Meta4NLP 2013

**The First Workshop on Metaphor in NLP**

**Proceedings of the Workshop**

13 June 2013
Atlanta, GA, USA

# Introduction

Characteristic to all areas of human activity (from poetic to ordinary to scientific) and, thus, to all types of discourse, metaphor becomes an important problem for natural language processing. Its ubiquity in language has been established in a number of corpus studies and the role it plays in human reasoning has been confirmed in psychological experiments. This makes metaphor an important research area for computational and cognitive linguistics, and its automatic identification and interpretation indispensable for any semantics-oriented NLP application.

The work on metaphor in NLP and AI started in the 1980s, providing us with a wealth of ideas on the structure and mechanisms of the phenomenon. The last decade witnessed a technological leap in natural language computation, whereby manually crafted rules gradually give way to more robust corpus-based statistical methods. This is also the case for metaphor research. In the recent years, the problem of metaphor modeling has been steadily gaining interest within the NLP community, with a growing number of approaches exploiting statistical techniques. Compared to more traditional approaches based on hand-coded knowledge, these more recent methods tend to have a wider coverage, as well as be more efficient, accurate and robust. However, even the statistical metaphor processing approaches so far often focused on a limited domain or a subset of phenomena. At the same time, recent work on computational lexical semantics and lexical acquisition techniques, as well as a wide range of NLP methods applying machine learning to open-domain semantic tasks, open many new avenues for creation of large-scale robust tools for recognition and interpretation of metaphor.

This workshop is the first one focused on modelling of metaphor using NLP techniques. Recent related events include workshops on Computational Approaches to Figurative Language (NAACL 2007) and on Computational Approaches to Linguistic Creativity (NAACL 2009, NAACL 2010). We received 14 submissions and accepted 10. Each paper was carefully reviewed by at least 3 members of the Program Committee. The selected papers offer explorations into the following directions: (1) creation of metaphor-annotated datasets; (2) identification of new features that are useful for metaphor identification; (3) cross-lingual metaphor identification.

The papers represent a variety of approaches to utilization and creation of datasets. While existing annotated corpora were used in some papers (Dunn, Tsvetkov et al), most papers describe creation of new annotated materials. Along with annotation guidelines adapted from the MIP and MIPVU procedures (Badryzlova et al), more intuitive annotation protocols are explored in Beigman Klebanov and Flor, Hovy et al, Heintz et al, Mohler et al, and Strzalkowski et al.

The papers present a number of novel and extended features for metaphor detection. Topic models, abstractness/concreteness, and semantic classifications based on an ontology are each used in multiple papers. Additional features include classes of named entities (Tsvetkov et al), WordNet examples and glosses (Wilks et al); suggestive evidence is presented regarding potential usefulness of a relationality feature (Jamrozik et al). A distinguishing characteristic of multiple submissions is the interest in cross-lingual approaches to metaphor identification. Accordingly, contributors explore features that can be supported by resources that exist in languages like Russian, Spanish, and Farsi (Strzalkowski et al., Tsvetkov et al, Heintz et al).

The program of the workshop also features two invited talks that complement the discussion by

addressing topics that are not addressed by this year's submissions, namely, the relationship between metaphor and action (Srini Narayanan), and interpretation of metaphors (John Barnden).

We wish to thank everyone who showed interest and submitted a paper, all of the authors for their contributions, the members of the Program Committee for their thoughtful reviews, the invited speakers for sharing their perspectives on the topic, and all the attendees of the workshop. All of these factors contribute to a truly enriching event!

Workshop co-chairs:
Ekaterina Shutova, University of California at Berkeley, USA
Beata Beigman Klebanov, Educational Testing Service, USA
Joel Tetreault, Nuance, USA
Zornitsa Kozareva, USC Information Sciences Institute, USA

**Organizers:**

Ekaterina Shutova, University of California, Berkeley, USA
Beata Beigman Klebanov, Educational Testing Service, USA
Joel Tetreault, Nuance, USA
Zornitsa Kozareva, USC Information Sciences Institute, USA

**Program Committee:**

Shlomo Argamon, Illinois Institute of Technology, USA
John Barnden, University of Birmingham, UK
Gemma Boleda, University of Texas at Austin, USA
Danushka Bollegala, University of Tokyo, Japan
Marisa Boston, Nuance, USA
David Bracewell, LCC, USA
Ted Briscoe, University of Cambridge, UK
Jaime Carbonell, CMU, USA
Stephen Clark, University of Cambridge, UK
Paul Cook, University of Melbourne, Australia
Gerard de Melo, University of California at Berkeley, USA
Alice Deignan, Leeds University, UK
Afsaneh Fazly, University of Toronto, Canada
Anna Feldman, Montclair State University, USA
Jerry Feldman, University of California at Berkeley, USA
Michael Flor, Educational Testing Service, USA
Marjorie Freedman, BBN, USA
Deidre Gentner, Northwestern University, USA
Yanfen Hao, Electronics Industry Research Institute, ShanXi, China
Jerry Hobbs, University of Southern California, USA
Eugenie Giesbrecht, Karlsruhe Institute of Technology, Germany
Valia Kordoni, Humboldt University Berlin, Germany
Anna Korhonen, University of Cambridge, UK
George Lakoff, University of California at Berkeley, USA
Alex Lascarides, University of Edinburgh, UK
Mark Lee, University of Birmingham, UK
Patricia Lichtenstein, University of California, Merced, USA
Katja Markert, University of Leeds, UK
James H. Martin,University of Colorado at Boulder, USA
Andreas Musolff, University of East Anglia, UK
Srini Narayanan, University of California at Berkeley, USA
Malvina Nissim, University of Bologna, Italy
Diarmuid Ó Séaghdha, University of Cambridge, UK
Gerard Steen, VU Amsterdam, The Netherlands

Thierry Poibeau, Ecole Normale Superieure and CNRS, France
Caroline Sporleder, Saarland University, Germany
Carlo Strapparava, Fondazione Bruno Kessler, Italy
Tomek Strzalkowski, SUNY Albany, USA
Marc Tomlinson, LCC, USA
Oren Tsur, Hebrew University, Israel
Peter Turney, National Research Council Canada, Canada
Tim van de Cruys, IRIT and CNRS, Toulouse, France
Tony Veale, Korean Advanced Institute for Science and Technology, Republic of Korea
Aline Villavicencio, Federal University of Rio Grande do Sul, Brazil and MIT, USA
Andreas Vlachos, University of Cambridge, UK
Yorick Wilks, Florida Institute of Human and Machine Cognition, USA

**Invited Speakers:**

Srini Narayanan, University of California, Berkeley, USA
John Barnden, University of Birmingham, UK

# Table of Contents

# Workshop Program

**Thursday, June 13, 2013**

9:00–9:10      Opening remarks

9:10–10:05      Invited talk: Srini Narayanan "From Metaphor to Action"

10:05–10:30      *What metaphor identification systems can tell us about metaphor-in-language*
Jonathan Dunn

10:30–11:00      Coffee break

11:00–11:25      *Argumentation-Relevant Metaphors in Test-Taker Essays*
Beata Beigman Klebanov and Michael Flor

11:25–11:45      *Relational words have high metaphoric potential*
Anja Jamrozik, Eyal Sagi, Micah Goldwater and Dedre Gentner

11:45–12:10      *Semantic Signatures for Example-Based Linguistic Metaphor Detection*
Michael Mohler, David Bracewell, Marc Tomlinson and David Hinote

12:10–13:40      Lunch

13:40–14:20      Invited talk: John Barnden "Computational Approaches to Metaphor Interpretation: Some Considerations arising from a Deep Reasoning System"

14:20–14:45      *Automatic Metaphor Detection using Large-Scale Lexical Resources and Conventional Metaphor Extraction*
Yorick Wilks, Adam Dalton, James Allen and Lucian Galescu

14:45–15:10      *Cross-Lingual Metaphor Detection Using Common Semantic Features*
Yulia Tsvetkov, Elena Mukomel and Anatole Gershman

15:10–15:30      *Identifying Metaphorical Word Use with Tree Kernels*
Dirk Hovy, Shashank Shrivastava, Sujay Kumar Jauhar, Mrinmaya Sachan, Kartik Goyal, Huying Li, Whitney Sanders and Eduard Hovy

15:30–16:00      Coffee break

16:00–16:25      *Automatic Extraction of Linguistic Metaphors with LDA Topic Modeling*
Ilana Heintz, Ryan Gabbard, Mahesh Srivastava, Dave Barner, Donald Black, Majorie Friedman and Ralph Weischedel

16:25–16:50   *Robust Extraction of Metaphor from Novel Data*
Tomek Strzalkowski, George Aaron Broadwell, Sarah Taylor, Laurie Feldman, Samira
Shaikh, Ting Liu, Boris Yamrom, Kit Cho, Umit Boz, Ignacio Cases and Kyle Elliot

16:50–17:15   *Annotating a Russian corpus of conceptual metaphor: a bottom-up approach*
Yulia Badryzlova, Natalia Shekhtman, Yekaterina Isaeva and Ruslan Kerimov

17:15–17:30   Closing remarks