2006

COLING • ACL

COLING • ACL 2006

21st International Conference on
Computational Linguistics
and
44th Annual Meeting of the
Association for Computational Linguistics

Proceedings of the Student Research Workshop

20 July 2006
Sydney, Australia

©2006 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

Table of Contents

Preface v

Organizers vii

Workshop Program ix

A Flexible Approach to Natural Language Generation for Disabled Children
Pradipta Biswas 1

Unsupervised Part-of-Speech Tagging Employing Efficient Graph Clustering
Chris Biemann 7

Sub-Sentential Alignment Using Substring Co-Occurrence Counts
Fabien Cromieres 13

Annotation Schemes and their Influence on Parsing Results
Wolfgang Maier 19

Modeling Human Sentence Processing Data with a Statistical Parts-of-Speech Tagger
Jihyun Park 25

Semantic Discourse Segmentation and Labeling for Route Instructions
Nobuyuki Shimizu 31

Investigations on Event-Based Summarization
Mingli Wu 37

Discursive Usage of Six Chinese Punctuation Marks
Ming Yue 43

Integrated Morphological and Syntactic Disambiguation for Modern Hebrew
Reut Tsarfaty 49

A Hybrid Relational Approach for WSD – First Results
Lucia Specia 55

On2L – A Framework for Incremental Ontology Learning in Spoken Dialog Systems
Berenike Loos 61

Focus to Emphasize Tone Structures for Prosodic Analysis in Spoken Language Generation
Lalita Narupiyakul 67

Extraction of Tree Adjoining Grammars from a Treebank for Korean
Jungyeul Park 73

Parsing and Subcategorization Data
Jianguo Li 79

Clavius: Bi-Directional Parsing for Generic Multimodal Interaction
Frank Rudzicz 85

Author Index 91

Preface

Welcome to the COLING/ACL 2006 Student Research Workshop! The Student Research Workshop is now an established tradition at ACL conferences. This year, we received 40 submissions ranging from 20 countries, and accepted 15 of them. 8 will be presented orally, and 7 as posters, during a common poster session with the main conference.

A total of 73 students and senior researchers agreed to serve on the program committee, which allowed us to assign 4 reviewers per paper. We would like to thank the reviewers for understanding the spirit of the Student Research Workshop and giving careful and constructive reviews. We hope their comments will be helpful to all the students who submitted their work.

All presenters received financial support from the U.S. National Science Foundation to assist them in their travel to Sydney.

We are very grateful to Rebecca Hwa, our faculty advisor, for her advice, constant support, and obtaining funding. Thanks to Stephen Wan and Chris Callison-Burch, co-chairs of the ACL-05 Student Research Workshop in Ann Arbor, who shared their invaluable experience. Finally, we would like to thank the program chairs of COLING/ACL 2006, Claire Cardie and Pierre Isabelle, the Publications chair Olivia Kwong, as well as the local organization committee, and Priscilla Rasmussen.

The COLING/ACL 2006 Student Research Workshop co-chairs:
Marine Carpuat and Kevin Duh

Organizers

Co-chairs:

Marine Carpuat, Hong Kong University of Science and Technology
Kevin Duh, University of Washington

Faculty advisor:

Rebecca Hwa, University of Pittsburgh

Senior members of the Program Committee:

Eneko Agirre, Basque Country University
Slaven Bilac, Google, Japan
Cem Bozsahin, Middle East Technical Univ
Stephen Clark, Oxford University
Simon Corston-Oliver, Microsoft Research
James Curran, Sydney
Ido Dagan, Bar Ilan University
Pablo Duboue, IBM
Eric Fosler-Lussier, Ohio State
Nizar Habash, Columbia University
Keith Hall, Johns Hopkins University
Mary Hearne, Dublin City University
Philipp Koehn, University of Edinburgh
Emiel Krahmer, Tilburg
Chin Yew Lin, USC/ISI
Bernardo Magnini, ITC irst
Diego Molla Aliod, Macquarie University
Grace Ngai, Hong Kong Polytechnic University
Patrick Pantel, USC/ISI
Cecile Paris, CSIRO
Eric Ringger, Brigham Young University
Graeme Ritchie, Aberdeen
Khalil Sima'an, University of Amsterdam
Michel Simard, NRC Institute for Information Technology
Caroline Sporleder, University of Edinburgh
Nicola Stokes, University of Melbourne
Matthew Stone, Rutgers
Keh-Yih Su, Behavior Design Corporation
Marc Swerts, Tilburg
Joel Tetreault, Pittsburgh
Simone Teufel, Cambridge
Leonoor van der Beek, Groningen
Richard Wicentowski, Swarthmore College
Florian Wolf, FW Consulting
Zhu Zhang, University of Arizona

Student members of the Program Committee:

Michele Banko, University of Washington
Colin Bannard, University of Edinburgh
John Blitzer, University of Pennsylvania
Phil Blunsom, University of Melbourne
Chris Callison-Burch, University of Edinburgh
Xavier Carreras, Technical University of Catalonia (UPC)
Yee Seng Chan, National University of Singapore
Trevor Cohn, University of Melbourne
Tiphaine Dalmas, University of Edinburgh
Hal Daume III, USC/ISI
Etienne Denoual, ATR
Elena Filatova, Columbia University
Erin Fitzgerald, Johns Hopkins University
Michael Fleischman, MIT
Ulrich Germann, Toronto
Teg Grenager, Stanford University
Declan Groves, Dublin City University
Kristy Hollingshead, OGI/Oregon Health and Science University
Matthew Lease, Brown University
Maria Liakata, University of Wales
Gideon Mann, Johns Hopkins University
Ryan McDonald, University of Pennsylvania
Ani Nenkova, Columbia University
Leif Nielsen, King's College
Bo Pang, Cornell
Jean-Philippe Prost, Macquarie University
Sarah Schwarm, University of Washington
Libin Shen, University of Pennsylvania
Yihai Shen, Hong Kong University of Science and Technology
Rion Snow, Stanford University
Charles Sutton, University of Massachussetts, Amherst
Paola Virga, Johns Hopkins University
Stephen Wan, Macquarie University
Ben Wellington, New York University
Theresa Wilson, University of Pittsburgh
Richard Zens, RWTH Aachen
Hao Zhang, University of Rochester
Bing Zhao, CMU

Workshop Program

Monday, 17 July 2006

17:30–19:30 Poster Session

Posters

A Flexible Approach to Natural Language Generation for Disabled Children
Pradipta Biswas

Unsupervised Part-of-Speech Tagging Employing Efficient Graph Clustering
Chris Biemann

Sub-Sentential Alignment Using Substring Co-Occurrence Counts
Fabien Cromieres

Annotation Schemes and their Influence on Parsing Results
Wolfgang Maier

Modeling Human Sentence Processing Data with a Statistical Parts-of-Speech Tagger
Jihyun Park

Semantic Discourse Segmentation and Labeling for Route Instructions
Nobuyuki Shimizu

Investigations on Event-Based Summarization
Mingli Wu

Thursday, 20 July 2006

Multilinguality

11:00–11:30 *Discursive Usage of Six Chinese Punctuation Marks*
Ming Yue

11:30–12:00 *Integrated Morphological and Syntactic Disambiguation for Modern Hebrew*
Reut Tsarfaty

12:00–12:30 *A Hybrid Relational Approach for WSD – First Results*
Lucia Specia

Spoken Language Processing

14:30–15:00 *On2L – A Framework for Incremental Ontology Learning in Spoken Dialog Systems*
Berenike Loos

15:00–15:30 *Focus to Emphasize Tone Structures for Prosodic Analysis in Spoken Language Generation*
Lalita Narupiyakul

Parsing

16:00–16:30 *Extraction of Tree Adjoining Grammars from a Treebank for Korean*
Jungyeul Park

16:30–17:00 *Parsing and Subcategorization Data*
Jianguo Li

17:00–17:30 *Clavius: Bi-Directional Parsing for Generic Multimodal Interaction*
Frank Rudzicz