

MATHEMATISCHES FORSCHUNGSINSTITUT OBERWOLFACH

T a g u n g s b e r i c h t 18/1987

Mathematical Methods in the Study of Natural and Computer Languages

(Mathematische Aspekte natürlicher und Computersprachen)

19.4. bis 25.4.1987

The conference on mathematics in the study of language was organized by Professors Barwise from Stanford, Fenstad (Oslo), Kamp (Austin) and Richter (Kaiserslautern) and brought together experts not only from a wide geographical area, but also of various research areas: mathematics, logic, philosophy, computer science, artificial intelligence and linguistics. This resulted in a lively scientific atmosphere and in many stimulating discussions about the lectures, between the talks and often all night long.

The chronological course of the meeting was beneficially planned and developed from the presentation of most recent works about theoretical foundations of the field into lectures about the application of these methods in solving concrete problems in natural language processing. Each day in turn began with a 'grammatical' lecture (they were also highly frequented by the participants of the logicians' conference, which was simultaneously held at Oberwolfach) and continued with three special talks. E.g., actual questions of theorem proving and problems around generalized quantifiers were one-day topics.

Main emphasis (in time and interest) was given to actual work on developing a mathematically well-founded framework for interpreting natural language sentences as constituents of larger discourses. The conference benefitted exceptionally from the fact that the internationally leading researchers of this field were among the organizers and lecturers.

All participants had the feeling that this conference showed that (computer) linguists can find many methods and results of existing mathematics that are useful for their work, and, on the other hand, that the questions that come up in natural language processing are suitable for mathematical investigations and of interest to the working mathematician. This meeting at Oberwolfach contributed to further progress in the study of language, and would be desirable to have a successor at the same lovely place that should focus on one of the topics of this important and fascinating research area.

Abstracts (in chronological order)

1. Johan van Benthem (Amsterdam):

Applications of mathematical logic in linguistics

1 The Question: Is there a significant use of mathematics in the study of natural language?
(Programming languages: only in asides)

2 Claim: There exist already enough connections to support a positive answer. The evidence is in the following list of examples:

- Syntax:
- 1) Formal Languages and Automata
 - 2) Categorical Grammar and Logical Proof Theory
 - 3) Fine-structure of Logical Syntax

- Semantics:
- 1) General Frameworks: Universal Algebra of Montague Grammar
Situation Theory
 - 2) General Features of Natural Language: Typefree modelling, Finite Model Theory, Logicality and Permutation Invariance (Connection with Lambda Calculus and Type Theory)
 - 3) Special Topics: e.g., determiners, temporal expressions

- Inference:
- 1) Natural Logic: large decidable fragments
 - 2) Non-standard Consequence: non-monotonicity
 - 3) Inference and representation: e.g., AI-discussion about computationally preferable representation of time, more generally: development of Qualitative Physics.

3 So, existing logic / mathematics is quite useful here - though often with a new twist: fine-structure of standard theories / alternative set-ups. But, also new mathematics will have to be developed (as in all important applications), notably in the study of larger linguistic structures (texts, discourse), as well as dynamic aspects of semantic processing.

2. Yiannis N. Moschovakis (Los Angeles):

Formal Properties of Parallel and Sequential Algorithms

The first half of the talk presented in outline the basic ideas of a theory of (parallel, no-side-effects, no-dependence on the state) pure algorithms. One studies certain set-theoretic objects called recursors which model pure algorithms on structures of a certain type; among all recursors, the algorithms of a structure are those definable in the Formal Language of Recursion, FLR. The technical part of the work comes down to the study of the syntax and the semantics of FLR. In this talk, only the first basic result about FLR was discussed, i.e. the axiomatizability of the binary relation on terms

(*) $s \sim t \Leftrightarrow s$ and t define the same algorithm on all structures.

In the second half of the talk the more general communicating structures and (both parallel and sequential, with side effects and possible dependence on the state) algorithms were introduced, and an alternative model of FLR was discussed, on these objects. The axioms of the reinterpretation of (*) still holds, so a new completeness theorem is obtained, for the new modelling.

The validity of the same axiomatizations for both models means that the formal properties of these notions of algorithms are the same.

3. Robert S. Boyer (Austin):

Quantification in Automatic Theorem Proving

We described a theorem-proving program which has been used to check such results as quadratic reciprocity and Gödel's incompleteness theorem. We asked for help in finding or creating a logic with

- (a) the power of set theory,
- (b) the conciseness of notation found in quantificational expressions such as $\forall x P(x)$, $\{x:p(x)\}$, and $\lambda x f(x)$,
- (c) no bound variables, and
- (d) a tractability that admits pleasant hand proofs.

The apparent contradiction between (b) and (c) may be resolved by consideration of the support that the von Neumann-Bernays-Gödel set theory provides for the $\{x;q(x)\}$ notation with a finite

axiomatization; however, this avenue violates condition (d). The technique described in "On Adding Bounded Quantification and non-Terminating Functions to A Computational Logic" (with J. Moore, Insitute for Computing Science, University of Texas) has been used to achieve (b), (c), and (d), but not (a).

4. Ulf R. Schmerl (München):

Resolution on Formula-Trees

We introduced a nonclausal resolution calculus on formula-trees which comprises classical resolution as a special case. The resolvents produced in this calculus are more structure preserving than in nonclausal resolution by Murray and Manna and Waldinger and simpler than in nested resolution by Traugott. Proofs of correctness and completeness were given. In some examples, first experiences made when implementing the calculus were discussed.

5. Jon Barwise (Stanford):

Semantics and Paradox

In this talk I discussed the problems semantical paradoxes pose for any mathematical framework for doing semantics of natural language. I illustrated one approach to the paradoxes by using Aczel's theory ZFC/AFA. I presented a simple formal language in which the paradoxes ca be expressed and then discussed two alternative semantics for it. In one semantics, the Russelian, some sentences cannot have thruth values. In the other semantics, the Austinian, sentences always express propositions with a determinate truth value. I then discussed a theorem which gives an Austinian characterization of those sentences that are paradoxical on the Russelian semantics. The talk was based on joint work with John Etchemendy.

6. *David Israel* (Stanford):

Background to Situation Theory

Motivation is given for developing a many-sorted first order theory of information-content and of the structures involved in informational relations.

The theory involves the postulation of n-ary relations as primitives and of propositions as structural complexes of a certain kind. A crucial requirement is that to the axioms of the theory . There should correspond objects in the universe of the theory.

7. Stanley Peters (Stanford):

Generalized Quantifiers and Anaphora

Different kinds of noun phrases in natural languages express different sorts of semantic content. For example, the following English noun phrases have the syntactic and the semantic types shown.

Noun Phrase	Syntactic variety	Content sort
John	proper name	individual
she	pronoun	individual
the lecturer	definite description	individual
a logician	indefinite description	individual
no linguist	quantified phrase	generalized quantifier
most tables	quantified phrase	generalized quantifier
two theorems	numeral phrase	generalized quantifier

The content of any phrase varies systematically with the circumstances in which the phrase is used; it is determined jointly by the meaning of the phrase and the context of utterance.

A generalized quantifier can be regarded, in extension, as a set of subsets of the domain of quantification. For instance, the content of "most tables" is

$$\{ A \subseteq D \mid R(T, A) \}$$

in a context where the determiner "most" expresses the relation R between sets, the noun "tables" corresponds to the set T , and D is the domain of quantification. (R might be the relation such that $R(X, Y)$ iff there is no function from $X \setminus Y$ onto $X \cap Y$.)

A pronoun can function either deictically (indexically) or anaphorically, depending on context. Deictic uses are, in effect, parameters whose values are fixed by context. Anaphoric uses of pronouns subdivide into two kinds.

(1) A co-parametric use of "his" in "John saw his watch." gives the sentence the content that

$$j \in \{x \mid x \text{ saw } j\text{'s watch}\}.$$

(2) A role-linking use of "his" gives the same sentence the content that

$$j \in \{x \mid x \text{ saw } x\text{'s watch}\}.$$

The difference between co-parametric and role-linking uses becomes clear in considering sentences like

"John saw his watch and so did Bill."

(1) $j \in \{x \mid x \text{ saw } j\text{'s watch}\}$ & $b \in \{x \mid x \text{ saw } j\text{'s watch}\}$.

(2) $j \in \{x \mid x \text{ saw } x\text{'s watch}\}$ & $b \in \{x \mid x \text{ saw } x\text{'s watch}\}$.

The different properties are ascribed to Bill.

An anaphorically used pronoun with a quantifying noun phrase such as "no logician" as its

antecedent can only serve for role-linking. E.g., "No logician saw his watch." can express $\{x \mid x \text{ saw } x\text{'s watch}\} \in \{A \subseteq D \mid \text{logician} \cap A = \emptyset\}$.

Accordingly, "No logician saw his watch, but John did." cannot express that John saw every logician's watch but no logician saw his own.

8. H. Volger (Passau):

Syntactical Characterization of Theories which admit Initial Structures

The question "Why universal Horn formulas matter in Computer Science?" (cf. Makowsky, LNCS 115) has a model theoretic and a proof theoretic answer. The model theoretic answer characterizes universal Horn theories as theories T which admit *uniformly* (i.e. for all consistent extensions by new facts (=insertions)) term models M which are generic for T (i.e. every element of M is denoted by a term and any fact is satisfiable in M iff its existential closure is derivable from T (=Closed World Assumption)). Note that a term model M is generic iff it is initial for T (i.e. there is a unique homomorphism into any other model of T). Moreover, the existence of a generic model for T is equivalent to an irreducibility property of the theory.

We obtain the known result for universal Horn theories (cf. Malcev, Algebraic Systems) from the more general characterization of pseudo-universal Horn theories (=limit theories in Volger, Math. Zeitschr. 166 (1979)) as theories T which *uniformly* admit initial structures. Note that a structure is initial iff it is a pseudo term structure which is generic for T . The missing link in Makowsky is the closure under equalizers of homomorphisms. This result helps to understand Malcev's result and it shows that up to a definitional extension by partial operations nothing can be gained using initial structures rather than term structures which are initial.

A related result characterizes the generic Horn theories which are axiomatized by formulas of the form $\forall x' (\alpha(x') \rightarrow \exists y' \beta(x', y'))$ with $\alpha, \beta \in \bigwedge \text{At}$. - Adding some restrictions we obtain a syntactical characterization which yields Prolog-Programs i.e. universal Horn theories which are strict and non-identifying. In this context Herbrand structures, where each element is denoted by a unique term, are used instead of term structures.

9. Jens Eric Fenstad (Oslo):

Natural Language Systems and Computational Semantics

A system for natural language analysis shall provide a framework for relating the linguistic form of utterances and their semantic interpretation. This calls for an extension of computational linguistics with its traditional emphasis of syntax and morphology to include a theory of computational semantics.

Basic to the approach which we presented in this talk is an algorithm for converting linguistic form to a format which we call a situation schema. The algorithm is in the spirit of current

unification-based approaches to grammar and exploit the idea of constraint propagation.

A situation schema has a well-defined (algebraic) structure, suggestive of "logical form"; but it is a structure different from the standard model-theoretic one. We argue that it is a structure better adapted for the analysis of the meaning relation in natural languages and that it provides a format useful for further processing.

In the first part of the paper we provided the necessary background from the model-theory of partial information (situation semantics); in particular, we reported briefly on some mathematical investigations into situational logic.

10. Hans Leiß (München):

Noun phrases and quantified terms

A formal language was defined whose syntax is very close to a fragment of German, covering noun phrases with determiners, restrictive relative clauses and locative prepositional phrases acting as modifiers on quantified nouns and verbs.

It was shown how to translate terms and formulas of this language into a typed first-order logic enriched by binary second order relations (i.e. generalized quantifiers) and predicate operators. Essentially, terms translate into quantifier blocks and prepositional phrases into bounds for quantifiers or into quantifier blocks plus predicate operators. It was sketched how to refine this translation when anaphoric expressions such as pronouns are included into the language.

Concerning semantics, it was strongly suggested to define a notion of 'coherent (or: formally understandable) text' as a refinement of 'consistent theory' in logic. In particular, it was proposed to modify the technical notion of 'consistency property' (which has been derived from Henkin's model construction of consistent theories) to deal with ambiguities, anaphoric expressions and the structure of texts. Recall that, roughly, consistency properties define legal construction steps in the process of building a tree of partial models.

It was sketched briefly how the above mentioned translation might enter into a modification of this construction which pays attention to at least the linear ordering of sentences (as opposed to sets of sentences in logic) and anaphora.

Essential aspects of text understanding and semantics were claimed to be represented in this process of constructing a tree of partial models rather than in the limit model itself.

A development of this proposal might lead to an abstract analogue of Hans Kamp's discourse representation theory for formal languages whose syntactic structure covers some aspects of natural language syntax.

11. Hans Kamp (Austin):

Discourse Representation Theory

The way we interpret natural language sentences that are constituents of larger coherent discourses or texts relies heavily on the information we have already extracted from the earlier part of the text or discourse. An adequate theory of natural language meaning must give an account of this contextual aspect of interpretation. Moreover, it must show how interpreting the sentence leads simultaneously to an incrementation of content and to a modified context with respect to which the next sentence should be interpreted.

Discourse Representation Theory gives a systematic description of this interpretation process. Central to the theory are so-called Discourse Representation Structures or DRS's. A DRS acts simultaneously as a representation of the joint content of the sentences that have been interpreted already and as the context for the sentence that comes next. (The content of a DRS can be characterized along familiar model-theoretic lines.) When a sentence is processed relative to a given DRS this will result in a new DRS which incorporates the contribution which the sentence makes to the content represented by the old DRS. Thus, abstractly, the process of sentence interpretation is a function from DRS's to DRS's. Sentence meaning should no longer be thought of in terms of the proposition a sentence expresses, but rather as the capacity of the sentence to modify a given DRS into a new one, typically one with stronger truth conditions (cf. the "file change potential" of I. Heim).

Discourse Representation Theory seems especially well equipped to handle intersentential connections such as pronominal or temporal anaphora. The talk presented a number of sample discourses and showed how the processing algorithm converts into DRS's with intuitively correct truth conditions.

12. Helle Frisak Sem (Oslo):

Correspondences between DRT, Situation Schema Theory and Situation Semantics

Over the past 10 years several new grammatical and semantical theories in the study of natural language have emerged: the Discourse Representation Theory (DRT) developed by Hans Kamp [1], the Situation Semantics by Jon Barwise and John Perry [2], and the Situation Schema Theory by Jens Erik Fenstad et al. [3].

In the first part of the lecture we discussed the relationship between DRS (Discourse Representation Structures) and Situation Schemata. Given a lexicon and a set of phrase structure rules we can for each sentence of the language construct an associated DRS which consists of a structured set of conditions, atomic conditions ($\alpha = u$, $\alpha(u)$, $\alpha(u,v)$) and complex conditions ($m_1 \rightarrow m_2$). The DRS represents a unique reading of the sentence with respect to scope order and coreference. From the corresponding situation schema of φ we can extract exactly the same atomic fact schemata, such that given a suitable Q-mode (quantifier scope reading) and coreferential conditions we get the same first order transcription of the DRS (φ) and SIT. φ .

In the second part of the talk we discussed a Situation Semantic interpretation of a DRS. In

Situation Semantics the meaning of a sentence φ is a conventional constraint between the utterance situation and the situation described:

$\llbracket \varphi \rrbracket$: involves, DC'_{φ} , S_{φ} ; 1.

From a DRS we can systematically construct the event-type of S_{φ} , the described situation, giving an interpretation of the DRS into situation semantics.

- [1] Kamp, H. (1981): "A Theory of Truth and Semantic Representation", in: Groenendijk, J. et al. (eds): Formal Methods in the Study of Language. Amsterdam.
- [2] Barwise, J., Perry, J. (1983): Situations and Attitudes. The MIT Press.
- [3] Fenstad, J.E., Halvorsen, P.K., Langholm, T., van Benthem, J. (1987): Situations, language, and Logic. Reidel.

13. Barbara H. Partee (Amherst):

Shifting Types in Natural Language Semantics

Montague's strong form of compositionality requires a homomorphism from the syntactic algebra to the semantic algebra (Montague, "Universal Grammar", 1970), including the assignment of a unique semantic type to each syntactic category. I argue from a combination of formal and empirical considerations that natural languages like English are better described with a framework which allows the assignment of a family of types to each syntactic category. One set of relevant facts includes transitive-verb conjunctions like "need and want" (intensional), "kick and hit" (extensional), "needed and bought" (mixed); their semantic behaviour is inconsistent with Montague's uniform type assignment, given the persuasiveness of what seems to be the best treatment of cross-categorial conjunctions. I argue for giving each lexical verb a basic lexical type plus introducing shifting rules to produce homonyms of higher types; together with an interpretation strategy of "try simplest types first", this gives the desired results.

Another set of phenomena involves the multiple interpretations of noun phrases, which may be of those types: entity, generalized quantifier, property. I discuss the problem of identifying "natural" type-shifting operations and describe several candidates for such operations. This also leads to a new perspective on the meanings of "be", "a", and "the". A call for further study, both formal and empirical, ends the presentation.

References:

- Partee (1987): "Noun Phrase Interpretation and Type-shifting Principles". in: Groenendijk et al. (eds.), Proc. of 5th Amsterdam Colloq. Foris Pub.
- Partee & Rooth (1983): "Generalized Conjunction and Type Ambiguity", in: Bäuerle et al (eds): Proc. of 1980 Konstanz Conf., de Gruyter.

14. Wolfgang Schönfeld (Heidelberg):

The outcome of LEX

LEX (Linguistic and logic based legal expert system) is an ongoing and nearly finished project at IBM Heidelberg Scientific Center. Its purpose is to develop an expert system which helps the lawyer to deal with cases of §142 StGB (hit-and-run at the scene of an accident). Moreover, it was designed to have a natural language front-end. We report on problems and successes in that project. Though we made great progress in some areas (lexicon, grammar, resolving of referents, inference, queries during inference), we experienced that the task as whole cannot yet be carried out. The reasons are:

1. There is no overall theory.
2. Even parts of it are underdeveloped:
Common sense knowledge, generation of logical form, dialog principles.

15. Michael M. Richter (Kaiserslautern):

Problems on the borderline between natural languages, computer languages and logic arising from expert systems

In expert systems the style of programming is mainly declarative which rises the question of knowledge representation. We distinguish roughly three levels: The cognitive level (connected with natural language), the logical level (knowledge representation language), and the computational level (programming languages). Some of the problems arising from the combination of these levels are discussed. On the computer language level we sketched difficulties arising from attempts to amalgamate logical, functional and object oriented languages together with (polymorphic) types. On other levels we presented some problems in hypothetical and default reasoning, the logic of question and answer. As an example of a very specific problem Allan's time interval logic was viewed from the viewpoint of reducing its complexity.

16. Wolfgang Wahlster (Saarbrücken):

VITRA: Discourse Domains and Conversational Setting

The aim of the project VITRA (VISual TRANslator) is the development of a computational theory of the relation between natural language and vision. In this talk, we will focus on the semantics of path prepositions (like 'along' or 'past') and their use for the description of trajectories of moving objects, the intrinsic and deictic use of spatial prepositions and the use of linguistic hedges to express various degrees of applicability of spatial relations.

First, we describe the implementation of the system CITYTOUR, a German question-answering system that simulates aspects of a fictitious sightseeing tour through a city. Then we show how the system was interfaced to an image sequence analysis system. From the top of a 35m high building, a stationary TV camera recorded an image sequence of a street crossing on videotape. In 130 selected frames the moving objects were automatically recognized by analyzing displacement vector

fields. Our system then answers natural language queries about the recognized events.

Finally, we discuss current extensions to the system for the generation of a report on a soccer game, which the system is watching. Here we focus on the problem of incremental, real-time text generation and the use of a re-representation component which models the assumed imagination of the listener.

17. Aravind K. Joshi (Philadelphia):

Locality and Linguistic Structure

Different grammatical formalisms are characterized by different domains of locality. For example, each rewriting rule in a context-free grammar (CFG) constitutes a domain of locality for CFG. Head Grammars (HG), Tree Adjoining Grammars (TAG), Categorical Grammars (CG), Indexed Grammars (IG), etc. all have different domains of locality. The particular domain of locality for a given grammatical formalism has implications for the specification of constituency, constraints (e.g. agreement), function-argument-relationships, word-order variation, and characterization of unification constraints. The elementary trees of a TAG provide a larger domain of locality as compared to CFG. This particular domain of locality enables one to localize all the so-called long-distance dependencies in natural languages. The long-distance nature of these dependencies then comes out as a byproduct of the operation of composition, called adjunction. This locality constrains the flow of information with respect to feature compatibility checking in unification. I described how TAGs can be embedded in the unification framework. This embedding results in a constrained unification based framework. The precise nature of the complexity results due to these constraints is still being worked. An exact semantics for this formalism can be given via a recursive transition network, as contrasted with a finite state automaton which has been used for specifying the semantics for feature structures by Ait-Kaci and Rounds and Kasper.

I also described results which show how a variety of systems such as HG, IG (restricted), and others are equivalent to TAG.

Berichterstatterin: C. Reddig-Siekmann

Tagungsteilnehmer

Prof.Dr. K. J. Barwise
Centre for Study Language and
Information
Stanford University
Ventura Hall

Stanford , CA 94305
USA

G. Görz
Regionales Rechenzentrum
Universität Erlangen-Nürnberg
Martensstraße 1

8520 Erlangen

Prof.Dr. J.F.A.K. van Benthem
Mathematisch Instituut
Universiteit van Amsterdam
Roetersstraat 15

NL-1018 WB Amsterdam

Prof.Dr. W. Hoepfner
Fachbereich Informatik
Universität Hamburg
Postfach 30 27 62

2000 Hamburg 36

Prof.Dr. R. Boyer
AI Program
MCC
9430 Research Blvd.

Austin , TX 78759
USA

Dr. D. J. Israel
SRI International
333 Ravenswood Avenue

Menlo Park , CA 94305
USA

E. Colban
Institute of Mathematics
University of Oslo
P. O. Box 1053 - Blindern

N-0316 Oslo 3

Prof.Dr. A. K. Joshi
Dept. of Computer and Information
Science
University of Pennsylvania
200 South 33rd Street

Philadelphia , PA 19104-6389
USA

Prof.Dr. J. E. Fenstad
Institute of Mathematics
University of Oslo
P. O. Box 1053 - Blindern

N-0316 Oslo 3

Prof.Dr. H. Kamp
Department of Philosophy
University of Texas at Austin

Austin , TX 78712
USA

Dr. H. Leiß
SIEMENS AG
KE ST 4
Rosenheimerstr. 116 b
8000 München 80

C. Reddig-Siekmann
SFB 314, FB 10 - Informatik IV
Universität des Saarlandes
6600 Saarbrücken 11

Prof.Dr. Y. N. Moschovakis
Dept. of Mathematics
University of California
Los Angeles , CA 90024
USA

Prof.Dr. M.M. Richter
Fachbereich Informatik
der Universität Kaiserslautern
Postfach 3049
6750 Kaiserslautern

Prof.Dr. A. Oberschelp
Fach Logik und Wissenschaftslehre
am Philosophischen Seminar
der Universität Kiel
Ohlshausenstr.40-60, Haus S12a
2300 Kiel 1

Prof.Dr. C. Rohrer
Linguistik
Fakultät 11
Universität Stuttgart
7000 Stuttgart 1

Prof.Dr. B. H. Partee
Department of Linguistics
South College
University of Massachusetts
Amherst , MA 01003
USA

Prof.Dr. E. Schinzel
Lehrgebiet Theoretische Informatik
der RWTH Aachen
Büchel 29-31
5100 Aachen

Professor Dr. S. Peters
Centre for Study Language and
Information
Stanford University
Ventura Hall
Stanford , CA 94305
USA

Dr. U. R. Schmerl
Mathematisches Institut
der Universität München
Theresienstr. 39
8000 München 2

Dr. W. Schönfeld
IBM Deutschland GmbH
Wissensch. Zentrum Heidelberg
Tiergartenstraße 15

6900 Heidelberg

Prof.Dr. H. F. Sem
Institute of Mathematics
University of Oslo
P. O. Box 1053 - Blindern

N-0316 Oslo 3

Prof.Dr. J. C. Shepherdson
Dept. of Mathematics
University of Bristol
University Walk

GB- Bristol , BS8 1TW

Prof. Dr. H. Volger
FB Informatik und Mathematik
Universität Passau
Graf-Salm-Str. 7, PF 2540

8390 Passau

Prof.Dr. W. Wahlster
Fachbereich 10 - Informatik
Universität des Saarlandes
Bau 36

6600 Saarbrücken