

ACL 2014

**52nd Annual Meeting of the
Association for Computational Linguistics**

Proceedings of the Student Research Workshop

June 22-27, 2014
Baltimore, Maryland, USA

©2014 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

ISBN 978-1-937284-79-4

Introduction

Welcome to the ACL 2014 Student Research Workshop.

Following the previous years' ACL Student Research Workshops, this year we have two different kinds of papers: research papers and thesis proposals. Thesis proposals are intended for advanced students who have decided on a thesis topic and wish to get feedback on their proposal and broader ideas for their continuing work, while research papers can describe completed work or work in progress with preliminary results.

We received 7 thesis proposals and 19 research papers this year. Out of these, we accepted 5 thesis proposals and 8 research papers leading to an acceptance rate of 71% for thesis proposals and 42% for research papers. This year's workshop also offered pre-submission mentoring for student authors wishing to improve the presentation of their papers prior to submission of the paper for the reviewing process. 6 students participated in this pre-submission mentoring program this year.

The SRW offers multiple avenues for the student authors to receive feedback at the conference.

This year's workshop features three sessions of student paper presentations. All the papers will be presented at the main conference poster session, giving the opportunity for students to interact and present their work to a large and diverse audience. We also have a separate oral presentation session for thesis proposal papers on the first day of the main conference. The third session is another oral presentation venue where student authors of the research papers briefly advertise their posters. This year, 15 students whose papers were accepted to the ACL main conference were chosen to also present their work at and be partially funded by SRW. These students' posters are also advertised during the poster highlight session.

In addition to oral presentation and poster sessions, each SRW paper and the 15 papers sponsored by the SRW is assigned a dedicated mentor. The mentor is an experienced researcher from academia or industry who will prepare in-depth comments and questions in advance for the presentation or poster session and will provide feedback to the student author.

Thanks to our funding sources, this year's SRW covers registration expenses and provides travel and lodging support to all student authors of the SRW papers. We gratefully acknowledge the support from the NSF, Google, Yahoo! and Baidu. In addition to the SRW papers, the NSF supports travel and registration for another 15 student authors of ACL main conference papers.

We are very grateful to our program committee members who gave constructive and detailed reviews for each of the student papers. Some of our PC members also participated in the pre-submission mentoring program and immensely helped student authors with writing and presentation of their papers. We also thank researchers who agreed to mentor and provide expert feedback on the student papers. We thank our faculty advisers Bill Byrne and Jordan Boyd-Graber for their guidance. We also thank the ACL 2014 organizing committee – Daniel Marcu, Kristina Toutanova, Hua Wu, Alexander Koller, Miyao Yusuke, David Yarowsky and Priscilla Rasmussen for their constant support and suggestions. Finally, we thank all students for their submissions and participation in the SRW.

Organizers:

Ekaterina Kochmar, University of Cambridge (UK)
Annie Louis, University of Edinburgh (UK)
Svitlana Volkova, Johns Hopkins University (USA)

Faculty Advisors:

Jordan Boyd-Graber, University of Maryland (USA)
Bill Byrne, University of Cambridge (UK)

Program Committee:

Gabor Angeli, Stanford University (USA)
Yoav Artzi, University of Washington (USA)
Chris Biemann, TU Darmstadt (Germany)
Arianna Bisazza, University of Trento (Italy)
Ted Briscoe, University of Cambridge (UK)
Shu Cai, University of Southern California (USA)
Asli Ceyilkamaz, Microsoft Research (USA)
Ilia Chetviorkin, Lomonosov Moscow State University (Russia)
Monojit Choudhury, Microsoft Research (India)
Trevor Cohn, University of Sheffield (UK)
Hal Daumé III, University of Maryland (USA)
Leon Derczynski, University of Sheffield (UK)
Aciel Eshky, University of Edinburgh (UK)
Kilian Evang, University of Groningen (Netherlands)
Annemarie Friedrich, Saarland University (Germany)
Thomas François, Université catholique de Louvain (Belgium)
Michael Gamon, Microsoft Research (USA)
Juri Ganitkevitch, Johns Hopkins University (USA)
Qin Gao, Microsoft Research (USA)
Liane Guillou, University of Edinburgh (UK)
Eva Hasler, University of Edinburgh (UK)
John Henderson, The MITRE Corporation (USA)
Ruihong Huang, University of Utah (USA)
Amjad Abu Jbara, University of Michigan, Ann Arbor (USA)
Beata Klebanov-Biegman, Educational Testing Service (USA)
Varada Kolhatkar, University of Toronto (Canada)
Jonathan Kummerfeld, University of California, Berkeley (USA)
Angeliki Lazaridou, University of Trento (Italy)
Angeliki Metallinou, University of Southern California (USA)
Mitch Marcus, University of Pennsylvania (USA)
Thomas Meyer, Google (Switzerland)
Meg Mitchell, Microsoft Research (USA)
Alessandro Moschitti, University of Trento (Italy)

Courtney Napoles, Johns Hopkins University (USA)
Patrick Pantel, Microsoft Research (USA)
Rashmi Prasad, University of Wisconsin, Milwaukee (USA)
Owen Rambow, University of Columbia (USA)
Roi Reichart, Technion - Israel Institute of Technology (Israel)
Philip Resnik, University of Maryland (USA)
Kairit Sirts, Tallinn University of Technology (Estonia)
Joel Tetreault, Yahoo! Labs (USA)
Ivan Titov, University of Amsterdam (Netherlands)
Eva Maria Vecchi, University of Cambridge (UK)
Stephan Vogel, Carnegie Mellon University (USA)
Jason Williams, Microsoft Research (USA)
Travis Wolfe, Johns Hopkins University (USA)
Bishan Yang, Cornell University (USA)
Xuchen Yao, Johns Hopkins University (USA)
Torsten Zesch, University of Duisburg, Essen (Germany)

Table of Contents

<i>Bayesian Kernel Methods for Natural Language Processing</i> Daniel Beck	1
<i>Extracting Temporal and Causal Relations between Events</i> Paramita Mirza	10
<i>Towards a discourse relation-aware approach for Chinese-English machine translation</i> Frances Yung	18
<i>Analyzing Positions and Topics in Political Discussions of the German Bundestag</i> Cécilia Zirn	26
<i>A Mapping-Based Approach for General Formal Human Computer Interaction Using Natural Language</i> Vincent Letard, Sophie Rosset and Gabriel Illouz	34
<i>An Exploration of Embeddings for Generalized Phrases</i> Wenpeng Yin and Hinrich Schütze	41
<i>Learning Grammar with Explicit Annotations for Subordinating Conjunctions</i> Dongchen Li, Xiantao Zhang and Xihong Wu	48
<i>Going beyond sentences when applying tree kernels</i> Dmitry Ilvovsky	56
<i>Multi-document summarization using distortion-rate ratio</i> Ulukbek Attokurov and Ulug Bayazit	64
<i>Disambiguating prepositional phrase attachment sites with sense information captured in contextualized distributional data</i> Clayton Greenberg	71
<i>Open Information Extraction for Spanish Language based on Syntactic Constraints</i> Alisa Zhila and Alexander Gelbukh	78
<i>Improving Text Normalization via Unsupervised Model and Discriminative Reranking</i> Chen Li and Yang Liu	86
<i>Semi-Automatic Development of KurdNet, The Kurdish WordNet</i> Purya Aliabadi	94

Conference Program

Monday, June 23, 2014

SRW Thesis Proposal Presentations

- 13:20–13:45 *Bayesian Kernel Methods for Natural Language Processing*
Daniel Beck
- 13:45–14:10 *Extracting Temporal and Causal Relations between Events*
Paramita Mirza
- 14:10–14:35 *Towards a discourse relation-aware approach for Chinese-English machine translation*
Frances Yung
- 14:35–15:00 *Analyzing Positions and Topics in Political Discussions of the German Bundestag*
Cäcilia Zirn

Poster Highlights

18:05–18:50 SRW Research Papers

A Mapping-Based Approach for General Formal Human Computer Interaction Using Natural Language
Vincent Letard, Sophie Rosset and Gabriel Illouz

An Exploration of Embeddings for Generalized Phrases
Wenpeng Yin and Hinrich Schütze

Learning Grammar with Explicit Annotations for Subordinating Conjunctions
Dongchen Li, Xiantao Zhang and Xihong Wu

Going beyond sentences when applying tree kernels
Dmitry Ilvovsky

Multi-document summarization using distortion-rate ratio
Ulukbek Attokurov and Ulug Bayazit

Disambiguating prepositional phrase attachment sites with sense information captured in contextualized distributional data
Clayton Greenberg

Monday, June 23, 2014 (continued)

Open Information Extraction for Spanish Language based on Syntactic Constraints

Alisa Zhila and Alexander Gelbukh

Improving Text Normalization via Unsupervised Model and Discriminative Reranking

Chen Li and Yang Liu

Poster Highlights

18:05–18:50 Main Conference Papers Sponsored by SRW

Poster Session

18:50–21:30 SRW Thesis Proposal and Research Papers

Poster Session

18:50–21:30 Main Conference Papers Sponsored by SRW