

Lecture Notes in Artificial Intelligence

1582

Subseries of Lecture Notes in Computer Science

Edited by J. G. Carbonell and J. Siekmann

Lecture Notes in Computer Science

Edited by G. Goos, J. Hartmanis and J. van Leeuwen

Springer

Berlin

Heidelberg

New York

Barcelona

Hong Kong

London

Milan

Paris

Singapore

Tokyo

Alain Lecomte François Lamarche
Guy Perrier (Eds.)

Logical Aspects of Computational Linguistics

Second International Conference, LACL '97
Nancy, France, September 22-24, 1997
Selected Papers



Springer

Series Editors

Jaime G. Carbonell, Carnegie Mellon University, Pittsburgh, PA, USA
Jörg Siekmann, University of Saarland, Saarbrücken, Germany

Volume Editors

Alain Lecomte
Université Pierre Mendès France (Grenoble 2)
1251 avenue Centrale, BP 47
F-38040 Grenoble Cedex, France
E-mail: Alain.Lecomte@upmf-grenoble.fr

François Lamarche
Guy Perrier
LORIA (Nancy)
615 rue de Jardin Botanique, BP 101
F-54602 Villers lès Nancy Cedex, France
E-mail: {Lamarche,Perrier}@loria.fr

Cataloging-in-Publication Data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Logic Aspects of computational linguistics : second international conference ; selected papers / LACL '97, Nancy, France, September 22 - 24, 1997. Alain Lecomte ... (ed.). - Berlin ; Heidelberg ; New York ; Barcelona ; Hong Kong ; London ; Milan ; Paris ; Singapore ; Tokyo : Springer, 1999
(Lecture notes in computer science ; Vol. 1582 : Lecture notes in artificial intelligence)
ISBN 3-540-65751-7

CR Subject Classification (1998): F.4, F.3, I.2.7, I.2.3, J.5

ISBN 3-540-65751-7 Springer-Verlag Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable for prosecution under the German Copyright Law.

© Springer-Verlag Berlin Heidelberg 1999
Printed in Germany

Typesetting: Camera-ready by author
SPIN: 10703147 06/3142 - 5 4 3 2 1 0 Printed on acid-free paper

Preface

This volume contains invited and selected papers of the Second International Conference on Logical Aspects of Computational Linguistics, held in Nancy, France, 22-24 September 1997.¹ Several perspectives are adopted in this volume and we will classify them according to their school of thought or to the main domain to which they seem to contribute. For instance, two invited papers, out of a total of four², are included, one by Jim LAMBEK and another by Denis BOUCHARD. LAMBEK's paper belongs to the logical tradition also represented in this volume by other contributions based on type theory, like the one by S. Shaumyan and P. Hudak³. It was a great honour to welcome Professor LAMBEK at this colloquium as the founding father of a new discipline, *language analysis by means of type theory*. All the participants knew how deeply indebted they have been to him since his first seminal papers of the early sixties on syntactic calculus. Besides his activities in pure mathematics, LAMBEK has continued to explore type logical systems, consistently looking for maximal simplicity. Here, he presents a new calculus: a bilinear formalism (therefore similar to the classical systems of linear logic) where multiplicative conjunction and disjunction are collapsed into a single operation/connector. What is appealing is the kind of rigorous and simple model that he can obtain for complex phenomena like verbal inflection. He can even give an account of what is analyzed in terms of Chomsky's theory of traces.

Denis BOUCHARD's contribution is completely different: it provides a purely linguistic analysis of *adjectival modification*, but this analysis must be recast in a more general perspective because it is treated like a case study in order to test the assumptions of the chomskyan minimalist program. Bouchard's purpose is to draw all the conclusions that can be drawn from the minimalist assumptions. He can therefore give us some guidelines for going deeper into a theory of language, and particularly for choosing adequate minimal concepts. He focuses here on the opposition between French and English relative to the question of word order in noun-adjective associations.

These two invited papers perfectly illustrate the theoretical situation of this conference between *logic* and *linguistics*. More precisely, we know that generative linguistics is oriented more and more towards operations of feature checking and

¹ The proceedings of the first conference, edited by Christian Retoré, were also published by Springer as volume 1328 of the LNCS/LNAI series. The reader may find in it a survey of works connecting logic and linguistics.

² The other two invited papers were : one by Michael MOORTGAT, published elsewhere, and another by Yves LAFONT, a tutorial on Linear Logic and its phase semantics

³ Of course all the papers connected to categorial grammars could be said to belong to this tradition. We put emphasis on these papers here because they advocate the use of pure type-logical systems

merging that can be thought of in terms of resource consumption and categorial selection. On the other hand, we know that contemporary logic has provided us with so-called resource conscious logics, like linear logic, and of course Lambek grammars. It was therefore natural and profitable to mix researchers coming from these various fields.

Another classification may be provided. It distinguishes papers according to their main domain of enquiry: philosophical, mathematical and foundational, linguistic-oriented, computational and applicational. The philosophical perspective concerns the philosophy of language and addresses the question of the place of logic with regards to language. This question is very controversial. Some authors (and Chomsky among them) argue that because *language is not logic*, all the convergence points between logical investigations and linguistic ones are mere accidents. For them, linguistic systems must be described by independent means and tools, logical formalization being spurious. This is of course not the attitude adopted here. Some contributors argue, on the contrary, that the system of language is very similar to Logic. For instance, S. SHAUMYAN and P. HUDAK present ideas that can be traced to the works of Haskell B. Curry on *combinatory logic*. Their view remains parallel to Montague's who claimed in 1970 that "there is no important theoretical difference between natural languages and the artificial languages of logicians." Of course, if we think of linguistics as an ordinary science, there is no reason to reject the idea of a mathematical formalization, and if that is the case, it seems natural that such a formalization be addressed by a particular branch of mathematics (not necessarily the same one that is used, say, for physics). Mathematical logic seems at present to be the best candidate for that, if only because it has taken the notion of language as one of its objects and has derived from it several suggestions for analysing language. Since it is not certain that we understand the same thing under the two notions of language, the linguistic and the logical, it seems the philosophical question will still remain open for a long time.

We can hope for more success by adopting a more technical perspective. A foundation for studying aspects of linguistics including categories and operations for combining signs can be found in the mathematics of categorial systems as they originate from Ajduckiewicz. Many authors in this volume share the view that modern techniques make it possible to generalize old categorial systems to new ones which are able to deal with several dimensions of signs at the same time. One basis for doing that consists in introducing several arbitrary products with their residuals, according to a line mainly explored by M. Moortgat, D. Oehrle and N. Kurtonina, but also by M. Hepple, G. Morrill. It is taken up in this volume by M. KANDULSKI, who demonstrates the strong equivalence of Ajduckiewicz and Lambek grammars when using these arbitrary products (we already had results from Buszkowski on the equivalence of Lambek and Ajduckiewicz grammars in the standard case). Such results must be classified among others concerning the generative capacity of grammars based on logical systems (and therefore categorial grammars). When we show that Ajduckiewicz and Lambek

grammars are equivalent to context-free ones and therefore equivalent to each other, we show that at least we don't lose anything when working in the logical framework compared to, say, a phrase structure grammar approach. But moreover, with the ability to use several products and thus to superpose several calculi onto a single one, we of course enrich the structural descriptions.

Another way of dealing with multi-dimensionality is *fibring*. In the past, there have been several attempts to combine a feature logic and a categorial one, as in the CUG and UCG models. The interest in Head Driven Phrase Structure Grammars has led to reversing the perspective and N. FRANCEZ proposes a conception according to which the categorial information “leaves-within” feature terms. Such technical approaches show us the rich power of logical techniques as soon as we dare to escape from strict orthodoxy and strict “logicism.” Of course, pure logicians can be frightened by the simple project of mixing several logics, but here the linguists are acting like physicists in making many trials and attempts before obtaining a “realistic” view of the field.

Another particularly relevant technical application of logics to computational linguistics concerns the use of automata for checking well-formedness with regards to some kind of theory expressed by a set of constraints. The principle and parameters approach provides such a theory. F. MORAWIETZ and T. CORNELL explore the power of automata for *monadic second order logic* in this perspective. By doing so, they implicitly show that the chomskyan theoreticisation is far from being out of the reach of logical investigation.

It is then interesting to see the impact of such logical investigations on the linguistic theory itself, this being the “linguistics-oriented” perspective in this volume, which is represented by works by H. HENDRIKS and D. HEYLEN.

It initially seems very strange that the logical framework is able to provide a very new and adequate tool for studying intonation, a dimension which seems to belong only to the phonological interpretation of a sentence. H. HENDRIKS shows the complexity of the phenomenon, which is not only part of phonology but also of *information packaging*. Traditional approaches to prosody based on the notion of constituent as defined in ordinary phrase structure grammars fail to provide a correct account of the association between intonational marks (like *pitch accent*) and the informational content. The proof-theoretic approach works better because it makes it possible to deal with multi-component signs in such a way that a syntactic analysis, a semantical form and a prosodic one are obtained at the same time. For instance, “intonation and word order are dealt with at one and the same level.” Such results are direct applications of well-known techniques from proof theory, like the Curry-Howard homomorphism.

The notion of *underspecification* has received much interest in computational linguistics for several years. It crucially occurs in the resolution of agreement problems. Johnson and Bayer have already shown that unification-based grammars give a less rich account of agreement than Lambek grammars. D. HEYLEN

goes further on this topic by showing how to more generally solve feature checking problems by entirely logical means. This involves the introduction of dual modalities, box and diamond, with their *residuation logic*. Each mode i represents some morphosyntactic feature, and underspecification is dealt with by assuming general modes that are related to specific instances by inclusion postulates.

Another kind of “linguistics-oriented” paper is provided by T. CORNELL who tries to give formal representations of minimalist grammars (as they have been defined by E. Stabler). After papers on the derivationalist view, CORNELL develops a representational one, keeping in mind the idea that the two views must always coexist simultaneously and that they stand in respect to each other like a deductive system and its proof representation. The derivationalist view could be compared to the (Gentzen) sequent presentation of a calculus and the representationalist one to its proof net syntax. By doing so, it can be shown that the approach in terms of movement and the one in terms of chains are not contradictory but complementary.

Logic is also applied to computation and implementation. Besides the paper by SHAUMYAN and HUDAK already mentioned, which shows the use of the programming language *Haskell* in order to implement applicative grammars, the reader will find a paper by J. HODAS, introducing the use of the language *Lolli*, based on linear logic. As it has been shown by D. Miller, linear logic can be used for programming according to different paradigms, which we can refer to as proof reduction (by means of the Curry-Howard homomorphism and the interpretation it provides in functional terms) and proof search (or logic programming). *Lolli* uses the later possibility and thus allows the implementation of logical grammars similar to DCGs. HODAS uses primitives of linear logic (exponentials and additive conjunction for instance) in a very innovative way in order to have derivations taking unbounded dependencies into account and to block some unwanted derivations. The trick consists in using rules which have nested implications. These implications are used locally, as they license empty nps in case of unbounded dependencies. But an np (for instance an np enclosed in a subject np) can be blocked from extraction (and therefore not realizable by an empty string) simply by marking it with the exponential “!”. We actually know that in this case, the “resource” can be produced only if all the consumed formulae are also marked with the same exponential, but the nested implicative formula allowing an empty np is not marked and therefore the deduction fails.

Questions of implementation are also treated by M. VILARES, M. A. ALONSO and D. CABRERO, who question efficiency of the DCG model and propose logical automata for parsing with DCGs according to an LALR method. Strategies for executing definite clause grammars are still often expressed directly as symbolic manipulations of terms and rules using backtracking, which does not constitute an adequate basis for efficient implementations. The strategy proposed by the

three authors is based on an evolution of the notion of *logical push-down automaton* introduced by B. Lang, “a push-down automaton that stores logical atoms and substitutions on its stack, and uses unification to apply transitions.”

Finally, a fruitful application of constructive type theory to the analysis of mathematical language is suggested by Z. LUO and P. C. CALLAGHAN. This work is related to others on the same topic, like Ranta and Coscoy’s, using different proof development systems (Alf, Coq or Lego). They all aim at using natural language in order to interact with proof systems. Of course, that necessitates a deep exploration of the language of mathematics (here called mathematical vernacular). The novelty of this paper is its treatment of *coercion* and *subtyping*.

On behalf of the Program Committee, I wish to thank the participants for the pleasant and stimulating atmosphere during the meeting. The Committee is especially grateful to the aforementioned invited speakers. I also wish to thank the two institutions that co-organized this event, namely INRIA-Lorraine and CRIN-CNRS, as well as the sponsors of this event: France-Télécom CNET, Xerox Research Center Europe, l’Institut National Polytechnique de Lorraine, l’Université Henri Poincaré, La Région Lorraine and La Communauté Urbaine du Grand Nancy.

More personally, I deeply thank the members of the Program Committee and the Organizing Committee, listed overleaf, for their hard work and support, and let us not forget the research project without which nothing would have happened *Calligramme*.⁴

I particularly thank François Lamarche, who proofread this preface and had to work hard to improve the presentation of some of the papers; Christian Retoré, who provided me with some templates from the first edition (LACL’96); and Guy Perrier, who was a very efficient organizer of this conference.

Grenoble, December 1998

Alain Lecomte (*Chairman*)

⁴ *Calligramme, Logique linéaire, réseaux de démonstration et grammaires catégorielles* (INRIA-Lorraine and CRIN-CNRS): D. Bechet, Ph. de Groote, F. Lamarche, A. Lecomte, J.-Y. Marion, G. Perrier, V. Antoine (secretary) and C. Retoré.

Programme Committee

Bob Carpenter	Bell Laboratories
Marc Dymetman	Rank-Xerox Research Center, Grenoble
Claire Gardent	Universität des Saarlandes, Saarbrücken
Philippe de Groote	INRIA-Lorraine & CRIN-CNRS, Nancy
Seth Kulick	University of Pennsylvania
François Lamarche	INRIA-Lorraine & CRIN-CNRS, Nancy
Alain Lecomte (Chair)	Université Grenoble 2 & INRIA-Lorraine
Michael Moortgat	OTS, Universiteit Utrecht
Glyn V. Morrill	Universitat Politècnica de Catalunya, Barcelona
Aarne Ranta	Universities of Helsinki & Tampere
Patrick Saint-Dizier	IRIT-Toulouse
Edward Stabler	UCLA, Los Angeles
Eric Villemonte de la Clergerie	INRIA-Rocquencourt

Organizing Committee

Valérie Antoine, Denis Bechet, Anne-Lise Charbonnier,
François Lamarche, Guy Perrier (Chair), Armelle Savary
INRIA-Lorraine & CRIN-CNRS, Nancy

Table of Contents

Invited papers

Type grammar revisited

Joachim Lambek 1

Optimal parameters

Denis Bouchard 28

Selected papers

Strong equivalence of generalized Ajduckiewicz and Lambek grammars

Maciej Kandulski 54

Linguistic, philosophical, and pragmatic aspects of type-directed natural language parsing

Sebastian Shaumyan and Paul Hudak 70

Derivational and representational views of minimalist transformational grammar

Thomas Cornell 92

The MSO logic-automaton connection in Linguistics

Frank Morawietz and Thomas Cornell 112

The logic of tune, a proof-theoretic analysis of intonation

Herman Hendriks 132

A linear logic treatment of phrase structure grammars for unbounded dependencies

Joshua Hodas 160

Underspecification in type-logical grammars

Dirk Heylen 180

On fibring feature logics with concatenation logics

Nissim Francez 200

An operational model for parsing definite clause grammars with infinite terms

Manuel Vilares, Miguel A. Alonso, and David Cabrero 212

Mathematical vernacular and conceptual well-formedness in mathematical language

Zhaohui Luo and Paul Callaghan 231

Author Index 251