Introduction

These are the proceedings of the NAACL-HLT 2007 Doctoral Consortium. Ten participants were selected from a total of eighteen applicants based on submission quality, advisor recommendation, and expected completion date for the PhD degree. The goals of this event are to provide these senior Ph.D. students with the opportunity to discuss and explore their research and career objectives with a panel of established researchers in the fields of natural language processing, speech technology, and information retrieval, and to develop the skills necessary to effectively communicate their research in preparation for future job talks.
Organizers:
Jackson Liscombe, Columbia University
Phillip Michalak, University of Rochester

Faculty Advisor:
Julia Hirschberg, Columbia University

Panelists:
James Allen
Chris Brew
Ciprian Chelba
Mona Diab
Graeme Hirst
Ed Hovy
Kevin Knight
Roland Kuhn
Gina Levow
Mitch Marcus
Bob Moore
Ani Nenkova
Mari Ostendorf
Michael Riley
Brian Roark
Stephen Robertson
Candace Sidner
Mark Steedman
Ellen Voorhees
Bonnie Webber
Ralph Weischedel
# Table of Contents

**Query Expansion Using Domain Information in Compounds**  
Karin Friberg ................................................................. 1

**Learning Structured Classifiers for Statistical Dependency Parsing**  
Qin Iris Wang ................................................................. 5

**Creating a Knowledge Base from a Collaboratively Generated Encyclopedia**  
Simone Paolo Ponzetto .................................................. 9

**Knowledge-Based Labeling of Semantic Relationships in English**  
Alicia Tribble ............................................................... 13

**Analysis of Summarization Evaluation Experiments**  
Marie-Josée Goulet ....................................................... 17

**Exploiting Event Semantics to Parse the Rhetorical Structure of Natural Language Text**  
Rajen Subba ................................................................. 21

**Dynamic Use of Ontologies in Dialogue Systems**  
Joana Paulo Pardal ....................................................... 25

**Semantic Frames in Romanian Natural Language Processing Systems**  
Diana Marie Trandabăt .................................................... 29

**Combining Evidence for Improved Speech Retrieval**  
J. Scott Olsson ............................................................. 33

**Unsupervised Natural Language Processing Using Graph Models**  
Chris Biemann ............................................................. 37
Doctoral Consortium Program

Sunday, April 22, 2007

8:30–9:00 Breakfast

9:00–9:40 *Query Expansion Using Domain Information in Compounds*
Karin Friberg

9:40–10:20 *Learning Structured Classifiers for Statistical Dependency Parsing*
Qin Iris Wang

10:20–11:00 *Creating a Knowledge Base from a Collaboratively Generated Encyclopedia*
Simone Paolo Ponzetto

11:00–11:40 *Knowledge-Based Labeling of Semantic Relationships in English*
Alicia Tribble

11:40–12:20 *Analysis of Summarization Evaluation Experiments*
Marie-Josée Goulet

12:20–13:20 Lunch

13:20–14:00 *Exploiting Event Semantics to Parse the Rhetorical Structure of Natural Language Text*
Rajen Subba

14:00–14:40 *Dynamic Use of Ontologies in Dialogue Systems*
Joana Paulo Pardal

14:40–15:20 *Semantic Frames in Romanian Natural Language Processing Systems*
Diana Marie Trandabăț

15:20–16:00 *Combining Evidence for Improved Speech Retrieval*
J. Scott Olsson

16:00–16:40 *Unsupervised Natural Language Processing Using Graph Models*
Chris Biemann

16:40–17:00 Coffee Break

17:00–18:00 Panel Discussion