

# Foreword

IJCNLP 2004 heralded a new era for computational linguistics and natural language processing in Asia, as it coincided with the launching of the Asian Federation of NLP Associations (AFNLP) in Sanya, Hainan, China. This was a timely development for the most populous and linguistically diversified region in the world, and there is little wonder that this conference managed to draw over 200 participants. Following an enthusiastic response to the call for papers from 19 regions and the vigorous and rigorous efforts of the Program Committee Co-chairs, the sizable collection of papers presented at this conference enabled AFNLP to be compared with related organizations such as COLING and ACL, which have been in existence for more than 40 years.

The organizers recognized that it is important for serious attempts to be made to increase and improve efforts in this area all across Asia. Towards this end, and to contribute to a better awareness of topical issues, IJCNLP 2004 included two panels of special interest: (1) Panel on Emerging Asian Language Processing Efforts, and (2) Panel on Multilingual NLP for Public Information Services, which was devoted to the NLP requirements envisioned by the organizers of the 2008 Olympics. There were also two thematic sessions on (a) Natural Language Technology in Mobile Information Retrieval and Text Processing User Interfaces, and (b) Text Mining in Biomedicine. The offerings of this conference were further enriched by tutorials and workshops as well as an Asian Symposium on Natural Language Processing to Overcome Language Barriers, organized as a satellite event by the Institute of Electronics, Information and Communication Engineers (IEICE) and the Communications Research Laboratory (CRL) of Japan.

The publication of the collection of papers from IJCNLP 2004 by Springer in the Lecture Notes in Artificial Intelligence (LNAI) series is a welcome move and further testifies to the standards of the research efforts reported here.

A conference of this size and scope is not possible without the contributions of many individuals. Special thanks must go to the International Advisory Committee (Nicoletta Calzolari, Eva Hajicova, Eduard Hovy, Mark Johnson, Aravind Joshi, Martin Kay, Bente Maegaard, Joseph Mariani, Makoto Nagao, Donia Scott, Hozumi Tanaka, Hans Uszkoreit, Tianshun Yao), and the Organizing Committee. In particular, we are especially grateful to the Program Committee Co-chairs, Keh-Yih Su and Jun'ichi Tsujii; Publicity Chair, Hitoshi Isahara; Satellite Events Chair, Key-Sun Choi; Finance Committee Chair and Vice Chair, Kam-Fai Wong and Tom Lai, who had to deal with a complex "one conference, two currencies" situation and separate budgeting; and the Publication Chair and Vice Chairs, Jong-Hyeok Lee, Olivia Kwong and Sadao Kurohashi. The Local Organizing Committee in Beijing undertook the challenging responsibility for local arrangements in Sanya, and we are grateful to its Chair, Youqi Cao, and its Vice Chair and External Liaison Person, Maosong Sun.

We must also thank Makoto Nagao, the Honorary Chair of the conference, and the three keynote speakers, Ching-Chun Hsieh, Mark Johnson and Hans Uszkoreit.

IJCNLP 2004 must gratefully acknowledge financial sponsorship from a number of sources: the Association for Computational Linguistics (ACL), the Association for Computational Linguistics and Chinese Language Processing (ACLCLP), the Association of Natural Language Processing (ANLP), the Chinese Information Processing Society of China (CIPS), the Korean NLP Society, Microsoft Japan for the workshop “Beyond Shallow Analyses – Formalisms and Statistical Modeling for Deep Analyses,” and the Chinese Language Computer Society (CLCS) for the workshop “Named Entity Recognition for Natural Language Processing Applications.” Through CIPS, additional contributions were obtained from the National Natural Science Foundation of China, Fujitsu R&D Center Co., Ltd., Institute of Automation, Chinese Academy of Sciences, Microsoft Research Asia, NEC Laboratories, China, Peking University, TRS Information Technology Limited, and Tsinghua University. Without such support this conference would not have been possible. We also wish to thank the Finance Office of the City University of Hong Kong for graciously taking on the responsibility as credit card clearinghouse, and the City University Press and the University Printing Office for advice. Last but not least, to the many unnamed individuals and students in Beijing, Hong Kong, Pusan, Taipei and Tokyo, whose contributions were no less important in making this conference possible, we say thank you.

The Sanya region of Hainan Island has gained considerable prominence in recent years, the latest global event being the Miss World Contest. It is hoped that in future years, reference will be made in NLP circles to AFNLP’s IJCNLP in Sanya as a watershed in NLP development.

August 2004

Benjamin K. Tsou  
President, AFNLP

# Preface

This book is a collection of selected papers presented at the 1st International Joint Conference on Natural Language Processing (IJCNLP 2004), the flagship conference organized by the Asian Federation of Natural Language Processing Associations (AFNLP) which took place on Hainan Island, China during 22–24 March, 2004.

A total of 211 papers from 19 regions were submitted to the conference. The program committee which consisted of 20 internationally well-respected experts acting as area chairs and organizers of the thematic sessions, selected 66 papers for oral presentation and 35 for poster presentation. Authors were invited to revise their papers to reflect the vigorous and constructive discussions at the conference, and we received 84 revised papers for this volume.

The articles in this volume cover diverse fields in NLP, from theoretical contributions in finite state technology, parsing, discourse modeling, machine learning, etc., to basic NLP technologies of semantic disambiguation, shallow parsing, sentence/text generation, etc., to NLP applications such as machine translation, text mining and information retrieval. We also have articles on topics crucial for Asian language processing, including word segmentation, chunkers, etc. The linguistic diversity in Asia has posed exciting challenges, not only to researchers who believe in the universality of human language and the technologies in computational linguistics, but also to those interested in practical applications of NLP for Asian languages. This volume should give them a good collection of the current and the most advanced results in such research.

At the same time, since the conference attracted contributions from active and respected researchers all over the world, the articles in this volume not only showcase language-specific research output, but also represent the general current achievements in our field, from both the theoretical and technological fronts. In particular, the two thematic sessions, “Text Mining in Biomedicine” and “Natural Language Technology in Mobile Information Retrieval and Text Processing User Interfaces,” solicited high-quality contributions on these emerging technological challenges.

We would like to thank the area chairs and the reviewers who carried out professionally rigorous reviews and selection of papers for the conference, as well as the authors who revised their already high-quality papers to reflect the discussion at the conference. Special thanks are also given to Jin-Dong Kim (University of Tokyo) and Christine Günther (Springer) who gave technical and editorial support in the preparation of this volume.

August 2004

Keh-Yih Su,      Jun’ichi Tsujii,  
Jong-Hyeok Lee,      Oi Yee Kwong