ACL HLT 2011

49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies

Proceedings of Student Session

19-24 June 2011
Portland, Oregon, USA

Production and Manufaturing by

Omnipress, Inc.
2600 Anderson Street
Madison, WI 53704
USA

©2011 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 978-1-932432-89-3

Introduction

Welcome to the ACL-HLT 2011 Student Session. This year, we have two separate submission types: Research Papers and Thesis Proposals. We accepted 22 out of the 57 submissions received from a wide variety of countries. Of the accepted submissions, 18 were Research Papers and 4 were Thesis Proposals. Each accepted paper is to be presented as a poster during the ACL-HLT 2011 Poster Session and assigned a mentor to provide specific feedback and guidance to our authors.

The overall quality of the submissions were high and we thank our Program Committee for their excellent feedback and reviews. We also thank our Faculty Advisors, Miles Osborne and Thamar Solorio, for their patience and guidance. We would also like to thank the entire ACL-HLT 2011 Organizing Committee, especially Dekang Lin, Brian Roark, Nate Bodenstab, Guodong Zhou, and Priscilla Rasmussen. All presenters were offered conference registration and travel support, thanks to generous support from the U.S. National Science Foundation (Award Number 1102435), the ACL Walker Student Fund, Google, and the European Association of Computational Linguistics. Finally, our highest gratitude goes to all the students who submitted to the Student Session. Thank you.

Student Program Committee Members:

Hamid Chinaei, *Laval University, Canada*
Markus Dreyer, *Johns Hopkins University, USA*
Jennifer Gillenwater, *University of Pennsylvania, USA*
Samer Hassan, *University of North Texas, USA*
Tom Kwiatkowski, *University of Edinburgh, UK*
Maider Lehr, *Oregon Health & Science University, USA*
Abby Levenberg, *University of Edinburgh, UK*
Ziheng Lin, *National University of Singapore, Singapore*
Prashanth Mannem, *Indian Institute of Technology Hyderabad, India*
Eric Morley, *Oregon Health & Science University, USA*
Nathan Schneider, *Carnegie Mellon University, USA*

Non-Student Program Committee Members:

Shane Bergsma, *Johns Hopkins University, USA*
Delphine Bernhard, *LIMSI-CNRS, France*
Steven Bethard, *Catholic University of Leuven, Belgium*
Phil Blunsom, *University of Oxford, UK*
Aoife Cahill, *Universität Stuttgart, Germany*
Ben Carterette, *University of Delaware, USA*
Yejin Choi, *Stony Brook University, USA*
Stephen Clark, *University of Cambridge, UK*
Berthold Crysmann, *University of Bonn, Germany*
James Curran, *University of Sydney, Australia*
Micha Elsner, *University of Edinburgh, UK*
Sharon Goldwater, *University of Edinburgh, UK*
Ben Hachey, *Macquarie University, Australia*
Julia Hirschberg, *Columbia University, USA*

Matt Huenerfauth, *The City University of New York, USA*
Michael Johnston, *AT&T Labs, USA*
Udo Kruschwitz, *University of Essex, UK*
Giridhar Kumaran, *Microsoft, USA*
Victor Lavrenko, *University of Edinburgh, UK*
Chin-Yew Lin, *Microsoft, USA*
Adam Lopez, *Johns Hopkins University, USA*
David McClosky, *Stanford University, USA*
Donald Metzler, *Univerity of Southern California, USA*
Taniya Mishra, *AT&T Labs, USA*
Saif Mohammad, *National Research Council Canada, Canada*
Sebastian Pado, *University of Stuttgart, Germany*
Katerina Pastra, *Institute for Language and Speech Processing, Greece*
Pavel Pecina, *Dublin City University, Ireland*
Hema Raghavan, *Yahoo, USA*
Jan Romportl, *University of West Bohemia, Czech Republic*
Helmut Schmid, *University of Stuttgart, Germany*
Benjamin Snyder, *University of Wisconsin-Madison, USA*
Veselin Stoyanov, *Johns Hopkins University, USA*
Yi Su, *Nuance, USA*
Joel Tetreault, *Educational Testing Service, USA*
Lucy Vanderwende, *Microsoft, USA*
Yorick Wilks, *University of Sheffield, UK*
Fan Yang, *Nuance, USA*
Alexander Yates, *Temple University, USA*
Emine Yilmaz, *Microsoft, USA*
Daniel Zeman, *Charles University in Prague, Czech Republic*

Student Co-Chairs:

Sasa Petrovic, University of Edinburgh
Emily Pitler, University of Pennsylvania
Ethan Selfridge, Oregon Health & Science University

Faculty Advisors:

Miles Osborne, University of Edinburgh
Thamar Solorio, University of Alabama at Birmingham

Table of Contents

Word Alignment Combination over Multiple Word Segmentation
Ning Xi, Guangchao Tang, Boyuan Li and Yinggong Zhao 1

Sentence Ordering Driven by Local and Global Coherence for Summary Generation
Renxian Zhang 6

Pre- and Postprocessing for Statistical Machine Translation into Germanic Languages
Sara Stymne 12

Exploring Entity Relations for Named Entity Disambiguation
Danuta Ploch 18

Extracting and Classifying Urdu Multiword Expressions
Annette Hautli and Sebastian Sulger 24

A Latent Topic Extracting Method based on Events in a Document and its Application
Risa Kitajima and Ichiro Kobayashi 30

Syntax-based Statistical Machine Translation using Tree Automata and Tree Transducers
Daniel Emilio Beck 36

ConsentCanvas: Automatic Texturing for Improved Readability in End-User License Agreements
Oliver Schneider and Alex Garnett 41

Disambiguating temporal-contrastive connectives for machine translation
Thomas Meyer 46

PsychoSentiWordNet
Amitava Das 52

Optimistic Backtracking - A Backtracking Overlay for Deterministic Incremental Parsing
Gisle Ytrestøl 58

An Error Analysis of Relation Extraction in Social Media Documents
Gregory Brown 64

Effects of Noun Phrase Bracketing in Dependency Parsing and Machine Translation
Nathan Green 69

Towards a Framework for Abstractive Summarization of Multimodal Documents
Charles Greenbacker 75

Sentiment Analysis of Citations using Sentence Structure-Based Features
Awais Athar 81

Combining Indicators of Allophony
Luc Boruta 88

Turn-Taking Cues in a Human Tutoring Corpus
Heather Friedberg 94

Predicting Clicks in a Vocabulary Learning System
Aaron Michelony 99

Exploiting Morphology in Turkish Named Entity Recognition System
Reyyan Yeniterzi 105

Social Network Extraction from Texts: A Thesis Proposal
Apoorv Agarwal 111

Automatic Headline Generation using Character Cross-Correlation
Fahad Alotaiby 117

K-means Clustering with Feature Hashing
Hajime Senuma 122

Conference Program

Monday, June 20, 2011

(12:00-2:00)Student Lunch

(6:00-8:30)Poster Session(Student Session)

Word Alignment Combination over Multiple Word Segmentation
Ning Xi, Guangchao Tang, Boyuan Li and Yinggong Zhao

Sentence Ordering Driven by Local and Global Coherence for Summary Generation
Renxian Zhang

Pre- and Postprocessing for Statistical Machine Translation into Germanic Languages
Sara Stymne

Exploring Entity Relations for Named Entity Disambiguation
Danuta Ploch

Extracting and Classifying Urdu Multiword Expressions
Annette Hautli and Sebastian Sulger

A Latent Topic Extracting Method based on Events in a Document and its Application
Risa Kitajima and Ichiro Kobayashi

Syntax-based Statistical Machine Translation using Tree Automata and Tree Transducers
Daniel Emilio Beck

ConsentCanvas: Automatic Texturing for Improved Readability in End-User License Agreements
Oliver Schneider and Alex Garnett

Disambiguating temporal-contrastive connectives for machine translation
Thomas Meyer

PsychoSentiWordNet
Amitava Das

Optimistic Backtracking - A Backtracking Overlay for Deterministic Incremental Parsing
Gisle Ytrestøl

An Error Analysis of Relation Extraction in Social Media Documents
Gregory Brown

Monday, June 20, 2011(continued)

Effects of Noun Phrase Bracketing in Dependency Parsing and Machine Translation
Nathan Green

Towards a Framework for Abstractive Summarization of Multimodal Documents
Charles Greenbacker

Sentiment Analysis of Citations using Sentence Structure-Based Features
Awais Athar

Combining Indicators of Allophony
Luc Boruta

Turn-Taking Cues in a Human Tutoring Corpus
Heather Friedberg

Predicting Clicks in a Vocabulary Learning System
Aaron Michelony

Exploiting Morphology in Turkish Named Entity Recognition System
Reyyan Yeniterzi

Social Network Extraction from Texts: A Thesis Proposal
Apoorv Agarwal

Automatic Headline Generation using Character Cross-Correlation
Fahad Alotaiby

K-means Clustering with Feature Hashing
Hajime Senuma