Application for a DFG Research Grant

Aλonzo: Higher-Order Reasoning Agents and Semantical Mediation of Mathematical Knowledge

Christoph Benzmüller

21. April 2004

1 General Information (Allgemeine Angaben)

1.1 Applicant (Antragsteller)

- Christoph E. Benzmüller, Dr.
- Assistant Professor (Hochschuldozent, C2)
- Date of Birth: 8. September 1968
  Nationality: German
- Other DFG funding:
  Collaborative Research Centre SFB 378, Project DIALOG and Project OMEGA
- Institution: Saarland University
  Department: Computer Science Department
- Postal Address:
  Saarland University
  Computer Science Department
  Postfach 151150
  66041 Saarbrücken
- Phone: 0681/302 4574
- Fax: 0681/302 5076
- E-mail: `chris@ags.uni-sb.de`
- Private Address:
  Landwehrplatz 6
  66111 Saarbrücken

- The curriculum vitae of the applicant is enclosed; see the appendix.

1.2 Topic (Thema)

Higher-Order Reasoning Agents and Semantical Mediation of Mathematical Knowledge[1]

1.3 Code name (reference) (Kennwort)

AλONZO

[1] This research proposal is written in English because of the proposed international collaborations, in particular, with the TPS project of Prof. Peter Andrews at Carnegie Mellon University. The applicant (who recently got promoted to C2 Professorship level at Saarland University) wants to create an own research group together with Chad Brown (still at CMU), who is currently the main developer of the TPS system.

1.4 Scientific discipline and field of work (Fachgebiet und Arbeitsrichtung)

Artificial Intelligence, Mathematical Knowledge Mediation, Semantic Web, Application-oriented Development of Higher-Order Theorem Proving

1.5 Scheduled total duration (Voraussichtliche Gesamtdauer)

4 years

1.6 Application period (Antragszeitraum)

2 years

1.7 First proposal (Bei Neuanträgen)

The proposed start of the project is 1. August 2004.

1.8 Summary (Zusammenfassung)

The long-term goal of my research is the (internationally joint) development of a mathematical assistance environment and its integration into the emerging mathematical semantic net[2]. Interactive proof systems are today routinely employed in industrial applications for safety and security verification and more recently they find applications in e-learning systems for mathematics.

The vision of a powerful mathematical assistance environment which provides computer-based support for most tasks of a mathematician has stimulated new projects and international research networks across the disciplinary and systems boundaries. Examples are the European CALCULEMUS[3] (Integration of Symbolic Reasoning and Symbolic Computation) and MKM[4] (Mathematical Knowledge Management, [BGH03]) initiatives, the EU projects MONET[5], OPENMATH and MOWGLI[6], and the American QPQ[7] repository of deductive software tools. Furthermore there are now numerous national projects in the US and Europe, which cover partial aspects of this vision, such as knowledge representation, deductive system support, user interfaces, mathematical publishing tools, etc.

Agent-based architectures and the semantic web provide means to link the tools (such as mathematical knowledge bases, theorem provers, decision procedures, computer algebra systems, etc.) developed, for instance, by the partners of the above research initiatives and they help to join resources.

[2] www.win.tue.nl/dw/monet/ and monet.nag.co.uk/mkm/
[3] www.calculemus.org
[4] monet.nag.co.uk/mkm/index.html
[5] monet.nag.co.uk/cocoon/monet/index.html
[6] www.mowgli.cs.unibo.it/
[7] www.qpq.org

The immediate goal of the Aλonzo project is (I) to provide an agent-based framework for context-sensitive higher-order theorem retrieval from mathematical knowledge bases (MKBs), e.g. in the mathematical semantic web, (II) to analyse and address the particular challenges imposed by this application for higher-order theorem proving systems, and (III) to exemplarily integrate and employ the agent-based knowledge retrieval tools within mathematical assistance environments currently developed in the MKM and CALCULEMUS research initiatives.

The applicant is the coordinator of the EU Research Training Network CALCULEMUS, of the successor proposal CALCULEMUS-II, and of the Saarland University node in MKM.

2 State of the art, preliminary work (Stand der Forschung, eigene Vorarbeiten)

2.1 State of the art (Stand der Forschung)

Mathematics and its foundational field mathematical logic have seen several phases of lasting scientific development and change in its long history. In the 17th century, for example, the development of mechanical calculators (e.g. Schickard 1623, Pascal 1642, Leibniz 1671) ultimately changed practice and societal perception of computations with numbers, which was still considered an “art” at the time of the romans. In the late 19th century and the early 20th century the foundations for the mechanisation of symbolic reasoning were laid (e.g. Frege 1882, Hilbert, Gödel 1930, Herbrand 1930, Gentzen 1933). While the last century provided the foundations and techniques for the mechanisation of symbolic reasoning, the vision of my research for the new millennium is to foster its all-embracing integration into mathematics research and education. To achieve this goal a new generation of mathematical software systems are currently under development — e.g. in the CALCULEMUS [8] (Integration of Symbolic Reasoning and Symbolic Computation) and MKM[9] (Mathematical Knowledge Management) initiatives — that provide integrated computer-based support for most working tasks of a mathematician — including computation and reasoning and mathematical knowledge representation — as well as for formal methods in computer science.

A large amount of mathematical knowledge is currently being formalised and made available on the Internet. Thus, semantical mediation of mathematical knowledge between consumers (e.g. mathematicians or mathematical assistance environments) and providers (e.g. mathematical knowledge bases) will become a central issue[10]. I propose to join the historic transition from pen and paper mathematics to modern, computer-supported mathematical knowledge management from its very beginning. The focus of the Aλonzo project therefore is on agent-based and semantic-web-based solutions for open and distributed mathematical assistance environments. A special interest will be on semantical mediation of mathematical knowledge with higher-order proof tools. Ideally the proposed mediators fruitfully cooperate with state-of-the-art search engines such as Google and other emerging mediation approaches exploiting meta-data and taxonomies.

The envisioned new generation of distributed and dynamically composed mathematical assistance environments will have a strong impact in many theoretical fields such as safety and security verification of computer software and hardware, theoretical physics and chemistry and other theoretical subjects. This new generation of mathematical assistance systems (MA) particularly contrasts the current situation characterized by partial and often non-interoperable solutions. Except for computer algebra systems these partial solutions have not yet reached sufficient acceptance and usage in mathematical practice.

[8] `www.calculemus.org`

[9] `monet.nag.co.uk/mkm/index.html`

[10] See `www.ima.umn.edu/complex/spring/searching.html`: This workshop proposal particularly illustrates the interest in improving search in MKBs by using meta-data, taxonomies, etc.

The situation in the unfortunately fragmented deduction systems area is similar to that of the AI field as a whole and criticised inter alia by Nils Nilsson (Kumagai Professor at Stanford, USA) in his speech at IJCAI 2003 where he received the IJCAI Research Excellence Award. Today many of the deduction system subareas even have separate conferences. As a consequence the ambitious goal of an integrated MA was very weakly represented at these conferences and in the deduction systems community until the end of the 90s. It is only very recently that this trend is reversed, with the CALCULEMUS and MKM communities as driving forces of this movement.

To foster the impact of our research it is necessary to educate the next generation of mathematicians and engineers with our tools and our mathematical assistance environments. An important issue therefore is to connect our assistance tools with the emerging MKBs, to integrate them with the emerging computer-based e-learning systems and to employ them in mathematics and engineering education. Computer-based e-learning systems are increasingly used in university and high school education and combining them in the domain of mathematics with our mathematical assistance systems is thus a consequent step.

It is my personal conviction that agent-based and semantic-web-based frameworks will play a key role in the build-up of this next generation mathematical assistance environments. Furthermore, in contrast to many other domains currently investigated in the semantic web context, mathematics is a well developed and well structured scientific discipline; formal representation languages with precise semantics and rich ontologies are already available. Not least for this reason, mathematics qualifies as an interesting test domain for semantic web solutions.

Higher-Order Logic Mathematical textbooks naturally employ higher-order logic (HOL) constructs and thus the currently fast evolving repositories of formalised mathematics also provide a high amount of mathematical knowledge encoded in HOL (either classical HOL or constructive type theory). Examples are the libraries of the MAs NuPrl[11], COQ[12], HOL[13], PVS[14], MIZAR[15], ISABELLE[16], and THEOREMA[17]. The knowledge of some of these libraries is currently translated into unified representations and merged in European projects such as MBASE[18] and HELM[19] or the very recent American LOGOSPHERE project[20]. Prof. Robert Constable (Cornell University, USA) in his invited lecture at the MKM workshop[21] in November 2003 at Heriot Watt University estimated that the LOGOSPHERE MKB contains approx. 70% HOL entries.

Unfortunately, the direct automation of HOL reasoning has widely been neglected in the last decades due to its assumed complexity and the confusion resulting from Gödels incompleteness

[11] www.cs.cornell.edu/Info/Projects/NuPrl
[12] coq.inria.fr
[13] hol.sourceforge.net
[14] pvs.csl.sri.com
[15] www.mizar.org
[16] www.cl.cam.ac.uk/Research/HVG/Isabelle
[17] www.theorema.org
[18] www.mathweb.org/mbase
[19] helm.cs.unibo.it
[20] www.logosphere.org
[21] www.macs.hw.ac.uk/˜fairouz/mkm-symposium03

results. Therefore, the deduction systems community today mainly concentrates on automation of first-order reasoning and thereby neglects the fact that first-order reasoning within, e.g., Zermelo-Fränkel set theory (as an alternative representation language for mathematics) can be considered as equally challenging and unsolved. Due to the current renaissance of HOL in many research areas it is *now* the time to further strengthen research on limited, application-oriented HOL reasoning and to adapt, if possible, successful techniques from first-order theorem proving to higher-order theorem proving. First-order theorem proving has highly benefitted from its comparably high research funding in the last decades and it has developed very efficient techniques that are not yet available in HOL theorem proving.

My own contributions as sketched in Section 2.2(D) and the recent work of Chad Brown [Bro02] w.r.t. the set variable problem are important milestones for the automation of HOL.[22]

Research on interoperability and coordination of different reasoning tools as proposed in Aλonzo is a core issue for the recent built-up of a new generation of MAs: rather than competing against first-order theorem provers our HOL proof agents shall cooperate with them within mathematical assistance environments (MAs).

Timeliness The 2503rd Council Meeting *Education, Youth and Culture* in Brussels, 5 and 6 May 2003, 8430/03 (Presse 114) states that: "In the area of mathematics, science and technology the European Union needs an adequate output of scientific specialists in order to become the most dynamic and competitive knowledge-based economy in the world. The need for more scientific specialists is underlined by the conclusion of the Barcelona European Council (2002) that overall spending on R&D and innovation in the union should be increased with the aim of approaching 3% of GDP by 2010".

The Aλonzo project directly addresses this recent objective of the European Union: Aλonzo proposes to develop interoperable knowledge-based reasoning techniques with applications in mathematics, engineering and e-learning, and it will contribute to the emerging Mathematical Semantic Web.

Aλonzo furthermore contributes directly and indirectly (via its involvement in the European research networks CALCULEMUS and MKM) to the education of students in exactly the connection between mathematics, computer science and knowledge-based technologies as envisioned by the European Commission.

2.2 Preliminary work, Progress report (Eigene Vorarbeiten)

(A) Mathematical Knowledge Management The mathematical knowledge management research initiative has the goal of revolutionising computer-based mathematics in the new millennium by a top-down approach starting from existing, mainly pen and paper based mathematical practice down to system support. At the core of MKM research is the transition from current forms of managing, maintaining, and accessing mathematical knowledge through its digitalisation and formalisation to fully computer-supported solutions.

[22] The anonymous JSL-referee of [J06-04] says: "This is a very significant paper which provides much needed foundations for further work in this area, ..."

My MKM relevant contributions include (i) an agent-based approach enabling distributed search for applicable assertions in repositories of formalised mathematical knowledge [C14-03, C13-03, W19-03, W16-02, W14-01, R20-03][23], (ii) a representation language for human-constructed mathematical proofs with under-specification [J08-03]; see also (F) below, (iii) a proposal for a human-oriented user interface based on so called *proof tasks* [J09-04, W24-03, W16-02], and (iv) the active involvement in the EU MKM network[24] as head of Saarland University node.

(B) Mathematical Assistance Environments The CALCULEMUS research community[25] is pursuing a bottom-up approach starting from existing tools and systems to building a new generation of mathematical assistance environments that provide integrated computer-based support for most tasks of a mathematician.

My contributions in this direction include (i) the development of the MA OMEGA (head of OMEGA group) [B01-03, C12-02, C11-02, C01-97], (ii) the coordination of the CALCULEMUS-II proposal [R24-03] in the EU 6th framework on *Systems for Computer-Supported Mathematical Knowledge Evolution* with 13 European project partners and further 18 academic and industrial collaborators — including SRI, NASA, and Intel, (iii) the agent-based suggestion mechanism OANTS for interactive theorem provers [C06-99, C04-98, W03-99, R12-99, R09-99], (iv) an approach for learning proof methods in proof-planning [J07-03, C09-00, W08-01, W06-01, R16-02, R14-01], (v) the analysis of shortcomings of proof-planning such as its dependency on the underlying logic layer [W12-01], (vi) an approach to agent-based expansion of proof tactics [W19-03], and (vii) the multi-modal user interface LOUI for MAs [J03-99, J02-99, W01-98].

(C) Integration of Reasoning Systems The integration of reasoning tools and the improvement of their interoperability is central to foster progress for (A) and (B).

My contributions include (i) the coordination of the CALCULEMUS-I research training network on *Systems for Integrated Symbolic Reasoning and Symbolic Computation* [W22-03, W21-03, E05-03, R22-03], (ii) my role as co-organiser and co-chair of the 2002 CALCULEMUS Autumn School in Pisa [E04-02, E03-02, E02-02, E01-02], (iii) a tactic-based approach for white-box integration of the higher-order proof assistants OMEGA and TPS [J01-99, W02-98], (iv) the extension of the OANTS suggestion mechanism into a heuristic, resource adaptive, automated, agent-based reasoning system supporting the integration of specialist reasoners [C10-01, C08-00, C07-00, C06-99, W15-02, W11-01, W10-01, W07-00, W04-99, R10-99], and (iv) the supervision of the PhD project of Jürgen Zimmer, the main developer and expert of the mathematical software bus MathWeb, on the automated coordination of reasoning tools within the MathWeb [Zim04].

[23]The bibliography uses a special bibliographystyle for the applicants own papers. These entries are labeled '[Xnn-mm] where 'X describes the type/category of publication ('B stands for Books and Chapters in Books, 'J for International Journals, 'E for Edited Proceedings and Books, 'C for International Conferences, 'W for International Workshops, 'T for Theses, and 'R for Technical Reports and Others), 'nn is a consecutive numbering in each category, and 'mm describes the year of publication.

[24]`monet.nag.co.uk/mkm/index.html`

[25]`www.calculemus.org`

(D) Higher-Order Logic —- Semantics and Mechanisation As a consequence of Gödels incompleteness theorems the scientific community interested in the automation of logical and mathematical reasoning has mainly concentrated on first-order logic while on the other hand the community developing interactive proof assistants has preferred higher-order logic as an elegant and expressive representation formalism. This discrepancy in the historic development, especially the lack of interest in the direct (appropriately limited!) automation of HOL reasoning is very unfortunate for the development of MAs. Amongst the frequently used arguments against research on the automation of HOL are: (a) the lack of an appropriate semantics and respective proof methods that can guide the development of complete HOL calculi[26], (b) the mechanisation of equality and extensionality reasoning is not sufficiently developed, and (c) the mechanisation of set variables is an open problem.

Problems (a) and (b) have been the main interest of my PhD project [T2-99] and I have contributed to (i) the development of a landscape of HOL model classes motivated by different roles extensionality and equality plays in them [J06-04, R18-03, R05-97], (ii) the extension of Smullyan's abstract consistency principle to this landscape of semantical notions [J06-04, R18-03, R05-97], (iii) investigation of the connection between cut-elimination and saturation in HOL [R19-03], (vi) the development of calculi for HOL reasoning with equality and extensionality: extensional resolution [J05-02, T2-99, C02-98, R06-97, R04-97], extensional paramodulation and RUE-resolution [T2-99, C05-99], natural deduction calculus [J06-04, R18-03], and sequent calculus [R19-03], and (v) the development of the HOL resolution theorem prover LEO [T2-99, C03-98].

(E) Semantical Mediators Mediation of mathematical knowledge based on syntactical and semantical filtering is required in order to make the fast emerging distributed repositories of formalised mathematical knowledge better accessible for MAs and humans.

My own work in this direction concentrated on (i) the proposal of a two layered architecture for semantic mathematical knowledge retrieval that combines syntactical pre-filtering with full semantic analyses supported by theorem provers [C13-03, W14-01] and (ii) the automation of full assertion level reasoning as a means for semantic filtering [C14-03].

(F) Natural Language Dialog with Proof Assistants The DIALOG project [R25-01] in the collaborative research centre SFB 378 *Resource-adaptive Cognitive Processes* aims at a mathematical tutoring system that employs an elaborate natural language dialogue component. To tutor mathematical proofs it supports a formally encoded mathematical theory including definitions and theorems along with their proofs, means of classifying the student's input in terms of the knowledge of the domain, and a theory of tutoring that should make use of hints. A main challenge is to couple natural language analysis with dynamic domain reasoning (supported by MAs and theorem provers) since the set of all valid mathematical proofs to be tutored cannot be statically modelled in general.

[26]According to G¨odel's theorem standard semantics does not allow for complete calculi. Weaker notions of semantics such as Henkin semantics [Hen50] are thus required.

My contributions include (i) an architecture for the DIALOG system [W20-03], (ii) a corpus of data gained in a Wizard of Oz case study [C15-04, W23-03], and (iii) an analysis of this corpus and consequences for dynamic domain reasoning with MAs in the DIALOG system context [W20-03]; this in particular includes the phenomenon of under-specification [J08-03, W25-03] in the representation of human constructed proofs and the challenge to analyse information completeness, accuracy, step-size, and relevance of uttered proof steps within the DIALOG system [J08-03].

3 Goals and work schedule (Ziele und Arbeitsprogramm)

3.1 Goals (Ziele)

The long-term goal of my research is the (internationally joint) development of a mathematical assistance environment and its integration into the emerging Mathematical Semantic Net[27]. Interactive proof systems are today routinely employed in industrial applications for safety and security verification and more recently they find applications in e-learning systems for mathematics.

The immediate goal of the Aλonzo project is (I) to provide a mediator framework for context-sensitive higher-order theorem retrieval from mathematical knowledge repositories in the mathematical semantic web, (II) to analyse and address the particular challenges imposed by this application for higher-order theorem proving systems, and (III) to exemplarily integrate and employ the agent-based knowledge retrieval tools within mathematical assistance environments of partners in the MKM and CALCULEMUS research initiatives.

Project Aλonzo has three work packages:

WP I **Semantic Mediators for Mathematical Knowledge Bases**

Knowledge acquisition and retrieval in the currently emerging huge repositories of formalised mathematical knowledge shall be supported by semantic mediators. These mediators shall, for instance, be capable of suggesting applicable theorems and lemmata in a given proof context within a MA. Informal mathematical practice and mathematical textbooks implicitly more often than not exploit higher-order arguments and constructs. Thus, most mathematical assistance environments as well as most of the emerging mathematical knowledge bases also provide HOL as their basic representation language. Therefore we propose to directly exploit HOL proof tools — in the range from HOL matching to fully extensional higher-order theorem proving — as core algorithms underlying the semantic mediators to be developed in this work package. Generally, mediation in this context may be supported also by state-of-the-art first-order theorem provers when employing higher-order to first-order transformation tools such as the OMEGA group's TRAMP system [CADE00-00]. Such an approach is currently pursued in a project of Larry Paulson at Cambridge University. We propose to consider this approach as a cooperating

[27] `www.win.tue.nl/dw/monet/`

solution (parallel to or as a pre-filter for the Aλonzo-agents) in our project. This is easily possible since OMEGA, TRAMP, and several state-of-the-art first-order theorem provers have been operating within the MathWeb system for several years (see a recent application in [B01-03]), i.e., the required infrastructure is already completely developed.

Why do we additionally propose to apply HOL proof agents for the mediation task? For instance, answer substitutions (instantiations of free/universal variables in the mediation context) are very interesting as a form of justification for the applicability of each retrieved lemma in a MKB. HOL answer substitutions can, of course, not easily be realised within a transformational first-order approach as sketched above and it would require several highly non-trivial modifications to the transformation module TRAMP and in particular to the employed first-order theorem provers (which are often only capable of producing *proof-found* or *no-proof-found* answers). And additionally extensionality reasoning is beyond the scope of such a transformational approach.

In addition to their application as semantic mediators within MAs the HOL proof agents to be developed shall also be employed as semantic search engines for mathematicians and they shall be integrated with e-learning systems, such as the Saarbrücken ActiveMath system. An application for the semantic mediation of mathematical knowledge for the DIALOG project (tutorial natural language dialog with a MA) is sketched in [C14-03].

1. Problem Analysis We propose the application of appropriately limited HOL reasoning for domain- and context-specific retrieval of mathematical knowledge from MKBs. For this we will adapt the two stage approach as discussed in [C13-03], which combines syntactically oriented pre-filters with semantic analysis. The pre-filters may employ efficiently processable criterions based on meta-data and ontologies in order to identify, e.g., sets of candidate theorems in the MKBs that are eventually applicable to a focused proof context within a MA. Then the HOL agents may be employed as post-filters (eventually concurrently to other systems such as first-order theorem provers) to exactly determine the applicable theorems of this set. Thereby additional information justifying the applicability of each determined theorem might be generated and provided. Exact semantic retrieval generally includes the following aspects: (i) Eventually logical transformations are required in order to see the connection between a theorem in a MKB and a focused subgoal in a MA; consider, e.g., a theorem of form $A \Leftrightarrow B$ and a subgoal of form $A \Rightarrow B \wedge (\neg A \Rightarrow \neg B)$. (ii) The variables of a theorem in a MKB may have to be instantiated with terms occurring in a focused subgoal; consider, e.g., a theorem $\forall X. is{-}square(X \times X)$ and a subgoal $is{-}square(2 \times 2)$. (iii) Free variables (meta-variables) may occur in a focused subgoal of a MA and they may have to be instantiated with terms occurring in a theorem in a MKB; consider, e.g., a subgoal $irrational(X)$ with metavariable X and a theorem $irrational(\sqrt{2})$. A particular challenge is that these three aspects may generally occur in combination and have to be addressed in combination.

Furthermore, focused subproblems within MAs usually have theory contexts. Thus, if the compatibility of a subproblem's context with a theorem's context cannot be established by simple and efficient means, additional theorem proving tasks may arise. This problem shall be further analysed and respective ways to combine it with the above challenges shall be explored.

2. Distributed, Resource-adaptive Architecture Based on the above problem analysis and the ideas in [C13-03] an architecture for semantic mediators employing HOL reasoning agents shall be developed. This architecture shall ideally exploit the syntactical and restricted semantical filter mechanism already provided by the emerging MKBs.

The emerging MKBs today already contain thousands of theorems and an enourmous increase can be predicted for the future. Even very strong syntactical pre-filters will therefore eventually still generate large sets of candidate theorems to be semantically processed by the proof agents. Thereby for each candidate theorem a generally non-decidable proof problem has to be analysed. This motivates mechanisms supporting resource-adaptive, distributed search with anytime character as proposed and partly realised in Ω-ANTS [C10-01, C08-00].

3. Search Engine for Mathematics In (1) and (2) we have discussed semantic mediation from the perspective of its exploitation within MAs. However, semantic mediation of mathematical knowledge may similarly directly support the mathematician as an improved search engine for mathematical theorems.

The project shall therefore investigate whether the approach can be succesfully coupled with state-of-the-art search engines such as Google (eventually such an integrated approach is also adequate and applicable for (2)).

In this application the issue of analysing the compatibility between the theorem's theory contexts and the usually not explicitly given context assumed by the requesting mathematician raises a particular challenge.

WP II Application-oriented Improvement of Higher-Order Theorem Provers

The requirements and challenges imposed by the application of higher-order theorem proving in WP I shall be analysed and addressed by the development of specialised calculi, strategies and heuristics. They will be realised within the existing HOL theorem provers LEO (Resolution) and TPS (Matrix). Special efficiency increasing techniques and tricks as developed and employed in first-order theorem proving shall, if possible, be additionally adapted to HOL theorem proving. To foster interoperability with MAs and other reasoning tools, a proof representation translation module shall be developed that transforms machine-oriented proofs generated by LEO and TPS into the semi-formal proof representation formats currently developed and investigated in MKM.

1. Semantics and Mechanisation of HOL The results in [J06-04, T2-99] shall be refined and extended. In addition to extensionality, the parameters that are relevant are the axiom of choice, the description operator, and the axiom of infinity. They are particularly relevant for mathematics and the representation of partial functions but they are not sufficiently investigated yet. For all introduced model classes corresponding abstract consistency principles (i.e., a set of respective abstract consistency conditions) shall be provided. These proof principles are crucial, e.g., for the analysis of calculi for higher order logic as illustrated in [J06-04] for HOL natural deduction calculi. Furthermore, for the introduced model classes machine-oriented calculi (resolution and matrix) and interaction-oriented calculi (natural deduction and sequent) shall be developed.

An important issue thereby is the automation of equality and extensionality reasoning in HOL. Completeness shall be analysed with the help of the developed techniques. These calculi (especially the machine-oriented) shall be refined and coupled with application-oriented strategies and heuristics. Some work hypotheses for completeness maintaining strategies already exist (see also [T2-99]).

The current trend in machine-oriented TP towards non-normal form calculi[28] is especially interesting for the automation of HOL reasoning. In fact, in my extensional resolution approach [T2-99, J05-02] unification and proof-search are integrated at one conceptual layer and they allow for mutual recursive calls. Thus, extensional HOL unification may employ proof search for comparing formulas which is an important ingredient for non-normal form reasoning and which is not supported in first-order unification. Note that unification of formulas, subsuming the question whether two formulas are equivalent, is particularly important for semantic mediation of knowledge as proposed in WP I.

2. Theorem Provers for HOL The machine-oriented calculi, strategies, and heuristics developed in (1.) shall be implemented and applied within the systems LEO (resolution theorem proving) and TPS (matrix theorem proving). Within case studies the strengths and weaknesses of both approaches shall be first analysed and then compared with the situation in first-order theorem proving. There resolution- resp. superposition-based approaches are currently dominating.

3. Adaption of Techniques from First-Order Theorem Proving Successful, efficiency increasing techniques developed for first-order theorem provers — term-indexing [RSV98], strong literal selection functions as employed in the superposition approach, etc. — shall, if possible, be adapted for the higher-order approaches and systems developed in Aλonzo.

However, this will cause non-trivial challenges: for instance, higher-order term-indexing modulo extensional equality generally requires full higher-order theorem proving within indexing and is therefore hardly feasible. Similarly, the adaptation of literal selection functions for higher-order logic causes a challenge since well-ordering criterions such as $\forall x.F > \{x \leftarrow T\}F$ for all terms T of the considered logic (instances of quantified formulas F are always smaller w.r.t. the considered order relation than F itself) are problematic in HOL: e.g., the HOL formula $\forall x_o \centerdot x_o$ has as instance the formula $\{x_o \leftarrow \forall x_o \centerdot x_o\}x_o$ which reduces to $\forall x_o \centerdot x_o$ again.

Approaches to higher-order term-indexing for the non-extensional case have recently been investigated in the PhD project of Brigitte Pientka at Carnegie Mellon University [Pie03]. We propose to investigate if these results can be adapted to the extensional case.

4. Proof Transformation In order to enable white-box integration of the HOL theorem provers developed in Aλonzo into a MA it is important to support proof transformations between

[28]An example for this trend is the approach *Superposition with Equivalence Reasoning and Delayed Clause Normal Form Transformation* [GS03]. I have illustrated in a recent presentation at the annual German meeting of the deduction systems community (see slides at `www.ags.uni-sb.de/~chris/papers/2003-deduktionstreffen-talk.pdf`) that the new rules and ideas applied in this approach are actually closely related to rules already provided my thesis work [T2-99] and as presented at CADE 1999 [C05-99].

the machine-oriented, resolution- and matrix-based proof representation formats and the user-oriented proof representation formats usually employed within MAs (e.g. natural deduction and sequent calculi or, as available in ΩMEGA, the less granular layer of assertion level proofs [CADE94-94] and natural language proofs [Fie01]).

The motivation thereby is: (a) a base system such as the MA ΩMEGA may require the translation of proofs constructed in external reasoners into its own proof representation format in order to be able to verify them, and (b) proof explanation systems such as P.rex [IJCAR01-01], which are capable of natural language proof presentations and restricted dialogs, usually pre-require the transformation of machine-oriented proof formats into human-oriented proof representations.

Aλonzo shall therefore investigate whether proof transformation mechanisms as already employed in TPS (non-extensional matrix proofs into natural deduction) and the first-order proof translator TRAMP [CADE00-00] can be fruitfully adapted to the calculi developed in Aλonzo. Alternatively an adaption of the tactic-based proof transformation approach as described in [J01-99, W02-98] may be chosen.

WP III Agent- or Web-Service based Integration and Coordination

The goal is to exemplarily employ the higher-order proof tools as proof agents (or Web services) for the MAs developed in CALCULEMUS and MKM (a first choice will be the MA ΩMEGA). Further research addresses cooperation and coordination of these proof agents with other reasoning agents in the scenario, such as first-order theorem provers, computer algebra systems or a mathematician. We aim at white-box solutions where proof objects generated by the higher-order proof tools can be translated upon request into proof representation formats of MAs in the scenario.

1. Agent or Web Service based Modelling The systems LEO and TPS shall be modelled as pro-active agents and exemplarily employed as proof agents working for MAs developed in CALCULEMUS and MKM (a first choice will be the MA ΩMEGA). Within these systems they shall be able to autonomously detect subgoals which potentially lie in their scope in order to attack them in the background. As integration infrastructure the systems Ω-ANTS and/or MathWeb shall be appropriately adapted and employed.

2. Coordination Cooperation between Aλonzo's HOL theorem provers and first-order theorem provers or computer algebra systems shall be investigated (e.g. in cooperation with the CALCULEMUS project). For the automated coordination of such cooperation three different guidance approaches are possible: (a) centralised guidance within a MA, (b) decentralised guidance within a network of cooperating agents, and (c) direct guidance of cooperations within the the HOL theorem provers itself.

Experience for option (a) exists in the ΩMEGA project in terms of the Ω-ANTS approach [C10-01] and proof planning [B01-03, MPS02, CMPS03] which both have been employed to coordinate cooperation of external specialist reasoning systems.

For option (b) the MathWeb system may be employed which is currently extended in the PhD project of Jürgen Zimmer in order to support intelligent brokering and coordination of reasoning systems; see [Zim03] and [Zim04].

Distributed architectures for (c) have been proposed in first-order theorem proving and they have been employed, for instance, in resolution-based systems. A respective taxonomy and an overview on the literature is given in [Bon00]. For HOL a stronger impact of such approaches seems plausible because HOL reasoning has many characteristics, such as alternative solutions in higher-order unification or alternatives for set variable instantiation, which naturally call for the application of distributed reasoning techniques.

3. White-Box Integration A proof transformation mechanism as investigated in WP II(4) shall be implemented and tested within case studies. With respect to white-box integration we will therefore first concentrate on ΩMEGA, which provides already good support for white-box integrations. Since the transformation of a found proof into another format may eventually require more computation time than proof search itself, it is an important issue to decouple both processes in our framework. Proof transformation shall thus be offered only upon request and subsequent to actual proof search.

4. User Interaction The users of MAs shall be fruitfully supported by the HOL proof agents within interactive proof construction. For this the HOL proof agents shall "silently" exploit available computation resources in the background in order to select and tackle subproblems in the interactive proof development. Successful proof attempts shall be signalled and communicated to the user in a human-oriented way. For this the proof transformation tool developed in (3) shall be coupled with the P.rex system in order to support natural language based proof representation and explanation.

3.1.1 Work Schedule (Arbeitsprogramm)

The work schedule is displayed in Figure 1.

In WP I problem analysis will be followed by modelling of the mediator architecture and a prototype implementation.

In WP II theoretical investigation and implementation will be pursued in parallel, i.e., in cycles of improvement. The adaptation of techniques from first-order theorem proving and the development of a proof transformation tool will start with some delay.

In WP III we will start with the modelling of the theorem provers LEO and TPS as proof agents; the encapsulated proof systems will then be periodically replaced according to the progress in WP II. With some delay we will address coordination with other proof agents and integration into MAs. Interaction of the proof agents with a MA user will be investigated in the last quarter of the project.

The proposed research is very ambitious. The scheduled duration of the project therefore is 4 years. In the first two years of the project as proposed above we will develop a full fledged prototype version of a mathematical knowledge mediator based on the higher-order theorem

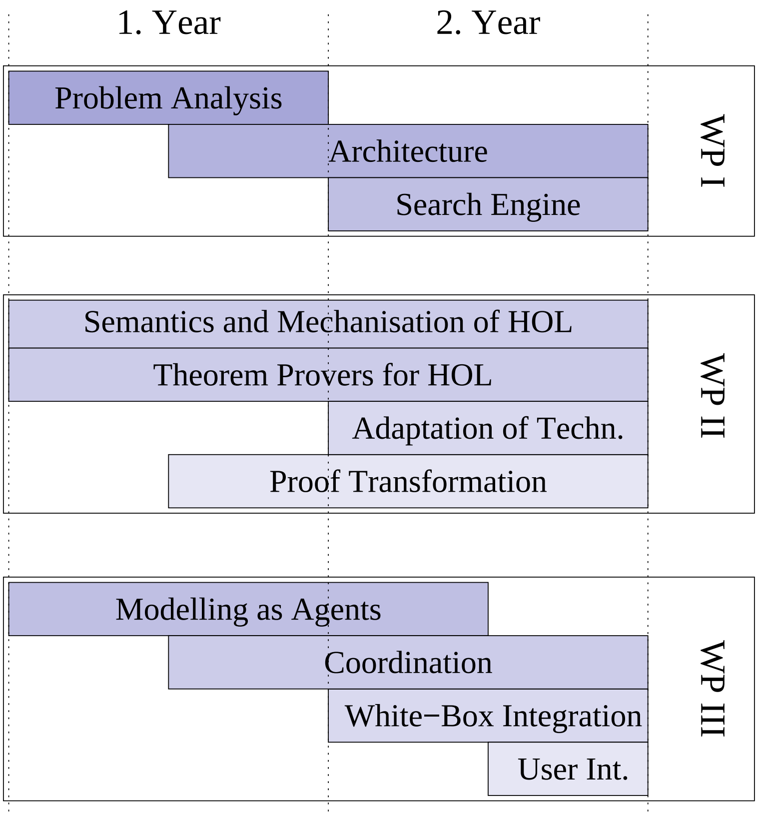

Abbildung 1: Work Schedule.

provers LEO and TPS and we will isolate and address some of the main research challenges. We expect that another two years will be needed to sufficiently complete and sum up the project.

We want to emphasize that the project does not start from scratch. The theorem provers LEO and TPS can be employed already in their current versions. The applicant is the developer of the LEO prover [T2-99, C03-98] and has worked with the TPS system [J01-99]. Chad Brown, who wants to join the project, is currently the main developer of the TPS prover. Thus, the project will join international resources in the area of higher-order theorem proving. An important issue of the project also is to foster the further development and use of the TPS system after the retirement of Peter Andrews.

3.1.2 Experiments with humans (Untersuchungen am Menschen)

— does not apply —

3.1.3 Experiments with animals (Tierversuche)

— does not apply —

3.1.4 Experiments with recombinant DNA (Gentechnische Experimente)

— does not apply —

4 Funds requested (Beantragte Mittel)

4.1 Staff (Personalbedarf)

1. *two* researchers BAT IIa/West for the complete application period.
2. *one* student research assistant (wiss. Hilfskraft, WHK)
   with degree (M.Sc.) working 19 hours per week for the complete application period.
3. *one* student research assistant (stud. Hilfskraft, SHK)
   without degree working 19 hours per week for the complete application period.

We are planning to employ Chad Brown (currently Carnegie Mellon University, Pittsburgh, USA) in the project. He is finishing his PhD study on the set variable instantiation problem under supervision of Prof. Peter Andrews in July this year and he then wants to join the Aλonzo project (see the enclosed statement). Chad Brown is an outstanding young researcher in the field and there already exists an ongoing cooperation with the applicant; see [J06-04, T2-99].

4.2 Scientific Instrumentation (Wissenschaftliche Geräte)

More than 10.000 Euro — does not apply —

Less than 10.000 Euro For project presentations at conferences, workhops, and research visits we request notebooks for the two researchers in the project. These notebooks will furthermore be used as the researchers *Arbeitsplatzrechner* at Saarland University. Saarland University has so far rejected all applications from our research group for notebooks in similar cases.

Requested Amount[29]: 2×3300 Euro

Fees — does not apply —

4.3 Consumables (Verbrauchsmaterial)

The consumables will be covered by the *Grundausstattung* of Prof. Siekmann at Saarland University.

4.4 Travel Expenses (Reisen)

The travel budget shall cover:

1. Dissemination of research results at international conferences and workshops such as:
   (a) IJCAI: International Joint Conference on Artificial Intelligence

[29] An offer is enclosed with this application.

(b) ECAI: European Conference on Artificial Intelligence

(c) IJCAR: International Joint Conference on Automated Reasoning

(d) CADE: Conference on Automated Deduction

(e) TPHOLs: Theorem Proving in Higher-Order Logics

(f) CALCULEMUS: Symposium on the Integration of Symbolic Reasoning and Symbolic Computation

(g) MKM: Symposium on Mathematical Knowldege Management

(h) FROCOS: Frontiers of Combining Systems

(i) KI, AIMSA, EPIA, etc.

Conference fees have increased considerably within the last years and now often range between 500 and 800 US Dollar.

Dissemination of research results on the automation of classical higher-order logic and on applications as sketched in this proposal shall foster the presence of this research field in the above mentioned research communities.

In addition to active conference and workshop participations we plan to closely collaborate with national and international research partners; see Sections 5.2 and 5.3. This includes regular research visits at the partner sides.

We thus request for each researcher and the project leader a sum of 3000 Euro per year.

$3 \times$ 3000 Euro/Year

2. Invitation of guest researchers / collaborators 1500 Euro/Year

4.5 Publication costs (Publikationskosten)

— does not apply —

4.6 Other costs (Sonstige Kosten)

— does not apply —

5 Preconditions for carrying out the project (Voraussetzungen für die Durchführung des Vorhabens)

5.1 Our team (Zusammensetzung der Arbeitsgruppe)

- Dr. C. Benzmüller (C2, Univ. des Saarlandes)
- Dr. S. Autexier (Deutsches Forschungszentrum für Künstliche Intelligenz)
- Dr. C.-P. Wirth (Bat IIa, Univ. des Saarlandes)

5.2 Cooperation with other scientists (Zusammenarbeit mit anderen Wissenschaftlern)

- WPs I and II: with the MBASE-Project of Prof. Michael Kohlhase at International University Bremen.
- WP III: with the ΩMEGA-Project (Jörg Siekmann) at Saarland University, the ACTIVEMATH-Project (Erica Melis) and the VSE-project (Dieter Hutter and Werner Stephan) at DFKI, Saarbrücken; with Prof. Christoph Kreitz at University of Potsdam.

5.3 Foreign contacts and cooperations (Arbeiten im Ausland und Kooperation mit ausländischen Partnern)

The Aλonzo project will be linked with the CALCULEMUS and MKM research initiatives. The aim is to provide mediator tools that support (not only) the MAs being developed by the partners of these networks.

The applicant is the coordinator of the EU Research Training Network CALCULEMUS, of the successor proposal CALCULEMUS-II, and of the Saarland University node in MKM.

A research collaboration is particularly planned with the following CALCULEMUS and MKM partners:

- WP II: with the TPS-Project of Prof. Peter Andrews and with Prof. Frank Pfenning at the Carnegie Mellon University, USA
- WP III: with the Ω-ANTS-Project of Volker Sorge and Manfred Kerber at the University of Birmingham; with the Dream group of Prof. Alan Bundy at University of Edinburgh, UK; with Prof. Bruno Buchberger at RISC Linz, Austria.

With all theses groups I have already cooperated substantially in the past, including joint publications.

5.4 Scientific equipment available (Apparative Ausstattung)

The scientific equipment (except for the applied notebooks) will be covered by Saarland University.

5.5 Saarland Universities general contribution (Laufende Mittel für Sachausgaben)

To be negotiated with Saarland University.

5.6 Other requirements (Sonstige Voraussetzungen)

Keine.

6 <u>Declarations</u> Erklärungen

6.1

The University of Edinburgh together with the applicant have applied for an European Young Investigator (EURYI[30]) Award in order to fund the Aλonzo project at Edinburgh University. The EURYI scheme aims to stimulate the best up and coming postdoctoral researchers in the world to pursue careers in academia in Europe based on scientific excellence. In case the EURYI application of Edinburgh University is successful I shall notify the DFG immediately, having the the options to possibly fund the project jointly (matching funds).

6.2

The *Vertrauensdozent* of Saarland University has been informed about this research grant application.

7 <u>Signature</u> (Unterschrift)

Dr. Christoph Benzmüller

8 <u>List of appendages</u> (Verzeichnis der Anlagen)

(1) Document 10_04

(2) CV

(3) Statement of Chad Brown to join the project (The enclosed statement has been written by Chad Brown for the EURYI proposal; he has already confirmed his interest to join the project in Saarbrücken before, e.g., in the context of my Aktionsplan Informatik proposal for the Aλonzo-project last year.)

(4) Notebook offer

[30] `www.dfg.de/internationales/nachwuchs/euryi_awards/index.html`

References

Note: The following bibliography uses a special bibliography style for my papers. These entries are labeled '[Xnn-mm]' where 'X' describes the type/category of publication ('B' stands for Books and Chapters in Books, 'J' for International Journals, 'E' for Edited Proceedings and Books, 'C' for International Conferences, 'W' for International Workshops, 'T' for Theses, and 'R' for Technical Reports and Others), 'nn' is a consecutive numbering in each category, and 'mm' describes the year of publication.

A complete list of publications can also be found at: `www.ags.uni-sb.de/~chris/cv-texmacs/cv-publications.html`.

[B01-03] Jörg Siekmann, Christoph Benzmüller, Armin Fiedler, Andreas Meier, Immanuel Normann, and Martin Pollet. *Proof Development in OMEGA: The Irrationality of Square Root of 2,* pp.271–314. Kluwer Applied Logic series (28). Kluwer Academic Publishers, 2003. ISBN 1-4020-1656-5.

[BGH03] Bruno Buchberger, Gaston Gonnet, and Michiel Hazewinkel. Special issue on mathematical knowledge management. *Annals of Mathematics and Artificial Intelligence*, 38(1-3):3–232, May 2003.

[Bon00] Maria Paola Bonacina. A taxonomy of parallel strategies for deduction. *Annals of Mathematics and Artificial Intelligence*, 29(1-4):223–257, 2000.

[Bro02] Chad E. Brown. Solving for set variables in higher-order theorem proving. In Andrei Voronkov, ed., *Proceedings of the 18th International Conference on Automated Deduction (CADE-19),* no.2392 in LNAI, pp.144–149, Copenhagen, Denmark, 2002. Springer.

[C01-97] Christoph Benzmüller, Lassaad Cheikhrouhou, Detlef Fehrer, Armin Fiedler, Xiaorong Huang, Manfred Kerber, Michael Kohlhase, Karsten Konrad, Erica Melis, Andreas Meier, Wolf Schaarschmidt, Jörg Siekmann, and Volker Sorge. OMEGA: Towards a mathematical assistant. In William McCune, ed., *Proceedings of 14th International Conference on Automated Deduction (CADE-14),* no.1249 in LNAI, pp.252–255, Townsville, Australia, 1997. Springer.

[C02-98] Christoph Benzmüller and Michael Kohlhase. Extensional higher-order resolution. In Claude Kirchner and Hélène Kirchner, eds., *Proceedings of the 15th International Conference on Automated Deduction (CADE-15),* no.1421 in LNAI, pp.56–71, Lindau, Germany, 1998. Springer.

[C03-98] Christoph Benzmüller and Michael Kohlhase. LEO – a higher-order theorem prover. In Claude Kirchner and Hélène Kirchner, eds., *Proceedings of the 15th International Conference on Automated Deduction (CADE-15),* no.1421 in LNAI, pp.139–143, Lindau, Germany, 1998. Springer.

[C04-98] Christoph Benzmüller and Volker Sorge. A blackboard architecture for guiding interactive proofs. In Fausto Giunchiglia, ed., *Proceedings of 8th International Conference on Artificial Intelligence: Methodology, Systems, Applications (AIMSA'98),* no.1480 in LNAI, pp.102–114, Sozopol, Bulgaria, 1998. Springer.

[C05-99] Christoph Benzmüller. Extensional higher-order paramodulation and RUE-resolution. In Harald Ganzinger, ed., *Proceedings of the 16th International Conference on Automated Deduction (CADE-16),* no.1632 in LNAI, pp.399–413, Trento, Italy, 1999. Springer.

[C06-99] Christoph Benzmüller and Volker Sorge. Critical agents supporting interactive theorem proving. In Pedro Borahona and Jose J. Alferes, eds., *Proceedings of the 9th Portuguese Conference on Artificial Intelligence (EPIA'99),* no.1695 in LNAI, pp.208–221, Evora, Portugal, 1999. Springer.

[C07-00] Christoph Benzmüller, Mateja Jamnik, Manfred Kerber, and Volker Sorge. Resource guided concurrent deduction. In Manfred Kerber and Michael Kohlhase, eds., *Symbolic Computation and Automated Reasoning*, pp.245–246. A.K.Peters, 2000.

[C08-00] Christoph Benzmüller and Volker Sorge. OANTS – an open approach at combining interactive and automated theorem proving. In Manfred Kerber and Michael Kohlhase, eds., *Symbolic Computation and Automated Reasoning*, pp.81–97. A.K.Peters, 2000.

[C09-00] Mateja Jamnik, Manfred Kerber, and Christoph Benzmüller. Towards learning new methods in proof planning. In Manfred Kerber and Michael Kohlhase, eds., *Symbolic Computation and Automated Reasoning*, pp.142–159. A.K.Peters, 2000.

[C10-01] Christoph Benzmüller, Mateja Jamnik, Manfred Kerber, and Volker Sorge. Experiments with an agent-oriented reasoning system. In Franz Baader, Gerhard Brewka, and Thomas Eiter, eds., *KI 2001: Advances in Artificial Intelligence, Joint German/Austrian Conference on AI, Vienna, Austria, September 19-21, 2001, Proceedings*, no.2174 in LNAI, pp.409–424. Springer, 2001.

[C11-02] Jörg Siekmann, Christoph Benzmüller, Vladimir Brezhnev, Lassaad Cheikhrouhou, Armin Fiedler, Andreas Franke, Helmut Horacek, Michael Kohlhase, Andreas Meier, Erica Melis, Markus Moschner, Immanuel Normann, Martin Pollet, Volker Sorge, Carsten Ullrich, Claus-Peter Wirth, and Jürgen Zimmer. Proof development with OMEGA. In Andrei Voronkov, ed., *Proceedings of the 18th International Conference on Automated Deduction (CADE-18)*, no.2392 in LNAI, pp.144–149, Copenhagen, Denmark, 2002. Springer.

[C12-02] Jörg Siekmann, Christoph Benzmüller, Armin Fiedler, A Meier, and M. Pollet. Proof development with OMEGA: Sqrt(2) is irrational. In Matthias Baaz and Andrei Voronkov, eds., *Logic for Programming, Artificial Intelligence, and Reasoning, 9th International Conference, LPAR 2002*, no.2514 in LNAI, pp.367–387. Springer, 2002.

[C13-03] Christoph Benzmüller, Andreas Meier, and Volker Sorge. Bridging theorem proving and mathematical knowledge retrieval. In *Festschrift in Honour of Jörg Siekmann*, LNAI, 2003. To appear.

[C14-03] Quoc Bao Vo, Christoph Benzmüller, and Serge Autexier. Assertion application in theorem proving and proof planning. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, Acapulco, Mexico, 2003. ISBN 0-127-05661-0.

[C15-04] M. Wolska, B. Quoc Vo, D. Tsovaltzi, I. Kruijff-Korbayova, E. Karagjosova, H. Horacek, M. Gabsdil, A. Fiedler, and C. Benzmüller. An annotated corpus of tutorial dialogs on mathematical theorem proving. In *Proceedings of International Conference on Language Resources and Evaluation (LREC 2004)*, Lisbon, Potugal, 2004. To appear.

[CADE00-00] A. Meier. TRAMP: Transformation of Machine-Found Proofs into Natural Deduction Proofs at the Assertion Level. In D. McAllester, ed., *Proceedings of the 17th Conference on Automated Deduction*, no.1831 in LNAI. Springer, 2000.

[CADE94-94] Xiaorong Huang. Reconstructing proofs at the assertion level. In Alan Bundy, ed., *Proceedings of the 12th Conference on Automated Deduction*, no.814 in LNAI, pp.738–752, Nancy, France, 1994. Springer.

[CMPS03] Arjeh Cohen, Scott H. Murray, Martin Pollet, and Volker Sorge. Certifying solutions to permutation group problems. In Franz Baader, ed., *Automated Deduction — CADE'19*, Vol.2741 of *LNCS*, pp.258–273, Miami Beach, Fl, USA, 2003. Springer-Verlag.

[E01-02] Jürgen Zimmer and Christoph Benzmüller (eds.). CALCULEMUS Autumn School 2002: Student Poster Abstracts. SEKI Technical Report SR-02-06, Fachbereich Informatik, Universität des Saarlandes, Saarbrücken, Germany, 2002.

[E02-02] Christoph Benzmüller and Regine Endsuleit (eds.). CALCULEMUS Autumn School 2002: Course Notes (Part I). SEKI Technical Report SR-02-07, Fachbereich Informatik, Universität des Saarlandes, Saarbrücken, Germany, 2002.

[E03-02] Christoph Benzmüller and Regine Endsuleit (eds.). CALCULEMUS Autumn School 2002: Course Notes (Part II). SEKI Technical Report SR-02-08, Fachbereich Informatik, Universität des Saarlandes, Saarbrücken, Germany, 2002.

[E04-02] Christoph Benzmüller and Regine Endsuleit (eds.). CALCULEMUS Autumn School 2002: Course Notes (Part III). SEKI Technical Report SR-02-09, Fachbereich Informatik, Universität des Saarlandes, Saarbrücken, Germany, 2002.

[E05-03] Christoph Benzmüller. Systems for integrated computation and deduction – interim report of the CALCULEMUS ihp network. SEKI Technical Report SR-03-05, Fachbereich Informatik, Universität des Saarlandes, Saarbrücken, 2003.

[Fie01] Armin Fiedler. *User-adaptive proof explanation*. Ph.D. thesis, Naturwissenschaftlich-Technische Fakultät I, Universität des Saarlandes, Saarbrücken, Germany, 2001.

[GS03] H. Ganzinger and J. Stuber. Superposition with equivalence reasoning and delayed clause normal form transformation. In F. Baader, ed., *Automated Deduction — CADE-19*, Vol.2741, pp.335–349, 2003.

[Hen50] Leon Henkin. Completeness in the theory of types. *Journal of Symbolic Logic*, 15(2):81–91, 1950.

[IJCAR01-01] A. Fiedler. *P.rex*: An interactive proof explainer. In R. Goré, A. Leitsch, and T. Nipkow, eds., *Automated Reasoning — 1st International Joint Conference, IJCAR 2001*, no.2083 in LNAI. Springer, 2001.

[J01-99] Christoph Benzmüller, Matt Bishop, and Volker Sorge. Integrating TPS and OMEGA. *Journal of Universal Computer Science*, 5:188–207, 1999.

[J02-99] Jörg Siekmann, Stephan Hess, Christoph Benzmüller, Lassaad Cheikhrouhou, Armin Fiedler, Helmut Horacek, Michael Kohlhase, Karsten Konrad, Andreas Meier, Erica Melis, Martin Pollet, and Volker Sorge. LOUI: Lovely OMEGA user interface. *Formal Aspects of Computing*, 11:326–342, 1999.

[J03-99] Jörg Siekmann, Helmut Horacek, Michael Kohlhase, Christoph Benzmüller, Lassaad Cheikhrouhou, Detlef Fehrer, Armin Fiedler, Stephan Hess, Karsten Konrad, Andreas Meier, Erica Melis, and Volker Sorge. An interactive proof development environment + anticipation = a mathematical assistant? *International Journal of Computing Anticipatory Systems (CASYS)*, 3:101–110, 1999.

[J05-02] Christoph Benzmüller. Comparing approaches to resolution based higher-order theorem proving. *Synthese, An International Journal for Epistemology, Methodology and Philosophy of Science*, 133(1-2):203–235, 2002.

[J06-04] Christoph Benzmüller, Chad Brown, and Michael Kohlhase. Higher order semantics and extensionality. *Journal of Symbolic Logic*, 2004. To appear.

[J07-03] Mateja Jamnik, Manfred Kerber, Martin Pollet, and Christoph Benzmüller. Automatic learning of proof methods in proof planning. *The Logic Journal of the IGPL*, 11(6):647–674, 2003.

[J08-03] Serge Autexier, Christoph Benzmüller, Armin Fiedler, Helmut Horacek, and Bao Quoc Vo. Assertion-level proof representation with under-specification. *Electronic in Theoretical Computer Science*, 93:5–23, 2003.

[J09-04] Malte Hübner, Serge Autexier, Christoph Benzmüller, and Andreas Meier. Interactive theorem proving with tasks. *Electronic Notes in Theoretical Computer Science*, 2004. To appear.

[MPS02] A. Meier, M. Pollet, and V. Sorge. Comparing Approaches to the Exploration of the Domain of Residue Classes. *Journal of Symbolic Computation, Special Issue on the Integration of Automated Reasoning and Computer Algebra Systems*, 34(4):287–306, October 2002. Steve Linton and Roberto Sebastiani, eds.

[Pie03] B. Pientka. Higher-order substitution tree indexing. In *Proceedings of the 19th International Conference on Logic Programming*, Mumbai, India, December 2003. To appear.

[R04-97] Christoph Benzmüller. A calculus and a system architecture for extensional higher-order resolution. Research Report 97-198, Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, USA, 1997.

[R05-97] Christoph Benzmüller and Michael Kohlhase. Model existence for higher-order logic. Seki-Report SR-97-09, Department of Computer Science, Saarland University, 1997.

[R06-97] Christoph Benzmüller and Michael Kohlhase. Henkin completeness of higher-order resolution. Seki-Report SR-97-10, Department of Computer Science, Saarland University, 1997.

[R09-99] Christoph Benzmüller and Volker Sorge. Resource adaptive agents in interactive theorem proving. Seki-Report SR-99-02, Department of Computer Science, Saarland University, 1999.

[R10-99] Christoph Benzmüller, Mateja Jamnik, Manfred Kerber, and Volker Sorge. Towards concurrent resource guided deduction. Seki-Report SR-99-07, Department of Computer Science, Saarland University, 1999.

[R12-99] Christoph Benzmüller, Volker Sorge, and John Byrnes. OANTS for interactive ATP. Draft, AG Siekmann, Saarland University, 1999.

[R14-01] Mateja Jamnik, Manfred Kerber, and Christoph Benzmüller. Automatic learning of proof methods in proof planning. Tech. rep.CSRP-01-08, University of Birmingham, School of Computer Science, 2001.

[R16-02] Mateja Jamnik, Manfred Kerber, Martin Pollet, and Christoph Benzmüller. Automatic learning of proof methods in proof planning. Tech. rep.CSRP-02-05, University of Birmingham, School of Computer Science, 2002.

[R18-03] Christoph Benzmüller, Chad E. Brown, and Michael Kohlhase. Higher order semantics and extensionality. Technical Report CMU-01-03, Carnegie Mellon University, Pittsburgh, PA, 2003.

[R19-03] Christoph Benzmüller, Chad E. Brown, and Michael Kohlhase. Semantic techniques for cut-elimination in higher order logic. Tech. rep., Saarland University, Saarbrücken, Germany and Carnegie Mellon University, Pittsburgh, USA, 2003. Manuscript.

[R20-03] Quoc Bao Vo, Christoph Benzmüller, and Serge Autexier. An approach to assertion application via generalized resolution. SEKI Report SR-03-01, Fachrichtung Informatik, Universität des Saarlandes, Saarbrücken, Germany, 2003.

[R22-03] Christoph Benzmüller and Corinna Hahn (editors). The CALCULEMUS Midterm Report. Unpublished EU Report, Saarland University, Saarbrücken, Germany, March 2003.

[R24-03] Christoph Benzmüller and Dieter Hutter. Calculemus-ii: Computer-supported mathematical knowledge evolution. Project proposal for a Marie Curie Research Training Network within the EU 6th framework, 2003.

[R25-01] Manfred Pinkal, Jörg Siekmann, and Christoph Benzmüller. Dialog: Tutorieller dialog mit einem mathematik assistenzsystem. Project proposal in the Collaborative Research Centre SFB 378 on Resource-adaptive Cognitive Processes, 2001.

[RSV98] I.V. Ramakrishnan, R. Sekar, and A. Voronkov. Term indexing. In J. Robinson and A. Voronkov, eds., *Handbook of Automated Reasoning*, Vol. 1. Elsevier Science, 1998.

[T2-99] Christoph Benzmüller. *Equality and Extensionality in Higher-Order Theorem Proving*. Ph.D. thesis, Naturwissenschaftlich-Technische Fakultät I, Universität des Saarlandes, 1999.

[W01-98] Jörg Siekmann, Stephan Hess, Christoph Benzmüller, Lassaad Cheikhrouhou, Detlef Fehrer, Armin Fiedler, Helmut Horacek, Michael Kohlhase, Karsten Konrad, Andreas Meier, Erica Melis, and Volker Sorge. A distributed graphical user interface for the interactive proof system. In *Proceedings of the International Workshop Üser Interfaces for Theorem Provers 1998 (UITP'98)*, pp.130–138, Eindhoven, Netherlands, 1998.

[W02-98] Christoph Benzmüller and Volker Sorge. Integrating TPS with ΩMEGA. In Jim Grundy and Malcolm Newey, eds., *Theorem Proving in Higher Order Logics: Emerging Trends*, Technical Report 98-08, Department of Computer Science and Computer Science Lab, The Australian National University, pp.1–18, Canberra, Australia, October 1998.

[W03-99] Christoph Benzmüller and Volker Sorge. Towards fine-grained proof planning with critical agents. In *Proceedings of the 6th Workshop on Automated Reasoning*, pp.19–20. Edinburgh College of Art & Divison of Informatics, University of Edinburgh, 1999.

[W04-99] Christoph Benzmüller, Mateja Jamnik, Manfred Kerber, and Volker Sorge. Agent based mathematical reasoning. In *Proceedings of the Calculemus Workshop: Systems for Integrated Computation and Deduction*, pp.1–12, July 1999.

[W06-01] Mateja Jamnik, Manfred Kerber, and Christoph Benzmüller. Towards learning new methods in proof planning. In *In Proceedings of the CADE-17 Workshop: Automated Deduction in the Context of Mathematics*, pp.1–12, 2001.

[W07-00] Christoph Benzmüller, Mateja Jamnik, Manfred Kerber, and Volker Sorge. Resource guided concurrent deduction. In *Proceedings of the AISB'2000 Symposium 'How to design a functioning mind'*, pp.137–138, Birmingham, England, 2000. Also in: Proceedings of the 7th Workshop on Automated Reasoning 'Bridging the Gap between Theory and Practice'.

[W08-01] Mateja Jamnik, Manfred Kerber, and Christoph Benzmüller. Learning proof methods in proof planning. In *Proceedings of the Eighth Workshop on Automated Reasoning, Bridging the Gap between Theory and Practice*, pp.5–6. University of York, 2001.

[W10-01] Christoph Benzmüller. An agent based approach to reasoning. In *Extended abstract for invited plenary talk at AISB'01 Convention 'Agents and Cognition*, pp.57–58. University of York, 2001.

[W11-01] Christoph Benzmüller, Mateja Jamnik, Manfred Kerber, and Volker Sorge. An agent-oriented approach to reasoning. In *Proceedings of the Calculemus Workshop 2001*, pp.48–63, Siena, Italy, 2001.

[W12-01] Christoph Benzmüller, Andreas Meier, Erica Melis, Martin Pollet, and Volker Sorge. Proof planning: A fresh start? In *Proceedings of the IJCAR 2001 Workshop: Future Directions in Automated Reasoning*, pp.25–37, Siena, Italy, 2001.

[W14-01] Christoph Benzmüller, Andreas Meier, and Volker Sorge. Distributed assertion retrieval. In *First International Workshop on Mathematical Knowledge Management RISC-Linz*, pp.1–7, Schloss Hagenberg, 2001.

[W15-02] Christoph Benzmüller and Volker Sorge. Agent-based theorem proving. In *Proceedings of the 9th Workshop on Automated Reasoning: Bridging the Gap between Theory and Practice*, pp.1–3, London, England, 2002.

[W16-02] Malte Hübner, Serge Autexier, and Christoph Benzmüller. Agent-based proof search with indexed formulas. In *Additonal Proceedings of 10th Symposium on the Integration of Symbolic Computation and Mechanized Reasoning (CALCULEMUS 2002)*, pp.11–20, Marseilles, France, 2002.

[W19-03] Bao Quoc Vo, Christoph Benzmüller, and Serge Autexier. Assertion application in theorem provinf and proof planning. In *Proceedings of the 10th Workshop on Automated Reasoning: Bridging the Gap between Theory and Practice*, Liverpool, England, 2003.

[W20-03] Christoph Benzmüller, Armin Fiedler, Malte Gabsdil, Helmut Horacek, Ivana Kruijff-Korbayova, Manfred Pinkal, Jörg Siekmann, Dimitra Tsovaltzi, Bao Quoc Vo, and Magdalena Wolska. Tutorial dialogs on mathematical proofs. In *Proceedings of IJCAI-03 Workshop on Knowledge Representation and Automated Reasoning for E-Learning Systems*, pp.12–22, Acapulco, Mexico, 2003.

[W21-03] Christoph Benzmüller. The CALCULEMUS research training network: A short overview. In *Proceedings of the 11th Symposium on the Integration of Symbolic Computation and Mechanized Reasoning (CALCULEMUS 2003)*, pp.1–16, Rome, Italy, 2003. MMIII ARACNE EDITRICE S.R.L. (ISBN 88-7999-545-6).

[W22-03] Christoph Benzmüller. The CALCULEMUS research training network: A short overview. In *Proceedings of the First QPQ Workshop on Deductive Software Components at CADE-19*, pp.13–27, Miami, USA, 2003.

[W23-03] Christoph Benzmüller, Armin Fiedler, Malte Gabsdil, Helmut Horacek, Ivana Kruijff-Korbayova, Manfred Pinkal, Jörg Siekmann, Dimitra Tsovaltzi, Bao Quoc Vo, and Magdalena Wolska. A wizard of oz experiment for tutorial dialogues in mathematics. In *Proceedings of AI in Education (AIED 2003) Workshop on Advanced Technologies for Mathematics Education*, Sydney, Australia, 2003.

[W24-03] Malte Hübner, Christoph Benzmüller, Serge Autexier, and Andreas Meier. Interactive proof construction at the task level. In *Proceedings of the Workshop User Interfaces for Theorem Provers (UITP 2003)*, pp.81–100, Rome, Italy, 2003. ARACNE EDITRICE S.R.L. (ISBN 88-7999-545-6). Also available as: Technical Report No. 189, Institut für Informatik, Albert-Ludwig-Universität, Freiburg.

[W25-03] Christoph Benzmüller, Armin Fiedler, Malte Gabsdil, Helmut Horacek, Ivana Kruijff-Korbayova, Manfred Pinkal, Jörg Siekmann, Dimitra Tsovaltzi, Bao Quoc Vo, and Magdalena Wolska. Language phenomena in tutorial dialogs on mathematical proofs. In *Proceedings of the 7th Workshop on the semantics and pragmatics of dialogue (DiaBruck)*, Wallerfangen, Germany, 2003.

[Zim03] Jürgen Zimmer. A new framework for reasoning agents. In Volker Sorge, Simon Colton, Michael Fisher, and Jeremy Gow, eds., *Proceedings of IJCAI-03 Workshop on Agents and Automated Reasoning*, pp.58–64, Acapulco, Mexico, 2003.

[Zim04] Jürgen Zimmer. A Framework for Agent-based Brokering of Reasoning Services. In Raul Monroy, Gustavo Arroyo-Figueroa, Luis Enrique Sucar, and Juan Humberto Sossa Azuela, eds., *MICAI*, Vol.2972 of *Lecture Notes in Computer Science*. Springer, 2004.

Deutsche Zusammenfassung

Das langfristige Ziel meiner Forschung ist die gemeinsame internationale Entwicklung eines leistungsfähigen mathematischen Assistenzsystems und dessen Integration und internationale Verwendung im entstehenden mathematischen semantischen Netz.[31] Interaktive Beweissysteme werden im Bereich der Formalen Methoden routinemäßig in der industriellen Praxis eingesetzt und werden heute erstmals in der Mathematikausbildung verwendet. Computer Algebra Systeme sind bereits seit längerem kommerziell erfolgreich.

Die Vision eines *komplexen und leistungsfähigen Gesamtsystem*, das verschiedene Deduktions- und Berechungswerkzeuge integriert, wird weltweit von mehreren Wissenschaftlern geteilt und hat zur Bündelung der notwendigen Expertise und Ressourcen im Rahmen international vernetzter Forschungsprojekte geführt. Beispiele sind das CALCULEMUS[32] Research Training Network (*Integration von Symbolischem Schließen und Symbolischer Berechnung*), oder das im Aufbau befindliche MKM[33] Netzwerk (*Mathematical Knowledge Management*); weiterhin zu nennen sind EU Projekte wie MONET[34] (Mathematics on the Net), OPENMATH[35], oder MOWGLI[36] (Mathematics on the Web), sowie die QPQ[37] Datenbank von Deduktionswerkzeugen am Stanford Research Institute in den USA. Zudem gibt es mittlerweile viele nationale Projekte in den USA und Europa die Einzelaspekte dieser Vision abdecken wie Wissensrepräsentation, deduktive Unterstützungstools, Benutzerschnittstellen, Mathematische Publikationswerkzeuge, usw.

Agenten-basierte Architekturen und das semantische Netz sin geeignete Ansätze, um die entstehenden Werkzeuge (wie mathematische Wissensbanken, Theorembeweiser, Entscheidungsprozeduren, Computeralgebra Systeme, usw.) miteinander zu verbinden und Resourcen zu bündeln.

Das unmittelbare Ziel des -Projekts ist es, Beweisagenten für die Logik höherer Ordnung zu entwickeln, deren Leistungsfähigkeit und Effizienz kontinuierlich zu verbessern, eine Koordinationsfähigkeit mit anderen Beweiserkomponenten im *Semantic Net* zu gewährleisten und diese exemplarisch für die Aλonzo/ΩMEGA-Systemumgebung zu realisieren.

Die Arbeitspakete erstrecken sich von (I) der Grundlagenforschung zur Semantik und Mechanisierung von Logik höherer Ordnung, über (II) die Entwicklung konkreter Beweissysteme und Anwendungen als (III) Beweisagenten in mathematischen Assistenzsystemen und als (IV) semantische Mediatoren für mathematische Wissensbanken, bis hin zum Erarbeiten von (V) einführender Literatur zum Themengebiet.

[31] http://www.win.tue.nl/dw/monet/
[32] http://www.eurice.de/calculemus/
[33] http://monet.nag.co.uk/mkm/
[34] http://monet.nag.co.uk/cocoon/monet/
[35] http://www.openmath.org/cocoon/openmath/projects/thematic/
[36] http://www.mowgli.cs.unibo.it/
[37] http://www.qpq.org