

Contents

Preface	ix
A New Life for Semantic Annotations?	
Bunt	1
PART I REGULAR PAPERS	3
Combining Knowledge-based Methods and Supervised Learning for Effective Italian Word Sense Disambiguation	
Basile, de Gemmis, Lops, and Semeraro	5
Semantic Representations of Syntactically Marked Discourse Status in Crosslinguistic Perspective	
Bender and Goss-Grubbs	17
High Precision Analysis of NPs with a Deep Processing Grammar	
Branco and Costa	31
Augmenting WordNet for Deep Understanding of Text	
Clark, Fellbaum, Hobbs, Harrison, Murray, and Thompson	45
How Well Do Semantic Relatedness Measures Perform? A Meta-Study	
Cramer	59
KnowNet: A Proposal for Building Highly Connected and Dense Knowledge Bases from the Web	
Cuadros and Rigau	71
Combining Word Sense and Usage for Modeling Frame Semantics	
De Cao, Croce, Pennacchiotti, and Basili	85
Answering Why-Questions in Closed Domains from a Discourse Model	
Delmonte and Pianta	103
Analyzing the Explanation Structure of Procedural Texts: Dealing with Advice and Warnings	
Fontan and Saint-Dizier	115

From Predicting Predominant Senses to Local Context for Word Sense Disambiguation	
Koeling and McCarthy	129
Automatic Fine-Grained Semantic Classification for Domain Adaptation	
Liakata and Pulman	139
Analysis of ASL Motion Capture Data towards Identification of Verb Type	
Malaia, Borneman, and Wilbur	155
The Idiom–Reference Connection	
McShane and Nirenburg	165
Resolving Paraphrases to Support Modeling Language Perception in an Intelligent Agent	
Nirenburg, McShane, and Beale	179
Everyday Language is Highly Intensional	
Ramsay and Field	193
Refining the Meaning of Sense Labels in PDTB: “Concession”	
Robaldo, Miltsakaki, and Hobbs	207
Connective-based Local Coherence Analysis: A Lexicon for Recognizing Causal Relationships	
Stede	221
Open Knowledge Extraction through Compositional Language Processing	
Van Durme and Schubert	239
PART II SHARED TASK	255
Introduction to the Shared Task on Comparing Semantic Representations	
Bos	257
Boeing’s NLP System and the Challenges of Semantic Representation	
Clark and Harrison	263
Wide-Coverage Semantic Analysis with Boxer	
Bos	277
Semantic and Pragmatic Computing with GETARUNS	
Delmonte	287
LXGram in the Shared Task “Comparing Semantic Representations” of STEP 2008	
Branco and Costa	299

Baseline Evaluation of WSD and Semantic Dependency in OntoSem	
Nirenburg, Beale, and McShane	315
The TEXTCAP Semantic Interpreter	
Callaway	327
Deep Semantic Analysis of Text	
Allen, Swift, and de Beaumont	343
PART III SHORT PAPERS	355
Textual Entailment as an Evaluation Framework for Metaphor Resolution: A Proposal	
Agerri, Barnden, Lee, and Wallington	357
Representing and Visualizing Calendar Expressions in Texts	
Battistelli, Couto, Minel, and Schwer	365
Addressing the Resource Bottleneck to Create Large-Scale Annotated Texts	
Chamberlain, Poesio, and Kruschwitz	375
A Resource-Poor Approach for Linking Ontology Classes to Wikipedia Articles	
Reiter, Hartung, and Frank	381
Top-Down Cohesion Segmentation in Summarization	
Tatar, Mihis, and Serban	389

Preface

Background and Motivation

Thanks to both statistical approaches and finite state methods, natural language processing (NLP), particularly in the area of robust, open-domain text processing, has made considerable progress in the last couple of decades. It is probably fair to say that NLP tools have reached satisfactory performance at the level of syntactic processing, be the output structures chunks, phrase structures, or dependency graphs. Therefore, the time seems ripe to extend the state-of-the-art and consider deep semantic processing as a serious task in wide-coverage NLP.

This is a step that normally requires syntactic parsing, as well as integrating named entity recognition, anaphora resolution, thematic role labelling and word sense disambiguation, and other lower levels of processing for which reasonably good methods have already been developed.

The goal of the STEP workshop is to provide a forum for anyone active in semantic processing of text to discuss innovative technologies, representation issues, inference techniques, prototype implementations, and real applications. The preferred processing targets are large quantities of texts — either specialised domains, or open domains such as newswire text, blogs, and wikipedia-like text. Implemented rather than theoretical work is emphasised in STEP.

Featuring in STEP 2008 workshop is a “shared task” on comparing semantic representations as output by state-of-the-art NLP systems. Participants were asked to supply a (small) text, before the workshop. The test data for the shared task is composed out of all the texts submitted by the participants, allowing participants to “challenge” each other. The output of these systems will be judged on a number of aspects by a panel of experts in the field, during the workshop.

Welcome to STEP 2008

STEP 2008 is organised as a three-day event at Ca’ Dolfin, at the Università Ca’ Foscari in Venice, Italy, taking place on September 22–24. In reply to our call for papers we received 40 submissions: 24 regular papers, 8 short papers, and 8 shared task papers. We accepted 30 of these: 18 regular papers, 5 short papers, and 7 shared task papers (yielding an overall acceptance rate of 75%). We would like to thank the referees and members of the programme committee for helping us to review and select the papers:

Roberto Basili (University of Rome “Tor Vergata”, Italy)
Johan Bos (University of Rome “La Sapienza”, Italy)
Ann Copestake (University of Cambridge, UK)

Rodolfo Delmonte (University of Venice “Ca’ Foscari”)
Nicola Guarino (ISTC-CNR, Trento, Italy)
Sanda Harabagiu (HLT, University of Texas, USA)
Alexander Koller (University of Edinburgh, UK)
Leonardo Lesmo (DI, University of Turin, Italy)
Katja Markert (University of Leeds, UK)
Eva Mok (ICSI, Berkeley, USA)
Dan Moldovan (HLT, University of Texas, USA)
Srini Narayanan (ICSI, Berkeley, USA)
Sergei Nirenburg (University of Maryland, USA)
Malvina Nissim (University of Bologna, Italy)
Vincenzo Pallotta (University of Freiburg, Switzerland)
Emanuele Pianta (ITC, Trento, Italy)
Massimo Poesio (University of Trento, Italy)
Stephen Pulman (Oxford University, UK)
Michael Schiehlen (IMS Stuttgart, Germany)
Bonnie Webber (University of Edinburgh, UK)

We also would like to thank Malvina Nissim and Daniel Bos for advice on formatting the proceedings, Gertjan Bos for designing the STEP logo, and Suhel Jaber for maintaining the STEP 2008 web site. We also thank Jane Spurr from College Publications for her general support and her advice on preparing the camera-ready version of this book. Finally, we’re very grateful to Harry Bunt (Tilburg University) and Sanda Harabagiu (University of Texas at Dallas) for giving invited presentations at STEP 2008.

Johan Bos & Rodolfo Delmonte
Italy, July 2008