Symposium on Logic and Algorithms in Computational Linguistics 2018 (LACompLing2018)

Stockholm, 28 – 31 August 2018
Department of Mathematics and Department of Philosophy
Stockholm University, Stockholm

Venue: Department of Mathematics
Kräftriket, house 5A (room 22), Stockholm

<http://staff.math.su.se/rloukanova/LACompLing2018-web/>

Abstracts of Talks

Table of Contents

Lasha Abzianidze (University of Groningen, Nederlands) Invited Talk
Compositional Semantics in the Parallel Meaning Bank
(joint work with Johan Bos)

ABSTRACT: The Parallel Meaning Bank (PMB) is a corpus of translations annotated with shared, formal meaning representations. The principle of compositionality lies at the heart of the corpus as it drives a derivation process of phrasal semantics and enables cross-lingual projection of meaning representations. The talk will present the PMB annotation pipeline and show how it leads to the formal, compositional semantics of translations. As a highlight, compositional treatment of several challenging semantic phenomena in English will be shown.

Krasimir Angelov (University of Gothenburg and Digital Grammars AB, Sweden) Invited Talk
A Parallel WordNet and Treebank in English, Swedish and Bulgarian

ABSTRACT: We present a work in progress about a parallel WordNet-like lexicon and a treebank for English, Swedish and Bulgarian. The lexicon uses the Princeton WordNet senses but in addition incorporates detailed morphological and syntactic information. Words accross languages with the same sense which are moreover frequent mutual translations are grouped together via a language-independent identifier. These features make the lexicon directly usable as a library in GF applications. As part of the development we also converted all examples from WordNet to a treebank parsed with the GF Resource Grammars. Thanks to that the examples are translated to Swedish and Bulgarian.

Rasmus Blanck and Aleksandre Maskharashvili (CLASP, FLOV, University of Gothenburg, Sweden)
From TAG to HOL Representations of AMRs via ACGs

ABSTRACT: We investigate a possibility of constructing an Abstract Categorial Grammar (ACG) that relates Tree Adjoining Grammar (TAG) and Higher Order Logic (HOL) formulas which encode Abstract Meaning Representations (AMRs). We also propose another ACG that relates TAG and HOL formulas expressing the neo-Davidsonian event semantics. Both of these encodings are based on the already existing ACG encoding of the syntax-semantics interface where TAG derivations are interpreted as HOL formulas representing Montague semantics. In particular, both of these encodings share the same abstract language coming from the ACG encoding of TAG with Montague semantics, which is second-order. For second-order ACGs, problems of parsing and generation are known to be of polynomial complexity. Thus we get the natural language generation and parsing with TAGs and HOL formulas modeling AMR for free.

Robin Cooper (University of Gothenburg, Sweden) Invited Talk
How to Play Games with Types
(joint work with Ellen Breitholtz)

ABSTRACT: This talk will discuss how the kind of game theory (GT) presented in the course by Heather Burnett and E. Allyn Smith at ESSLLI 2017

(https://www.irit.fr/esslli2017/courses/6) and Burnett's paper "Signalling Games, Sociolinguistic Variation and the Construction of Style" (http://www.heatherburnett.net/uploads/9/6/6/0/96608942/burnett_smgs.pdf) could be connected to work on TTR, a type theory with records, and Ginzburg's KOS, a formal approach to conversational semantics. Here are some points I will consider:

1. Recasting GT in TTR. They both talk about types (of action) and when GT talks about possible worlds it is really what TTR would call types of situations. (The same holds of the use of the term "possible worlds" in probability theory). I will sketch an example of how it might look.

2. But what might doing (1) add to a linguistic theory? KOS/TTR might provide a framework for dealing with issues like choosing which games to play, misunderstandings between two agents about what game is being played or accommodating a game on the basis of another agent's behaviour. There is a notion of game in my paper "How to do things with types" (https://www.cisuc.uc.pt/ckfinder/userfiles/files/TR%202014-02.pdf). There is more detail in my book draft (https://sites.google.com/site/typetheorywithrecords/drafts) and also in Ellen Breitholtz's work on enthymemes and topoi in her thesis and book in preparation. Ginzburg's work on genre and conversation types is related. The games in this literature are very simple from the perspective of GT. They are defined in terms of a string type for a string of events on the gameboard which is traversed by an agent trying to realize the types. We have nothing to say about how you would make choices in a non-deterministic game, but GT would add that. It could be extremely productive to embed game theory in a theory of dialogue — one even begins to imagine metagames, games you play about concerning which game to play. We can perhaps supply a way of connecting GT to dialogue and grammar in a formal setting.

3. We could view this as making a connection between games and a general theory of action along the lines of "How to do things with types". The assumption seems to be that you compute utility and then perform the action that has highest utility for you. But you could think of other strategies: e.g. cooperative (make the move that has the highest utility irrespective of player), altruistic (maximize the utility of the other player). If you think of games as assigning utilities to event types at a given state of play, perhaps exploiting techniques from our work on probabilistic TTR (http://csli-lilt.stanford.edu/ojs/index.php/LiLT/article/view/52) you could have a superordinate theory of action which would tell you what you might do depending on which strategy you are using.

Hercules Dalianis (DSV-Stockholm University, Sweden) Invited Talk

HEALTH BANK — A Workbench for Data Science Applications in Healthcare

ABSTRACT: Healthcare has many challenges in form of monitoring and predicting adverse events as healthcare associated infections or adverse drug events. All this can happen while treating a patient at the hospital for their disease. The research question is: When and how many adverse events have occurred, how

can one predict them? Nowadays all information is contained in the electronic patient records and are written both in structured form and in unstructured free text. This talk will describe the data used for our research in HEALTH BANK — Swedish Health Record Research Bank containing over 2 million patient records from 2007–2014. Topics are detection of symptoms, diseases, body parts and drugs from Swedish electronic patient record text, including deciding on the certainty of a symptom or disease and detecting adverse (drug) events. Future research are detecting early symptoms of cancer and de-identification of electronic patient records for secondary use.

Philippe de Groote (Directeur de Recherche, Inria, France) Invited Talk
    New Progress in Continuation-Based Dynamic Logic
    ABSTRACT: In this talk, we revisit the type-theoretic dynamic logic introduced by de Groote (2006) and developed by Lebedeva (2012). We show how a slightly richer notion of continuation allows new dynamic connectives and quantifiers to be defined in a systematic way.

Marie Duzi (VSB-Technical University of Ostrava, Czech Republic) Invited Talk
    Negation, Presupposition and Truth-Value Gaps
    ABSTRACT: There are many kinds of negation and denial. Perhaps the most common is Boolean negation 'not' that applies to propositions-in-extension, i.e. truth-values. The others are, inter alia, the property of propositions of not being true which applies to propositions; the complement function which applies to sets; privation which applies to properties; negation as failure applied in logic programming; negation as argumentation ad absurdum, and many others. I am going to deal with negation of propositions that come attached with a presupposition that is entailed by the positive as well as negated form of a given proposition. However, there are two kinds of negation, namely internal and external negation, which are not equivalent. I will prove that while the former is presupposition-preserving, the latter is presupposition-denying. This issue has much in common with the difference between topic and focus articulation within a sentence. Whereas articulating the topic of a sentence activates a presupposition, articulating the focus frequently yields merely an entailment. While the Russellian wide-scope (external) negation gets the truth-conditions of a sentence right for a subject occurring as a focus, Strawsonian narrow-scope (internal) negation is validly applicable for a subject occurring as the topic. My background theory is Transparent Intensional Logic (TIL). It is an expressive logic apt for the analysis of sentences with presuppositions, because in TIL we work with partial functions, in particular with propositions with truth-value gaps. Moreover, procedural semantics of TIL makes it possible to uncover the hidden semantic features of sentences, make them explicit and logically tractable.

Tim Fernando (Trinity College Dublin, Ireland)
    Intervals and Events with and without Points
    ABSTRACT: Intervals and events are examined in terms of strings with and without the requirement that certain symbols occur uniquely. Allen interval re-

lations, Dowty's aspect hypothesis and inertia are understood against strings, compressed into canonical forms, describable in Monadic Second-Order logic. See: https://www.scss.tcd.ie/Tim.Fernando/stock.pdf

Annie Foret (IRISA - University of Rennes 1, France) Invited Talk

   On Categorial Grammatical Inference and Logical Information Systems

   ABSTRACT: We shall consider several classes of categorial grammars and discuss their learnability. We consider learning as a symbolic issue in an unsupervised setting, from raw or from structured data, for some variants of Lambek grammars and of categorial dependency grammars. In that perspective, we discuss for these frameworks different type constructors and structures, some limitations (negative results) but also some algorithms (positive results) under some hypothesis.

   On the experimental side, we also consider the Logical Information Systems approach, that allows for navigation, querying, updating, and analysis of heterogeneous data collections where data are given (logical) descriptors. Categorial grammars can be seen as a particular case of Logical Information System.

Jonathan Ginzburg (Laboratoire de Linguistique Formelle, Université Paris-Diderot and Laboratoire d'Excellence LabEx-EFL, France) Invited Talk

   Combining Verbal and Non-Verbal Interaction in Dialogue

   ABSTRACT: The talk will provide detailed motivation, contrary to received wisdom until recently, as to the mutual interaction between non-verbal social signals such as laughter, smiling, frowning etc and content emanating from verbal material. In particular, I will argue that such non-verbal social signals bear propositional content and can participate in own and other communication management (e.g., clarification requests and corrections). I will show how the content emanating from non-verbal social signals can be integrated in type theoretic accounts of dialogue interaction by combining work in existing frameworks with psychological and computational approaches to emotion appraisal and to common sense reasoning.

Justyna Grudzinska (University of Warsaw, Poland) Invited Talk

   Taking Scope with Continuations and Dependent Types

   (joint work with Marek Zawadowski)

   ABSTRACT: Dependent type theoretical frameworks have been used to model linguistic phenomena of central importance, e.g., unbound anaphora (Ranta 1994, Cooper 2004, Bekki 2014, Grudzinska et al. 2014), lexical phenomena such as selectional restrictions and coercions (Asher 2011, Luo 2012), adjectival and adverbial modification (Luo et al. 2017). Continuations have been used for an influential in situ analysis of quantifier scope ambiguities (Barker 2002). In my talk I will present a semantic system combining continuations and dependent types (joint work with Marek Zawadowski) that is sufficient to account for a broad range of existing readings for multi-quantifier sentences, including simple sentences and more complex syntactic environments such as inverse linking.

Harald Hammarström (Uppsala University, Sweden) Invited Talk

Intelligibility in Natural Languages and Formal Languages

ABSTRACT: A common view holds that intelligibility in natural languages is a gradient property. In this view intelligibility can be anywhere in the continuum between none, partial or total, with no natural border that would divide it into a binary yes/no property. We argue that this view is premature. With some simple assumptions on what consitutes a language, we show that intelligbility can be defined as a yes/no property in a simple formal way, without imposing an arbitrary threshold. The indication from this result is that natural languages, can, at least in theory, be discretized and counted in a systematic manner.

M. Dolores Jiménez López (GRLMC-Research Group on Mathematical Linguistics, Universitat Rovira i Virgili, Tarragona, Spain) Invited Talk

Complexity, Natural Language and Machine Learning

ABSTRACT: The talk focuses on linguistic complexity. Are all languages equally complex? Does it make sense to compare the complexity of languages? Can languages differ in complexity? Complexity is a controversial concept in linguistics. Until recently, natural language complexity has not been widely researched and and it is still not clear how complexity has to be defined and measured. It is necessary to provide an objective and meaningful method to calculate linguistic complexity. In order to reach this goal, an interdisciplinary solution — where computational models should be taken into account — is needed. Linguistics must propose tools for the analysis of natural language complexity, since the results obtained from these studies may have important implications both from a theoretical and from a practical point of view.

Ron Kaplan (Stanford University, US) KeyNote Talk

An Architecture for Structured Ambiguity Management

ABSTRACT: A pipeline for full-fledged natural language understanding consists of components that deal with information at different levels of remove from the elements that make up an utterance. Computing across the full pipeline is difficult because complex patterns (at all levels) may overlap in different ways, giving rise to ambiguities that feed from one component to the next. A typical approach is to apply probabilistic or heuristic preferences within each component so as to reduce the number of candidates that it feeds forward to the next. This has an obvious disadvantage: ambiguity resolution based on local information may eliminate the only candidate that gives the best result when all later components are taken into account. An alternative approach is to organize representations so as to "manage" the way ambiguous structures are propagated rather than attempting to resolve ambiguity at each level. The final result can then be globally optimal with respect to the whole pipeline. The trick is to do this without blowing up the computation.

Yusuke Kubota (University of Tsukuba, Japan) Invited Talk

Type-Logical Grammar and Natural Language Syntax

ABSTRACT: In this talk, I will first briefly sketch my recent work, which focused on developing a particular version of Type-Logical Grammar with emphasis on linguistic application. I will then speculate on what (I think) is still missing in my own research and what still needs to be done and whether now is a good time to start addressing these issues seriously. While I believe that my previous work has revealed some interesting points of comparison between Type-Logical Grammar and mainstream Chomskian syntax, it has also raised (or at least made me aware of) many issues pertaining to the relationship between theoretical linguistics and computational linguistics. I will touch on these issues and speculate on future directions.

Shalom Lappin (University of Gothenburg, Sweden) Invited Talk

Towards a Computationally Viable Framework for Semantic Representation

ABSTRACT: Most formal semantic theories proposed since Montague (1974) employ possible worlds to model intensions and modality. Classical theories of knowledge representation also use worlds to represent epistemic states and reasoning. If worlds are construed as equivalent to ultrafilters in a lattice of propositions (maximal consistent sets of propositions), then they pose serious problems of tractable representability. In addition, traditional worlds-based semantic theories are unable to accommodate vagueness, which is a pervasive feature of predication. They also do not explain semantic learning, and it is not clear how they could be naturally extended to incorporate such an explanation. To offer a cognitively plausible system for interpreting expressions in natural language a semantic theory should generate tractable representations, handle vagueness of predication, and provide the basis for an account of semantic learning. In this paper I discuss the problem of computational tractability of semantic representation. I suggest a probabilistic Bayesian alternative to classical worlds-based semantics, and I indicate how it can deal with intensions, modality, vagueness, epistemic states, and semantic learning.

Hans Leiß (Ludwig-Maximilians-Universität München, Germany) Invited Talk

Predication with Sentential Subject in GF

ABSTRACT: The resource grammar library of the Grammatical Framework of Ranta et al. distinguishes binary or ternary verbs with nominal or prepositional objects from verbs whose objects have the form of a sentence, a question or an infinitive. No such distinction is made for the subject position of verbs. We introduce syntactic categories for verbs, adjectives and verb phrases with sentential subjects and extend the predication grammar of Ranta (EACL, 2014) so that sentential subjects can only be combined with verb phrases of appropriate types (which may arise by passivizing verbs with sentential objects). We also report on the price in computational complexity that has to be paid for the gain in linguistic accuracy.

Zhaohui Luo (Royal Holloway, University of London, UK) Invited Talk

Formal Semantics in Modern Type Theories: An Overview

ABSTRACT: I'll give an overview, and report some recent developments, of Formal Semantics in Modern Type Theories (MTT-semantics for short). MTT-semantics is a semantic framework for natural language, in the tradition of Montague's semantics. However, while Montague's semantics is based on Church's simple type theory (and its models in set theory), MTT-semantics is based on dependent type theories, which we call modern type theories, such as Martin-Lof's type theory (MLTT) and the Unifying Theory of dependent Types (UTT). Thanks to recent development, MTT-semantics has become not only a full-blown alternative to Montague's semantics, but also a very attractive framework with a promising future for linguistic semantics.

In this talk, MTT-semantics will be explicated, and its advantages explained, by focussing on the following:

1. The rich structures in MTTs, together with subtyping, make MTTs a nice and powerful framework for formal semantics of natural language.

2. MTT-semantics is both model-theoretic and proof-theoretic and hence very attractive, both theoretically and practically.

By explaining the first point, we'll introduce MTT-semantics and, at the same time, show that the use and development of coercive subtyping play a crucial role in making MTT-semantics viable. The second point shows that MTTs provide a unique and nice semantic framework that was not available before for linguistic semantics. Being model-theoretic, MTT-semantics provides a wide coverage of various linguistic features. Being proof-theoretic, its foundational languages MTTs have proof-theoretic meaning theory based on inferential uses (appealing philosophically and theoretically) and it establishes a solid foundation for practical reasoning in natural languages based on proof assistants such as Coq (appealing practically). Altogether, this strengthens the argument that MTT-semantics is a promising framework for formal semantics, both theoretically and practically.

Mehdi Mirzapour, Jean-Philippe Prost, and Christian Retoré (LIRMM, Montpellier University CNRS, 161 Rue Ada, France)

Categorial Proof Nets and Dependency Locality: A New Metric for Linguistic Complexity

ABSTRACT: This work provides a quantitative computational account of why a sentence has harder parse than some other one, or that one analysis of a sentence is simpler than another one. We take for granted Gibson's results on human processing complexity, and we provide a new metric which uses (Lambek) Categorial Proof Nets. In particular, we correctly model Gibson's account in his Dependency Locality Theory. The proposed metric correctly predicts some performance phenomena such as structures with embedded pronouns, garden pathing, unacceptability of center embedding, preference for lower attachment and passive paraphrases acceptability. Our proposal extends existing distance-based proposals on Categorial Proof Nets for complexity measurement while it opens the door to include semantic complexity, because of the syntax-semantics interface in categorial grammars.

Aarne Ranta (University of Gothenburg and Digital Grammars AB, Sweden)
Invited Talk

Concept Alignment for Compositional Translation

ABSTRACT: Translation between natural languages is not compositional in a naive word-to-word sense. But many problems can be solved by using higher-level concepts, implementable as abstract syntax constructors in type theory together with compositional linearization functions in Grammatical Framework (GF). The question then arises: what are these constructors for a given set of languages? A whole spectrum of possibilities suggests itself: word senses (as in WordNet), multiword phrases (as in statistical machine translation), predication frames (as in FrameNet), syntactic deep structures (as in GF Resource Grammar Library), and lexico-syntactic constructions (as in Construction Grammar). The talk will study the problem in the light of experiences for building a cross-lingual lexicon of concepts in the General Data Protection Regulation (GDPR) in five languages. We have identified over 3000 concepts of varying complexity. A lot of manual work has been needed in the process, but some ideas have emerged toward a computational approach that generates concept alignment candidates by automated analysis.

Frank Richter (Goethe University Frankfurt a.M., Germany) Invited Talk

Computational Semantics: Representations and Reasoning

ABSTRACT: Computing with classical meaning representations of formal semantics encounters two major problems (with many sub-problems): How do we compose logical representations for natural language expressions in a computationally feasible grammar, and how do we actually reason with the sophisticated logical representations that theoretical linguists devise? This talk revisits the construction of logical representations in a few empirically and theoretically challenging areas of grammar, and presents a treatment of formulae of higher-order logic which makes it possible to use first order model builders and theorem provers to reason with them, with special attention to the emerging overall architecture.

Mehrnoosh Sadrzadeh (Queen Mary University of London, UK) Invited Talk

Lambdas, Vectors, and Dynamic Logic

(This is joint work with Reinhard Muskens and is supported by a Royal Society International Exchange Award.)

ABSTRACT: Vector models of language are based on the contextual aspects of language, the distributions of words and how they co-occur in text. Truth conditional models focus on the logical aspects of language, compositional properties of words and how they compose to form sentences. In the truth conditional approach, the denotation of a sentence determines its truth conditions, which can be taken to be a truth value, a set of possible worlds, a context change potential, or similar. In the vector models, the degree of co-occurrence of words in context determines how similar the meanings of words are. In this talk, we put these two models together and develop a vector semantics based on the simply typed lambda calculus models of natural language. We provide two types of vector

semantics: a static one that uses techniques familiar from the truth conditional tradition of Montague and a dynamic one based on a form of dynamic interpretation inspired by Heim's context change potentials. We show how the dynamic model revokes a dynamic logic whose implication can be applied to admittance of a sentence by a corpus, and provide examples.

Manfred Sailer (Goethe University Frankfurt a.M., Germany) Invited Talk
    Contraint-Based Underspecified Semantic Combinatorics
    ABSTRACT: In this talk, I will review a number of challenges of the syntax-semantics interface for a standard concept of compositionality. Such phenomena include: scope ambiguity, negative concord, discontinuous semantic contribution, polyadic quantification, and incomplete utterances. I will argue that a constraint-based underspecified semantic combinatorics, as pursued in Lexical Resource Semantics (LRS), allows for a natural and interesting analysis of such phenomena. A system like LRS combines insights and techniques of computational and formal semantics and, as such, continues the tradition of fruitful interaction between computational and theoretical linguistics.

Satoshi Tojo (School of Information Science, Japan Advanced Institute of Science and Technology (JAIST), Japan) Invited Talk
    Linear Algebraic Representation of Knowledge State of Agent
    ABSTRACT: We first propose a linear algebraic representation for the frame property, that is the accessibility in possible worlds as adjacency matrix. We show that the product between an adjacency matrix and a column vector of valuation results in possibility modality, and translate also the necessity modality, employing Boolean operations. Then, we apply the method to agent communication; we represent the belief change of agents by dynamic epistemic logic (DEL), and show that the belief change can also be shown by a sequence of linear transformation on accessibility matrix. Finally, we discuss the requirements for the formal presentation of 'who knows what at which time'.

Adrià Torrens Urrutia (Universitat Rovira i Virgili, Tarragona, Spain)
    A Proposal to Describe Fuzziness in Natural Language
    ABSTRACT: In this presentation, we propose formal models that consider grammaticality as a gradient property instead of the categorical view of grammaticality defended in theoretical linguistics. Given that deviations from the norm are inherent to the spontaneous use of language, linguistic analysis tools should account for different levels of grammaticality.

Christian Wurm (University of Düsseldorf, Germany) Invited Talk
    Reasoning with Ambiguity
    ABSTRACT: Ambiguity is often considered to be a nemesis of logical reasoning. Still, when addressing natural language semantics with formal logic, we somehow have to address it: we can "lose it in translation" by saying all ambiguity is syntactic and we interpret unambiguous syntactic derivations; we can use meta-formalisms in order to represent it; but the fact remains that humans

usually can perfectly reason with ambiguous statements. Hence it seems to be an interesting idea to include ambiguity into logic itself. In this talk, I will present the results of my pursuit of this idea, which are partly very surprising and odd, but in the very end (I hope) provide us with a deeper understanding of ambiguity and maybe even the nature of meaning.

Yuan Xie (Utrecht University, The Netherlands)

Referential Dependencies in Chinese: A Syntax- Discourse Processing Model

ABSTRACT: I am proposing a syntax-discourse processing model for the representation and interpretation of referential dependencies in Chinese. Chinese referentially dependent expressions (e.g. pronouns, reflexives, certain full noun phrases) are different from those in many indo-European languages and rely more on discourse (e.g. using bare noun phrases to express definiteness–lacking overt article the; sentence-free reflexive ziji (self-N)– referring to the speaker), for this reason, this model, taking both the morphosyntactic and discourse features of the referentially dependent expressions into consideration, reflects the view that referentially dependent nominal expressions and their antecedents are information units that are stored in our working memory system and the referential dependencies are established through the interactions of those information units in our working memory system.

Robert Östling (Stockholm University, Sweden) Invited Talk

Language Structure from Parallel Texts

ABSTRACT: Some texts have been translated into thousands of languages, a fact that allows us to compare the structures of language from a bird's-eye view. This information can then be used to study the evolutionary forces driving language change. I will discuss some of our results in this area, as well as current models for formalizing the phenomenon of human language on a global scale.