Contents

1 Schedule .......................................................... 1
2 Welcome to IWCS 2013 ........................................ 3
3 Organization ......................................................... 4
4 Main Conference ................................................... 7
   Schedule .................................................................. 7
   Abstracts .................................................................. 12
5 Workshop: Annotation of Modal Meaning in Natural Language ............................................. 21
6 Workshop: Computational Semantics in Clinical Text .......................................................... 27
7 Workshop: Computational Models of Spatial Language Interpretation and Generation ........... 31
8 Workshop: Towards a Formal Distribution Semantics .......................................................... 35
9 Workshop: Interoperable Semantic Annotation ................................................................. 39
10 Getting around ....................................................... 45
   Venue .................................................................... 45
   Around the city ..................................................... 45
   Hotel Mercure ....................................................... 47
   Where to eat ......................................................... 47
   Banquet .................................................................. 50
11 Further Information ............................................... 53
   Internet and WiFi .................................................. 53
   Sightseeing .......................................................... 53
   Contact info ......................................................... 53
12 S7 train timetable .................................................... 54
13 Schedule ............................................................. 56
Welcome to IWCS 2013 in Potsdam! We are looking forward to a great conference with all of you.

The aim of IWCS is to provide a forum for researchers interested in all areas of computational semantics: computation, annotation, extraction, and representation of meaning in natural language, through approaches that are symbolic, statistical, or anything in between. And in particular, the aim of IWCS is not only to provide a forum, but to bring together researchers from all these different areas and foster discussions that bridge the gaps between areas. The IWCS 2013 conference continues this eclectic tradition, as its list of sessions illustrates: We have papers on discourse, textual entailment, semantic annotation, distributional models, semantic roles, corpus studies, semantic relations, formal semantics, natural language generation, inferences, and ontologies.

The conference program consists of 25 long papers (presented as talks), and 16 short papers (presented as posters and lightning talks). These presentations were selected from 60 long and 22 short submissions. With 3 papers that were submitted as long but accepted as short, the overall acceptance rate is 42% for long papers, and 64% for short. In addition, we are experimenting with an explicitly discussion-oriented format: Thursday afternoon is set aside for an Open Space event, also known as an Unconference. This will be a time at which you can propose any topic for discussion that you wish. We look forward to seeing all the topics you’ll come up with!

We thank all the people who helped organize IWCS 2013. We were delighted to have such a wonderful program committee, who helped put together an excellent conference program. We were also lucky to work with a local organizing team that has done an amazing job with setting up the conference venue and preparing the website, proceedings, and conference handbook. They are listed in detail in this handbook; be sure to show them your appreciation. Thanks, guys: This conference would not have happened without you.

Now all that remains for us to do is to wish you an enjoyable conference!

Katrin Erk and Alexander Koller
IWCS 2013 Co-Chairs
ORGANIZATION

Main Conference

PROGRAM CHAIR

Katrin Erk  University of Texas at Austin

PROGRAM COMMITTEE

Tim Baldwin  Tim Baldwin
Marco Baroni  Marco Baroni
Anya Belz  Anya Belz
Jonathan Berantz  Jonathan Berantz
Raffaella Bernardi  Raffaella Bernardi
Patrick Blackburn  Patrick Blackburn
Gemma Boleda  Gemma Boleda
Johan Bos  Johan Bos
Antonio Branco  Antonio Branco
Chris Brew  Chris Brew
Susan Brown  Susan Brown
Paul Buitelaar  Paul Buitelaar
Harry Bunt  Harry Bunt
Aljoscha Burchardt  Aljoscha Burchardt
Nicoletta Calzolari  Nicoletta Calzolari
Rui Chaves  Rui Chaves
Philipp Cimiano  Philipp Cimiano
Ariel Cohen  Ariel Cohen
Robin Cooper  Robin Cooper
Ann Copestake  Ann Copestake
Montse Cuadros  Montse Cuadros
Dipanjan Das  Dipanjan Das
Rodolfo Delmonte  Rodolfo Delmonte
Dmitriy Dligach  Dmitriy Dligach
Markus Egg  Markus Egg
Raquel Fernandez  Raquel Fernandez
Anette Frank  Anette Frank
Claire Gardent  Claire Gardent
Dan Garrette  Dan Garrette
Jonathan Ginzburg  Jonathan Ginzburg
Edward Grefenstette  Edward Grefenstette
Ed Hovy  Ed Hovy
Lauri Karttunen  Lauri Karttunen
Rohit Kate  Rohit Kate
Ralf Klabunde  Ralf Klabunde
Emiel Kraher  Emiel Kraher
Shalom Lappin  Shalom Lappin
Alex Lascarides  Alex Lascarides
Piroska Lendvai  Piroska Lendvai
Leonardo Lesmo  Leonardo Lesmo
Will Lowe  Will Lowe
Paul Mc Kevitt  Paul Mc Kevitt
Diana McCarthy  Diana McCarthy
Marie-Francine Moens  Marie-Francine Moens
Raymond Mooney  Raymond Mooney
Alessandro Moschitti  Alessandro Moschitti
Roberto Navigli  Roberto Navigli
Malvina Nissim  Malvina Nissim
Diarmaid O Seaghdha  Diarmuid O Seaghdha
Katya Ovchinnikova  Katya Ovchinnikova
Sebastian Pado  Sebastian Pado
Vincenzo Pallotta  Vincenzo Pallotta
Alexis Palmer  Alexis Palmer
Anselmo Penas  Anselmo Penas
Marco Pennachiotti  Marco Pennachiotti
Manfred Pinkal  Manfred Pinkal
Paul Piwek  Paul Piwek
Richard Power  Richard Power
Stephen Pulman  Stephen Pulman
Allan Ramsay  Allan Ramsay
Michaela Regneri  Michaela Regneri
Stephen Roller  Stephen Roller
Rolf Schwitter  Rolf Schwitter
Jennifer Spenader  Jennifer Spenader
Caroline Sporleder  Caroline Sporleder
Manfred Stede  Manfred Stede
Mary Swift  Mary Swift
Idan Szpektor  Idan Szpektor
Stefan Thater  Stefan Thater
Peter Turney  Peter Turney
Tim van de Cruys  Tim van de Cruys
Kees van Deemter  Kees van Deemter
Benjamin Van Durme  Benjamin Van Durme
Jan van Eijck  Jan van Eijck
Josef van Genabith  Josef van Genabith
Carl Vogel  Carl Vogel
Fabio Zanzotto  Fabio Zanzotto
## Local organization

### LOCAL CHAIR

- **Alexander Koller**  
  University of Potsdam

### LOCAL ORGANIZERS

- **Ines Mauer**  
  University of Potsdam
- **Konstantina Garoufi**  
  University of Potsdam

### WEBSITE

- **Dan Garrette**  
  University of Texas at Austin
- **Nikos Engonopoulos**  
  University of Potsdam

### PUBLICATIONS

- **Thomas Hanneforth**  
  University of Potsdam

### CONFERENCE HANDBOOK

- **Martín Villalba**  
  University of Potsdam
main conference

The main conference will take place in Seminar Room H02 on the ground floor.

Schedule

Tuesday, March 19

Workshops

<table>
<thead>
<tr>
<th>Room</th>
<th>Title</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Room 12</td>
<td>Annotation of Modal Meanings in Natural Language (WAMM)</td>
<td>March 19, 9:30 – 17:20</td>
</tr>
<tr>
<td>Room 14</td>
<td>Computational Semantics in Clinical Text (CSCT 2013)</td>
<td>March 19, 14:00 – 17:30</td>
</tr>
<tr>
<td>Room 15</td>
<td>Computational Models of Spatial Language Interpretation and Generation (CoSLI-3)</td>
<td>March 19, 9:10 – 13:30</td>
</tr>
<tr>
<td>Room 16</td>
<td>Towards a formal distributional semantics</td>
<td>March 19, 9:15 – 17:30</td>
</tr>
<tr>
<td>Room 23</td>
<td>ISA-9 Joint ACL – ISO Workshop on Interoperable Semantic Annotation</td>
<td>March 19, 9:15– 17:00, March 20, 10:05–17:00</td>
</tr>
</tbody>
</table>

17:30 – 19:30 Welcome reception
Conference Venue

Wednesday, March 20

08:45 – 09:00 Opening remarks

Invited Talk

09:00 – 10:00 Truth-conditional, distributional, compositional: On the way to a unified framework for Computational Semantics?
Manfred Pinkal
**DISCOURSE**

Session Chair: Gemma Boleda

10:00 – 10:30  **What excludes an alternative in coherence relations**  
Bonnie Webber

10:30 – 11:00  **Graph-based classification of explicit and implicit discourse relations**  
Yannick Versley

11:00 – 11:30  ☕️  Coffee break

**RTE/Annotation**

Session Chair: Marco Baroni

11:30 – 12:00  **Parsimonious semantic representations with projection pointers**  
Noortje J. Venhuizen, Johan Bos, and Harm Brouwer

12:00 – 12:30  **A search task dataset for German textual entailment**  
Britta Zeller and Sebastian Padó

12:30 – 13:00  **Semantic annotation of textual entailment**  
Assaf Toledo, Stavroula Alexandropoulou, Sophia Katrenko, Heidi Klockmann, Pepijn Kokke and Yoad Winter

13:00 – 14:00  🍽️  Lunch

**Distributional Approaches**

Session Chair: Sebastian Pado

14:00 – 14:30  **Multi-step regression learning for compositional distributional semantics**  
Edward Grefenstette, Georgiana Dinu, Yao-Zhong Zhang, Mehrnoosh Sadrzadeh and Marco Baroni

14:30 – 15:00  **Evaluating topic coherence using distributional semantics**  
Nikolaos Aletras and Mark Stevenson

15:00 – 15:30  **Towards a semantics for distributional representations**  
Katrin Erk

15:30 – 16:00  **Intensionality was only alleged: on adjective-noun composition in distributional semantics**  
Gemma Boleda, Marco Baroni, Nghia The Pham and Louise McNally

16:00 – 16:30  ☕️  Coffee break
### SEMANTIC ROLES

**Session Chair:** Mary Swift

**16:30 – 17:00**  
**Sources of evidence for implicit argument resolution**  
Egoitz Laparra and German Rigau

**17:00 – 17:30**  
**Towards weakly supervised resolution of null instantiations**  
Philip Gorinski, Josef Ruppenhofer and Caroline Sporleder

### CORPUS STUDIES

**Session Chair:** Mary Swift

**17:30 – 18:00**  
**A corpus study of clause combination**  
Olga Nikitina and Sebastian Padó

**18:00 – 18:30**  
**Regular meaning shifts in German particle verbs: a case study**  
Sylvia Springorum, Jason Utt and Sabine Schulte im Walde

**19:30**  
Conference dinner

---

**Thursday, March 21**

### SEMANTIC RELATIONS

**Session Chair:** Anette Frank

**09:00 – 09:30**  
**Recognizing spatial containment relations between event mentions**  
Kirk Roberts, Michael Skinner and Sanda Harabagiu

**09:30 – 10:00**  
**Temporal relation classification based on temporal reasoning**  
Francisco Costa and António Branco

**10:00 – 10:30**  
**The impact of selectional preference agreement on semantic relational similarity**  
Bryan Rink and Sanda Harabagiu

**10:30 – 11:00**  
Coffee break

### SEMANTICS

**Session Chair:** Louise McNally

**11:00 – 11:30**  
**UCCA: A semantics-based grammatical annotation scheme**  
Omri Abend and Ari Rappoport

**11:30 – 12:00**  
**Empirical validation of Reichenbach’s Tense framework**  
Leon Derczynski and Robert Gaizauskas
12:00 – 12:30  Probabilistic induction for an incremental semantic grammar  
Arash Eshghi, Matthew Purver and Julian Hough

12:30 – 13:30  Lunch

13:30 – 18:30  Open Space afternoon

Friday, March 22

INVITED TALK

09:00 – 10:00  Structures for machine translation: strings, trees, and graphs  
Kevin Knight

NATURAL LANGUAGE GENERATION

Session Chair: Tatjana Scheffler

10:00 – 10:30  Domain adaptable semantic clustering in statistical NLG  
Blake Howald, Ravikumar Kondadadi and Frank Schilder

10:30 – 11:00  Generating natural language from linked data: unsupervised template extraction  
Daniel Duma and Ewan Klein

POSTER SESSION

11:00 – 11:30  Lightning talks

A corpus-based taxonomy of question responses  
Pawel Lupkowski and Jonathan Ginzburg

A pilot experiment in knowledge authoring as dialogue  
Artemis Parvizi, Caroline Jay, Christopher Mellish, Jeff Pan, Yuan Ren, Robert Stevens and Kees van Deemter

Finite state temporality and context free languages  
Derek Kelleher and Carl Vogel

Fitting, not clashing! A distributional semantic model of logical metonymy  
Alessandra Zarcone, Alessandro Lenci, Sebastian Padó and Jason Utt

Gamification for word sense labeling  
Noortje J. Venhuizen, Valerio Basile, Kilian Evang and Johan Bos

Learning semantic representations in a bigram language model  
Jeff Mitchell

Logic programs vs. first-order formulas in textual inference  
Yuliya Lierler and Vladimir Lifschitz

Predicate-specific annotations for implicit role binding: corpus annotation, data analysis and evaluation experiments  
Tatjana Moor, Michael Roth and Anette Frank

Pre-processing MRSES  
Tore Bruland
Scope disambiguation as a tagging task  
Kilian Evang and Johan Bos

Semantic similarity computation for abstract and concrete nouns using network-based distributional semantic models  
Elias Iosif, Alexandros Potamianos, Maria Giannoudaki and Kalliopi Zervanou

Situated utterances and discourse relations  
Matthew Stone, Una Stojnic and Ernest Lepore

Towards a tight integration of syntactic parsing with semantic disambiguation by means of declarative programming  
Yuliya Lierler and Peter Schüller

The annotation of quantification and its interpretation  
Harry Bunt

Using network approaches to enhance the analysis of cross-linguistic polysemy  
Johann-Mattis List, Anselm Terhalle and Matthias Urban

What is a text, what isn’t, and what this has to do with lexical semantics  
Aurelie Herbelot

11:30 – 13:00  🍽 Posters and coffee

13:00 – 14:00  🍽️ Lunch

Inference and Ontologies

Session Chair: Graeme Hirst

14:00 – 14:30  Sentiment composition using a parabolic model  
Baptiste Chardon, Farah Benamara, Yvette Mathieu, Vladimir Popescu and Nicholas Asher

14:30 – 15:00  Recognising sets and their elements: tree kernels for entity instantiation identification  
Andrew McKinlay and Katja Markert

15:00 – 15:30  Automatically deriving event ontologies for a commonsense knowledge base  
James Allen, Jansen Orfan, Will de Beaumont, Choh Man Teng, Lucian Galescu and Mary Swift

15:30 – 16:00  Learning corpus patterns using finite state automata  
Octavian Popescu

16:00 – 16:15  Closing remarks
Abstracts

Wednesday Morning 1: Invited Talk
(09:00 – 10:00)

Truth-conditional, distributional, compositional: On the way to a unified framework for Computational Semantics?
Manfred Pinkal

Truth-conditional semantics for natural language provides for clear concepts of denotation, compositionality, and entailment. While it has formed the basis for a deep and detailed description of the structure of meaning in natural language, it falls dramatically short when it comes to robustness and coverage. Distributional semantics, on the other hand, enables efficient and comfortable acquisition of wide-coverage lexical semantic information from raw text corpora, but it does not easily lend itself to the modeling of structured information about meaning. Nor does it offer a straightforward story about the relationship between language and the world.

The truth-conditional and distributional paradigms are clearly complementary in their strengths and weaknesses. It is not at all obvious, however, how the views on semantics and semantic processing provided by the two respective frameworks might combine into a comprehensive and consistent picture.

In this talk, I will inspect recent approaches that aim to interleave the two paradigms, and I will discuss how much progress we have made towards a unified framework for computational semantics.

Wednesday Morning 2: Discourse
(10:00 – 11:00)

What excludes an alternative in coherence relations?
Bonnie Webber

This paper identifies features that occur frequently in coherence relations labeled CHOSEN ALTERNATIVE. This achieves two goals: (1) to identify evidence for an argument being considered an alternative excluded from further consideration, and (2) to contribute to the automatic identification of coherence relations and their arguments. It is shown that the simplest of these features occur significantly more often in implicit CHOSEN ALTERNATIVE relations than in explicit CHOSEN ALTERNATIVE relations, where a connective helps signal this sense.

Subgraph-based classification of explicit and implicit disclosure relations
Yannick Versley

Current approaches to recognizing discourse relations rely on a combination of shallow, surface-based features (e.g., bigrams, word pairs), and rather specialized hand-crafted features. As a way to avoid both the shallowness of word-based representations and the lack of coverage of specialized linguistic features, we use a graph-based representation of discourse segments, which allows for a more abstract (and hence generalizable) notion of syntactic (and partially of semantic) structure. Empirical evaluation on a hand-annotated corpus of German discourse relations shows that our graph-based approach not only provides a suitable representation for the linguistic factors that are needed in disambiguating discourse relations, but also improves results over a strong state-of-the-art baseline by more accurately identifying Temporal, Comparison and Reporting discourse relations.
Wednesday Morning 3: RTE/annotation
(11:30 – 13:00)

Parsimonious semantic representations with projection pointers
Noortje J. Venhuizen, Johan Bos, and Harm Brouwer

The influential idea by van der Sandt (1992) to treat presuppositions as anaphora in the framework of Discourse Representation Theory (DRT, Kamp and Reyle, 1993) has inspired a lot of debate as well as elaborations of his account. In this paper, we propose an extension of DRT, called Projective DRT, which adds pointers to all DRT referents and conditions, indicating their projection site. This means that projected content need not be moved from the context in which it is introduced, while it remains clearly discernible from asserted content. This approach inherits the attractive properties from van der Sandt’s approach to presupposition, but precludes a two-step resolution algorithm by treating projection as variable binding, which increases compositionality and computational efficiency. The result is a flexible representational framework for a descriptive theory of projection phenomena.

A search task dataset for german textual entailment
Britta Zeller and Sebastian Padó

We present the first freely available large German dataset for Textual Entailment (TE). Our dataset builds on posts from German online forums concerned with computer problems and models the task of identifying relevant posts for user queries (i.e., descriptions of their computer problems) through TE. We use a sequence of crowdsourcing tasks to create realistic problem descriptions through summarisation and paraphrasing of forum posts. The dataset is represented in RTE-5 Search task style and consists of 172 positive and over 2800 negative pairs. We analyse the properties of the created dataset and evaluate its difficulty by applying two TE algorithms and comparing the results with results on the English RTE-5 Search task. The results show that our dataset is roughly comparable to the RTE-5 data in terms of both difficulty and balancing of positive and negative entailment pairs. Our approach to create task-specific TE datasets can be transferred to other domains and languages.

Semantic annotation of textual entailment
Assaf Toledo, Stavroula Alexandropoulou, Sophia Katrenko, Heidi Klockmann, Pepijn Kokke and Yoad Winter

We introduce a new formal semantic model for annotating textual entailments, that describes restrictive, intersective and appositive modification. The model contains a formally defined interpreted lexicon, which specifies the inventory of symbols and the supported semantic operators, and an informally defined annotation scheme that instructs annotators in which way to bind words and constructions from a given pair of premise and hypothesis to the interpreted lexicon. We explore the applicability of the proposed model to the Recognizing Textual Entailment (RTE) 1-4 corpora and describe a first-stage annotation scheme based on which manual annotation work was carried out. The constructions we annotated were found to occur in 80.65% of the entailments in RTE 1-4 and were annotated with cross-annotator agreement of 68% on average. The annotated RTE corpora are publicly available for the research community.
Multi-step regression learning for compositional distributional semantics
Edward Grefenstette, Georgiana Dinu, Yao-Zhong Zhang, Mehrnoosh Sadrzadeh and Marco Baroni

We present a model for compositional distributional semantics related to the framework of Coecke et al. (2010), and emulating formal semantics by representing functions as tensors and arguments as vectors. We introduce a new learning method for tensors, generalising the approach of Baroni and Zamparelli (2010). We evaluate it on two benchmark data sets, and find it to outperform existing leading methods. We argue in our analysis that the nature of this learning method also renders it suitable for solving more subtle problems compositional distributional models might face.

Evaluating topic coherence using distributional semantics
Nikolaos Aletras and Mark Stevenson

This paper introduces distributional semantic similarity methods for automatically measuring the coherence of a set of words generated by a topic model. We construct a semantic space to represent each topic word by making use of Wikipedia as a reference corpus to identify context features and collect frequencies. Relatedness between topic words and context features is measured using variants of Pointwise Mutual Information (PMI). Topic coherence is determined by measuring the distance between these vectors computed using a variety of metrics. Evaluation on three data sets shows that the distributional-based measures outperform the state-of-the-art approach for this task.

Towards a semantics for distributional representations
Katrin Erk

Distributional representations have recently been proposed as a general-purpose representation of natural language meaning, to replace logical form. There is, however, one important difference between logical and distributional representations: Logical languages have a clear semantics, while distributional representations do not. In this paper, we propose a semantics for distributional representations that links points in vector space to mental concepts. We extend this framework to a joint semantics of logic and distributions by linking intensions of logical expressions to mental concepts.

Intensionality was only alleged: on adjective-noun composition in distributional semantics
Gemma Boleda, Marco Baroni, Nghia The Pham and Louise McNally

Distributional semantics has very successfully modeled semantic phenomena at the word level, and recently interest has grown in extending it to capture the meaning of phrases via semantic composition. We present experiments in adjective-noun composition which (1) show that adjectival modification can be successfully modeled with distributional semantics, (2) show that composition models inspired by the semantics of higher-order predication fare better than those that perform simple feature union or intersection, (3) contrary to what the theoretical literature might lead one to expect, do not yield a distinction between intensional and non-intensional modification, and (4) suggest that head noun polysemy and whether the adjective corresponds to a typical attribute of the noun are relevant factors in the distributional representation of adjective phrases.
Sources of evidence for implicit argument resolution
Egoitz Laparra and German Rigau

Traditionally, semantic role labelling systems have focused on searching the fillers of those explicit roles appearing within sentence boundaries. However, when the participants of a predicate are implicit and can not be found inside sentence boundaries, this approach obtains incomplete predicative structures with null arguments. Previous research facing this task have coincided in identifying the implicit argument filling as a special case of anaphora or coreference resolution. In this work, we review a number of theories that model the behaviour of discourse coreference and propose some adaptations to capture evidence for the implicit argument resolution task. We empirically demonstrate that exploiting such evidence our system outperforms previous approaches evaluated on the SemEval-2010 task 10 dataset. We complete our study identifying those cases that traditional coreference theories can not cover.

Towards weakly supervised resolution of null instantiations
Philip Gorinski, Josef Ruppenhofer and Caroline Sporleder

This paper addresses the task of finding antecedents for locally uninstantiated arguments. To resolve such null instantiations, we develop a weakly supervised approach that investigates and combines a number of linguistically motivated strategies that are inspired by work on semantic role labeling and coreference resolution. The performance of the system is competitive with the current state-of-the-art supervised system.

A corpus study of clause combination
Olga Nikitina and Sebastian Padó

We present a corpus-based investigation of cases of clause combination that can be expressed both through coordination or with subordination. We analyse the data with a two-step computational model which first distinguishes subordination from coordination and then determines the direction for cases of subordination. We find that a wide range of features help with the prediction, notably frequency of predicate participants, presence of adjuncts and sharing of participants between the clause predicates.

Regular meaning shifts in german particle verbs: a case study
Sylvia Springorum, Jason Utt and Sabine Schulte im Walde

This paper provides a corpus-based study on German particle verbs. We hypothesize that there are regular mechanisms in meaning shifts of a base verb in combination with a particle that do not only apply to the individual verb, but across a semantically coherent set of verbs. For example, the syntactically similar base verbs *brummen* ‘hum’ and *donnern* ‘rumble’ both describe an irritating, displeasing loud sound. Combined with the particle *auf*, they result in near-synonyms roughly meaning ‘forcefully assigning a task’ (in one of their senses). Covering 6 base verb groups and 3 particles with 4 particle meanings, we demonstrate that corpus-based information on the verbs’ subcategorization frames plus conceptual properties of the nominal complements is a sufficient basis for defining such meaning shifts. While the paper is considerably more extensive than earlier related work, we view it as a case study toward a more automatic approach to identify and formalize meaning shifts in German particle verbs.
Thursday Morning 1: Semantic relations
(09:00 – 10:30)

Recognizing spatial containment relations between event mentions
Kirk Roberts, Michael Skinner and Sanda Harabagiu

In this paper, we present an approach for recognizing spatial containment relations that hold between event mentions. Event mentions refer to real-world events that have spatio-temporal properties. While the temporal aspect of event relations has been well-studied, the spatial aspect has received relatively little attention. The difficulty in this task is the highly implicit nature of event locations in discourse. We present a supervised method that is designed to capture both explicit and implicit spatial relation information. Our approach outperforms the only known previous method by a 14 point increase in F$_1$-measure.

Temporal relation classification based on temporal reasoning
Francisco Costa and António Branco

The area of temporal information extraction has recently focused on temporal relation classification. This task is about classifying the temporal relation (precedence, overlap, etc.) holding between two given entities (events, dates or times) mentioned in a text. This interest has largely been driven by the two recent TempEval competitions. Even though logical constraints on the structure of possible sets of temporal relations are obvious, this sort of information deserves more exploration in the context of temporal relation classification. In this paper, we show that logical inference can be used to improve or sometimes dramatically improve existing machine learned classifiers for the problem of temporal relation classification.

The impact of selectional preference agreement on semantic relational similarity
Bryan Rink and Sanda Harabagiu

Relational similarity is essential to analogical reasoning. Automatically determining the degree to which a pair of words belongs to a semantic relation (relational similarity) is greatly improved by considering the selectional preferences of the relation. To determine selectional preferences, we induced semantic classes through a Latent Dirichlet Allocation (LDA) method that operates on dependency parse contexts of single words. When assigning relational similarities to pairs of words, if the agreement of selectional preferences is considered alone, a correlation of 0.334 is obtained against the manual ranking outperforming the previously best reported score of 0.229.

Thursday Morning 2: Semantics
(11:00 – 12:30)

UCCA: A semantics-based grammatical annotation scheme
Omri Abend and Ari Rappoport

Syntactic annotation is an indispensable input for many semantic NLP applications. For instance, Semantic Role Labelling algorithms almost invariably apply some form of syntactic parsing as preprocessing. The categories used for syntactic annotation in NLP generally reflect the formal patterns used to form the text. This results in complex annotation schemes, often tuned to one language or domain, and unintuitive to non-expert annotators. In this paper we propose a different approach and advocate substituting existing syntax-based approaches with semantics-based grammatical annotation. The rationale of this approach is to use manual labor where there is no substitute for it (i.e., annotating semantics), leaving the detection of formal regularities to automated statistical algorithms. To this end, we propose a simple semantic annotation scheme, UCCA for Universal Conceptual Cognitive Annotation. The scheme covers many of the most important
elements and relations present in linguistic utterances, including verb-argument structure, optional adjuncts such as adverbials, clause embeddings, and the linkage between them. The scheme is supported by extensive typological cross-linguistic evidence and accords with the leading Cognitive Linguistics theories.

**Empirical validation of Reichenbach’s Tense framework**
*Leon Derczynski and Robert Gaizauskas*

There exist formal accounts of tense and aspect, such as that detailed by Reichenbach (1947). Temporal semantics for corpus annotation are also available, such as TimeML. This paper describes a technique for linking the two, in order to perform a corpus-based empirical validation of Reichenbach’s tense framework. It is found, via use of Freksa’s semi-interval temporal algebra, that tense appropriately constrains the types of temporal relations that can hold between pairs of events described by verbs. Further, Reichenbach’s framework of tense and aspect is supported by corpus evidence, leading to the first validation of the framework. Results suggest that the linking technique proposed here can be used to make advances in the difficult area of automatic temporal relation typing and other current problems regarding reasoning about time in language.

**Probabilistic induction for an incremental semantic grammar**
*Arash Eshghi, Matthew Purver and Julian Hough*

We describe a method for learning an incremental semantic grammar from a corpus in which sentences are paired with logical forms as predicate-argument structure trees. Working in the framework of Dynamic Syntax, and assuming a set of generally available compositional mechanisms, we show how lexical entries can be learned as probabilistic procedures for the incremental projection of semantic structure, providing a grammar suitable for use in an incremental probabilistic parser. By inducing these from a corpus generated using an existing grammar, we demonstrate that this results in both good coverage and compatibility with the original entries, without requiring annotation at the word level. We show that this semantic approach to grammar induction has the novel ability to learn the syntactic and semantic constraints on pronouns.

**Friday Morning 1: Invited Talk**
*(09:00 – 10:00)*

**Structures for machine translation: strings, trees, and graphs**
*Kevin Knight*

String acceptors and transducers are critical technologies for natural language and speech systems. They flexibly capture many kinds of stateful, left-to-right substitution; simple transducers can be composed into more complex ones; and they are trainable. Tree acceptors and transducers provide even more transformational power. Still, strings and trees are both weak at representing linguistic structure involving semantics and reference (“who did what to who”). Viewing semantic structures as directed acyclic graphs, we take a look at probabilistic acceptors and transducers for them, demonstrate some linguistic transformations, and point toward a foundation for semantics-based statistical machine translation.
Friday Morning 2: Natural Language Generation
(10:00 – 11:00)

Domain adaptable semantic clustering in statistical NLG
Blake Howald, Ravikumar Kondadadi and Frank Schilder

We present a hybrid natural language generation system that utilizes Discourse Representation Structures (DRSs) for statistically learning syntactic templates from a given domain of discourse in sentence “micro” planning. In particular, given a training corpus of target texts, we extract semantic predicates and domain general tags from each sentence and then organize the sentences using supervised clustering to represent the “conceptual meaning” of the corpus. The sentences, additionally tagged with domain specific information (determined separately), are reduced to templates. We use a SVM ranking model trained on a subset of the corpus to determine the optimal template during generation. The combination of the conceptual unit, a set of ranked syntactic templates, and a given set of information, constrains output selection and yields acceptable texts. Our system is evaluated with automatic, non–expert crowdsourced and expert evaluation metrics and, for generated weather, financial and biography texts, falls within acceptable ranges. Consequently, we argue that our DRS driven statistical and template–based method is robust and domain adaptable as, while content will be dictated by a target domain of discourse, significant investments in sentence planning can be minimized without sacrificing performance.

Generating natural language from linked data: unsupervised template extraction
Daniel Duma and Ewan Klein

We propose an architecture for generating natural language from Linked Data that automatically learns sentence templates and statistical document planning from parallel RDF datasets and text. We have built a proof-of-concept system (LDDS) trained on un-annotated text from the Simple English Wikipedia and RDF triples from DBpedia, focusing exclusively on factual, non-temporal information. The goal of the system is to generate short descriptions, equivalent to Wikipedia stubs, of entities found in Linked Datasets. We have evaluated the LDDS system against a simple generate-from-triples baseline and human-generated output. In evaluation by humans, LDDS significantly outperforms the baseline on two of three measures: non-redundancy and structure and coherence.

Friday Afternoon 1: Inference and ontologies
(14:00 – 16:00)

Sentiment composition using a parabolic model
Baptiste Chardon, Farah Benamara, Yvette Mathieu, Vladimir Popescu and Nicholas Asher

In this paper, we propose a computational model that accounts for the effects of negation and modality on opinion expressions. Based on linguistic experiments informed by native speakers, we distil these effects according to the type of modality and negation. The model relies on a parabolic representation where an opinion expression is represented as a point on a parabola. Negation is modelled as functions over this parabola whereas modality through a family of parabolas of different slopes; each slope corresponds to a different certainty degree. The model is evaluated using two experiments, one involving direct strength judgements on a 7-point scale and the other relying on a sentiment annotated corpus. The empirical evaluation of our model shows that it matches the way humans handle negation and modality in opinionated sentences.
Recognising sets and their elements: tree kernels for entity instantiation identification
Andrew McKinlay and Katja Markert

We apply tree kernels to entity instantiations. An entity instantiation is an entity relationship, in which a set of entities is mentioned, and then a member or subset of this set is introduced. We present the first reliably annotated intrasentential entity instantiation corpus, along with an extension to the intersentential annotations in McKinlay and Markert (2011). We then apply tree kernels to both inter- and intrasentential entity instantiations, showing comparable results to an extensive set of unstructured features. The combination of tree kernels and unstructured features leads to significant improvements over either method in isolation.

Automatically deriving event ontologies for a commonsense knowledge base
James Allen, Jansen Orfan, Will de Beaumont, Choh Man Teng, Lucian Galescu and Mary Swift

We describe work aimed at building commonsense knowledge by reading word definitions using deep understanding techniques. The end result is a knowledge base allowing complex concepts to be reasoned about using OWL-DL reasoners. We show that we can use this system to automatically create a mid-level ontology for WordNet verbs that has good agreement with human intuition with respect to both the hypernym and causality relations. We present a detailed error analysis that reveals areas of future work needed to enable high-performance learning of conceptual knowledge by reading.

Learning corpus patterns using finite state automata
Octavian Popescu

Words get their meaning in context and Harris’s Distributional Hypothesis has been used in computational linguistics in order to identify the relationship between co-occurring words and their senses. In general, the local context contains the necessary information for word sense disambiguation (Stevenson & Wilks 2001). However, the exact extent of the local context varies significantly. To cope with this problem, previous research has shown that the regularity of word usage in natural language can be exploited (Pustejovsky & Hanks 2001). Many times, words are used in phrases with a patternable structure. On the basis of corpus evidence (Popescu & Magnini 2007), or on the basis of the lexicographer’s intuition on the normal usage (Hanks 2005) a set of patterns can be built which makes the link between context and word senses. In this paper we focus on patterns centered on verbs. We show that their structure is learnable and by employing a learning algorithm we are able to build a recognizer able to match such patterns against previously unseen text.
The Workshop on Annotation of Modal Meaning in Natural Language (WAMM) will take place in Room S12 on Tuesday, March 19 from 09:30 to 17:30.

Organization

PROGRAM CHAIRS

Paul Portner
Georgetown University
Aynat Rubinstein
Georgetown University
Graham Katz
CACI International

Pranav Anand
University of California Santa Cruz
Mona Diab
Columbia University
Ferdinand de Haan
Oracle
Valentine Hacquard
University of Maryland
Iris Hendrickx
University of Lisbon
Marie-Catherine de Marneffe
Stanford University
Lori Levin
Carnegie Mellon University
Christopher Potts
Stanford University
James Pustejovsky
Brandeis University
Ines Rehbein
Potsdam University
Josef Ruppenhofer
University of Hildesheim
Roser Sauri
Barcelona Media
Janyce Wiebe
University of Pittsburgh
Schedule

09:30 – 09:45  Welcome

09:45 – 10:30  Invited talk – Annotation of modal meanings: where formal semantics can help and where it cannot
Anette Frank

10:30 – 10:45  Discussion

10:45 – 11:00  🍵 Coffee Break

11:00 – 11:20  Challenges in modality annotation in a Brazilian Portuguese spontaneous speech corpus
Luciana Avila and Heliana Mello

11:20 – 11:30  Discussion

11:30 – 11:50  Influence of modality markers on the conditional interpretation of the German preposition "ohne"
Claudia Roch

11:50 – 12:00  Discussion

12:00 – 12:30  Analyzing modal and enunciative discursive heterogeneity: how to combine semantic resources and a syntactic parser analysis
Delphine Battistelli and Marine Damiani

12:30 – 12:40  Discussion

12:40 – 14:15  🍽️ Lunch

14:15 – 14:45  Toward fine-grained annotation of modality in text
Aynat Rubinstein, Hillary Harner, Elizabeth Krawczyk, Daniel Simonson, Graham Katz, and Paul Portner

14:45 – 14:55  Discussion

Marta Carretero and Juan Rafael Zamorano-Mansilla

15:25 – 15:35  Discussion

15:35 – 16:00  🍵 Coffee Break
<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>16:00 – 16:30</td>
<td><strong>Annotating modal expressions in the Chinese treebank</strong>&lt;br&gt;Yanyan Cui and Ting Chi</td>
</tr>
<tr>
<td>16:30 – 16:40</td>
<td><strong>Discussion</strong></td>
</tr>
<tr>
<td>16:40 – 17:10</td>
<td><strong>Distant annotation of Chinese tense and modality</strong>&lt;br&gt;Nianwen Xue, Yuchen Zhang, and Yaqin Yang</td>
</tr>
<tr>
<td>17:10 – 17:20</td>
<td><strong>Discussion</strong></td>
</tr>
<tr>
<td>17:30</td>
<td><strong>IWCS 2013 Welcome reception</strong></td>
</tr>
</tbody>
</table>
Abstracts

Invited talk – Annotation of modal meanings: where formal semantics can help and where it cannot
Anette Frank

Modality is a well-studied phenomenon in philosophy of language and formal semantics. In computational linguistics, the distinction between factual and non-factual contexts and the classification of different types of modal interpretations (such as epistemic, bouletic, or circumstantial modality) is highly relevant for classical and novel NLP applications, such as information extraction, sentiment analysis, or the computational analysis of discourse.

The study of modality in formal semantics mainly focuses on inferential properties. Some work, especially in the framework of DRT, examines special discourse properties found with modal constructions, such as extended operator scope, known as the phenomenon of modal subordination, the interpretation of graded modals and their accessibility in discourse, or the interaction of epistemic and non-epistemic modality in conditionals. Modal systems differentiate various types of modality (e.g. epistemic vs. deontic), yet there is in general no clear conceptual definition for different types of modal contexts.

The emerging computational treatment of modality has achieved considerable progress: There are first annotation schemes for (non-)veridicality along with annotated data sets. This led to a number of first computational approaches for annotating (non-)veridicality, which show that the task is feasible. The situation seems more difficult for the annotation of different types of modality, in particular non-epistemic modality, graded modality and modal subordination phenomena.

In my talk I will confront the insights formal semantics can offer for the interpretation of modality with the challenges of empirical, corpus-based semantic annotation of modality. I will focus in particular on the annotation of non-epistemic modality, its interplay with epistemic modality and their interaction with context.

Challenges in modality annotation in a Brazilian Portuguese spontaneous speech corpus
Luciana Avila and Heliana Mello

This short paper introduces the first notes about a modality annotation system that is under development for a spontaneous speech Brazilian Portuguese corpus (C-ORAL-BRASIL). We indicate our methodological decisions, the points which seem to be well resolved and two issues for further discussion and investigation.

Influence of modality markers on the conditional interpretation of the German preposition "ohne"
Claudia Roch

This paper investigates the impact of modality markers on the conditional interpretation of the German preposition ohne ('without'). It tackles the question whether it is the preposition itself that possesses a conditional sense or whether it may be due to a modal context that the interpretation arises. The paper presents an annotation study for modality factors (e.g. mood, modal auxiliary verbs, modal adjectives, modal adverbs, modal infinitives, negation) in the context of these sentences. The statistical analysis of the data has been carried out by means of a correspondence analysis in order to identify the relevant factors for the conditional interpretation. The results suggest that primarily the verb mood has an influence.
Analyzing modal and enunciative discursive heterogeneity: how to combine semantic resources and a syntactic parser analysis
Delphine Battistelli and Marine Damiani

This paper introduces our methodology for annotating variations in enunciative and modal commitment in a text. We first present the theoretical background of the study which puts the emphasis on the close interaction between time, aspect, modality and evidentiality (TAME) categories (and also markers). We then present our semantic resources which encompass not only lexical items, but also morphological inflections and syntactic constructions. We finally describe the first step of our global natural language processing (NLP) workflow which uses a syntactic analysis parser.

Toward fine-grained annotation of modality in text
Aynat Rubinstein, Hillary Harner, Elizabeth Krawczyk, Daniel Simonson, Graham Katz, and Paul Portner

We present a linguistically-informed schema for annotating modal expressions and describe its application to a subset of the MPQA corpus of English texts (Wiebe et al. 2005). The annotation is fine-grained in two respects: (i) in the range of expressions that are defined as modal targets and (ii) in the amount of information that is annotated for each target expression. We use inter-annotator reliability results to support a two-way distinction between priority and non-priority modality types.

An analysis of disagreement-provoking factors in the analysis of epistemic modality and evidentiality: the case of English adverbials
Marta Carretero and Juan Rafael Zamorano-Mansilla

This paper reports on a series of annotation experiments carried out on a number of English adverbials. The experiments, based on occurrences obtained from the British National Corpus, focused on the distinction of epistemic and evidential meanings from other kinds of meanings. The results led to the conclusion that many of the cases of inter-annotator disagreement were due to certain syntactic and semantic factors. Some of these factors will be described in detail, together with the decisions made in each case for prospective annotation.

Annotating modal expressions in the Chinese treebank
Yanyan Cui and Ting Chi

This paper reports an effort to annotate modality in the Penn Chinese Treebank. We introduce the modals and features that were annotated, and describe the phases of our working process. Along with this, we address the issues in the preparation of annotation guidelines, and present the preliminary results of the first pass. Finally, we analyze the types of disagreement, and propose directions to improve consistency.

Distant annotation of Chinese tense and modality
Nianwen Xue, Yuchen Zhang, and Yaqin Yang

In this paper we describe a “distant annotation” method by which we mark up tense and modality of Chinese eventualities via a word-aligned parallel corpus. We first map Chinese verbs to their English counterpart via word alignment, and then annotate the resulting English text spans with coarse-grained tense and modality categories that we believe apply to both English and Chinese. Because English has richer morpho-syntactic indicators for tense and modality than Chinese, we hope this distant annotation approach will yield more consistent annotation than if we annotate the Chinese side directly. We report experimental results that show this expectation is largely borne out.
WORKSHOP: COMPUTATIONAL SEMANTICS IN CLINICAL TEXT

The workshop Computational semantics in clinical text (CSCT) will take place in Room S14 on Tuesday, March 19 from 13:00 to 17:00.

Organization

PROGRAM CHAIRS

| Stephen Wu       | Mayo Clinic |
| Nigam Shah       | Stanford University |
| Kevin Bretonnel Cohen | University of Colorado School of Medicine |

PROGRAM COMMITTEE

| Sophia Ananiadou       | University of Manchester |
| Steven Bethard         | University of Colorado - Boulder |
| Olivier Bodenreider    | NLM |
| Rebecca Crowley        | University of Pittsburgh |
| Trevor Cohen           | University of Texas – Houston |
| Dina Demner-Fushman    | NLM |
| Henk Harkema           | Nuance |
| Yang Huang             | Kaiser Permanente |
| Hongfeng Liu           | Mayo Clinic |
| Tim Miller             | Childrens Hospital Boston and Harvard University |
| Rodney Nielsen         | University of North Texas |
| Serguei Pakhomov       | University of Minnesota |
| Martha Palmer          | University of Colorado – Boulder |
| Ehud Reiter            | University of Aberdeen |
| Thomas Rindflesch      | NLM |
| Guergana Savova        | Childrens Hospital Boston and Harvard University |
| Martijn Schuemie       | Erasmus Medical Center |
| Melissa Tharp          | University of California – San Diego |
| Luca Toldo             | Merck KGaA |
| Ozlem Uzuner           | MIT |
| Kees van Deemter       | University of Aberdeen |
## Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Session Title</th>
<th>Presenters</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:00 – 14:10</td>
<td><strong>Introduction</strong></td>
<td>Stephen Yu</td>
</tr>
<tr>
<td>14:10 – 14:55</td>
<td><strong>Keynote address</strong></td>
<td>Pierre Zweigenbaum</td>
</tr>
<tr>
<td>14:55 – 15:15</td>
<td><strong>Towards converting clinical phrases into SNOMED CT expressions</strong></td>
<td>Rohit Kate</td>
</tr>
<tr>
<td>15:15 – 15:30</td>
<td><strong>A framework to generate sets of terms from large scale medical vocabularies for natural language processing</strong></td>
<td>Salah Aït-Mokhtar, Caroline Hagège and Pajolma Rupi</td>
</tr>
<tr>
<td>15:30 – 16:00</td>
<td><strong>Coffee Break</strong></td>
<td></td>
</tr>
<tr>
<td>16:00 – 16:20</td>
<td><strong>Analysis of cross-institutional medication information annotations in clinical notes</strong></td>
<td>Sunghwan Sohn, Cheryl Clark, Scott Halgrim, Sean Murphy, Siddhartha Jonnalagadda, Kavishwar Waghohikar, Stephen Wu, Christopher Chute and Hongfang Liu</td>
</tr>
<tr>
<td>16:20 – 16:35</td>
<td><strong>Figurative language in Swedish clinical texts</strong></td>
<td>Dimitrios Kokkinakis</td>
</tr>
<tr>
<td>16:35 – 16:55</td>
<td><strong>Investigating topic modelling for therapy dialogue analysis</strong></td>
<td>Christine Howes, Matthew Purver and Rose McCabe</td>
</tr>
<tr>
<td>16:55 – 17:10</td>
<td><strong>The VeriCliG project: extraction of computer interpretable guidelines via syntactic and semantic annotation</strong></td>
<td>Camilo Thorne, Marco Montali, Elena Cardillo, Claudio Eccher and Diego Calvanese</td>
</tr>
<tr>
<td>17:10 – 17:30</td>
<td><strong>Evaluating the use of empirically constructed lexical resources for named entity recognition</strong></td>
<td>Siddhartha Jonnalagadda, Trevor Cohen, Stephen Wu, Hongfang Liu and Graciela Gonzalez</td>
</tr>
<tr>
<td>17:30</td>
<td><strong>IWCS 2013 Welcome reception</strong></td>
<td></td>
</tr>
</tbody>
</table>
Abstracts

Towards converting clinical phrases into SNOMED CT expressions
Rohit Kate

Converting information contained in natural language clinical text into computer-amenable structured representations can automate many clinical applications. As a step towards that goal, we present a method which could help in converting novel clinical phrases into new expressions in SNOMED CT, a standard clinical terminology. Since expressions in SNOMED CT are written in terms of their relations with other SNOMED CT concepts, we formulate the important task of identifying relations between clinical phrases and SNOMED CT concepts. We present a machine learning approach for this task and using the dataset of existing SNOMED CT relations we show that it performs well.

A framework to generate sets of terms from large scale medical vocabularies for natural language processing
Salah Aît-Mokhtar, Caroline Hagège and Pajolma Rupi

In this paper we present our current work on integrating large-scale terminological information into NLP tools. We focus on the problem of selecting and generating a set of suitable terms from the resources, based on cleaning and addition rules. We propose a general framework in which the raw data of the resources are first loaded into a knowledge base (KB). The selection and generation rules are then defined in a declarative way using query templates in the query language of the KB system. We illustrate the use of this framework to select and generate term sets from two different UMLS datasets.

Analysis of cross-institutional medication information annotations in clinical notes
Sunghwan Sohn, Cheryl Clark, Scott Halgrim, Sean Murphy, Siddhartha Jonnalagadda, Kavishwar Wagholikar, Stephen Wu, Christopher Chute and Hongfang Liu

A large amount of medication information resides in unstructured text in electronic medical records, which requires advanced techniques to be properly mined. In clinical notes, medication information follows certain semantic patterns (e.g., medication, dosage, frequency, mode, etc.). Some medication descriptions contain additional word(s) between medication attributes. Therefore, it is essential to understand the semantic patterns as well as the patterns of the context interspersed among them (i.e., context patterns) to effectively extract comprehensive medication information. In this paper we examined both semantic and context patterns and compared those found in Mayo Clinic and i2b2 challenge data. We found that some variations exist between the institutions but the dominant patterns are common.

Figurative language in Swedish clinical texts
Dimitrios Kokkinakis

Automated processing of clinical texts is commonly faced with various less exposed, and not so regularly discussed linguistically complex problems that need to be addressed. One of these issues concerns the usage of figurative language. Figurative language implies the use of words that go beyond their ordinary meaning, a linguistically complex and challenging problem and also a problem that causes great difficulty for the field of natural language processing (NLP). The problem is equally prevalent in both general language and also in various sublanguages, such as clinical medicine. Therefore we believe that a comprehensive model of e.g. clinical language processing needs to account for figurative language usage, and this paper provides a description, and preliminary results towards this goal. Since the empirical, clinical data used in the study is limited in size, there is no formal distinction made between different sub-classifications of figurative language. e.g., metaphors, idioms or simile. We illustrate several types of figurative expressions in
the clinical discourse and apply a rather quantitative and corpus-based level analysis. The main research questions that this paper asks are whether there are traces of figurative language (or at least a subset of such types) in patient-doctor and patient-nurse interactions, how can they be found in a convenient way and whether these are transferred in the electronic health records and to what degree.

Investigating topic modelling for therapy dialogue analysis
Christine Howes, Matthew Purver and Rose McCabe

Previous research shows that aspects of doctor-patient communication in therapy can predict patient symptoms, satisfaction and future adherence to treatment (a significant problem with conditions such as schizophrenia). However, automatic prediction has so far shown success only when based on low-level lexical features, and it is unclear how well these can generalise to new data, or whether their effectiveness is due to their capturing aspects of style, structure or content. Here, we examine the use of topic as a higher-level measure of content, more likely to generalise and to have more explanatory power. Investigations show that while topics predict some important factors such as patient satisfaction and ratings of therapy quality, they lack the full predictive power of lower-level features. For some factors, unsupervised methods produce models comparable to manual annotation.

The VeriCliG project: extraction of computer interpretable guidelines via syntactic and semantic annotation
Camilo Thorne, Marco Montali, Elena Cardillo, Claudio Eccher and Diego Calvanese

We consider the problem of extracting formal process representations of the therapies defined by clinical guidelines, viz., computer interpretable guidelines (CIGs), based on UMLS and semantic and syntactic annotation. CIGs enable the application of formal methods (such as model checking, verification, conformance assessment) to the clinical domain. We argue that, while minimally structured, correspondences among clinical guideline syntax and discourse relations and clinical process constructs should however be exploited to successfully extract CIGs. We review work on current clinical syntactic and semantic annotation, pinpointing their limitations, and discuss a CIG extraction methodology based on recent efforts on business process modeling notation (BPMN) model extraction from natural language text.

Evaluating the use of empirically constructed lexical resources for named entity recognition
Siddhartha Jonnalagadda, Trevor Cohen, Stephen Wu, Hongfang Liu and Graciela Gonzalez

In this work, we evaluate what value may lie in automatically generated features based on distributional semantics when using machine-learning named entity recognition (NER). The features we generated and experimented with include n-nearest words, support vector machine (SVM)-regions, and term clustering, all of which are considered semantic (or distributional semantic) features. The addition of n-nearest words feature resulted in a greater increase in F-score than adding a manually constructed lexicon to a baseline system that extracts medical concepts from clinical notes. Although the need for relatively small annotated corpora for retraining is not obviated, lexicons empirically derived from unannotated text can not only supplement manually created lexicons, but replace them. This phenomenon is observed in extracting concepts both from biomedical literature and clinical notes.
The Workshop on **Computational Models of Spatial Language Interpretation and Generation (CoSLI-3)** will take place in Room S15 on Tuesday, March 19 from 09:00 to 13:00.

**Organization**

**Program Chairs**

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>John Kelleher</td>
<td>Dublin Institute of Technology</td>
</tr>
<tr>
<td>Simon Dobnik</td>
<td>University of Gothenburg</td>
</tr>
<tr>
<td>Robert Ross</td>
<td>Dublin Institute of Technology</td>
</tr>
</tbody>
</table>

**Programme Committee**

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michel Aurnague</td>
<td>University of Toulouse-Le Mirail</td>
</tr>
<tr>
<td>Kenny Coventry</td>
<td>University of East Anglia</td>
</tr>
<tr>
<td>Max Egenhofer</td>
<td>University of Maine</td>
</tr>
<tr>
<td>Carola Eschenbach</td>
<td>University of Hamburg</td>
</tr>
<tr>
<td>Nick Hawes</td>
<td>University of Birmingham</td>
</tr>
<tr>
<td>Matthew Klenk</td>
<td>Palo Alto Research Center</td>
</tr>
<tr>
<td>Alexander Klipple</td>
<td>Penn State</td>
</tr>
<tr>
<td>Valia Kordoni</td>
<td>University of Saarland</td>
</tr>
<tr>
<td>Kate Lockwood</td>
<td>California State University</td>
</tr>
<tr>
<td>Philippe Muller</td>
<td>University Paul Sabatier</td>
</tr>
<tr>
<td>Ekaterina Ovchinnikova</td>
<td>University of Southern California</td>
</tr>
<tr>
<td>Ehud Reitter</td>
<td>University of Aberdeen</td>
</tr>
<tr>
<td>Stefanie Tellex</td>
<td>Massachusetts Institute of Technology</td>
</tr>
<tr>
<td>Kees van Deemter</td>
<td>University of Aberdeen</td>
</tr>
<tr>
<td>Emile van der Zee</td>
<td>University of Lincoln</td>
</tr>
<tr>
<td>Martijn van Otterlo</td>
<td>Radboud University Nimegen</td>
</tr>
<tr>
<td>Hendrik Zender</td>
<td>Nuance Communications</td>
</tr>
<tr>
<td>Joost Zwartz</td>
<td>University of Utrecht</td>
</tr>
<tr>
<td>Time</td>
<td>Session</td>
</tr>
<tr>
<td>------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>09:10 – 09:15</td>
<td><strong>Opening</strong></td>
</tr>
</tbody>
</table>
| 09:15 – 09:40 | *Where things happen: on the semantics of event localization*  
|             | James Pustejovsky                                                      |
| 09:40 – 10:05 | *Spatial descriptions in type theory with records*                           
|             | Simon Dobnik and Robin Cooper                                           |
| 10:05 – 10:30 | *Clock-modeled ternary spatial relations for visual scene analysis*        
|             | Joanna Isabelle Olszewska                                               |
| 10:30 – 11:00 | 🍵 **Coffee Break**                                                     |
| 11:00 – 11:25 | *Human evaluation of conceptual route graphs for interpreting spoken route descriptions*  
|             | Raveesh Meena, Gabriel Skantze and Joakim Gustafson                     |
| 11:25 – 11:50 | *Deriving salience models from human route directions*                   
|             | Jana Götze and Johan Boye                                               |
| 11:50 – 12:30 | *Invited talk – Modularity and abstraction in spatial semantics: dividing up the cake*  
|             | by Prof. John Bateman, the University of Bremen                         |
| 12:30 – 14:00 | ☕️ **Lunch**                                                            |
| 17:30       | IWCS 2013 Welcome reception                                             |
Abstracts

Where things happen: on the semantics of event localization
James Pustejovsky

The problem of temporally situating events in language has been approached by a number of philosophical techniques, including Davidson's particularist theory of event individuation and Kim's property exemplification theory. Both of these theories have been developed within linguistic semantic traditions as well, and others. However, the problem of event localization (spatially situating events) has not been discussed as extensively in the semantics literature. In this paper, I discuss the procedures for identifying where events, as expressed in natural language, are located in space. Aspects of the semantics of event localization have been recently proposed, including the notion of the "shape" of a movement, as well as treating movement verbs as "path creation" predicates. In this paper, I build on these and some additional observations to outline a more general semantics of event localization. I then outline a procedure that extends the path metaphor used for motion predicates, distinguishing between the event locus and the spatial aspect of an event. In the process, I discuss how localization is supervenient upon the participants in the events.

Spatial descriptions in type theory with records
Simon Dobnik and Robin Cooper

We present how TTR (Type Theory with Records) can model both geometric perception and conceptual (world) knowledge relating to the meaning of spatial descriptions for a robotic agent.

Clock-modeled ternary spatial relations for visual scene analysis
Joanna Isabelle Olszewska

The analysis and the description of complex visual scenes characterized by the presence of many objects of interests involve reasoning on spatial relations such as "above", "below", "before", "after" and "between". In this context, we have defined these semantic concepts in terms of ternary spatial relations and we have formalized them using the clock model which is based on the clock-face division and the semantic notions of hours to describe relative spatial positions. The presented approach has been efficiently applied for the automated understanding of spatial relations between multiple objects in real-world computer vision image datasets.

Human evaluation of conceptual route graphs for interpreting spoken route descriptions
Raveesh Meena, Gabriel Skantze and Joakim Gustafson

We present a human evaluation of the usefulness of conceptual route graphs (CRGs) when it comes to route following using spoken route descriptions. We describe a method for data-driven semantic interpretation of route descriptions into CRGs. The comparable performances of human participants in sketching a route using the manually transcribed CRGs and the CRGs produced on speech recognized route descriptions indicate the robustness of our method in preserving the vital conceptual information required for route following despite speech recognition errors.

Deriving salience models from human route directions
Jana Gotze and Johan Boye

We present an approach to derive individual preferences in the use of landmarks for route instructions in a city environment. Each possible landmark that a person can refer to in a given situation is modelled as a feature vector, and the preference (or salience) associated with the landmark can be computed as a weighted sum of these features. The weight vector, representing the person's personal salience model, is automatically derived from the person's own route descriptions. Experiments show that the derived salience models can correctly predict the user's choice of landmark in 69% of the cases.
Modularity and abstraction in spatial semantics: dividing up the cake

John Bateman

Several approaches to the interpretation of spatial language in natural texts assume that it is necessary to engage in knowledge-rich, deep reasoning in order to ascertain what is intended. Such approaches make few distinctions between different kinds of semantics and employ highly expressive formalisms for capturing what in many cases appear to be relatively simple uses of natural language. In this talk I work through several examples that suggest that this may be unnecessary. By imposing distinct levels of abstraction, each subject to rather different motivations, I argue that it is more revealing to divide the problem so that each level of abstraction can operate more or less independently of the others. This has both theoretical and practical benefits. Theoretically, I propose that this is closer to how people actually process spatial language: only the degree of expressivity and granularity of distinctions assumed necessary for a particular problem need to be invoked; practically, the relative independence of levels of abstractions allows them to be treated as modules, which can then be replaced or compared with other modules as individual research or implementation tasks might prefer.
WORKSHOP: TOWARDS A FORMAL DISTRIBUTION SEMANTICS

The workshop Towards a formal distribution semantics (TFDS) will take place in Room S16 on Tuesday, March 19 from 09:00 to 17:30.

Organization

PROGRAM CHAIRS

Aurélie Herbelot  Université de Potsdam
Roberto Zamparelli  Center for Mind/Brain Sciences (CIMeC)
Gemma Boleda  The University of Texas at Austin

PROGRAM COMMITTEE

Nicholas Asher  Université Paul Sabatier / The University of Texas at Austin
Marco Baroni  University of Trento
David Beaver  The University of Texas at Austin
Raffaella Bernardi  University of Trento
Stephen Clark  University of Cambridge
Ann Copestake  University of Cambridge
Katrin Erk  The University of Texas at Austin
Mohan Ganesalingam  University of Cambridge
Ed Grefenstette  University of Oxford
Louise McNally  Universitat Pompeu Fabra
Dáire Ó Séaghdha  University of Cambridge
Sebastian Pado  Universitäts Heidelberg
Manfred Pinkal  Saarland University
Stephen Pulman  University of Oxford
Mehrnoosh Sadrzadeh  University of Oxford
Mark Steedman  University of Edinburgh
Stefan Thater  Saarland University
Jason Utt  University of Stuttgart
Eva Maria Vecchi  University of Trento
Schedule

09:15 – 09:30  Introduction

09:30 – 10:30  Invited talk – Formal and distributional semantics: from romance to relationship
Louise McNally

10:30 – 11:00  ☕️  Coffee Break

11:00 – 11:40  Sentence paraphrase detection: when determiners and word order make the difference
Nghia Pham, Raffaella Bernardi, Yao Zhong Zhang and Marco Baroni

11:40 – 12:20  The curious case of metonymic verbs: a distributional characterization
Jason Utt, Alessandro Lenci, Sebastian Padó and Alessandra Zarcone

12:20 – 13:00  Semantic transparency: challenges for distributional semantics
Melanie Bell and Martin Schäfer

13:00 – 14:00  🍽️  Lunch

14:00 – 14:40  Can distributional approaches improve on good old-fashioned lexical semantics?
Ann Copestake

14:40 – 15:40  Invited talk: combining logic-based and distributional representations for inferences over text
Katrin Erk

15:40 – 16:00  ☕️  Coffee Break

16:00 – 16:30  General discussion

16:30 – 17:30  Panel on compositional distributional semantics
Led by Marco Baroni (COMPOSES Project)

17:30  IWCS 2013 Welcome reception
Abstracts

Formal and distributional semantics: from romance to relationship
Louise McNally

When I first learned about distributional semantics, I got very excited because I am very concerned with a problem about which "traditional" formal semantics methods have little to say, namely how to make sense of the distinction, assuming one exists, between linguistic content and world knowledge, and how these come together in the resolution of the context-dependent dimension of semantic composition (for example, how we model the effect of adjectival modifiers on nouns). I also confess that I have a hard time buying into the idea of a logic-like "language of thought", and I thought distributional representations, appropriately enriched, might ultimately lead to better approximations of the mental representation of semantic content than what we currently have. Since distributional semantic representations and the operations used to combine them lend themselves naturally to problems like word sense disambiguation, I was willing to set aside all my other concerns about these representations (like how they would deal with function words) in the hopes of making a significant advance in my understanding of context-dependent meaning. A few years and a few modeling experiences later (Boleda et al., 2012; Boleda et al., 2013), I have to say the romance is over: I don’t expect to be abandoning formal semantics any time soon. However, I won’t be abandoning distributional semantics, either. In this talk, I’ll briefly discuss some of the things that I’ve learned that have most surprised and interested me, some of the obstacles that I consider most significant, and therefore most urgent, if the distributional semantics community is to attract the collaboration of formal semanticists, something that I think would greatly benefit both communities, and some problems that I think could be addressed now and that formal semanticists could profitably contribute to and learn from.

Sentence paraphrase detection: when determiners and word order make the difference
Nghia Pham, Raffaella Bernardi, Yao Zhong Zhang and Marco Baroni

Researchers working on distributional semantics have recently taken up the challenge of going beyond lexical meaning and tackle the issue of compositionality. Several Compositional Distributional Semantics Models (CDSMs) have been developed and promising results have been obtained in evaluations carried out against data sets of small phrases and as well as data sets of sentences. However, we believe there is the need to further develop good evaluation tasks that show whether CDSM truly capture compositionality. To this end, we present an evaluation task that highlights some differences among the CDSMs currently available by challenging them in detecting semantic differences caused by word order switch and by determiner replacements. We take as starting point simple intransitive and transitive sentences describing similar events, that we consider to be paraphrases of each other but not of the foil paraphrases we generate from them. Only the models sensitive to word order and determiner phrase meaning and their role in the sentence composition will not be captured into the foils’ trap.

The curious case of metonymic verbs: a distributional characterization
Jason Utt, Alessandro Lenci, Sebastian Padó and Alessandra Zarcone

Logical metonymy combines an event-selecting verb with an entity-denoting noun (e.g., The writer began the novel), triggering a covert event interpretation (e.g., reading, writing). Experimental investigations of logical metonymy must assume a binary distinction between metonymic (i.e. eventselect- ing) verbs and non-metonymic verbs to establish a control condition. However, this binary distinction (whether a verb is metonymic or not) is mostly made on intuitive grounds, which introduces a potential confounding factor. We describe a corpus-based approach which characterizes verbs in terms of their behavior at the syntax-semantics interface. The model assesses the extent to which transitive verbs prefer event-denoting objects over entity-denoting objects. We
then test this “eventhood” measure on psycholinguistic datasets, showing that it can distinguish not only metonymic from non-metonymic verbs, but that it can also capture more fine-grained distinctions among different classes of metonymic verbs, putting such distinctions into a new graded perspective.

**Semantic transparency: challenges for distributional semantics**

*Melanie Bell and Martin Schäfer*

Using data from Reddy et al. (2011), we present a series of regression models of semantic transparency in compound nouns. The results indicate that the frequencies of the compound constituents, the semantic relation between the constituents, and metaphorical shift of a constituent or of the compound as a whole, all contribute to the overall perceived level of transparency. While not proposing an actual distributional model of transparency, we hypothesise that incorporating this information into such a model would improve its success and we suggest some ways this might be possible.

**Can distributional approaches improve on good old-fashioned lexical semantics?**

*Ann Copestake*

In this position paper, I discuss some linguistic problems that computational work on lexical semantics has attempted to address in the past and the implications for alternative models which incorporate distributional information. I concentrate in particular on phenomena involving count/mass distinctions, where older approaches attempted to use lexical semantics in their models of syntax. I outline methods by which the earlier models allowed the transmission of information between lexical items (regular polysemy and inheritance) and address the possibility that similar techniques could usefully be incorporated into distributional models.

**Combining logic-based and distributional representations for inferences over text**

*Katrin Erk*

Distributional models have recently been suggested as an alternative to logic-based sentence semantics because of their ability to model word and phrase similarity and to model gradience in meaning, and because they can be learned automatically from data. For me, an important additional property is their ability to describe the nuances of meaning in context without recourse to sense lists. However, the model that we have proposed for sentence meaning is not a purely distributional one, but a combined distributional and logic-based representation, as we view the two approaches as complementary. Logic-based approaches have well-known mechanisms for representing function words and representing the semantics of long, complex expressions. But they typically have an impoverished representation of the meaning of content words, and cannot typically be used to talk about word and phrase similarity. The strengths and weaknesses of distributional models are exactly the opposite. In this talk, I will describe recent work on using a joint logical and distributional approach to judge sentence similarity within a project geared at deep language understanding. I will also mention first steps towards defining a more systematic connection between the two frameworks.

**Panel: Compositional distributional semantics**

*Led by Marco Baroni*

Semantic composition has been the core operation studied by formal semanticists since at least Frege’s seminal work, and recently compositionality has attracted much interest among distributional semanticists as well. The topic of how to achieve compositionality with distributional models constitutes thus a natural area of convergence (or clash) between the two fields. In this panel, we will try to assess what is the current state of development in the area of compositional semantics, what are the most important obstacles to be faced, and what are the implications of all this (if any) for the classic formal approach to compositionality.
WORKSHOP: INTEROPERABLE SEMANTIC ANNOTATION

The Ninth Joint ACL - ISO Workshop on Interoperable Semantic Annotation (ISA-9) will take place in Room S23 on Tuesday from 09:15 to 17:00, and on Wednesday from 09:00 to 17:00.

Organization

Program Chairs

Harry Bunt  
Tilburg University  
Kiyong Lee  
Korea University  
James Pustejovsky  
Brandeis University  
Laurent Romary  
INRIA and HU Berlin

Programme Committee

Jan Alexandersson  
Alex Fang  
Koiti Hasida  
Michael Kipp  
Alessandro Lenci  
Martha Palmer  
Andrei Popescu-Belis  
James Pustejovsky  
Claudia Soria  
Piek Vossen  
Thierry Declerck  
Robert Gaizauskas  
Nancy Ide  
Kiyong Lee  
Inderjeet Mani  
Volha Petukhova  
Rashmi Prasad  
Laurent Romary  
Thorsten Trippel
Schedule

TUESDAY, MARCH 19

09:15 – 09:20 Welcome, opening

09:20 – 09:55 Which units for modality annotation? Caterina Mauri, Malvina Nissim, Paola Pietrandrea and Andrea Sanso

09:55 – 10:30 Multi-layered annotation of non-textual data for spatial information Kiyong Lee

10:30 – 11:00 ☕ Coffee Break

11:00 – 11:35 Capturing motion in ISO-SpaceBank James Pustejovsky

11:35 – 12:30 Overview and status report of project ISO-Space James Pustejovsky

12:30 – 14:00 🍽️ Lunch

14:00 – 14:25 Interoperability in the Australian National Corpus Steve Cassidy

14:25 – 15:00 Conceptual and representational choices in defining an ISO standard for semantic role annotation Harry Bunt and Martha Palmer

15:00 – 15:30 Overview and status report of ISO-Semantic roles Harry Bunt

15:30 – 16:00 ☕ Coffee Break

16:00 – 16:30 Veridicity annotation in the lexicon? A look at factive adjectives Annie Zaenen and Lauri Karttunen

16:30 – 17:00 Discussion on possible ISO project on veridicity annotation Annie Zaenen

17:30 IWCS 2013 Welcome reception
**Wednesday, March 20**

10:10 – 10:50  **Issues in the addition of ISO-compliant annotations to the Switchboard corpus**  
Harry Bunt, Alex Fang, Jin Cao, Xiaoyue Liu and Volha Petukhova

11:00 – 11:30  ☕ Coffee Break

11:30 – 12:05  **More than only noun-noun compounds: towards and annotation scheme for the semantic modeling of other noun compound terms**  
Ben Verhoeven and Gerard B. van Huysstee

12:05 – 12:20  **Overview and status report of project ISO-DRel (semantic relations in discourse)**  
Rashmi Prasad

12:20 – 13:45  🍽️ Lunch

13:45 – 14:15  **Overview and status report of project ISO-Basics**  
Harry Bunt

14:15 – 14:40  **Annotation for annotation - toward eliciting implicit linguistic knowledge through annotation**  
Takenobu Tokunaga, Ryu Iida and Koh Mitsuda

14:40 – 15:15  **Inference patterns with intensional adjectives**  
James Pustejovsky

15:15 – 16:00  **Discussion: Interoperability of ISO semantic annotation frameworks, either published or under development**  
Harry Bunt, Kiyong Lee, James Pustejovsky and Laurent Romary

16:00 – 16:30  ☕ Coffee Break

16:30 – 17:00  **ISO TC 37/SC 4 WGs plenary meeting**
Abstracts

Which units for modality annotation?
*Caterina Mauri, Malvina Nissim, Paola Pietrandrea and Andrea Sansò*

We present an annotation model of modality which is (i) cross-linguistic, relying on a wide, strongly typologically motivated approach, and (ii) hierarchical and layered, accounting for both factuality and speaker’s attitude, while modelling these two aspects through separate annotation schemes. Modality is defined through cross-linguistic categories, but the classification of actual linguistic expressions is language-specific. This makes our annotation model a powerful tool for investigating linguistic diversity in the field of modality on the basis of real language data, being thus also useful from the perspective of machine translation systems.

Multi-layered annotation of non-textual data for spatial information
*Kiyong Lee*

Spatial and spatio-temporal information is often carried by non-textual data such as maps, diagrams, tables, or pictures, both still and moving, either embedded in a text or standalone. The annotation of nontextual data raises the following questions: (i) what are the markables and how should they be coded? (ii) how should relevant information be inferred which is implicit in the data? We answer these questions with a multilayered approach.

Capturing motion in ISO-SpaceBank
*James Pustejovsky*

This paper presents the first description of the motion subcorpus of ISO-SpaceBank (MotionBank) and discusses how motion events are represented in ISO-Space 1.5, a specification language for the representation of spatial information in language. We present data from this subcorpus with examples from the pilot annotation, focusing specifically on the annotation of motion events and their various participants. These data inform further discussion of outstanding issues concerning semantic annotation, such as quantification and measurement. We address these questions briefly as they impact the design of ISO-Space.

Interoperability in the Australian National Corpus
*Steve Cassidy*

The Australian National Corpus (AusNC) provides a technical infrastructure for collecting and publishing language resources representing Australian language use. As part of the project we have ingested a wide range of resource types into the system, bringing together the different meta-data and annotations into a single interoperable database. This paper describes the initial collections in AusNC and the procedures used to parse a variety of data types into a single unified annotation store.

Conceptual and representational choices in defining an ISO standard for semantic role annotation
*Harry Bunt and Martha Palmer*

This paper presents two essential elements of the ISO standard for semantic role annotation which is under development (ISO CD 24617-5:2012), namely (a) the metamodel, which describes the types of concepts that may occur in semantic role annotation and their conceptual relations, and (b) an annotation language for expressing semantic role annotations, with its abstract syntax, XML-based concrete syntax, and semantics.
Veridicity annotation in the lexicon? A look at factive adjectives
Annie Zaenen and Lauri Karttunen

In this note, we look at the factors that influence veridicity judgments with factive predicates. We show that more context factors play a role than is generally assumed. We propose to use crowdsourcing techniques to understand these factors better and briefly discuss the consequences for the associate of lexical signatures with items in the lexicon.

Issues in the addition of ISO-compliant annotations to the switchboard corpus
Harry Bunt, Alex Fang, Jin Cao, Xiaoyue Liu and Volha Petukhova

This paper analyzes the issues that arise when trying to add annotations to the dialogues in the Switchboard corpus according to ISO standard 24617-2, exploiting the existing SWBD-DAMSL annotations. These issues relate to differences between the two tag sets; to the highly multidimensional view that underlies the ISO standard; to differences in segmenting the dialogues into functional units; to the use of in-line markups for certain phenomena in Switchboard, and to the use of intra-dialogue dependence relations as defined in the ISO standard. The analysis is supplemented by a discussion of how the existing annotations may be helpful to semi-automatically create a fully-fledged ISO standard annotation alongside the existing SWBD-DAMSL annotation.

More than only noun-noun compounds: towards and annotation scheme for the semantic modeling of other noun compound terms
Ben Verhoeven and Gerard B. van Huysstee

The computational processing of compound semantics poses several interesting challenges. Up to now, the processing of nominal compounds with non-noun left-hand constituents (henceforth XN compounds) has not received any attention, despite the fact that these also seem to be rather productive in Germanic languages. In our research project, we aim to fill this hiatus by investigating various kinds of compounds in Afrikaans and Dutch, develop annotation protocols and data sets, and model the semantics of such compounds. In this publication we present the alpha version of an annotation protocol that was designed for both descriptive linguistic and computational linguistic purposes. We describe the protocol development and discuss the current version.

Annotation for annotation - toward eliciting implicit linguistic knowledge through annotation
Takenobu Tokunaga, Ryu Iida and Koh Mitsuda

The last two decades witnessed a great success of revived empiricism in NLP research. However, there are still several NLP tasks that are not successful enough. As one of many directions for going beyond the revived empiricism, this paper introduces a project for annotating annotations with annotators’ rationales behind them. As a first step of this enterprise, the paper particularly focuses on data collection during the annotation and discusses their potential uses. Finally a preliminary experiment for data collection is described with the data analysis.

Inference patterns with intensional adjectives
James Pustejovsky

In this paper we report on an ongoing multi-institution effort to encode inferential patterns associated with adjective modification in English. We focus here on a subset of intensional adjectives typically referred to as "non-subsective" predicates. This class includes adjectives such as "alleged", "supposed", "so-called", and related modally subordinating predicates. We discuss the initial results of corpus-based investigations to discriminate the patterns of inference associated with these adjectives. Based on these studies, we have created an initial annotation specification that we are using to create a corpus of adjective-related inferences in English.
GETTING AROUND

Venue

The conference and the workshops will take place in the Griebnitzsee campus of the University of Potsdam, located right next to the Griebnitzsee train station (Bhf Griebnitzsee). All sessions will take place in Building 6, marked in the map below.

The Griebnitzsee campus is accessible from both Potsdam and Berlin via the S7 train line, which runs every 10 minutes:

- From Potsdam Hbf, S7 direction Ahrensfelde (duration: 6 minutes)
- From Berlin Hbf, S7 direction Potsdam (duration: 32 minutes)

The same train also serves several other train stations in Potsdam and Berlin.

Around the city

Your name badge doubles as a public transportation ticket for the Berlin ABC zone, which encompasses the cities of Berlin and Potsdam and includes Griebnitzsee and both airports. This ticket allows you to travel for free using public transportation, as many times as you like, during the days of the conference.

The Berlin/Brandenburg metropolitan area, which includes Potsdam, has a dense public transportation network. We recommend you consult the VBB Fahrinfo website¹ for your trips. Smartphone users can download the free apps Fahrinfo Mobil (Android) or iFahrinfo (iPhone), or alternatively access the route planner².

¹http://www.vbb.de/en/index.html
²http://mobil.bvg.de
You should have also received a map of the City of Potsdam with all bus and tram lines which details all their stops. During your stay in Potsdam, you are likely to use trains from the S-Bahn line, and either buses or trams (streetcars):

The **S-Bahn S7** train (yellow-red trains) connects Berlin and Potsdam, with many intermediate stations: Griebnitzsee is the second in the Potsdam – Berlin direction (less than 10 minutes away), while Berlin Central Station (Berlin Hauptbahnhof) is the 12th stop, 38 minutes away. You can always identify the S-Bahn departure/arrival platforms by their green S logo. The timetable of the S7 train has been included on page 54.

**Regional trains** (RE, RB) are red, and may offer faster (but less regular) alternatives than the S-Bahn in some cases.

Downtown Potsdam is serviced by **trams** (streetcars), all of them pass at some point through the Potsdam Central Station (Potsdam Hauptbahnhof).

The areas not serviced by trams are typically covered by **buses**. They usually have a narrower schedule at night, although some S-Bahn and U-Bahn lines are substituted by them after midnight.

The **U-Bahn** subway network covers a large section of Berlin.

Please note that your pass **does not** cover high speed trains (e.g. IC, ICE, EC).

The central interchange point for public transportation is **Potsdam Central Station** (written as either **Potsdam Hauptbahnhof** or **Potsdam Hbf**). Trams and buses stop outside the station, a few meters away from the South Entrance, while trains stop at their designated platforms inside. In particular, platforms 6 and 7 are dedicated to arriving and departing S-Bahn trains. Also, inside the station you'll find lots of shops covering a wide range of interests, including among others food, telephony, clothing, a cinema and a supermarket. You can check your train's departure time at the main electronic board, **except** the S-Bahn trains – those depart every 10 minutes, so they are not displayed. Trams and buses stop outside the Station's south entrance (the one underneath the main board).
Hotel Mercure

Many participants will be staying at the Hotel Mercure. The Mercure is an easy 5-10 minute walk from the central station. When you exit the station through the west entrance, you will already see the "Hotel Mercure" sign on top of the tall hotel building. You will need to cross a bridge (Lange Brücke) and then turn left.

Alternatively, you can take a tram for one stop. From the central station, take the tram 92 towards Kirschallee, tram 93 towards Glienicker Brücke, or tram 96 towards Viereckremise and get off at the stop "Alter Markt" (the first stop after the central station). The tram stop "Alter Markt" is across the road from the hotel, on the other side of the bridge from the station.

Where to eat

Next to the location of the scheduled coffee breaks, you will find a small coffee shop called "Die Bohne". You are welcome to purchase additional hot drinks there. "Die Bohne" is open until 14:30 each day (14:00 on Fridays).

Regarding lunch, during IWCS 2013 the best alternative will be the Mensa Cafeteria, located in the same building as the Conference. You will find Mensa tokens in your conference materials, which you can redeem there for a free meal including a main dish and either a dessert or a piece of fruit. If you also desire a drink, there will be plenty available for purchase. The Cafeteria’s menu is composed of up to four possible dishes (Angebot I to IV), they are always published at the monitor located over the entrance, and there’s also a display inside the cafeteria where you can actually look at what each dish looks like. English translations of the menu will be posted near the Mensa entrance each day.

As an alternative, you can also try one of the following places, near Griebnitzsee Station. Given that Rudolf-Breitscheid-Straße is located on the opposite side of the train station, the most direct way would be through the Station’s underground passage.

- **Albers** Rudolf-Breitscheid-Str. 201 (< €10, German plain)
- **Avendi Hotel** Rudolf-Breitscheid-Str. 190 (> €20, German, fancy)
- **Piazza Toscana** Rudolf-Breitscheid-Str. 177 (~ €15, Italian)

You may also wish to visit downtown Potsdam, where you’ll find plenty of varied shops and restaurants. You can get there from Potsdam’s Central Station (Potsdam Hauptbahnhof) by taking either Tram 92 towards Kirschallee or Tram 96 towards Viereckremise, and descending in both cases at the Brandenburger Straße stop.

Some of the available restaurants in this area are detailed on the map on the next page.
1. **Café Hundertwasser** Kurfürstenstraße 52 (€ 10 – € 15, German, Austrian, Mediterranean)

2. **Potsdamer Kulturcafé** Benkertstraße 23 (< € 10, vegetarian)

3. **Café Heider** Friedrich-Ebert-Straße 29 (€ 10 – € 15, German)

4. **Restaurant Zum Fliegenden Holländer** Benkertstraße 5 (€ 10 – € 15, fish, German)

5. **Restaurant Ma Cuisine** Hebbelstraße 54 (> € 15, French)

6. **Juliette** Jägerstraße 39 (> € 15, French)

7. **Brasserie zu Gutenberg** Jägerstraße 6 (€ 10 – € 15, German, French, plain)

8. **Fischrestaurant ”Der Butt”** Gutenbergstraße 25 (> € 15, fish, German)

9. **Pfeffer & Salz** Brandenburger Straße 47 (> € 15, Italian)

10. **My Keng** Brandenburger Straße 20 (< € 10, Sushi, Cambodian)

11. **Noidue** Lindenstraße 6 (€ 10 – € 15, Italian, gourmet)

12. **La Madeleine** Lindenstraße 20 (< € 10, Creperie, French)

13. **Café Kieselstein** Hegelallee 23 (€ 10 – 15, vegan/vegetarian)

14. **Matador** Brandenburger Str. 2 (€ 10 – € 15, American, Italian, Steakhouse)

15. **Alter Stadwächter** Schopenhauerstraße 33 (€ 10 – € 15, German)

16. **Waage** Am Neuen Markt. 12 (> € 15, Italian, Mediterranean)
Banquet

The conference dinner takes place on Wednesday, March 20, at 19:30 in the Bornstedt Crown Estate. Your registration fee includes a buffet of regional specialties and one drink to get you started.

Getting there

The Bornstedt Crown Estate is located in the northern part of Park Sanssouci, on the far side of the city center from Potsdam Hbf. To get there, take the S7 to Potsdam Hbf and change to the Tram 92 (towards Kirschallee). Get off at its final stop, Kirschallee.

1. Walk around the tram and go down a set of small green stairs to the street. This is Pappelallee street, which runs parallel to the tram tracks. Turn right and walk towards the traffic lights.

2. At the traffic lights, you should see a white street sign pointing left towards "Krongut" (= Crown Estate). Follow the sign, down Kirschallee street. You will keep a historic building with a blue light on top to your right. Ignore the street sign pointing to parking lots for "Park Sanssouci" and "Krongut Bornstedt".

3. Cross Potsdamer Straße. On the next intersection to your right, you will find an old white building with a sign that says "Schülertreff Ribbeck-Eck".

4. Turn left at the Ribbeck-Eck into Ribbeckstraße. Follow Ribbeckstraße for about 500 meters. You will find the Bornstedt Crown Estate on your left.

On your way back, follow the instructions in reverse. Be sure to loop around the right-hand side of the historic building. Note that the tram leaves from the northern tram stop (5) on its way back into town (direction "Marie-Juchacz-Straße" via S Hauptbahnhof). The tram leaves at :19, :39, :59 after each hour until late.
Krongut - The Crown Estate of Bornstedt

The most famous touristic landmark of Potsdam is the Sanssouci palace with its surrounding gardens. The palace was designed by King Frederick the Great of Prussia in the 1740s, who made it one of the cultural centers of Europe, entertaining such guests as Voltaire and Bach. It was expanded further by subsequent Prussian Kings and German Emperors. The entire landscape includes smaller palaces, temples, and follies, and has been a UNESCO World Heritage Site since 1990.

The conference dinner will take place at one of the more intimate palaces in the Sanssouci ensemble, the Bornstedt Crown Estate. Located about 400 meters north of Sanssouci itself, Bornstedt was purchased by the Prussian royal family in 1664. Its initial use was as a brewery and distillery. After a fire destroyed the old buildings in 1846, King Frederick William IV rebuilt Bornstedt in the Italian style. It was the home of Emperor Frederick III and his wife, Princess Victoria (daughter of Queen Victoria of Britain), while he was Crown Prince. After the end of the German monarchy in 1919, the Crown Estate was converted into apartments; among others, the son of Rudolf Diesel (inventor of the Diesel engine) lived here for a while.

Today, the Bornstedt Crown Estate houses a restaurant with its own attached brewery and bakery. During the day, a number of shops around the central courtyard offer handicraft products that are manufactured on the premises, including hats, jewelry, and clothes. Each winter, a Christmas market takes place in the courtyard. We hope you will enjoy our dinner at the Bornstedt Crown Estate.
FURTHER INFORMATION

Internet and WiFi

Participants with laptops and mobile devices will be able to access the internet through the wireless network. Depending on your location within the building, there are two main ways to connect to the network:

- **Ground floor:** Participants on the ground floor can access the network named *Universität Potsdam*. The network is WPA protected; a password will be provided to you on the printed version of this booklet.

- **First floor:** Those attending workshops on the first floor can access the network named *UP-Conference*. When trying to access a webpage for the first time, you will be prompted for a password on your browser. You have received a wi-fi voucher with your personal password as part of your conference materials.

- **Eduroam:** Participants with access to the Eduroam network can also make use of this option. Please note that the signal is rather weak on the ground floor, but you shouldn’t have any access problems on the first floor.

By connecting to a wireless network of the University of Potsdam, you agree to the Terms of Use of the *Zentrale Einrichtung für Informationsverarbeitung und Kommunikation (ZEIK)* of the University of Potsdam (http://www.uni-potsdam.de/u/ambek/ambek698.htm#name3), along with the terms in the National Telecommunications Act.

Sightseeing

Both Potsdam and Berlin are popular tourist destinations. Your bus passes give you access to all public transportation within Potsdam and Berlin, and we encourage you to explore. Some sightseeing options are listed on the conference website.

Guided tours of Potsdam depart from Potsdam Central Station every day at 14:00 and cost around € 18. Tours last around 2:30 hours, and the guides typically speak both English and German.

Contact info

During the conference, the local organizers and student volunteers will be easily identified by their bright red conference t-shirts. They will be glad to help you with any problem and/or request you might have.

In the printed version of this booklet, you’ll be also provided with phone numbers to call in case of emergency.
### S7 Train Timetable

#### Schedule from S Potsdam Hbf to Ahrensfelde

<table>
<thead>
<tr>
<th>Departure</th>
<th>Arrival</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00</td>
<td>06:30</td>
</tr>
<tr>
<td>06:10</td>
<td>06:40</td>
</tr>
<tr>
<td>06:20</td>
<td>06:50</td>
</tr>
<tr>
<td>06:30</td>
<td>07:00</td>
</tr>
<tr>
<td>06:40</td>
<td>07:10</td>
</tr>
<tr>
<td>06:50</td>
<td>07:20</td>
</tr>
<tr>
<td>07:00</td>
<td>07:30</td>
</tr>
<tr>
<td>07:10</td>
<td>07:40</td>
</tr>
<tr>
<td>07:20</td>
<td>07:50</td>
</tr>
<tr>
<td>07:30</td>
<td>08:00</td>
</tr>
<tr>
<td>07:40</td>
<td>08:10</td>
</tr>
<tr>
<td>07:50</td>
<td>08:20</td>
</tr>
<tr>
<td>08:00</td>
<td>08:30</td>
</tr>
<tr>
<td>08:10</td>
<td>08:40</td>
</tr>
<tr>
<td>08:20</td>
<td>09:00</td>
</tr>
<tr>
<td>08:30</td>
<td>09:10</td>
</tr>
<tr>
<td>08:40</td>
<td>09:20</td>
</tr>
<tr>
<td>09:00</td>
<td>09:30</td>
</tr>
<tr>
<td>09:20</td>
<td>09:50</td>
</tr>
<tr>
<td>09:30</td>
<td>10:00</td>
</tr>
<tr>
<td>10:00</td>
<td>10:30</td>
</tr>
<tr>
<td>10:30</td>
<td>11:00</td>
</tr>
<tr>
<td>11:00</td>
<td>11:30</td>
</tr>
<tr>
<td>11:30</td>
<td>12:00</td>
</tr>
<tr>
<td>12:00</td>
<td>12:30</td>
</tr>
<tr>
<td>12:30</td>
<td>13:00</td>
</tr>
<tr>
<td>13:00</td>
<td>13:30</td>
</tr>
<tr>
<td>13:30</td>
<td>14:00</td>
</tr>
<tr>
<td>14:00</td>
<td>14:30</td>
</tr>
<tr>
<td>14:30</td>
<td>15:00</td>
</tr>
<tr>
<td>15:00</td>
<td>15:30</td>
</tr>
<tr>
<td>15:30</td>
<td>16:00</td>
</tr>
<tr>
<td>16:00</td>
<td>16:30</td>
</tr>
<tr>
<td>16:30</td>
<td>17:00</td>
</tr>
<tr>
<td>17:00</td>
<td>17:30</td>
</tr>
<tr>
<td>17:30</td>
<td>18:00</td>
</tr>
<tr>
<td>18:00</td>
<td>18:30</td>
</tr>
<tr>
<td>18:30</td>
<td>19:00</td>
</tr>
<tr>
<td>19:00</td>
<td>19:30</td>
</tr>
<tr>
<td>19:30</td>
<td>20:00</td>
</tr>
<tr>
<td>20:00</td>
<td>20:30</td>
</tr>
<tr>
<td>20:30</td>
<td>21:00</td>
</tr>
<tr>
<td>21:00</td>
<td>21:30</td>
</tr>
</tbody>
</table>

#### Schedule from Ahrensfelde to S Potsdam Hbf

<table>
<thead>
<tr>
<th>Departure</th>
<th>Arrival</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00</td>
<td>06:30</td>
</tr>
<tr>
<td>06:10</td>
<td>06:40</td>
</tr>
<tr>
<td>06:20</td>
<td>06:50</td>
</tr>
<tr>
<td>06:30</td>
<td>07:00</td>
</tr>
<tr>
<td>06:40</td>
<td>07:10</td>
</tr>
<tr>
<td>06:50</td>
<td>07:20</td>
</tr>
<tr>
<td>07:00</td>
<td>07:30</td>
</tr>
<tr>
<td>07:10</td>
<td>07:40</td>
</tr>
<tr>
<td>07:20</td>
<td>07:50</td>
</tr>
<tr>
<td>07:30</td>
<td>08:00</td>
</tr>
<tr>
<td>07:40</td>
<td>08:10</td>
</tr>
<tr>
<td>07:50</td>
<td>08:20</td>
</tr>
<tr>
<td>08:00</td>
<td>08:30</td>
</tr>
<tr>
<td>08:10</td>
<td>08:40</td>
</tr>
<tr>
<td>08:20</td>
<td>09:00</td>
</tr>
<tr>
<td>08:30</td>
<td>09:10</td>
</tr>
<tr>
<td>08:40</td>
<td>09:20</td>
</tr>
<tr>
<td>09:00</td>
<td>09:30</td>
</tr>
<tr>
<td>09:20</td>
<td>09:50</td>
</tr>
<tr>
<td>09:30</td>
<td>10:00</td>
</tr>
<tr>
<td>10:00</td>
<td>10:30</td>
</tr>
<tr>
<td>10:30</td>
<td>11:00</td>
</tr>
<tr>
<td>11:00</td>
<td>11:30</td>
</tr>
<tr>
<td>11:30</td>
<td>12:00</td>
</tr>
<tr>
<td>12:00</td>
<td>12:30</td>
</tr>
<tr>
<td>12:30</td>
<td>13:00</td>
</tr>
<tr>
<td>13:00</td>
<td>13:30</td>
</tr>
<tr>
<td>13:30</td>
<td>14:00</td>
</tr>
<tr>
<td>14:00</td>
<td>14:30</td>
</tr>
<tr>
<td>14:30</td>
<td>15:00</td>
</tr>
<tr>
<td>15:00</td>
<td>15:30</td>
</tr>
<tr>
<td>15:30</td>
<td>16:00</td>
</tr>
<tr>
<td>16:00</td>
<td>16:30</td>
</tr>
<tr>
<td>16:30</td>
<td>17:00</td>
</tr>
<tr>
<td>17:00</td>
<td>17:30</td>
</tr>
<tr>
<td>17:30</td>
<td>18:00</td>
</tr>
<tr>
<td>18:00</td>
<td>18:30</td>
</tr>
<tr>
<td>18:30</td>
<td>19:00</td>
</tr>
<tr>
<td>19:00</td>
<td>19:30</td>
</tr>
<tr>
<td>19:30</td>
<td>20:00</td>
</tr>
<tr>
<td>20:00</td>
<td>20:30</td>
</tr>
<tr>
<td>20:30</td>
<td>21:00</td>
</tr>
<tr>
<td>21:00</td>
<td>21:30</td>
</tr>
</tbody>
</table>

### Notes
- The timetable includes trains running from S Potsdam Hbf to Ahrensfelde and vice versa.
- Times are in 24-hour format.
- The schedule runs from 06:00 to 21:30 on a daily basis.
<table>
<thead>
<tr>
<th>Time</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00</td>
<td>Intro</td>
<td>Paper 1</td>
<td>Invited Talk</td>
</tr>
<tr>
<td>10:00</td>
<td>Intro</td>
<td>Paper 2</td>
<td>Discourse</td>
</tr>
<tr>
<td>11:00</td>
<td>Intro</td>
<td>Paper ISO-DRel overview</td>
<td>RTE/Annotation</td>
</tr>
<tr>
<td>12:00</td>
<td>Invited Talk</td>
<td>ISO Basics</td>
<td>Semantic Relations</td>
</tr>
<tr>
<td>13:00</td>
<td>Plenary</td>
<td>Papers 3-4</td>
<td>Semantic Roles</td>
</tr>
<tr>
<td>14:00</td>
<td>Keynote Address</td>
<td>ISO TC 37/SC 4</td>
<td>Corpus Studies</td>
</tr>
<tr>
<td>15:00</td>
<td>Invited Talk</td>
<td>Distributional Approaches</td>
<td>Conf. Dinner</td>
</tr>
<tr>
<td>16:00</td>
<td>Paper 4</td>
<td>Semantic Roles</td>
<td></td>
</tr>
<tr>
<td>17:00</td>
<td>Invited Talk</td>
<td>Papers 4-5</td>
<td></td>
</tr>
<tr>
<td>18:00</td>
<td>Paper 6</td>
<td>ISO SRI Overview</td>
<td></td>
</tr>
<tr>
<td>19:00</td>
<td>Paper 6</td>
<td>Presentation</td>
<td></td>
</tr>
<tr>
<td>20:00</td>
<td>Welcome reception</td>
<td>Paper 6</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:**
- WAMM (S12)
- CSCT (S14)
- CoSLI (S15)
- TFDS (S16)
- ISA-9 (S23)

**WS topics:**
- Natural Language Generation
- Inference and Ontologies
- Closing remarks

**Days:**
- Tuesday
- Wednesday
- Thursday
- Friday

**Times:**
- 9:00 AM - 20:00 PM