IN PARTNERSHIP WITH:
CNRS

Université de Lorraine

Activity Report 2011

Team Sémagramme

Semantic Analysis of Natural Language

IN COLLABORATION WITH: Laboratoire lorrain de recherche en informatique et ses applications (LORIA)

RESEARCH CENTER
Nancy - Grand Est

THEME
Audio, Speech, and Language Processing

Table of contents

Team Sémagramme

Keywords: Logics, Natural Language, Semantics, Linguistics

1. Members

Research Scientists
Philippe de Groote [Team Leader, Senior Researcher, INRIA, HdR]
Bruno Guillaume [Junior Researcher, INRIA]
Sylvain Pogodalla [Junior Researcher, INRIA]

Faculty Members
Maxime Amblard [Associate Professor, Université de Lorraine]
Guy Perrier [Professor, Université de Lorraine, HdR]

Technical Staff
Paul Masson [until September 30]

PhD Students
Ekaterina Lebedeva [INRIA Cordi Grant]
Mathieu Morey [until September 30, MENRT]
Florent Pompigne [Contrat doctoral]
Sai Qian [Contrat doctoral]
Shohreh Tabatabayi Seifi [INRIA Cordi Grant]

2. Overall Objectives

2.1. Overall objectives

Computational linguistics is a discipline at the intersection of computer science and linguistics. On the theoretical side, it aims to provide computational models of the human language faculty. On the applied side, it is concerned with natural language processing and its practical applications.

From a structural point of view, linguistics is traditionally organized into the following sub-fields:

- Phonology, the study of language abstract sound systems.
- Morphology, the study of word structure.
- Syntax, the study of language structure, i.e., the way words combine into grammatical phrases and sentences.
- Semantics, the study of meaning at the levels of words, phrases, and sentences.
- Pragmatics, the study of the ways in which the meaning of an utterance is affected by its context.

Computational linguistics is concerned by all these fields. Consequently, various computational models, whose application domains range from phonology to pragmatics, have been developed. Among these, logic-based models play an important part, especially at the "higher" levels.

At the level of syntax, generative grammars [29] may be seen as basic inference systems, while categorial grammars [38] are based on substructural logics specified by Gentzen sequent calculi. Finally, model-theoretic grammars [49] amount to sets of logical constraints to be satisfied.

At the level of semantics, the most common approaches derive from Montague grammars, [41], [42], [43] which are based on the simply typed λ-calculus and Church's simple theory of types. [30] In addition, various logics (modal, hybrid, intensional, higher- order...) are used to express logical semantic representations.

At the level of pragmatics, the situation is less clear. The word *pragmatics* has been introduced by Morris [45] to designate the branch of philosophy of language that studies, besides linguistic signs, their relation to their users and the possible contexts of use. The definition of pragmatics was not quite precise, and for a long time several authors have considered (and some authors are still considering) pragmatics as the wastebasket of syntax and semantics. [23] Nevertheless, as far as discourse processing is concerned (which includes pragmatic problems such as pronominal anaphora resolution), logic-based approaches have also been successful. In particular, Kamp's Discourse Representation Theory [36] gave rise to sophisticated 'dynamic' logics. [34] The situation, however, is less satisfactory than it is at the semantic level. On the one hand, we are facing a kind of logical "tower of Babel". The various pragmatic logic-based models that have been developed, while sharing underlying mathematical concepts, differ in several respects and are too often based on *ad hoc* features. As a consequence, they are difficult to compare and appear more as competitors than as collaborative theories that could be integrated. On the other hand, several phenomena related to discourse dynamics (e.g., context updating, presupposition projection and accommodation, contextual reference resolution...) are still lacking deep logical explanations. We strongly believe, however, that this situation can be improved by applying to pragmatics the same approach Montague applied to semantics, using the standard tools of mathematical logic.

Accordingly:

> *The overall objective of the Sémagramme project is to design and develop new unifying logic-based models, methods, and tools for the semantic analysis of natural language utterances and discourses. This includes the logical modelling of pragmatic phenomena related to discourse dynamics. Typically, these models and methods will be based on standard logical concepts (stemming from formal language theory, mathematical logic, and type theory), which should make them easy to integrate.*

The project is organized along three research directions (i.e., *Syntax-semantics interface*, *Discourse dynamics*, and *Common basic resources*), which interact as explained in the following paragraphs.

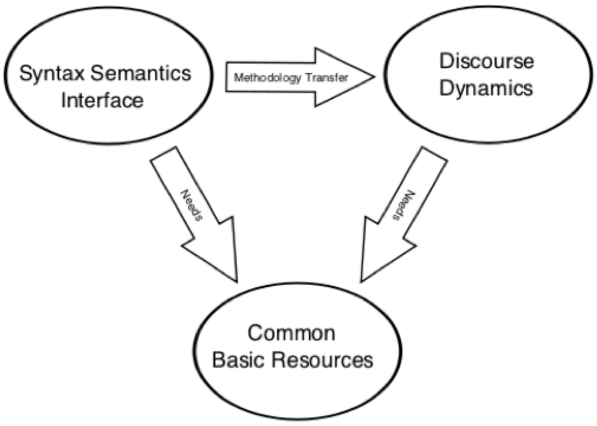

Figure 1.

2.1.1. Syntax-semantics interface

The Sémagramme project intends to focus on the semantics of natural languages (in a wider sense than usual, including some pragmatics). Nevertheless, the semantic construction process is syntactically guided, that is, the constructions of logical representations of meaning is based on the analysis of the syntactic structures.

We do not want, however, to commit ourselves to such or such specific theory of syntax. Consequently, our approach should be based on an abstract generic model of the syntax-semantic interface.

Here, an important idea of Montague comes into play, namely, the “homomorphism requirement”: semantics must appear as a homomorphic image of syntax. While this idea is almost a truism in the context of mathematical logic, it remains challenged in the context of natural languages. Nevertheless, Montague’s idea has been quite fruitful, especially in the field of categorial grammars, where van Benthem showed how syntax ans semantics could be connected using the Curry-Howard isomorphism. [58] This correspondence is the keystone of the syntax-semantics interface of modern type-logical grammars. [44] It also motivated the definition of our own Abstract Categorial Grammars. [52]

Technically, an Abstract Categorial Grammar consists simply of a (linear) homomorphism between two higher-order signatures. Extensive studies have shown that this simple model allows several grammatical formalisms to be expressed, providing them with a syntax-semantics interface for free. [53], [55], [56], [47], [37], [48]

We intend to carry on with the development of the Abstract Categorial Grammar framework. At the foundational level, we will define and study possible type theoretic extensions of the formalism, in order to increase its expressive power and its flexibility. At the implementation level, we will continue the development of an Abstract Categorial Grammar support system.

As said above, to consider the syntax-semantics interface as the starting point of our investigations allows us not to be committed to some specific syntactic theory. The Montagovian syntax-semantics interface, however, cannot be considered to be universal. In particular, it does not seem to be that well adapted to dependency and model-theoretic grammars. Consequently, in order to be as generic as possible, we intend to explore alternative models of the syntax-semantics interface. In particular, we will explore relational models where several distinct semantic representations can correspond to a same syntactic structure.

2.1.2. Discourse dynamics

It is well known that the interpretation of a discourse is a dynamic process. Take a sentence occurring in a discourse. On the one hand, it must be interpreted according to its context. On the other hand, its interpretation affects this context, and must therefore result in an updating of the current context. For this reason, discourse interpretation is traditionally considered to belong to pragmatics. The cut between pragmatics and semantics, however, is not that clear.

As we mentioned above, we intend to apply to some aspects of pragmatics (mainly, discourse dynamics) the same methodological tools Montague applied to semantics. The challenge here is to obtain a completely compositional theory of discourse interpretation, by respecting Montague’s homomorphism requirement. We think that this is possible by using techniques coming from programming language theory, in particular, continuation semantics [51], [24], [25], [50] and the related theories of functional control operators. [31], [32]

We have indeed successfully applied such techniques in order to model the way quantifiers in natural languages may dynamically extend their scope. [54] We intend to tackle, in a similar way, other dynamic phenomena (typically, anaphora and referential expressions, presupposition, modal subordination...).

What characterize these different dynamic phenomena is that their interpretations need information to be retrieved from a current context. This raises the question of the modeling of the context itself. At a foundational level, we have to answer questions such as the following. What is the nature of the information to be stored in the context? What are the processes that allow implicit information to be inferred from the context? What are the primitives that allow a context to be updated? How does the structure of the discourse and the discourse relations affect the structure of the context? These questions also raise implementation issues. What are the appropriate datatypes? How can we keep the complexity of the inference algorithms sufficiently low?

2.1.3. Common basic resources

Even if our research primarily focuses on semantics and pragmatics, we nevertheless need syntax. More precisely, we need syntactic trees to start with. We consequently need grammars, lexicons and parsing algorithms to produce such trees. During the last years, we have developped the notion of interaction grammar [35] as a model of natural language syntax. This includes the development of grammar for French, [46] together with morpho-syntactic lexicons. We intend to continue this line of research and development. In particular, we want to increase the coverage of our French grammar, and provide our parser with more robust algorithms.

Further primary resources are needed in order to put at work a computational semantic analysis of utterances and discourses. As we want our approach to be as compositional as possible, we must develop lexicons annotated with semantic information. This opens the quite wide research area of lexical semantics.

Finally, when dealing with logical representations of utterance interpretations, the need for inference facilities is ubiquitous. Inference is needed in the course of the interpretation process, but also to exploit the result of the interpretation. Indeed, an advantage of using formal logic for semantic representations is the possibility of using logical inference to derive new information. From a computational point of view, however, logical inference may be highly complex. Consequently, we need to investigate which logical fragments can be used efficiently for natural language oriented inference.

2.2. Highlights

Sylvain Pogodalla chaired the international conference Logical Aspects of Computational Linguistics (LACL 2011) organized in Montpellier.

3. Scientific Foundations

3.1. Fondation

The present proposal relies on deep mathematical foundations. We intend to develop models based on well-established mathematics. We seek two main advantages from this approach. On the one hand, by relying on mature theories, we have at our disposal sets of mathematical tools that we can use to study our models. On the other hand, developing various models on a common mathematical background will make them easier to integrate, and will ease the search for unifying principles.

The main mathematical domains on which we rely are formal language theory, symbolic logic, and type theory.

3.1.1. Formal language theory

studies the purely syntactic and combinatorial aspects of languages, seen as sets of strings (or possibly trees or graphs). Formal language theory has been especially fruitful for the development of parsing algorithms for context-free languages. We use it, in a similar way, to develop parsing algorithms for formalisms that go beyond context-freeness. Language theory also appears to be very useful in formally studying the expressive power and the complexity of the models we develop.

3.1.2. Symbolic logic

(and, more particularly, proof-theory) is concerned with the study of the expressive and deductive power of formal systems. In a rule-based approach to computational linguistics, the use of symbolic logic is ubiquitous. As we previously said, at the level of syntax, several kinds of grammars (generative, categorial...) may be seen as basic deductive systems. At the level of semantics, the meaning of an utterance is capture by computing (intermediate) semantic representations that are expressed as logical forms. Finally, using symbolic logics allows one to formalize notions of inference and entailment that are needed at the level of pragmatics.

3.1.3. Type theory and typed λ-calculus

Among the various possible logics that may be used, Church's simply typed λ-calculus and simple theory of types (a.k.a. higher-order logic) play a central part. On the one hand, Montague semantics is based on the simply typed λ-calculus, and so is our syntax-semantics interface model. On the other hand, as shown by Gallin, [33] the target logic used by Montague for expressing meanings (i.e., his intensional logic) is essentially a variant of higher-order logic featuring three atomic types (the third atomic type standing for the set of possible worlds).

4. Application Domains

4.1. Application Domains

The present proposal focuses on the semantics of natural language, including the semantic analysis of discourses. Consequently, our applicative domains concern natural language processing applications that rely on a deep semantic analysis. For instance, one may cite the following ones:

- textual entailment and inference;
- dialogue systems;
- semantic-oriented query systems;
- content analysis of unstructured documents;
- (semi) automatic knowledge acquisition.

In fact, the need for semantics seems to be ubiquitous. There is, however, a challenge here. We need to find applications for which a deep semantic analysis results in a real improvement over non semantic-based techniques.

Nevertheless, the possible applications one may imagine are numerous, but we do not want to be too specific about it, at this stage. We intend to develop applications in the framework of collaborations. Therefore, the actual applicative developments we will undertake will depend of the partners we are currently seeking.

5. Software

5.1. LEOPAR

Participants: Bruno Guillaume [correspondant], Guy Perrier, Mathieu Morey, Paul Masson.

5.1.1. Software description

LEOPAR is a parser for natural languages which is based on the formalism of Interaction Grammars [35]. It uses a parsing principle, called "electrostatic parsing" which consists in neutralizing opposite polarities. A positive polarity corresponds to an available linguistic feature and a negative one to an expected feature.

Parsing a sentence with an Interaction Grammar consists in first selecting a lexical entry for each of its words. A lexical entry is an underspecified syntactic tree, a tree description in other words. Then, all selected tree descriptions are combined by partial superposition guided by the aim of neutralizing polarities: two opposite polarities are neutralized by merging their support nodes. Parsing succeeds if the process ends with a minimal and neutral tree. As IGs are based on polarities and under-specified trees, LEOPAR uses some specific and non-trivial data-structures and algorithms.

The electrostatic principle has been intensively considered in LEOPAR. The theoretical problem of parsing IGs is NP-complete; the nondeterminism usually associated to NP-completeness is present at two levels: when a description for each word is selected from the lexicon, and when a choice of which nodes to merge is made. Polarities have shown their efficiency in pruning the search tree:

- In the first step (tagging the words of the sentence with tree descriptions), we forget the structure of descriptions, and only keep the bag of their features. In this case, parsing inside the formalism is greatly simplified because composition rules reduce to the neutralization of a negative feature-value pair $f \longleftarrow v$ by a dual positive feature-value pair $f \longrightarrow v$. As a consequence, parsing reduces to a counting of positive and negative polarities present in the selected tagging for every pair (f, v): every positive occurrence counts for $+1$ and every negative occurrence for -1, the sum must be 0.

- Again in the tagging step, original methods were developped to filter out bad taggings. Each unsaturated polarity p in the grammar induces constraints on the set of contexts in which it can be used: the unsaturated polarity p must find a *companion* (*i.e.* a tree description able to saturated it); and the set of companions for the polarity p can be computed statically from the grammar. Each lexical selection which contains an unsaturated polarity without one of its companions can be safely removed.
- In the next step (node-merging phase), polarities are used to cut off parsing branches when their trees contain too many non neutral polarities.

5.1.2. Current state of the implementation

LEOPAR is presented and documented at http://leopar.loria.fr; an online demonstration page can be found at http://leopar.loria.fr/demo.

It is open-source (under the CECILL License http://www.cecill.info) and it is developed using the InriaGforge platform (http://gforge.inria.fr/projects/semagramme/)

The main features of current software are:

- automatic parsing of a sentence or a set of sentences,
- dependency and parse-tree representation of sentences,
- interactive parsing (the user chooses the couple of nodes to merge),
- visualization of grammars produced by XMG or of sets of description trees associated to some word in the linguistic resources,

During 2011, with the help of an engineer, the LEOPAR software was improved in several ways:

- A new graphical interface (using GTK) was designed
- New algorithms for the super-tagging step of the parsing process were implemented. These algoritms are described in[9].

5.2. ACG Development Toolkit

In order to support the theoretical work on ACG, we have been developing a support system. The objectives of such a system are twofold:

1. to make possible to implement and experiment grammars the modeling of linguistic phenomena;
2. to make possible to implement and experiment results related to the ACG formalisms. Such results can concern parsing algorithms, type extensions, language extensions, etc.

The current version of the the ACG development toolkit prototype[1] issues from a first release published in October 2008. Further releases have been published before the ESSLLI 2009 course on ACG. It focuses on providing facilities to develop grammars. To this end, the type system currently implemented is the linear core system plus the (non-linear) intuitionistic implication, and a special attention has been paid to type error management. As a major limitation, this version only considers transformation from abstract terms to object terms, and not the other way around.

Enabling transformation from the object terms to the abstract terms is the first step of future development for the ACG support system. A parsing algorithm based on [37]'s methods is being implemented for second-order ACGs. It is based on a translation of ACG grammars into Datalog programs and is well-suited to fine-grained optimization. A summer internship from ENS Cachan, Clovis Eberhart (L3) has been implementing the translation from the higher-order signatures and terms data structures to the Datalog clauses data structures. It still remains to be integrated to the main branch.

[1] Available at http://acg.gforge.inria.fr with a CeCILL license.

In order to allow for a larger character set as input, another extension implemented this summer by another internship from École des Mines de Nancy, Grégoire Brenon (M1) was to extend the lexer and the parser for the data files with UTF-8 capabilities (OCaml lacks such a built-in capability).

However, since we're interested not only by recognizability (hence whether some fact is provable) but also by the parsing structure (hence the proof), the Datalog solver requires further adaptations. Note however that in the general case, the decidability of translating an object term to an abstract one is still an open problem.

5.3. GREW

Participants: Bruno Guillaume [correspondant], Guy Perrier, Mathieu Morey, Paul Masson.

Grew is a Graph Rewriting tools dedicated to applications in NLP. It was developed as a support tool during the PhD thesis of Mathieu Morey.

It is freely-available (from the page http://wikilligramme.loria.fr/doku.php?id=grew:grew) and it is developed using the InriaGforge platform (http://gforge.inria.fr/projects/semagramme/)

We list below some of the major specificities of the GREW software.

- Graph structures can use a build-in notion of feature structures.
- The left-hand side of a rule is described by a graph called a pattern; injective graph morphisms are used in the pattern matching algorithm.
- Negative pattern can be used for a finer control on the left-hand side of rules.
- The right-hand side or rules is described by a sequence of atomic commands that describe how the graph should be modified during the rule application.
- Subset of rules are grouped in modules; the full rewriting process being a sequence of module applications.
- The GREW software has support both for confluent and non-confluent modules; when a non-confluent modules is used, all normal forms are returned and then ambiguity is handled in a natural way.
- GREW can be used on Corpus mode with statistics about rules usage or with an a Graphical User Interface which can show all intermediate graphs used during the rewriting process (useful either to debug rewriting system of for demonstrations).

During the last 18 months, the GREW software were used for several kind of applications manipulating syntactic and/or semantic graph representations:

- to build DMRS semantic representation from syntactic dependency trees ([26], [14], [9]);
- to enrich surface syntactic structures ([13], [9]);
- to detect annotation errors in the French Treebank.

5.4. Other developments

Participants: Bruno Guillaume [correspondant], Paul Masson.

Other peripheral developments of the team are available either as web service of as downloadable code:

- A concordancer named CONDOR. The main features of this tool are:
  - It is usable online: http://condor.loria.fr;
  - It is possible to search for all inflexions (given by a lexicon) of some words;
  - It is possible to combine two searches and to search for a couple of words to find collocations.
- A program (named DEP2PICT) to build graphical representations of dependency structures.
  - it is presented and documented at: http://dep2pict.loria.fr;
  - it is usable online at http://dep2pict.loria.fr/demo;
  - it can produce PNG, SVG and PDF output formats;
  - it can be used to represented dependency structures with chunks;
  - it support CONLL input format.

6. New Results

6.1. Coordination Parsing

Participants: Bruno Guillaume, Guy Perrier.

In the development of the French grammar, FRIGRAM, Joseph Le Roux and Guy Perrier have tackled the difficult problem of modelling and parsing coordination [39]. They have enriched FRIGRAM with a module expressing different syntactic constructions with coordination. An important drawback of this approach is the number of elementary constructions that have to introduced to obtain a reasonable coverage of the phenomenon.

In the continuation of his Master thesis, Valmi Dufour-Lussier with Bruno Guillaume and Guy Perrier worked on a different approach. They propose to process coordination at the parsing level as a linguistic performance issue, outside the grammar, rather than as a matter of competence [15]. They apply a specific algorithm to combine coordinated syntactic structures that were partially parsed using a coordination-less grammar, resulting in a directed acyclic parse graph in which constituent sharing appears sharply. They have experimented the algorithm within the framework of Tree-Adjoining Grammars (although it can be adapted to other formalisms) on a small subset of the Penn Treebank [2]. They have shown that it is able to handle many types of coordinative constructions, including left and right node raising, argument clusters, and verb gapping.

6.2. Graph Rewriting

Participants: Bruno Guillaume, Mathieu Morey, Guy Perrier.

Guillaume Bonfante (from CARTE team), Bruno Guillaume, Mathieu Morey and Guy Perrier have improved their graph rewriting calculus, experimenting it in two directions. Taking an asynchronous perspective on the syntax-semantics interface, they have designed a modular graph rewriting system to produce underspecified semantic representations from a syntactic dependency graph [14]. They experimentally validated this approach on a set of sentences extracted from the French Treebank annotated with syntactic dependencies [27]. The results open the way for the production of underspecified semantic dependency structures from corpora annotated with syntactic dependencies and, more generally, for a broader use of modular rewriting systems for computational linguistics.

[2] http://www.cis.upenn.edu/~treebank/

In a second application, they show how to enrich a syntactic dependency annotation of the French Treebank, using graph rewriting, in order to compute its semantic representation [13]. The rewriting system is composed of grammatical and lexical rules structured in modules. The lexical rules use a control information extracted from Dicovalence, a lexicon of French verbs [3].

6.3. ACG Type System

Participants: Philippe de Groote, Sylvain Pogodalla, Florent Pompigne.

In order to extend the flexibility and the expressiveness of the the ACG framework, [57] proposed a type-system extension. However, the formal properties of the system have to be proved. In his PhD work, Florent Pompigne is proposing alternate η-rules and commutative conversions in order to get the desirable properties. This work, currently in progress, relates to former proposals for a linear calculus with dependent types [28] and a calculus for extensionality with variants [40].

6.4. Logic and Grammars

Participant: Maxime Amblard.

Maxime Amblard has presented an extension of Minimalist Categorial Grammars (MCG) to encode Chomsky's phases in [11]. These grammars are based on Partially Commutative Logic (PCL) and encode properties of Minimalist Grammars (MG) of Stabler. The first implementation of MCG were using both non-commutative properties (to respect the linear word order in an utterance) and commutative ones (to model features of different constituents). Here, we propose to adding Chomsky's phases with the non-commutative tensor product of the logic.

6.5. Discourse dynamics

Participants: Maxime Amblard, Sai Qian.

Sai Qian and Maxime Amblard has presented a framework which constructs an event-style discourse semantics, [17]. The discourse dynamics are encoded in continuation semantics [54] and various rhetorical relations are embedded in the resulting interpretation of the framework. They assume that discourse and sentence are distinct semantic objects, that play different roles in meaning evaluation. Moreover, two sets of composition functions, for handling different discourse relations, are introduced.

6.6. Modeling pathological discourse

Participant: Maxime Amblard.

Maxime Amblard starts a conjoint work with a psychologist Michel Musiol (IntePsy) and a philosopher Manuel Rebuschi (Archives Poincaré) about developing a formal analysis of pathological conversations involving schizophrenic speakers [18]. Such conversations give rise to manifest incongruities or ruptures that can be seen as mere contradictions by any "normal" speaker. Our construal relies both on semantic and pragmatic features of conversation. After an overview on the making of the corpus, we propose a SDRT-inspired account of pathological conversations, and we apply it to two relevant excerpts. We conclude with a short discussion about the localization of incoherencies by schizophrenics, either in semantics or in pragmatics, and its importance for our understanding of thought disorders.

[3] http://bach.arts.kuleuven.be/dicovalence/

7. Partnerships and Cooperations

7.1. Regional Initiatives

7.1.1. DiaRaFor

DiaRaFor,*Dialogues, rationalités et formalismes. Etudes croisées logique / psychologie / épistémologie* is a MSH-Lorraine project.

Schizophrenia is well-known among mental illnesses for the strength of the thought disorders it involves, and for their widespread and spectacular manifestations: from deviant social behavior to delusion, not to speak about affective and sensitive distortions. In the present paper we expose the first steps of a scientific research about one specific manifestation, namely disorders in conversational speech. This is an interdisciplinary research, both empirical and theoretical from several domains, namely psychology, philosophy, linguistic and informatics.

Maxime Amblard is envolved in this project which ended in 2011. A new application on this topic is send in 2011 to the MSH with Maxime Amblard as leader : SLAM - Schizophrenia and Language : Analysis and Modeling.

8. Dissemination

8.1. Animation of the scientific community

- Maxime Amblard was part of the scientific committee of TALN 11 (Montpellier, France) and WoSS8 8th Workshop on Syntax & Semantics (Paris, France).
- Maxime Amblard is head of the master *Sciences Cognitives et Applications* of University Nancy 2.
- Maxime Amblard is head of the master 2 speciality *Traitements Automatiques des Langues* of the master SCA University Nancy 2.
- Maxime Amblard has prepared the revised version for the master SCA from 2013-2018.
- Maxime Amblard was president of the selection committee (recruitment of a permanent lecturer), for a PRCE and for a PRAG at University Nancy 2.
- Maxime Amblard was curator of the exhibit *Fascination ou aversion pour le numérique : encodage/décodage* at University Nancy 2.
- Maxime Amblard was member of the organization comity of the *forum des Sciences Cognitives*.
- Maxime Amblard joined the editorial board of InterStice)i(.
- Maxime Amblard is vice-treasurer of the *Association pour le Traitement Automatique des Langues* (ATALA)
- Maxime Amblard is member *Opération Poste* team.
- Guy Perrier is member of the editorial board of the journal *Traitement Automatique des Langues*.
- Guy Perrier attended the thesis defense of Mathieu Morey, Université Nancy 2, November 3, as the supervisor of the thesis.
- Guy Perrier was the local coordinator of the Erasmus Mundus Master program *Language and Communication Technologies* for the University of Nancy 2 until August. He organized the graduation ceremony for the whole consortium in Nancy May 26.
- Bruno Guillaume is a member of the COMIPERS (Comité de recrutement INRIA Lorraine/LORIA des personnels scientifiques contractuels) of the Nancy INRIA Research Center.

- Bruno Guillaume is a member of the CUMIR (Commission des Utilisateurs de Moyens Informatiques) of the Nancy INRIA Research Center.
- Sylvain Pogodalla is local coordinator of the Erasmus Mundus Master program *Language and Communication Technologies* for the University of Nancy 2 starting from September 1st.
- Sylvain Pogodalla chaired the international conference Logical Aspects of Computational Linguistics (LACL 2011) organized in Montpellier.
- Sylvain Pogodalla is member of the editorial board of the journal *T.A.L.* (*Traitement automatique des Langues*.
- Sylvain Pogodalla reviewed papers for *Transaction on Computational Logic*, for *Research on Language and Computation* and for *Synthese*.
- Sylvain Pogodalla was part of the scientific committees of: TALN 2011 (Montpellier), Mathematics of Language (MOL 2011), Constraints in Discourse (CID 2011), International Conference in Computational Semantics (IWCS 2011).
- Sylvain Pogodalla is head of the Commission des Développements Technologiques of the Nancy INRIA Research Center.
- Sylvain Pogodalla participated in the jury of the thesis defense of Bruno Méry (Bordeaux University, SIGNES INRIA project team)

8.2. Teaching

Licence: *Penser les algorithmes*, 20h, L1, Université Nancy 2, France (Maxime Amblard);

Licence: *Introduction au Traitement Automatique des Langues*, 20h, L2, Université Nancy 2, France (Maxime Amblard);

Licence: *Algorithmique avancée*, 50h, L2, Université Nancy 2, France (Maxime Amblard);

Licence: *Formalismes de représentation et raisonnement*, 20h, L3, Université Nancy 2, France (Maxime Amblard);

Licence: *Outils conceptuels*, 30h, L3, Université Nancy 2, France (Guy Perrier);

Master: *Algorithmique pour l'Intelligence Artificielle*, 37h, M1, Université Nancy 2, France (Maxime Amblard);

Master: *Discourse and Dialogue*, 20h, M2, Université Nancy 2, France (Maxime Amblard);

Master: *tools and algorithms for NLP*, 34h,M2, Université Nancy 2, France (Guy Perrier);

Master: *initiation to NLP*, 30h, M1, Université Nancy 2, France (Guy Perrier);

Master: *programming for NLP* 30h, M1, Université Nancy 2, France (Guy Perrier);

Master: *Grammatical formalisms*, 20h, M2, Université Nancy 2, France (Bruno Guillaume);

Master: *Computational Semantics*, 20h, M2, Université Nancy 2, France (Sylvain Pogodalla);

PhD: Mathieu Morey, *Étiquetage grammatical symbolique et interface syntaxe-sémantique des formalismes grammaticaux lexicalisés polarisés* [9], defended on November 3rd;

PhD in progress: Sai Qian, *Investigation of variable accessibility in natural language semantic modelization*, since september 2009, Philippe de Groote, Maxime Amblard;

PhD in progress: Florent Pompigne, *Logical Modeling of Natural Language and ACG*, since september 2008, Philippe de Groote, Sylvain Pogodalla;

Master 1: Geoffrey Mougel and Cédric Beuzit, *Interface et vérification pour la modélisation de conversations schizophréniques*, Maxime Amblard;

Master 1: Lorraine Tosi and Pierre Kimmel, *LEOPAR au Collège*, Bruno Guillaume.

9. Bibliography

Major publications by the team in recent years

[1] N. ASHER, S. POGODALLA. *A Montagovian Treatment of Modal Subordination*, in "20th Semantics and Linguistic Theory conference - SALT2010", Vancouver, Canada, N. LI, D. LUTZ (editors), 2011, http://hal.inria.fr/inria-00565616/en.

[2] G. BONFANTE, B. GUILLAUME, M. MOREY, G. PERRIER. *Modular Graph Rewriting to Compute Semantics*, in "9th International Conference on Computational Semantics - IWCS 2011", Oxford, Royaume-Uni, J. BOS, S. PULMAN (editors), January 2011, p. 65–74, http://hal.inria.fr/inria-00579244/en/.

[3] G. BONFANTE, B. GUILLAUME, G. PERRIER. *Analyse syntaxique électrostatique*, in "Traitement Automatique des Langues (TAL)", 2003, vol. 44, n^{o} 3, p. 93-120, Article dans revue scientifique avec comité de lecture. nationale., http://hal.inria.fr/inria-00099582/en/.

[4] B. GUILLAUME, G. PERRIER. *Interaction Grammars*, in "Research on Language & Computation", 2009, vol. 7, p. 171–208.

[5] P. DE GROOTE. *Towards abstract categorial grammars*, in "Association for Computational Linguistics, 39th Annual Meeting and 10th Conference of the European Chapter", Toulouse, France, none, July 2001, p. 148-155, Colloque avec actes et comité de lecture. internationale., http://hal.inria.fr/inria-00100529/en.

[6] P. DE GROOTE, S. POGODALLA. *On the expressive power of abstract categorial grammars: Representing context-free formalisms*, in "Journal of Logic, Language and Information", 2004, vol. 13, n^{o} 4, p. 421–438.

[7] P. DE GROOTE, S. POGODALLA, C. POLLARD. *About Parallel and Syntactocentric Formalisms: A Perspective from the Encoding of Convergent Grammar into Abstract Categorial Grammar*, in "Fundamenta Informaticae", 2011, vol. 106, n^{o} 2-4, p. 211-231 [*DOI :* 10.3233/FI-2011-384], http://hal.inria.fr/inria-00565598/en.

[8] P. DE GROOTE. *Towards a Montagovian Account of Dynamics*, in "16th Semantics and Linguistic Theory conference - SALT2006", Tokyo, Japan, M. GIBSON, J. HOWELL (editors), 2006, http://elanguage.net/journals/index.php/salt/article/view/16.1.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[9] M. MOREY. *Étiquetage grammatical symbolique et interface syntaxe-sémantique des formalismes grammaticaux lexicalisés polarisés*, Université de Lorraine, November 2011, http://hal.inria.fr/tel-00640561/en.

Articles in International Peer-Reviewed Journal

[10] P. DE GROOTE, S. POGODALLA, C. POLLARD. *About Parallel and Syntactocentric Formalisms: A Perspective from the Encoding of Convergent Grammar into Abstract Categorial Grammar*, in "Fundamenta Informaticae", 2011, vol. 106, n^{o} 2-4, p. 211-231 [*DOI :* 10.3233/FI-2011-384], http://hal.inria.fr/inria-00565598/en.

International Conferences with Proceedings

[11] M. AMBLARD. *Encoding Phases using Commutativity and Non-commutativity in a Logical Framework*, in "6th International Conference on Logical Aspect of Computational Linguistic - LACL 2011", Montpellier, France, S. POGODALLA, J.-P. PROST (editors), Lecture Notes in Artificial Intelligence, Springer, June 2011, vol. 6736, p. 1–16, ISBN : 978-3-642-22220-7 The original publication is available at www.springerlink.com [*DOI :* 10.1007/978-3-642-22221-4_1], http://hal.inria.fr/hal-00601621/en.

[12] N. ASHER, S. POGODALLA. *A Montagovian Treatment of Modal Subordination*, in "20th Semantics and Linguistic Theory conference - SALT2010", Vancouver, Canada, N. LI, D. LUTZ (editors), 2011, http://hal.inria.fr/inria-00565616/en.

[13] G. BONFANTE, B. GUILLAUME, M. MOREY, G. PERRIER. *Enrichissement de structures en dépendances par réécriture de graphes*, in "Traitement Automatique des Langues Naturelles (TALN)", Montpellier, France, 2011, http://hal.inria.fr/inria-00579251/en.

[14] G. BONFANTE, B. GUILLAUME, M. MOREY, G. PERRIER. *Modular Graph Rewriting to Compute Semantics*, in "9th International Conference on Computational Semantics - IWCS 2011", Oxford, United Kingdom, J. BOS, S. PULMAN (editors), January 2011, p. 65–74, http://hal.inria.fr/inria-00579244/en.

[15] V. DUFOUR-LUSSIER, B. GUILLAUME, G. PERRIER. *Parsing Coordination Extragrammatically*, in "5th Language & Technology Conference", Poznan, Poland, 2011, http://hal.inria.fr/hal-00639929/en.

[16] C. GARDENT, Y. PARMENTIER, G. PERRIER, S. SCHMITZ. *Lexical Disambiguation in LTAG using Left Context*, in "5th Language & Technology Conference - LTC'11", Poznan, Poland, November 2011, http://hal.inria.fr/hal-00629902/en.

[17] S. QIAN, M. AMBLARD. *Event in Compositional Dynamic Semantics*, in "6th International Conference on Logical Aspect of Computational Linguistic - LACL 2011", Montpellier, France, S. POGODALLA, J.-P. PROST (editors), Lecture Notes in Artificial Intelligence, Springer, June 2011, vol. 6736, p. 219–234, ISBN : 978-3-642-22220-7 The original publication is available at www.springerlink.com [*DOI :* 10.1007/978-3-642-22221-4_15], http://hal.inria.fr/hal-00601620/en.

National Conferences with Proceeding

[18] M. AMBLARD, M. MUSIOL, M. REBUSCHI. *Une analyse basée sur la S-DRT pour la modélisation de dialogues pathologiques*, in "Actes de la 18e conférence sur le Traitement Automatique des Langues Naturelles - TALN 2011", Montpellier, France, M. LAFOURCADE, V. PRINCE (editors), Laboratoire d'Informatique de Robotique et de Microélectronique, June 2011, 6, http://hal.inria.fr/hal-00601622/en.

Scientific Books (or Scientific Book chapters)

[19] M. AMBLARD. *Minimalist Grammars and Minimalist Categorial Grammars, definitions toward inclusion of generated languages*, in "Logic and Grammar: Essays Dedicated to Alain Lecomte on the Occasion of His 60th Birthday", S. POGODALLA, M. QUATRINI, C. RETORÉ (editors), Lecture Notes in Artificial Intelligence, springer, June 2011, p. 61–80, ISBN : 978-3-642-21489-9 The original publication is available at www.springerlink.com [*DOI :* 10.1007/978-3-642-21490-5_4], http://hal.inria.fr/hal-00617040/en.

[20] N. ASHER, S. POGODALLA. *SDRT and Continuation Semantics*, in "New Frontiers in Artificial Intelligence JSAI-isAI 2010 Workshops, LENLS, JURISIN, AMBN, ISS, Tokyo, Japan, November 18-19, 2010, Revised

Selected Papers", T. ONADA, D. BEKKI, E. MCCREADY (editors), LNCS, Springer, 2011, vol. 6797, p. 3–15 [*DOI :* 10.1007/978-3-642-25655-4_2], http://hal.inria.fr/inria-00565744/en.

Books or Proceedings Editing

[21] S. POGODALLA, J.-P. PROST (editors). *Logical Aspects of Computational Linguistics, 6th International Conference, LACL 2011*, LNCS/LNAI, Springer, 2011, vol. 6736 [*DOI :* 10.1007/978-3-642-22221-4], http://hal.inria.fr/inria-00607874/en.

[22] S. POGODALLA, M. QUATRINI, C. RETORÉ (editors). *Logic and Grammar. Essays Dedicated to Alain Lecomte on the Occasion of His 60th Birthday*, LNCS/LNAI, Springer, 2011, vol. 6700 [*DOI :* 10.1007/978-3-642-21490-5], http://hal.inria.fr/inria-00607880/en.

References in notes

[23] Y. BAR-HILLEL. *Out of the pragmatic wastebasket*, in "Linguistic Inquiry", 1971, vol. 2, n° 3, p. 401–407.

[24] C. BARKER. *Continuations and the Nature of Quantification*, in "Natural Language Semantics", 2002, vol. 10, n° 3, p. 211–242.

[25] C. BARKER. *Continuations in Natural Language*, in "Proceedings of the Fourth ACM SIGPLAN Continuations Workshop (CW'04)", 2004.

[26] G. BONFANTE, B. GUILLAUME, M. MOREY, G. PERRIER. *Réécriture de graphes de dépendances pour l'interface syntaxe-sémantique*, in "Traitement Automatique des Langues Naturelles - TALN 2010", Montréal, Canada, July 2010, http://hal.inria.fr/inria-00526437/en/.

[27] M. CANDITO, B. CRABBÉ, P. DENIS. *Statistical French dependency parsing: treebank conversion and first results*, in "Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC 2010)", La Valletta, Malte, European Language Resources Association (ELRA), May 2010, p. 1840-1847, http://hal.archives-ouvertes.fr/hal-00495196/en/.

[28] I. CERVESATO, F. PFENNING. *A Linear Logical Framework*, in "Information and Computation", 2002, vol. 179, n° 1, p. 19–75, http://dx.doi.org/10.1006/inco.2001.2951.

[29] N. CHOMSKY. *Three models for the description of language*, in "Information Theory, IRE Transactions on", 1956, vol. 2, n° 3, p. 113–124.

[30] A. CHURCH. *A formulation of the simple theory of types*, in "The journal of symbolic logic", 1940, vol. 5, n° 2, p. 56–68.

[31] M. FELLEISEN, D. P. FRIEDMAN, E. KOHLBECKER, B. DUBA. *A syntactic theory of sequential control*, in "Theor. Comput. Sci.", June 1987, vol. 52, p. 205–237, http://dx.doi.org/10.1016/0304-3975(87)90109-5.

[32] M. FELLEISEN, R. HIEB. *The revised report on the syntactic theories of sequential control and state*, in "Theoretical Computer Science", 1992, vol. 103, n° 2, p. 235 - 271 [*DOI :* 10.1016/0304-3975(92)90014-7], http://www.sciencedirect.com/science/article/pii/0304397592900147.

[33] D. GALLIN. *Intensional and higher-order modal logic: with applications to Montague semantics*, North-Holland mathematics studies, North-Holland, 1975, http://books.google.fr/books?id=jkiNo5Rok7cC.

[34] J. GROENENDIJK, M. STOKHOF. *Dynamic predicate logic*, in "Linguistics and philosophy", 1991, vol. 14, n° 1, p. 39–100.

[35] B. GUILLAUME, G. PERRIER. *Interaction Grammars*, in "Research on Language & Computation", 2009, vol. 7, p. 171–208.

[36] H. KAMP, U. REYLE. *From discourse to logic: Introduction to modeltheoretic semantics of natural language, formal logic and discourse representation theory*, Kluwer Academic Dordrecht,, The Netherlands, 1993, vol. 42.

[37] M. KANAZAWA, S. SALVATI. *Generating Control Languages with Abstract Categorial Grammars*, in "Proceedings of the 12th conference on Formal Grammar", 2007.

[38] J. LAMBEK. *The mathematics of sentence structure*, in "The American Mathematical Monthly", 1958, vol. 65, n° 3, p. 154–170.

[39] J. LE ROUX, G. PERRIER. *La coordination dans les grammaires d'interaction*, in "Traitement Automatique des Langues", 2006, vol. 47, n° 3, p. 89–113.

[40] S. LINDLEY. *Extensional Rewriting with Sums.*, in "TLCA'07", LNCS, 2007, vol. 4583, p. 255–271.

[41] R. MONTAGUE. *English as a formal language*, in "Linguaggi nella societae nella tecnica", 1970, p. 189–224.

[42] R. MONTAGUE. *Universal grammar*, in "Theoria", 1970, vol. 36, n° 3, p. 373–398.

[43] R. MONTAGUE. *The Proper Treatment of Quantification in Ordinary English*, in "Formal Semantics", 1973, p. 221–242.

[44] M. MOORTGAT. *Categorial type logics*, in "Handbook of logic and language", 1997, p. 93–177.

[45] C. MORRIS. *Foundations of the Theory of Signs*, University of Chicago Press, 1938, vol. 1.

[46] G. PERRIER. *A French Interaction Grammar*, in "International Conference on Recent Advances in Natural Language Processing - RANLP 2007", Borovets, Bulgarie, G. ANGELOVA, K. BONTCHEVA, R. MITKOV, N. NICOLOV, K. SIMOV (editors), INCOMA Ltd, Shoumen, Bulgaria, 2007, p. 463-467, http://hal.inria.fr/inria-00184108/en/.

[47] S. POGODALLA. *Computing semantic representation: Towards ACG abstract terms as derivation trees*, in "Proceedings of the Seventh International Workshop on Tree Adjoining Grammar and Related Formalisms (TAG+7)", 2004.

[48] C. RETORÉ, S. SALVATI. *A Faithful Representation of Non-Associative Lambek Grammars in Abstract Categorial Grammars*, in "Journal of Logic Language and Information", April 2010, vol. 19, n° 2, p. 185–200, http://hal.inria.fr/inria-00409557/en/.

[49] J. ROGERS. *A model-theoretic framework for theories of syntax*, in "Proceedings of the 34th annual meeting on Association for Computational Linguistics", Association for Computational Linguistics, 1996, p. 10–16.

[50] C.-C. SHAN. *Delimited continuations in natural language: quantification and polarity sensitivity*, in "CoRR", 2004, vol. cs.CL/0404006.

[51] C. STRACHEY, C. P. WADSWORTH. *Continuations: A Mathematical Semantics for Handling Full Jumps*, in "Higher-Order and Symbolic Computation", 1974, vol. 13, p. 135–152.

[52] P. DE GROOTE. *Towards abstract categorial grammars*, in "Proceedings of the 39th Annual Meeting on Association for Computational Linguistics", Association for Computational Linguistics, 2001, p. 252–259.

[53] P. DE GROOTE. *Tree-Adjoining Grammars as Abstract Categorial Grammars*, in "TAG+6, Proceedings of the sixth International Workshop on Tree Adjoining Grammars and Related Frameworks", Università di Venezia, 2002, p. 145–150.

[54] P. DE GROOTE. *Towards a montagovian account of dynamics*, in "Semantics and Linguistic Theory (SALT) 16", M. GIBSON, J. HOWELL (editors), 2006, p. 1-16.

[55] P. DE GROOTE, S. POGODALLA. *m-linear context-free rewriting systems as abstract categorial grammars*, in "oceedings of Mathematics of Language - MOL-8", 2003.

[56] P. DE GROOTE, S. POGODALLA. *On the expressive power of abstract categorial grammars: Representing context-free formalisms*, in "Journal of Logic, Language and Information", 2004, vol. 13, n^o^ 4, p. 421–438.

[57] P. DE GROOTE, S. MAAREK. *Type-theoretic extensions of Abstract Categorial Grammars*, in "New Directions in Type-Theoretic Grammars, proceedings of the workshop", 2007, p. 18–30, http://let.uvt.nl/general/people/rmuskens/ndttg/ndttg2007.pdf.

[58] J. VAN BENTHEM. *Essays in logical semantics*, Reidel, 1986, n^o^ 29.