

ExProM 2016

**Extra-Propositional Aspects of Meaning
in Computational Linguistics**

Proceedings of the Workshop

December 12, 2016

Osaka, Japan

Funding for student travel grants was provided by the National Science Foundation under Grant No. 1523586. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Copyright of each paper stays with the respective authors (or their employers).

ISBN978-4-87974-718-1

Preface

During the last decade, semantic representation of text has focused on extracting propositional meaning, i.e., capturing who does what to whom, how, when and where. Several corpora are available, and existing tools extract this kind of knowledge, e.g., semantic role labelers trained on PropBank, NomBank or FrameNet. But propositional semantic representations disregard significant meaning encoded in human language. For example, while sentences (1-2) below share the same propositional meaning regarding verb *carry*, they do not convey the same overall meaning. In order to truly capture what these sentences mean, extra-propositional aspects of meaning (ExProM) such as uncertainty, negation and attribution must be taken into account.

1. Thomas Eric Duncan likely contracted the disease when he carried a pregnant woman sick with Ebola.
2. Thomas Eric personally told me that he never carried a pregnant woman with Ebola.

The Extra-Propositional Aspects of Meaning (ExProM) in Computational Linguistics Workshop focuses on a broad range of semantic phenomena beyond propositional meaning, i.e., beyond linking propositions and their semantic arguments with relations such as AGENT (who), THEME (what), LOCATION (where) and TIME (when).

ExProM is pervasive in human language and, while studied from a theoretical perspective, computational models are scarce. Humans use language to describe events that do not correlate with a real situation in the world. They express desires, intentions and plans, and also discuss events that did not happen or are unlikely to happen. Events are often described hypothetically, and speculation can be used to explain why something is a certain way without a strong commitment. Humans do not always (want to) tell the (whole) truth: they may use deception to hide lies. Devices such as irony and sarcasm are employed to play with words so that what is said is not what is meant. Finally, humans not only describe their personal views or experiences, but also attribute statements to others. These phenomena are not exclusive of opinionated texts. They are ubiquitous in language, including scientific works and news as exemplified below:

- A better team might have prevented this infection.
- Some speculate that this was a failure of the internal communications systems.
- Infected people typically don't become contagious until they develop symptoms.
- Medical personnel can be infected if they don't use protective gear, such as surgical masks and gloves.
- You cannot get it from another person until they start showing symptoms of the disease, like fever.
- You can only catch Ebola from coming into direct contact with the bodily fluids of someone who has the disease and is showing symptoms.
- We've never seen a human virus change the way it is transmitted.
- There is no reason to believe that Ebola virus is any different from any of the viruses that infect humans and have not changed the way that they are spread.

In its 2016 edition, the Extra-Propositional Aspects of Meaning (ExProM) in Computational Linguistics Workshop was collocated with the 26th International Conference on Computational Linguistics (COLING 2016) in Osaka, Japan. The workshop took place on December 12, 2016, and the program consisted of six papers and an invited talk by Preslav Nakov (Qatar Computing Research Institute, HBKU).

ExProM 2016 is a follow-up of three previous events: the 2010 Negation and Speculation in Natural Language Processing Workshop (NeSp-NLP 2010), ExProM 2012 and ExProM 2015. We would like to thank the authors of papers for their interesting contributions, the members of the program committee for their insightful reviews, and Preslav Nakov for being the invited speaker. We are also grateful to the National Science Foundation for a grant to support student travel to the workshop.

Eduardo Blanco, Roser Morante, and Roser Saurí.

Organisers

Eduardo Blanco, University of North Texas, USA
Roser Morante, VU University Amsterdam, The Netherlands
Roser Saurí, Oxford University Press, UK

Programme Committee

Mithun Balakrishna - Lymba Corporation
Emily M. Bender - University of Washington
Cosmin Adrian Bejan - Vanderbilt University
Tommaso Caselli - VU University Amsterdam
Marie-Catherine de Marneffe - The Ohio State University
Iris Hendrickx - Radboud University
Lori Levin - Carnegie Mellon University
Erwin Marsi - Norwegian University of Science and Technology
Malvina Nissim - University of Groningen
Christopher Potts - Stanford University
Sampo Pyysalo - University of Cambridge
German Rigau - UPV/EHU
Ellen Riloff - University of Utah
Paolo Rosso - Universitat Politècnica de Valencia
Erik Velldal - University of Oslo
Bonnie Webber - University of Edinburgh

Invited Speaker

Preslav Nakov - Qatar Computing Research Institute, HBKU

Table of Contents

<i>Program</i>	ix
<i>‘Who would have thought of that!’: A Hierarchical Topic Model for Extraction of Sarcasm-prevalent Topics and Sarcasm Detection</i> Aditya Joshi, Prayas Jain, Pushpak Bhattacharyya and Mark Carman	1
<i>Detecting Uncertainty Cues in Hungarian Social Media Texts</i> Veronika Vincze	11
<i>Detecting Level of Belief in Chinese and Spanish</i> Juan Pablo Colomer, Keyu Lai and Owen Rambow	22
<i>Contradiction Detection for Rumorous Claims</i> Piroska Lendvai and Uwe Reichel	31
<i>Invited Talk: Negation and Modality in Machine Translation</i> Preslav Nakov	41
<i>Problematic Cases in the Annotation of Negation in Spanish</i> Salud María Jiménez-Zafra, Maite Martín, L. Alfonso Ureña Lopez, Toni Martí and Mariona Taulé 42	
<i>Building a Dictionary of Affixal Negations</i> Chantal van Son, Emiel van Miltenburg and Roser Morante	49

Workshop Program

Monday December 12, 2016

9:00–9:10 *Opening remarks*

9:10–9:45 *'Who would have thought of that!': A Hierarchical Topic Model for Extraction of Sarcasm-prevalent Topics and Sarcasm Detection*
Aditya Joshi, Prayas Jain, Pushpak Bhattacharyya and Mark Carman

09:45–10:20 *Detecting Uncertainty Cues in Hungarian Social Media Texts*
Veronika Vincze

10:20–10:40 **Coffee break**

10:40–11:15 *Detecting Level of Belief in Chinese and Spanish*
Juan Pablo Colomer, Keyu Lai and Owen Rambow

11:15–11:50 *Contradiction Detection for Rumorous Claims*
Piroska Lendvai and Uwe Reichel

11:50–14:00 **Lunch break**

Invited talk

14:00–15:00 *Negation and Modality in Machine Translation*
Preslav Nakov

15:00–15:20 *Coffee break*

15:20–15:55 *Problematic Cases in the Annotation of Negation in Spanish*
Salud María Jiménez-Zafra, Maite Martín, L. Alfonso Ureña Lopez, Toni Martí and Mariona Taulé

15:55–16:30 *Building a Dictionary of Affixal Negations*
Chantal van Son, Emiel van Miltenburg and Roser Morante

16:30–16:50 *Discussion and closing remarks*

