellent. All the participants were divided into groups and were assigned projects in Machine translation, Morphological Analysis and Natural Language Parsing under the supervision and Guidance of Mentors. There were regular project meetings, discussions among the participants and with their respective mentors during the 2-week winter school span.

On 7th January, 2008, at NLP Winter School Symposium, participants presented their project work results. It was a marvelous experience and I am thankful to the organizers for providing such a golden opportunity to all the researchers in NLP. My special thanks to Dr. Amba Kulkarni who had helped me during the winter school in my project.



At the final presentation for Winter School

What the participants said...

"I learnt a lot from the project. A brilliant program (with minor faults which can be easily ignored)! Thanks to IIIT-Hyderabad."
-Ram Raj Lohani (Tribhuvan University, Nepal)

"We enjoyed it very much. And it was really fruitful, especially the project."
- Asanka Wasala (University of Colombo, Sri Lanka)

The IJCNLP-2008 experience

Rafiya Begum

Our paper (co-authored with Samar Husain, Arun Dhvaj, Dipti Misra Sharma, Lakshmi Bai, and Rajeev Sangal) titled **Dependency Annotation Scheme for Indian Languages** introduces a major effort which is currently underway to develop a large scale tree bank for Indian Languages. It is an indispensable resource for linguistic investigations and also necessary for the development of efficient natural language tools and applications. The present work is on Hindi.

Paninian grammar, which is a dependency grammar is followed as the annotation scheme. The tool being used for the annotation is part of Sanchay, which is a collection of tools and APIs for South Asian languages. In this paper we have mentioned about certain issues like conjunct verbs, conjuncts and Ellipses. The components which make up a conjunct verb are chunked separately, but semantically they constitute a single unit. The head of the two or more conjoined element lies in the conjunct. A null element marked with a special tag 'NULL' is introduced in cases of ellipses. To ensure that the quality of the annotated corpus is good, the annotators cross-validate each other's work. This process is temporary because we are still in the process of revising the guidelines. Some experiments were also conducted to see whether the syntactic cues can be exploited for better machine learnability, whether certain generalization can be made for a constraint parser and how far would the automatic annotation help the annotators?

The response for the paper was very positive. Many people found the work very interesting. There were many people who were also dealing with the same kind of work. There were also people who were dealing with Prague Dependency Bank. They found our work very similar to theirs. It was a learning experience being part of such a big conference.



Interacting with the delegates at the Poster session



Participants at the poster session



Tutorial in progress

URL: <http://lrc.iiit.ac.in/newsletter>

Contact: Ashwini Vaidya
(ashwini_vaidya@research.iiit.ac.in)

IJCNLP IN PICTURES

Dr Aravind Joshi delivering the keynote address



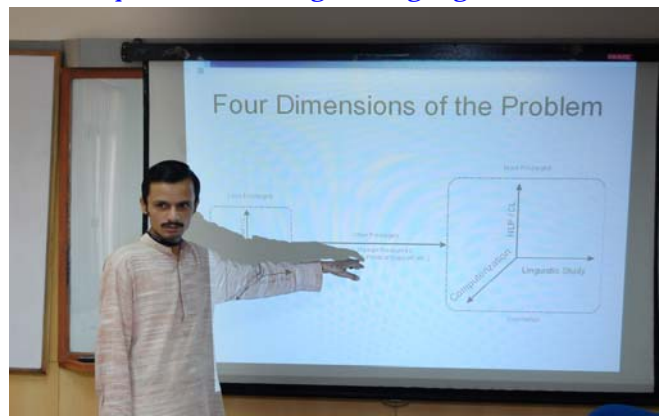
Sessions in progress at Khemka Auditorium, ISB



The sessions in progress in ISB



Workshop on Less Privileged Languages



Lunch at ISB



Workshop on NER for South & South Asian Languages



Banquet dinner and musical evening



Some organizing committee members and volunteers

