

# Program the 2nd Workshop on South and Southeast Asian Natural Language Processing

**Tuesday, November 8, 2011**

8:30–8:45 Opening Remarks

8:45–10:00 Invited Talk by Pushpak Bhattacharyya

10:00–10:30 Break

## **WSSANLP Session I**

10:30–11:00 *Hybrid Inflectional Stemmer and Rule-based Derivational Stemmer for Gujarati*  
Kartik Suba, Dipti Jiandani and Pushpak Bhattacharyya

11:00–11:30 *Improving Persian-English Statistical Machine Translation: Experiments in Domain Adaptation*

Mahsa Mohaghegh, Abdolhossein Sarrafzadeh and Tom Moir

11:30–12:00 *Thai Word Segmentation Verification Tool*

Supon Klaitin, Kanyanut Kriengkhet, Sitthaa Phaholphinyo and Krit Kosawat

12:00–12:30 *The Semi-Automatic Construction of Part-Of-Speech Taggers for Specific Languages by Statistical Methods*

Tomohiro YAMASAKI, Hiromi WAKAKI and Masaru SUZUKI

12:30–14:00 Lunch Break

## **WSSANLP Session II**

14:00–14:30 *Towards a Malay Derivational Lexicon: Learning Affixes Using Expectation Maximization*

Suriani Sulaiman, Michael Gasser and Sandra Kuebler

14:30–15:00 *Punjabi Language Stemmer for nouns and proper names*

Vishal Gupta and Gurpreet Singh Lehal

15:00–15:30 *Challenges in Urdu Text Tokenization and Sentence Boundary Disambiguation*

Zobia Rehman, Waqas Anwar and Usama Ijaz Bajwa

15:30–16:00 Break

**Tuesday, November 8, 2011 (continued)**

**WSSANLP Session III**

- 16:00–16:30 *Challenges in Developing a Rule based Urdu Stemmer*  
Sajjad Ahmad Khan, Waqas Anwar and Usama Ijaz Bajwa
- 16:30–17:00 *Developing a New System for Arabic Morphological Analysis and Generation*  
Mourad Gridach and Nouredine Chenfour
- 17:00–17:15 Closing Remarks