

ACL-05

# **Empirical Modeling of Semantic Equivalence and Entailment**

**Proceedings of the Workshop**

30 June 2005  
University of Michigan  
Ann Arbor, Michigan, USA

Production and Manufacturing by  
*Omnipress Inc.*  
*Post Office Box 7214*  
*Madison, WI 53707-7214*

Sponsorship gratefully received from  
*Microsoft Research*  
*One Microsoft Way*  
*Redmond, Washington 98052, USA*

©2005 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)  
75 Paterson Street, Suite 9  
New Brunswick, NJ 08901  
USA  
Tel: +1-732-342-9100  
Fax: +1-732-342-9339  
[acl@aclweb.org](mailto:acl@aclweb.org)

## Introduction

The last few years have seen a surge in interest in modeling techniques aimed at measuring semantic equivalence and entailment, with work on paraphrase acquisition/generation, WordNet-based expansion, distributional similarity, supervised learning of semantic variability in information extraction, and the identification of patterns in template-based QA. Being able to identify when two strings "mean the same thing" or that one entails the other are crucial abilities for a broad range of NLP-related applications, ranging from question answering to summarization.

These proceedings contain a rich variety of papers centered on the problem of modeling semantic overlap between linguistic strings. This is a difficult problem space, encompassing issues of lexical choice, syntactic alternation, semantic inference, and reference/discourse structure.

We were pleased by the strong level of interest in the workshop, which resulted in a number of high-quality submissions. Each paper was blind-reviewed by 2-3 members of the Program Committee, and we were forced to make some difficult choices in determining the final schedule.

This workshop is intended to bring together people working on empirical, application-independent approaches to the practical problems of semantic inference. While different applications face similar underlying semantic problems, these problems have typically been addressed in an application-specific manner. In the absence of a generic evaluation framework, it is difficult to compare semantic methods that were developed for different applications. We are particularly hopeful that the workshop will help foster discussion around common datasets and evaluation strategies that will help guide future work in this area.

We would like to express our deepest gratitude to the hard-working members of the program committee. We'd also like to thank Mirella Lapata, Jason Eisner, Philipp Koehn, and Dragomir Radev for their organizational help.

We hope you enjoy this workshop!

Bill Dolan and Ido Dagan

**Organizers:**

Bill Dolan, Microsoft Research  
Ido Dagan, Bar Ilan University

**Program Committee:**

Srinivas Bangalore, AT&T Research  
Regina Barzilay, MIT  
Chris Brockett, Microsoft Research  
Pascale Fung, Hong Kong University of Science and Technology  
Oren Glickman, Bar Ilan University  
Cyril Goutte, Xerox Research Centre Europe  
Ed Hovy, ISI  
Kentaro Inui, Nara Institute of Science and Technology  
Dekang Lin, University of Alberta  
Daniel Marcu, ISI  
Kathy McKeown, Columbia University  
Dan Moldovan, University of Texas at Dallas  
Chris Quirk, Microsoft Research  
Maarten de Rijke, University of Amsterdam  
Hinrich Schuetze, University of Stuttgart  
Satoshi Sekine, New York University  
Peter Turney, National Research Council of Canada

**Invited Speaker:**

Dan Roth, University of Illinois at Urbana-Champaign

## Table of Contents

<i>Classification of Semantic Relations by Humans and Machines</i> Erwin Marsi and Emiel Krahmer .....	1
<i>The Distributional Similarity of Sub-Parses</i> Julie Weeds, David Weir and Bill Keller .....	7
<i>Measuring the Semantic Similarity of Texts</i> Courtney Corley and Rada Mihalcea .....	13
<i>Training Data Modification for SMT Considering Groups of Synonymous Sentences</i> Hideki Kashioka .....	19
<i>Recognizing Paraphrases and Textual Entailment Using Inversion Transduction Grammars</i> Dekai Wu .....	25
<i>Local Textual Inference: Can it be Defined or Circumscribed?</i> Annie Zaenen, Lauri Karttunen and Richard Crouch .....	31
<i>Discovering Entailment Relations Using "Textual Entailment Patterns"</i> Fabio Massimo Zanzotto, Maria Teresa Pazienza and Marco Pennacchiotti .....	37
<i>A Probabilistic Setting and Lexical Cooccurrence Model for Textual Entailment</i> Oren Glickman and Ido Dagan .....	43
<i>Generating an Entailment Corpus from News Headlines</i> John Burger and Lisa Ferro .....	49
<i>Definition and Analysis of Intermediate Entailment Levels</i> Roy Bar-Haim, Idan Szpektor and Oren Glickman .....	55



## Conference Program

### Thursday, June 30, 2005

- 9:00–9:15 Opening Remarks by Bill Dolan and Ido Dagan
- 9:15–9:40 *Classification of Semantic Relations by Humans and Machines*  
Erwin Marsi and Emiel Krahmer
- 9:40–10:05 *The Distributional Similarity of Sub-Parses*  
Julie Weeds, David Weir and Bill Keller
- 10:05–10:30 *Measuring the Semantic Similarity of Texts*  
Courtney Corley and Rada Mihalcea
- 10:30–11:00 Break
- 11:00–11:25 *Training Data Modification for SMT Considering Groups of Synonymous Sentences*  
Hideki Kashioka
- 11:25–12:25 Invited Talk by Dan Roth, University of Illinois at Urbana-Champaign
- 12:25–2:00 Lunch
- 2:00–2:25 *Recognizing Paraphrases and Textual Entailment Using Inversion Transduction Grammars*  
Dekai Wu
- 2:25–2:50 *Local Textual Inference: Can it be Defined or Circumscribed?*  
Annie Zaenen, Lauri Karttunen and Richard Crouch
- 2:50–3:15 *Discovering Entailment Relations Using "Textual Entailment Patterns"*  
Fabio Massimo Zanzotto, Maria Teresa Pazienza and Marco Pennacchiotti
- 3:15–3:40 *A Probabilistic Setting and Lexical Cooccurrence Model for Textual Entailment*  
Oren Glickman and Ido Dagan
- 3:40–4:00 Break

**Thursday, June 30, 2005 (continued)**

4:00–4:25     *Generating an Entailment Corpus from News Headlines*  
John Burger and Lisa Ferro

4:25–4:50     *Definition and Analysis of Intermediate Entailment Levels*  
Roy Bar-Haim, Idan Szpektor and Oren Glickman

4:50–5:40     Panel Discussion