Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics

Proceedings of the Student Research Workshop

June 2, 2010
Los Angeles, California
Introduction

Welcome to the 2010 NAACL-HLT Student Research Workshop in Los Angeles, California! The purpose of this workshop is to provide feedback to students whose work may be in its early stages and to help prepare these students to future academic and professional pursuits. During this workshop, each presentation will be followed by a discussion on the work by a panel of senior researchers.

This year, we received 20 submissions from 5 countries. We thank all authors who submitted. We are grateful to the program committee for their time, effort, and professional consideration. We also thank the panelists in advance for their time and helpful feedback. We also owe a debt of gratitude to the 2010 NAACL-HLT main conference organizers and to the National Science Foundation whose support makes many aspects of this workshop possible.
Organizers:
Adriane Boyd, The Ohio State University
Mahesh Joshi, Carnegie Mellon University
Frank Rudzicz, University of Toronto

Faculty Advisors:
Julia Hockenmaier, University of Illinois
Diane Litman, University of Pittsburgh

Program Committee:
Hua Ai, Carnegie Mellon University
Abhishek Arun, University of Edinburgh
Srinivas Bangalore, AT&T
S.R.K. Branavan, Massachusetts Institute of Technology
Chris Brew, The Ohio State University
Chris Callison-Burch, Johns Hopkins University
Nathaneal Chambers, Stanford University
Steve DeNeefe, Information Sciences Institute
Markus Dickinson, Indiana University
Micha Elsner, Brown University
Jenny Finkel, Stanford University
Eric Fosler-Lussier, The Ohio State University
Timothy Fowler, University of Toronto
Oana Frunza, University of Ottawa
Roxana Girju, University of Illinois at Urbana-Champaign
Dan Goldwasser, University of Illinois at Urbana-Champaign
Mark Hasegawa-Johnson, University of Illinois at Urbana-Champaign
Sandra Kuebler, Indiana University
Yang Liu, The University of Texas at Dallas
Daniel Marcu, Information Sciences Institute
Jon May, Information Sciences Institute
Cosmin Munteanu, University of Toronto
Douglas O’Shaughnessy, Institut national de la recherche scientifique
Ted Pedersen, University of Minnesota
Rashmi Prasad, University of Pennsylvania
Partha Pratim Talukdar, University of Pennsylvania
Owen Rambow, Columbia University
Sujith Ravi, Information Sciences Institute
Hannah Rohde, Northwestern University
William Schuler, The Ohio State University
Matthew Stone, Rutgers University
Marilyn Walker, University of California, Santa Cruz
Xiaodan Zhu, University of Toronto
## Table of Contents

*Improving Syntactic Coordination Resolution using Language Modeling*
  Philip Ogren ................................................................. 1

*On Automated Evaluation of Readability of Summaries: Capturing Grammaticality, Focus, Structure and Coherence*
  Ravikiran Vadlapudi and Rahul Katragadda ........................................ 7

*Detecting Novelty in the context of Progressive Summarization*
  Praveen Bysani ............................................................... 13

*Extrinsic Parse Selection*
  David Goss-Grubbs .......................................................... 19

*Towards a Matrix-based Distributional Model of Meaning*
  Eugenie Giesbrecht ......................................................... 23

*Distinguishing Use and Mention in Natural Language*
  Shomir Wilson ............................................................... 29

*A Learning-based Sampling Approach to Extractive Summarization*
  Vishal Juneja, Sebastian Germesin and Thomas Kleinbauer ..................... 34

*Temporal Relation Identification with Endpoints*
  Chong Min Lee ............................................................. 40

*Identifying Opinion Holders and Targets with Dependency Parser in Chinese News Texts*
  Bin Lu .............................................................. 46

*A Data Mining Approach to Learn Reorder Rules for SMT*
  Avinesh PVS ............................................................... 52

*Fine-Tuning in Brazilian Portuguese-English Statistical Transfer Machine Translation: Verbal Tenses*
  Lucia Silva ............................................................... 58
Workshop Program

Wednesday, June 2, 2010

Session 1

10:40–11:10  Improving Syntactic Coordination Resolution using Language Modeling
              Philip Ogren

11:10–11:40  On Automated Evaluation of Readability of Summaries: Capturing Grammaticality,
              Focus, Structure and Coherence
              Ravikiran Vadlapudi and Rahul Katragadda

11:40–12:10  Detecting Novelty in the context of Progressive Summarization
              Praveen Bysani

12:10–2:00  Lunch

Session 2

2:00–2:30  Extrinsic Parse Selection
              David Goss-Grubbs

2:30–3:00  Towards a Matrix-based Distributional Model of Meaning
              Eugenie Giesbrecht

3:00–3:30  Distinguishing Use and Mention in Natural Language
              Shomir Wilson

6:30–8:30  Poster Session

A Learning-based Sampling Approach to Extractive Summarization
Vishal Juneja, Sebastian Germesin and Thomas Kleinbauer

Temporal Relation Identification with Endpoints
Chong Min Lee

Identifying Opinion Holders and Targets with Dependency Parser in Chinese News
Texts
Bin Lu
A Data Mining Approach to Learn Reorder Rules for SMT
Avinesh PVS

Fine-Tuning in Brazilian Portuguese-English Statistical Transfer Machine Translation: Verbal Tenses
Lucia Silva

Note: The following posters were also given as oral presentations.

Improving Syntactic Coordination Resolution using Language Modeling
Philip Ogren

On Automated Evaluation of Readability of Summaries: Capturing Grammaticality, Focus, Structure and Coherence
Ravikiran Vadlapudi and Rahul Katragadda

Extrinsic Parse Selection
David Goss-Grubbs

Towards a Matrix-based Distributional Model of Meaning
Eugenie Giesbrecht

Distinguishing Use and Mention in Natural Language
Shomir Wilson