Lambek Calculus
Multimodal and Polymorphic Extensions

Michael Moortgat
(editor)

DYANA-2

Dynamic Interpretation of Natural Language
ESPRIT Basic Research Project 6852

Deliverable R1.1.B
September 1994

Lambek Calculus
Multimodal and Polymorphic Extensions

Michael Moortgat
Universiteit Utrecht
(editor)

authors:
Michael Moortgat
Mark Hepple
Natasha Kurtonina and Michael Moortgat
Dick Oehrle
Martin Emms

DYANA-2

Dynamic Interpretation of Natural Language
ESPRIT Basic Research Project 6852

Deliverable R1.1.B
September 1994

ESPRIT Basic Research Project 6852 DYANA-2

Dynamic Interpretation of Natural Language

Universiteit van Amsterdam
Institute for Logic, Language and Computation (ILLC)

University of Edinburgh
Centre for Cognitive Science (ECCS)

Universität München
Centrum für Informations- und Sprachforschung (CIS)

Universitetet i Oslo
Department of Linguistics and Philosophy (ILF)

Universität Stuttgart
Institut für Maschinelle Sprachverarbeitung (IMS)

Universität Tübingen
Seminar für Sprachwissenschaft (SfS)

Universiteit Utrecht
Research Institute for Language and Speech (OTS)

For copies of reports, updates on project activities and other DYANA-related information, contact:

The DYANA-2 Project Administrator
ILLC/Department of Philosophy
University of Amsterdam
Nieuwe Doelenstraat 15
NL-1012 CP Amsterdam

Contents

Introduction

This deliverable collects a number of papers relating to the subtasks 'Logics of structured resources' and 'Model-theory for categorial polymorphism'. The papers are technical: they concern the logical foundations of the type-logical grammar architecture. In this introduction we situate the contributions within the broader context of the research efforts of the Grammar Architecture area, and we point to related work on linguistic applications making use of the technical results presented here.

Part 1. Combining logics: substructural communication

Categorial type logics are systems of inference designed for reasoning about structured linguistic resources — signs. The resources exhibit structure in dimensions of linear order, hierarchical grouping and dependency, to name just a few prominent aspects of linguistic structure that play a role in determining well-formedness. One obtains a family of categorial type inference systems by systematically modulating the granularity of discrimination with respect to the various dimensions of linguistic structure.

In the linguistic applications developed within this Area, one characterizes the grammar for (a fragment of) a language in terms of a set of lexical type declarations. In general, it will not be possible to restrict oneself to one particular type logic to carry out this task: the grammar writer will need access to the combined inferential capacities of the different logics. For this to be possible one needs a theory of systematic communication between type systems. The contributions in this part of the deliverable both address this issue of communication.

The paper 'Residuation in mixed Lambek systems' considers two types of mixed categorial logics. The first type is obtained by generalizing residuation to families of n-ary connectives, and by putting together the different arities in one logic. The paper focuses on residuation for *unary* connectives, $\Diamond, \Box^{\downarrow}$, and their combination with the usual binary connectives. Parallel to the treatment of the binary connectives, the paper starts from the pure logic of residuation for the pair $\Diamond, \Box^{\downarrow}$, gradually adding structural postulates for applications where a coarser sense of discrimination is appropriate. This development provides a general logical framework for the study of the various unary operators that have been introduced in the recent categorial literature.

For the $\Diamond, \Box^{\downarrow}$ extension a number of fundamental logical results is established. Completeness with respect to mixed (2,3) Kripke-style relational frame models is obtained by an extenstion of Došen's [?] canonical model construction for the case of simple ternary frames. Gentzen style presentation is given and shown equivalent to the axiomatic presentation used in establishing the completeness result. The Gentzen presentation is shown to enjoy Cut Elimination.

The second type of mixed system studied in this contribution is obtained by combining a number of unimodal systems into one multimodal logic. The combined multimodal logic is set up in such a way that the individual resource management properties of the constituting logics are preserved. But the inferential capacity of the mixed logic is greater than the sum of its component parts through the addition of frame conditions with the corresponding interaction postulates regulating the communication between the component logics.

The paper 'Controlling resource management' puts the mixed framework of inference to the test in a study of modal embeddings. Following the logical tradition, the issue of communication has already been addressed in the categorial literature in terms of embeddings of stronger (less discriminating) logics within weaker ones. Embedding theorems of this type show that it is possible to gain full access to the flexibility of a logic with a more liberal resource management regime from within a system with a more constrained form of resource management. Systematic treatment of embeddings in this direction can be found in Venema's contribution to Deliverable R1.1.A [**?**] and in Kurtonina [**?**]. The paper by Kurtonina and Moortgat is interested in the complementary question, which is at least as relevant from the linguistic point of view: can one systematically recover full *control* over structural discrimination (with respect to the parameters of order, dominance, dependency) from within a coarser logic where such discrimination is lacking? The paper gives a positive answer, thus providing a general logical perspective on constraints in the dimensions of order, dominance and dependency. In his comments on the paper, Dick Oehrle clarifies the proposed embeddings by studying them from the more general perspective of composition of residuated mappings and structure-building operations.

The contributions in this part of the deliverable raise two fundamental questions for future research. One concerns the 'direction' of the flow of information when juxtaposing notions of linguistic composition with different resource management properties. The view proposed in the paper 'Residuation in mixed Lambek systems' is that from the more discriminating notion of composition one can infer the less discriminating one. As pointed out in the comments on this contribution, Mark Hepple in a series of papers has developed a framework of multimodal communication where the flow of information runs in the opposite direction: from the less discriminating to the more discriminating resource management. Oehrle in his contribution to this deliverable discusses these contrastive views, and suggests a strategy of modal embeddings for their resolution. We hope to explore this line of research as part of next year's efforts on Task 1.1.

The second question is suggested by the embedding results discussed in Kurtonina and Moortgat's contribution. The standard embeddings of strong logics within weaker ones, after the model of Girard's embedding of IL in LL, are based on modal decoration in terms of an S4 universal modality $\Box$. In contrast, the embeddings of the paper by Kurtonina and Moortgat make *minimal* resource management assumptions, using just the pure logic of residuation for $\Diamond, \Box^{\downarrow}$. The question then arises whether it would be possible to obtain an embedding strategy for *licensing* structural relaxation also in terms of simply the residuation assumptions for $\Diamond, \Box^{\downarrow}$. A positive answer would provide a more delicate alternative to the brute force S4 modalisation — an alternative which would be completely in line with the general 'minimalistic' program underlying the categorial work in this Area.

Relations with ongoing work

The technical results in these contributions are connected in various ways to descriptive and computational work within Area 1. The paper 'Residuation in mixed Lambek systems' shows that the Hepple-Hendriks uniform head-driven search regime discussed in Hendriks' contribution to Deliverable R1.1.A can be enforced via a translation in terms of $\Diamond, \Box^{\downarrow}$, thus providing a declarative alternative to the procedural sequent annotation of the Hepple-Hendriks normalisation strategy. In a series of papers ([**?**, **?**, **?**]), Moortgat, Morrill and Oehrle use the program of multimodal grammatical inference to present analyses of head-adjunction phenomena and locality constraints. Hendriks' contribution to Deliverable R1.3.B treats the different forms of information packaging in Catalan and English by means of an interpreted (intonationally/syntactically and semantically/informationally) version of the dependency logic of [**?**] enriched with the $\Diamond, \Box^{\downarrow}$ connectives proposed in the present Deliverable.

Part 2. Polymorphic Lambek calculus: completeness results

The second part of the deliverable is concerned with model-theory for $\mathrm{L}^{(/,\backslash,\forall,\exists)}$— Lambek calculus extended with universal and existential category quantifiers. The 2nd Order Lambek calculus defines a particular set of sequents. The question arises as to whether the same set can be picked out by a semantics which defines them as the universally valid sequents. The contribution by Emms studies this question by taking analogs of model classes for the Lambek calculus $\mathrm{L}^{(/,\backslash)}$: ternary frame semantics, residuated semi-groups and language models (string semantics), in order of increasing specialisation.

$\mathrm{L}^{(/,\backslash)}$ is complete with respect to the *string-semantic models*, in which categories are interpreted as sets of strings. A *polymorphic string-semantic model* specifies some set of string-sets as the range of quantification, and various settings for this quantifier range are considered. The simplest condition is that the range should be the set of all string-sets. It is shown, however, that $\mathrm{L}^{(/,\backslash,\forall,\exists)}$ is then *incomplete*. Another possible condition is that the range should be the set of category values (thus allowing some string sets to be excluded) and completeness of $\mathrm{L}^{(/,\backslash,\forall)}$ under this condition is shown. A stronger condition is that the range should be the set of values of *closed* categories. Examples of models meeting this condition are given and it is indicated what would suffice for completeness of $\mathrm{L}^{(/,\backslash,\forall)}$ under this condition. It is also noted that just as the addition of product complicates the proof of string-semantic completeness, so does the addition of $\exists$.

$\mathrm{L}^{(/,\backslash,\bullet)}$ is complete with respect to *residuated semi-groups* [**?**], in which three operations $/$, $\backslash$, and $\bullet$ relate to a partial order via the residuation scheme $b \leq a\backslash c$ iff $a\bullet b \leq c$ iff $a \leq c/b$. In the polymorphic extension, the quantifiers are interpreted by g.l.b.'s and l.u.b.'s, and completeness of $\mathrm{L}^{(/,\backslash,\bullet,\forall,\exists)}$is shown. It is also shown that $\mathrm{L}^{(/,\backslash,\bullet,\forall,\exists)}$ is not a conservative extension of $\mathrm{L}^{(/,\backslash,\forall,\exists)}$, which thereby blocks a simple entailment to completeness for $\mathrm{L}^{(/,\backslash,\forall,\exists)}$.

Finally, the *associative ternary frame* semantics is considered, with respect to which $\mathrm{L}^{(/,\backslash)}$ is complete [**?**]. In the polymorphic version, two different ways are shown in which the accessibility relationship can approximate aspects of concatenation and which lead to incompleteness when quantifiers range over arbitrary sets of worlds.

Relations with ongoing work

Again, this technical contribution should be read in conjunction with related descriptive and computational work that has been produced during the second year of the project. We refer the reader to [**?**, **?**] for a systematic comparison (with respect to linguistic expressivity and recognizing capacity) of polymorphic and operator-based extensions of the basic categorial framework.

References

[1] Michele Abrusci, Claudia Casadio, and Michael Moortgat (eds) (1994), *Linear Logic and Lambek Calculus*. Proceedings First Rome Workshop, OTS/DYANA, Utrecht/Amsterdam.

[2] Wojczech Buszkowski (1986), 'Completeness results for Lambek Syntactic Calculus'. In *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, **32**, pp 13–28.

[3] Paul Dekker and Martin Stokhof (eds) (1994) *Proceedings 9th Amsterdam Colloquium*. ILLC, A'dam.

[4] Došen, K. (1992) 'A brief survey of frames for the Lambek calculus'. Zeitschr. f. math. Logik und Grundlagen d. Mathematik **38**, 179–187.

[5] Martin Emms (1994), 'Movement in polymorphic and labelled calculi'. In [**?**].

[6] Martin Emms (1994), 'Extraction covering extensions of the Lambek calculus are not context-free'. In [**?**], pp 268–286.

[7] Herman Hendriks (1993), 'Lambek semantics'. In Deliverable R1.1.A.

[8] Mark Hepple (1994), 'Labelled deduction and discontinuous constituency'. In [**?**].

[9] Natasha Kurtonina (1994), 'Modal extensions of Lambek calculi'. In [**?**].

[10] Michael Moortgat and Glyn Morrill (1994), 'Heads and phrases. Type calculus for dependency and constituent structure'. In [**?**].

[11] Michael Moortgat and Dick Oehrle (1994), 'Adjacency, dependency and order' In [**?**], pp 447–466.

[12] Glyn Morrill (1994), 'Structural facilitation and structural inhibition'. In [**?**].

[13] Yde Venema (1993), 'Meeting a modality? Restricted permutation for the Lambek calculus'. In Deliverable R1.1.A.

Michael Moortgat
Universiteit Utrecht

Task 1.1, subtask 2

Logics of Structured Resources

Residuation in Mixed Lambek Systems

Michael Moortgat
(Universiteit Utrecht)

Comments on Multimodal Systems

Mark Hepple
(University of Sheffield)

Task 1.1, subtask 2

Logics of Structured Resources

Controlling Resource Management

Natasha Kurtonina and Michael Moortgat
(Universiteit Utrecht)

Comments on Kurtonina and Moortgat

Modal Embeddings, Residuation, and Composition

Dick Oehrle
(University of Arizona)

Task 1.1, subtask 3

Model-theory for Categorial Polymorphism

Completeness Results for Polymorphic Lambek Calculus

Martin Emms
(Universität München)