Antrag auf Gewährung einer Sachbeihilfe

Aλonzo

Deduktionsagenten höherer Ordnung für mathematische Assistenzsysteme

Dr. Christoph Benzmüller

20. Oktober 2003

Aλonzo

Management Zusammenfassung (Executive Summary)

Das langfristige Ziel meiner Forschung ist die Entwicklung eines großen mathematischen Assistenzsystems und dessen Integration und internationale Verwendung im entstehenden mathematischen semantischen Netz.[1] 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[2] Research Training Network (*Integration von Symbolischem Schließen und Symbolischer Berechnung*), oder das im Aufbau befindliche MKM[3] Netzwerk (*Mathematical Knowledge Management*); weiterhin zu nennen sind EU Projekte wie MONET[4] (Mathematics on the Net), OPENMATH[5], oder MOWGLI[6] (Mathematics on the Web), sowie die QPQ[7] Datenbank von Deduktionswerkzeugen am Stanford Research Institute in den USA.

Das unmittelbare Ziel dieses 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.

[1] `http://www.win.tue.nl/dw/monet/`
[2] `http://www.eurice.de/calculemus/`
[3] `http://monet.nag.co.uk/mkm/`
[4] `http://monet.nag.co.uk/cocoon/monet/`
[5] `http://www.openmath.org/cocoon/openmath/projects/thematic/`
[6] `http://www.mowgli.cs.unibo.it/`
[7] `http://www.qpq.org`

1 Allgemeine Angaben

1.1 **Antragsteller**

1.1.1
- Dr. Christoph E. Benzmüller
- Hochschulassistent (C1), Projektleiter
- 8. September 1968, deutsch
- Universität des Saarlandes
  Fachrichtung Informatik
  Postfach 1150
  66041 Saarbrücken
  Tel.: 0681/302 4574
  Fax: 0681/302 5076
  e-mail: `chris@ags.uni-sb.de`
- Privatanschrift:
  Landwehrplatz 6
  66111 Saarbrücken

1.2 **Thema**: Deduktionsagenten höherer Ordnung für Mathematische Assistenzsysteme

1.3 **Kennwort**: AλONZO

1.4 **Fachgebiet und Arbeitsrichtung** Informatik, Künstliche Intelligenz, Deduktionssysteme, Klassische Logik höherer Ordnung

1.5 **Voraussichtliche Gesamtdauer**: 5 Jahre.

1.6 **Antragszeitraum**: 5 Jahre

1.7 **Gewünschter Beginn der Förderung**: 1. Januar 2004

1.8 **Zusammenfassung**:

Das langfristige Ziel meiner Forschung ist die Entwicklung eines großen mathematischen Assistenzsystems und dessen Integration und internationale Verwendung im entstehenden mathematischen semantischen Netz.[8]

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 (CAS) sind bereits seit längerem kommerziell erfolgreich.

Die wichtigsten Merkmale des von mir mittelfristig geplanten mathematischen Assistenzsystems sind:

(a) *Eine offene und erweiterbare Systemarchitektur in der die Modellierung der integrierten Service-Module als kooperierende Agenten im mathematischen semantischen Netz erfolgt.*

(b) *Eine enge Integration symbolischer Berechnung (CAS) und symbolischen Schließens (Deduktion).*

(c) *Die Einbindung von Spezialbeweisern, Entscheidungsprozeduren, Modellgenerierern, etc. als Subsysteme des mathematischen Assistenzsystems.*

[8] `http://www.win.tue.nl/dw/monet/`

(d) *Eine enge Verzahnung mit den großen zur Zeit entstehenden mathematischen Wissensbanken zur Akquisition und Verwaltung mathematischen Wissens.*[9]

(e) *Unterstützungsmechanismen zur Exploration, Validierung, und Verwaltung domänenspezifischen Wissens.*

(f) *Expressive und benutzerorientierte Kommunikationssprache(n) auf Benutzerseite mit semiformalen und natürlichsprachlichen Interaktionsmöglichkeiten in einer leistungsfähigen, multi-modalen Benutzeroberfläche.*

(g) *Transformationsmechanismen zwischen der expressiven Wissensrepräsentation im mathematischen Assistenzsystem und den optimierten aber oft eingeschränkten Repräsentationen der eingebundenen Subsysteme.*

(h) *Werkzeuge zur publikationsfähigen Aufbereitung des gewonnen Wissens und, umgekehrt, zur Aufnahme und Verifikation bereits publizierten Wissens.*

(i) *Integration oder Kooperation mit Tutorsystemen zur Unterstützung der Mathematikausbildung.*

Der Fortschritt wird nach meiner Überzeugung zwar durch Verbesserungen in jedem einzelnen der genannten Bereiche aber vor allem durch die Zusammenführung der entstehenden Komponenten dieser Bereiche zu einem *komplexen und leistungsfähigen Gesamtsystem* erreicht.

Diese Vision wird weltweit von mehreren Wissenschaftlern geteilt und hat zur Bündelung der notwendigen Expertise und Ressourcen im Rahmen internationaler Forschungsprojekte geführt. Beispiele sind das CALCULEMUS[10] Research Training Network (*Integration von Symbolischem Schließen und Symbolischer Berechnung*), oder das von Bruno Buchberger (RISC Linz) initierte und im Aufbau befindliche MKM[11] Netzwerk (*Mathematical Knowledge Management*); weiterhin zu nennen sind EU Projekte wie MONET[12] (Mathematics on the Net), OPENMATH[13], oder MOWGLI[14] (Mathematics on the Web: Get it by Logics and Interfaces), sowie die QPQ[15] Datenbank von Deduktionswerkzeugen am Stanford Research Institute SRI in den USA.

Unter Ausnutzung meiner internationalen Beziehungen (z.B. als aktives Mitglied in mehreren der genannten internationalen Netzwerken) möchte ich eine eigene Arbeitsgruppe aufbauen, die an dem skizzierten Ziel *leistungsfähige und kommerziell verwertbare Assistenzsysteme für die Mathematik im mathematischen semantischen Netz* arbeitet. Der besondere Schwerpunkt meiner Forschung liegt in der Mechanisierung der klassischen Logik höherer Ordnung (HOL) im obigen Szenario. Ein expressiver Repräsentationsformalismus, der über die Logik erster Ordnung hinausgeht, ist für die Mathematik außerordentlich wichtig und HOL wird daher zunehmend in mathematischen Wissensbanken, in der Kommunikation mit dem Benutzer eines mathematischen Assistenzsystems und in Spezialbeweisern und Beweisplanungssystemen verwendet. HOL erlebt derzeit eine Renaissance auch im logischen Programmieren (λ-Prolog) und dem Gebiet der Formalen Methoden.

Als Ausgangspunkt meiner Forschung möchte ich die ΩMEGA-Beweisentwicklungsumgebung verwenden, die in der Arbeitsgruppe von Prof. Siekmann an der Universität des Saarlandes entwickelt

---

[9] Bereits existierende Beispiele sind die Wissensbanken MIZAR und MBASE, sowie die ausgelagerten Wissensbanken der Beweisassistenten COQ, ISABELLE/HOL, und TPS. Erwähnenswert sind auch die europäischen Forschungsinitiativen MKM und Euler in denen mathematische Wissensbanken entstehen sollen.

[10] `http://www.eurice.de/calculemus/`

[11] `http://monet.nag.co.uk/mkm/`; der erste MKM Workshop fand 2001 am RISC Linz statt, siehe [Buchberger *et al*, 2003].

[12] `http://monet.nag.co.uk/cocoon/monet/`

[13] `http://www.openmath.org/cocoon/openmath/projects/thematic/`

[14] `http://www.mowgli.cs.unibo.it/`

[15] `http://www.qpq.org`

wird; dieses *Bottom-up Vorgehen* ausgehend von einem existierenden System begründet auch die Wahl von Saarbrücken als idealen (deutschen) Standort für meine geplante Forschungsgruppe. Das weitere Vorgehen orientiert sich dann einerseits an den bereits bestehenden Forschungsaktivitäten im expandierenden ΩMEGA-Umfeld (Saarbrücken, Birmingham, Pittsburgh bzw. Bremen) andererseits aber insbesondere auch an den bestehenden, engen Kooperationen im CALCULEMUS und MKM Netzwerk — u.a. mit Edinburgh (Alan Bundy), Linz (Bruno Buchberger), Eindhoven (Arjeh Cohen), und Genua (Alessandro Armando). Das Augenmerk liegt dabei auf der Integrierbarkeit und Austauschbarkeit der insgesamt entstehenden Systeme.

Aus dem großen mittelfristig geplanten Ziel möchte ich mit diesem Antrag die folgenden konkreten Arbeitspunkte herausgreifen:

AP I (Grundlagen) Semantik und Mechanisierung von HOL

AP II (Systementwicklung 1) Beweiser für HOL

AP III (Systementwicklung 2) Agentenbasierte Integration und Koordination

AP IV (Anwendung) Semantische Mediatoren (basierend auf HOL)

AP V (Dissemination) Einführende Literatur und soziologische Ziele

Deutschland hat heute auf dem Gebiet der Deduktionssysteme eine internationale Führungsposition inne und der Aufbau einer eigenen Arbeitsgruppe im Rahmen des Aktionsplans Informatik würde es mir ermöglichen den HOL Ansatz weltweit erstmals in ein mathematischen Assistenzsystem mit den skizzierten Eigenschaften zu integrieren. Dies würde zudem eine nachhaltige und von der ΩMEGA-Gruppe unabhängige Forschung zu diesem Thema garantieren.

2 Stand der Forschung, eigene Vorarbeiten

2.1 Stand der Forschung

2.1.1 Deduktionssysteme und Mathematische Assistenzsysteme

Formalisierung und Mechanisierung mathematischen Denkens und Problemlösens sind ein alter Menschheitstraum, nicht zuletzt wegen der Hoffnung diese Art der Formalisierung dann auf andere Bereiche des menschlichen Denkens übertragen zu können. In der griechischen Antike beschrieb vor allem Aristoteles die Schlußregeln des menschlichen Denkens und hinterließ uns eine begrenzte Theorie des logischen Schließens bestehend aus Axiomen und Regeln (z.B. der Modus Ponens). Das Ziel der Formalisierung menschlichen Schließens, das allen Logikentwicklungen zugrunde liegt, wurde von Gottfried Wilhelm Leibniz im 18. Jahrhundert wieder aufgegriffen und um den Gedanken der Mechanisierbarkeit ergänzt. Seine Idee war die einer *Lingua Charakteristika* und eines *Calculus Ratiocinator*, also einer universellen Sprache zusammen mit einem allgemeinen Beweiskalkül, in der sich alle Probleme und Fragen durch *bloßes Ausrechnen* (*calculemus*, lat., imp. pl. — Lasst uns rechnen!) klären lassen könnten.

Auch der vorliegende Antrag ist durch diese Idee motiviert. Neben den großen Fortschritten in der mathematischen Logik mußten im vorigen Jahrhundert jedoch auch herbe Rückschläge hingenommen werden: So zeigen Gödels fundamentale Unentscheidbarkeits- und Unvollständigkeitsresultate [Gödel, 1931] die Grenzen von Hilbert's Programm, das heißt der Mechanisierbarkeit mathematischen Schließens, und aus dem Paradies — nämlich einen allumfassenden Kalkül, der in der Lage ist, sämtliche *Wahrheiten* einfach auszurechnen — sind wir vertrieben worden.

Auf der anderen Seite haben automatische Beweissysteme, basierend auf Herbrand's fundamentalem Resultat der Widerlegungsvollständigkeit und Semientscheidbarkeit [Herbrand, 1930] in einzelnen Anwendungsdomänen eine beachtliche Leistungsstärke erreicht und werden heute auch erfolgreich in industrienahen Anwendungen, wie der Programm- und Hardwareverifikation, eingesetzt. In der Mathematik selbst konnten nicht-triviale Sätze erstmals mit dem Computer bewiesen werden: Der Gruppe um Bill McCune am Argonne National Laboratory, Illinois, USA, gelang es 1997 das seit 60 Jahren unbewiesene Robbins Problem mit dem automatischen Beweiser EQP zu lösen [McCune, 1997] und viele fundamentale Theoreme der Mathematik und Metamathematik sind erstmals mit Rechnern im letzten Jahrzehnt bewiesen worden.

Aufgrund der enormen Komplexität automatischen Beweisens erscheint es aber wenig realistisch, daß automatische Beweissysteme bereits in wenigen Jahrzehnten eine den heutigen Schachcomputern[16] vergleichbare Leistungsfähigkeit im Vergleich zum Menschen erreichen werden. Allein durch *Brute-Force* Suche in Kombination mit schneller Hardware, wie in den meisten Schachcomputern, wird ein solcher Erfolg im automatischen Beweisen nicht wiederholbar sein. Vielmehr wird die Integration und Kombination leistungsfähiger traditioneller Deduktionssysteme für Logik erster Ordnung (z.B. Vampire [Rizanow & Voronkov, 2002, Riazanov & Voronkov, 2001], E-Setheo [Stenz & Wolf, 1999], SPASS [Weidenbach *et al*, 1999], OTTER [McCune & Wos, 1997, McCune, 1994], EQP [McCune, 1997], E [Schulz, 2002], WALDMEISTER [Hillenbrand & Löchner, 2002, Hillenbrand *et al*, 1999]) und höherer Ordnung (z.B. TPS [Andrews *et al*, 2000, Andrews *et al*, 1996], HOL [Gordon & Melham, 1993], PVS [Owre *et al*, 1996], NUPRL [Allen *et al*, 2000], ISABELLE/HOL [Paulson, 1994, Nipkow *et al*, 2002], COQ [COQ, 2003], ΩMEGA [Siekmann *et al*, 2003, Siekmann *et al*, 2002a, Siekmann *et al*, 2002b, Benzmüller *et al*, 1997]) mit Beweisplanungstechniken [Melis & Siekmann, 1999, Bundy, 1988] und anderen mehr kognitiv orientierten Verfahren notwendig sein, um entscheidende Fortschritte zu erzielen. Ebenfalls ist die Integration von Algorithmen und Spezialverfahren aus Computeralgebrasystemen

[16] Der Schachcomputer Deep Blue von IBM hat im Mai 1997 mit 3.5 zu 2.5 Punkten den Weltmeister Gary Kasparov geschlagen.

(z.B. MAPLE [Redfern, 1996], MATHEMATICA [Wolfram, 1988, Kaufmann, 1992]), Constraint-Lösern und Entscheidungsprozeduren wichtig.

Eine Kontrolle für die automatische Koordination solcher Integration kann beispielsweise auf der Beweisplanungsebene erreicht werden, siehe z.B. [Meier *et al*, 2002, Cohen *et al*, 2003]. Dieser Ansatz erlaubt es, zunächst grobe Beweispläne aufbauend auf einem Satz theoriespezifischer Beweismethoden zu entwickeln, um diese dann anschließend zu detaillierten Beweisen auf Kalkülebene zu verfeinern. Die integrierten Subsysteme können die Beweisplanung unterstützen, indem sie Unterbeweise führen, Nebenrechnungen ausführen oder die Erfüllbarkeit von Constraints überprüfen. Auch bei der Verfeinerung abstrakter Beweispläne können die eingebauten Subsysteme zur Lösung wichtiger Unterprobleme eingesetzt werden.

Die Weiterentwicklung und Perfektionierung des traditionellen automatischen Beweisen, insbesondere im weitgehend vernachlässigten Bereich der Logik höherer Ordnung, ist demnach ein wichtiges Unterziel bei der Entwicklung eines leistungsfähigen Assistenzsystems für die Mathematik, das auch nichttriviale Probleme autonom lösen kann.

Seit de Bruijn's Pionierprojekt AUTOMATH [de Bruijn, 1968, de Bruijn, 1980] gibt es auf dem Gebiet der mathematischen Assistenzsysteme verschiedene ambitionierte Projekte zur Automatisierung bzw. Semi-Automatisierung der mathematischen Wissensverarbeitung und des mathematischen Beweisens. Zu den wichtigsten gehören NUPRL, COQ, HOL, PVS, ISABELLE/HOL, TPS und die Saarbrücker Systeme ΩMEGA und VSE [Autexier *et al*, 2000]. Eine treibende und visionäre Kraft im Gebiet ist auch Bruno Buchberger und sein THEOREMA Projekt am RISC Linz [Buchberger, 2001a, Buchberger, 2001b]; die Idee das mathematische Assistenzsysteme zukünftig auch die Exploration neuen mathematischen Wissens unterstützen sollen wird erstmals in [Buchberger, 1999, Buchberger, 2000, Buchberger, 2003] vorgeschlagen. Die zur Zeit beeindruckendste Sammlung formal repräsentierten und bewiesenen mathematischen Wissens hat das polnische MIZAR-Projekt[17].

2.1.2 Logik höherer Ordnung

In der Logik wurden zahlreiche formale Repräsentationsformalismen für die Mathematik entwickelt, untersucht und angewendet, wobei die wichtigste Unterscheidung die zwischen dem Prädikatenkalkül erster Ordnung und Logiken höherer Ordnung ist.

Logiken erster Ordnung lassen sich effizient implementieren (z.B. ist die zentrale Operation der Unifikation entscheidbar und linear lösbar), während dies für Logiken höherer Stufe leider nicht gilt. Daher wurden in den vergangenen vierzig Jahren in erster Linie Deduktionssysteme für Logiken erster Stufe entwickelt und HOL galt als zu schwierig.

Ein typisches Beispiel eines formalen Repräsentationsformalismus für die Mathematik in Logik erster Ordnung ist die axiomatische Beschreibung der Mengenlehre, z.B. die Axiomatisierungen von Zermelo-Fränkel oder von Neumann-Bernays-Gödel. Neben der wenig natürlichen Repräsentation und der mangelhaften Intuitivität in diesem Ansatz gibt es weitere grundsätzliche Probleme, z.B. hinsichtlich der Mechanisierung von Mengenkomprehension, der Extensionalität oder der mathematischen Induktion, die eine Ausdrucksmächtigkeit höherer Stufe erfordert. Die Komprehensionsaxiome und Extensionalitätsaxiome müssen dem Suchraum bei der Beweissuche in axiomatischer Mengenlehre explizit hinzugefügt werden und um Konsistenz zu gewährleisten, müssen die Komprehensionsaxiome zusätzlich eingeschränkt und eine Mengen-Klassen Unterscheidung eingeführt werden.

Einen weitaus expressiveren Ansatz zur Formalisierung bietet die klassische Logik höherer Ordnung — auch klassische Typtheorie genannt. Diese Sprache baut auf Alonzo Church's einfach getypten λ-Kalkül [Church, 1940] auf und reichert die Signatur um logische Konnektive mit fester Semantik an; eine

[17] `http://mizar.org`

Quantifikation über Funktions- und Mengenvariablen (beliebigen Typs in der Typhierarchie) ist in dieser Sprache möglich. In ihr lassen sich bereits ganz einfache mathematische Sachverhalte, wie zum Beispiel die Assoziatitivität der funktionalen Verknüpfung, sehr natürlich ausdrücken: $\forall F_{\alpha\to\beta}, G_{\beta\to\gamma}, H_{\gamma\to\delta}\textbf{.}(F \circ G) \circ H = F \circ (G \circ H)$; wobei $\circ$ wiederum definiert ist als $\circ = \lambda F_{\alpha\to\beta}, G_{\beta\to\gamma}, X_{\alpha}\textbf{.}G\ (F\ X)$. Dieser Ansatz zur Formalisierung der Mathematik und zur Lösung der Grundlagenproblematik (Russel u.a.) erscheint heute der vielversprechendste. Eine empfehlenswerte Einführung in die Logik höherer Ordnung geben die Textbücher von Peter Andrews [Andrews, 2002, Andrews, 1986], dem Schüler von A. Church, der in diesem Jahr den prestigeträchtigen *Herbrand Award* für seine Arbeiten zur Mechanisierung von Logik höherer Ordnung erhalten hat.

Ein alternativer Ansatz zur klassischen Logik höherer Ordnung ist der *Calculus of Constructions* [Coquand & Huet, 1985], ein getypter λ-Kalkül für intuitionistische Logik höherer Ordnung. Diese Sprache kann auch als Symbiose aus Girard's System F_w (einem getypten λ-Kalkül höherer Ordnung) und einem getypten λ-Kalkül erster Ordnung im Stile von deBruijn's AUTOMATH oder Martin-Löf's intuitionistischer Typtheorie [Martin-Löf, 1994] angesehen werden. EineÜberblick zu dieser Forschungsrichtung gibt Herman Geuvers in [Geuvers, 2003] und Systeme die auf diesem Ansatz basieren, wie NUPRL und COQ, werden in [Barendregt & Geuvers, 2001] vorgestellt.

Zu den weiteren Ansätzen, die hier nicht weiter vertieft werden, gehören auch Quine's *New Foundations* [Van Quine, 1973, Holmes, 1999] und das Beweissystem Watson [Holmes & Foss, 2001].

In meinen bisherigen Arbeiten habe ich mich auf die klassische Logik höherer Ordnung im Sinne von Church und Andrews konzentriert, wohlwissend das eine gegenseitige Befruchtung mit alternativen Ansätzen zu erwarten ist und entsprechend untersucht werden sollte.

Theoretische Grundlagen, Semantik Ein wesentliches Charakteristikum der Standardsemantik für klassische Logiken höherer Ordnung ist, daß jedem funktionalen Typ $\alpha \to \beta$, der in der getypten Sprache betrachtet wird, das jeweils volle Funktionsuniversum gebildet über den zugrundeliegenden Universen für die Typen α und β zugeordnet wird. Die Wahl der Universen zu den gewählten Basistypen (üblicherweise ι und o, wobei als Universum für o die zweielementige Menge der Wahrheitswerte gewählt wird) legt demnach auch bereits alle Funktionsuniversen für die funktionalen Typen in dieser Sprache fest. Der Übergang zur Henkin Semantik bei der Analyse von Kalkülen ist durch die Gödelschen Unvollständigkeitsresultate [Gödel, 1931] motiviert, die aufzeigen, daß in einer klassischen Logik höherer Ordnung mit Standardsemantik keine vollständigen Kalküle möglich sind. Leon Henkin hat den auf vollen Funktionsuniversen aufbauenden Begriff der Standardsemantik in [Henkin, 1950, Henkin, 1996, Andrews, 1972] verallgemeinert, indem er auch partielle Funktionsuniversen zuläßt; allerdings unter Einhaltung der *Denotatpflicht*, die sicherstellt, daß die gewählten Universen reichhaltig genug sind, um jedem Ausdruck der Sprache auch ein entsprechendes Denotat zuzuordnen. Die Henkin'sche Verallgemeinerung hat zur Folge, daß sich die Menge der allgemeingültigen Formeln verringert, weil mehr Gegenmodelle konstruiert werden können, und diese Generalisierung geht genau soweit, daß vollständige Kalküle möglich werden.

Die Henkin Semantik erfüllt die Eigenschaften der funktionalen Extensionalität (zwei Funktionen sind genau dann gleich wenn sie punktweise gleich sind) und Boole'schen Extensionalität (zwei Aussagen sind genau dann gleich wenn sie äquivalent sind) und eignet sich deshalb grundsätzlich auch als semantischer Bezugsrahmen für die Mathematik. Leider erweist sich eine direkte Vollständigkeitsanalyse von Kalkülen mit der Henkin Semantik als äußerst schwierig und eine vereinheitlichte Methodik für solche Analysen fehlte bisher. Die abstrakte Konsistenzmethode für υ-Komplexe von Peter Andrews [Andrews, 1971] ist eine Adaption von Smullyan's abstrakter Konsistenzmethode [Smullyan, 1963] für die erste Ordnung und war bisher das einzige, oftmals aber inadäquate Hilfsmittel. υ-Komplexe bilden im Vergleich zur Henkin Semantik eine weitere starke Abschwächung der Bedingungen zur Modellkonstruktion und die beiden Extensionalitätsprinzipien gelten beispielsweise nicht mehr.

Durch die eigenen Vorarbeiten [Benzmüller, 1999a, Benzmüller *et al*, 2003a, Benzmüller *et al*, 2003b], und aktuelle Arbeiten von Chad Brown (CMU, Pittsburgh, USA) [Brown, 2003] wird diese Lücke jedoch geschlossen: Die Methode der abstrakten Konsistenz (die syntaktische Kriterien als Werkzeug zur Vollständigkeitsanalyse bereitstellt und dadurch eine weitaus kompliziertere direkte semantische Argumentation hinfällig macht) steht nun nicht nur für die Henkin Semantik zur Verfügung sondern für eine ganze Landschaft von Semantiken zwischen Henkin Semantik und Andrews' υ-Komplexen. Diese Semantiken sind motiviert durch unterschiedliche Einschränkungen hinsichtlich Gleichheit und Extensionalität. Diese aktuellen Vorarbeiten sind nicht nur als semantischer Bezugsrahmen für die Mathematik geeignet, sondern auch als Bezugsrahmen für andere Einsatzgebiete der Logik höherer Ordnung wie Anwendungen in der Linguistik, dem Gebiet der Programmiersprachen und der Programmverifikation. Diese Gebiete zeichnen sich untereinander und gegenüber der Mathematik gerade durch unterschiedliche semantische Anforderungen hinsichtlich Gleichheit und Extensionalität aus (siehe die Beispiele in der Einleitung zu [Benzmüller *et al*, 2003a]).

Kalküle Die Geschichte zur Mechanisierung von Logiken höherer Ordnung ist fast so alt wie das Gebiet der Deduktion selbst, das Mitte der 50er Jahre mit ersten Systementwicklungen begann.

Alan Robinson entwickelt in [Robinson, 1968, Robinson, 1969] einen der ersten Ansätze zur Mechanisierung der Logik höherer Ordnung. Dieser basiert auf dem Tableauxverfahren und greift Ideen von Schütte [Schütte, 1960] und Takeuti [Takeuti, 1953] auf. Wichtige Pionierarbeiten folgten dann kurze Zeit später mit Peter Andrews' Vorschlag zur Resolution höherer Ordnung [Andrews, 1971], Jensen und Pietrzykowski's Ansatz [Jensen & Pietrzykowski, 1972] und speziell Gerard Huet's *Constrained Resolution* [Huet, 1972, Huet, 1973a].

Eine der größten Herausforderungen für die Mechanisierung von Logiken höherer Ordnung ist die Unentscheidbarkeit und Komplexität der Unifikation höherer Stufe [Lucchesi, 1972, Huet, 1973b, Goldfarb, 1981]. Erste nicht-vollständige Verfahren zur vollen Unifikation höherer Ordnung werden in [Gould, 1966, Darlington, 1971, Ernst, 1971] präsentiert und die ersten vollständigen Verfahren gehen zurück auf [Huet, 1972] und auf [Jensen & Pietrzykowski, 1972]. Ein generelles Problem für die Mechanisierung ist, daß unendliche viele allgemeinste Unifikatoren für ein Unifikationsproblem in höherer Ordnung existieren können.

Während Andrews' Resolutionsansatz noch keine Unifikation vorsieht, schlägt Huet in seiner *Constrained Resolution* eine geschickte Richtung ein: statt die potentiellen Unifikationspaare beim Resolvieren sofort zu lösen werden diese, als Unifikationsconstraints kodiert, der Resolvente hinzugefügt. Ihre Lösung wird verzögert bis zum Ende der Beweissuche, d.h. bis eine leere Klausel erzeugt wurde, die nur noch durch ein zu lösendes Unifikationsproblem beschränkt ist. Huet's Ansatz profitierte dann zusätzlich von seiner Entdeckung der Prä-Unifikation höherer Ordnung. Diese vermeidet das Instanzenraten in *FlexFlex*-Situationen[18] und bietet dennoch ein vollständiges und semi-entscheidbares Verfahren zur Beantwortung der Lösbarkeit der aufgesammelten Unifikationsconstraints.

In den 80er Jahren hat Peter Andrews die Konnektionsmethode [Andrews, 1976, Andrews, 1981, Bibel, 1983] als Ausgangspunkt zur Mechanisierung der Logik höherer Ordnung in seinem TPS-Projekt entwickelt. Unifikation und Prä-Unifikation höherer Ordnung haben sich dabei als leistungsfähige Deduktionswerkzeuge erwiesen und ihre Mächtigkeit im Rahmen der Automatisierung einfacher mathematischer Theoreme wird beispielsweise durch [Andrews *et al*, 1996] belegt.

Der Resolutionsansatz wurde erst Anfang der 90er wieder aufgegriffen. Michael Kohlhase integriert in seiner Dissertation [Kohlhase, 1994] Sorten in die Logik höherer Ordnung und entwickelt eine sortierte Unifikation höherer Ordnung, die er dann als Motor seines sortierten Resolutionskalküls einsetzt.

[18] Die *FlexFlex*-Regel ist die problematischste, weil unendlich verzweigende Regel in der Unifikation höherer Ordnung. Sie verwendet blindes Raten bei der Unifikation zweier Terme mit jeweils freien Variablen an Kopfposition.

Die Dissertation von Kohlhase (sowie [Kohlhase, 1995]) lieferte auch erste Anstöße zur Verbesserung der Extensionalitäts- und Gleichheitsbehandlung, sowie zur Adaption des Beweisprinzips der abstrakten Konsistenz für Henkin Semantik; diese Probleme wurden dann vom Antragsteller in seiner Dissertation gelöst.

Bei zwei Kernproblemen zur Automatisierung von Logiken höherer Ordnung wurden in den vergangenen Jahren erhebliche Fortschritte gemacht: (I) in der zielgerichteten Mechanisierung von Gleichheit und Extensionalität, und (II) der zielgerichteten Instanziierung von Mengenvariablen. Diese beiden bisher ungelösten Probleme führten bis dato dazu, daß das Schließen in Logiken höherer Ordnung einen sehr hohen Anteil an blinder Beweissuche hatte. Problem (I) wird in der Dissertation des Antragstellers [Benzmüller, 1999a] behoben: statt die Extensionalitätsaxiome blind dem Suchraum hinzuzufügen, erfolgen gegenseitige und zielgerichtete Aufrufe zwischen Unifikation und Beweissuche. Problem (II) wird aktuell in der Dissertation von Brown adressiert [Brown, 2002]: statt eine blinde Aufzählung von Formelschemata für Mengenvariablen durchzuführen (wie in den Systemen TPS und LEO) sammelt ein Constraintsystem Bedingungen über die zu instanziierenden Mengenvariablen auf, generiert daraus zielgerichtet Instanziierungsvorschläge und propagiert diese dann an das Beweissystem. Beide Arbeiten haben gezeigt, daß die Anteile blinder Suche in der Mechanisierung von Logiken höherer Ordnung drastisch reduziert werden können.

Interessanterweise ergeben sich insbesondere bei vielen einfachen Problemen positive Effekte für den Suchraum durch den expressiveren Repräsentationsformalismus. In der naiven Mengentheorie beispielsweise erweisen sich einfache Probleme nach Expansion der enthaltenen Definitionen höherer Ordnung oft als rein propositionale Probleme, während sie in einer Kodierung in axiomatischer Mengenlehre häufig nicht-triviale Suchprobleme ergeben. Mit dem Beweiser LEO [Benzmüller & Kohlhase, 1998b] wurde dies durch eine große Anzahl von Theoremen der Mengenlehre (der TPTP Datenbank des CASC-Beweiserwettbewerb[19] für Beweiser erster Ordnung) unter Verwendung höherstufiger Kodierungen demonstriert; siehe Abbildung 1 für eine Illustration der Beweissuche in LEO.

Systeme Die wichtigsten mathematischen Assistenzsysteme, die auf klassischer Logik höherer Ordnung, aufbauen sind HOL, ISABELLE/HOL, und PVS. Bei den Systemen die eine vollständige Automatisierung anstreben nimmt das TPS System [Andrews *et al*, 2000, Andrews *et al*, 1996], das an der CMU unter der Leitung von Prof. Peter Andrews seit Beginn der 80er Jahre entwickelt wurde, eine führende Stellung ein. Der Kalkül von TPS basiert auf der Konnektions- bzw. “Mating”-methode [Andrews, 1976, Andrews, 1981, Bibel, 1983].

Als eigene Vorarbeit des Antragstellers wurde der automatische Beweiser LEO entwickelt [Benzmüller & Kohlhase, 1998b]; dies ist der erste maschinenorientierte Ansatz für die Logik höherer Ordnung, der eine zielgerichtete Gleichheits- und Extensionalitätsbehandlung einführt und dessen Kalkül Henkinvollständigkeit erreicht.

Die eindrucksvollste Leistungsbilanz zur praktischen Verwendung von Systemen für die Logik höherer Ordnung wird auf der internationalen TPHOLs Konferenz[20] vorgestellt. Der Schwerpunkt dieser Konferenz sind Anwendungen von HOL im Bereich der formalen Methoden mit interaktiven und semiautomatischen Beweisumgebungen wie PVS, HOL, ISABELLE/HOL, oder NUPRL. Ein besonders eindrucksvolles Beispiel für den praktischen Einsatz von HOL bieten die Arbeiten von John Harrison (z.B. zur Verifikation von Fließkommaoperationen) bei INTEL [Harrison, 2000].

[19] `http://www.cs.miami.edu/~tptp/CASC/`; in dem jährlich die besten Beweissysteme auf einer Datenbank von einigen tausend Testbeispielen (TPTP) im Wettbewerb getestet werden und das beste System mit einem Preis ausgezeichnet wird.

[20] `http://tphols.informatik.uni-freiburg.de/`

Die Beweissuche im Kalkül der extensionalen Resolution höherer Stufe in LEO soll anhand der Distributivität von Schnitt und Vereinigung illustriert werden:

$$\forall B_{\alpha\to o}, C_{\alpha\to o}, D_{\alpha\to o} \centerdot B \cup (C \cap D) = (B \cup C) \cap (B \cup D)$$

In [Ganzinger & Stuber, 2003] gehört dieses Theorem (welches in TPTP unter der Referenz *set171+3* enthalten ist) zur Menge der Evaluationsbeispiele für den neuen Ansatz der *Superposition with Equivalence Reasoning and Delayed Clause Normal Form Transformation*, einer Erweiterung der Superposition erster Stufe in Richtung eines nicht-Normalform Systems; dieses ist motiviert durch die offensichtliche Schwäche[a] der Beweiser erster Stufe für solche Probleme. Der erweiterte Superpositionsansatz übertägt (wie man leicht erkennt) Ideen des Paramodulationsansatzes aus [Benzmüller, 1999b, Benzmüller, 1999a] in den Kontext der ersten Stufe und er ist implementiert im Saturate System. Dieses erzeugt 159 Klauseln bei der Beweissuche für obiges Theorem und benötigt 2.900ms auf einem 2Ghz Notebook für den Widerlegungsbeweis.
Bei Kodierung in Logik höherer Stufe und Verwendung von LEO ergibt sich folgender Lösungsweg. Die Aussage wird negiert (Widerlegungsansatz) und die Definitionen $\cup = \lambda A_{\alpha\to o}, B_{\alpha\to o}, x_\alpha \centerdot (A\ x) \lor (B\ x)$ und $\cap = \lambda A_{\alpha\to o}, B_{\alpha\to o}, x_\alpha \centerdot (A\ x) \land (B\ x)$ werden expandiert. Die resultierende Formel wird dann normalisiert zur Klausel C_1 bestehend aus einem einzigen negierten Literal (in der Notation $[\ldots]^F$) — in diesem Fall eine negierte Gleichung zweier funktionaler Terme, d.h. ein extensionaler Unifikationsconstraint:

$$C_1 : [\lambda X_\alpha \centerdot (b\,X) \lor ((c\,X) \land (d\,X)) = \lambda X_\alpha \centerdot ((b\,X) \lor (c\,X)) \land ((b\,X) \lor (d\,X)))]^F$$

b, c, und d sind dabei neue Skolemkonstanten für die Variablen A, B und C.
Die zielgerichtete funktionale und Bool'sche Extensionalitätsbehandlung in LEO reduziert diese initiale Klausel zu

$$C_2 : [(b\,x) \lor ((c\,x) \land (d\,x)) \Leftrightarrow ((b\,x) \lor (c\,x)) \land ((b\,x) \lor (d\,x)))]^F$$

wobei x ist eine weitere neue Skolemkonstante ist.
Die sich anschließende Klauselnormalisierung ergibt dann eine rein propositionale, d.h. entscheidbare, Menge von Klauseln und nur diese Klauseln befinden sich nun noch im Suchraum von LEO. Insgesamt werden lediglich 33 Klauseln erzeugt und auf einem 2,5GHz schnellen PC werden 820ms für den Wiederlegungsbeweis benötigt, d.h. eine Repräsentation in einer getypten Logik höherer Stufe ist effizienter und nicht — wie häufig vermutet — aufwendiger als eine mengentheoretische Darstellung in erster Ordnung.

[a] Figure 1 in [Ganzinger & Stuber, 2003] zeigt, das dieses Problem durch den Gewinner des CASC Wettbewerbs 2002, Vampire 5.0, nicht gelöst wird und auch E-Setheo csp02 schafft den Beweis nicht.

Abbildung 1: Beweissuche in HOL für einfache Theoreme der naiven Mengenlehre.

Lehrbücher Peter Andrews' Buch [Andrews, 1986] und die vollständig überarbeitete Neuauflage [Andrews, 2002] sind derzeit die besten Lehrbücher für Logik höherer Ordnung; sie wenden sich aber eher an Graduierte bzw. erfahrene Wissenschaftler und sind als einführende Literatur nur bedingt geeignet. Die neuesten Entwicklungen auf dem Gebiet der Semantik für Logik höherer Ordnung sind noch nicht enthalten, ebensowenig wie der Aspekt maschinenorientierter Kalküle zur Automatisierung von Logik höherer Ordnung.

Während also für den Bereich der Logik erster Ordnung didaktisch hervorragend geeignete und das gesamte Gebiet — von den logischen Kalkülen bis zur Systementwicklung — umfassende Literatur in umfangreicher Zahl publiziert ist, besteht im Bereich der Logik höherer Ordnung noch ein starker Bedarf. Ohne geeignete Lehrbücher ist es insbesondere schwer den wissenschaftlichen Nachwuchs an das Gebiet heranzuführen bzw. zu gewinnen.

2.1.3 Integration von Beweis- und Berechnungssystemen

Die Integration verschiedener Spezialsysteme in ein heterogenes Gesamtsystem hat seit Mitte der 90er Jahre stark an Bedeutung gewonnen. Das EU Netzwerk CALCULEMUS[21] wurde mit dem Ziel gegründet, bessere mathematische Assistenzsysteme durch die Integration von symbolischem Schließen (Deduktion) und symbolischer Berechnung (Computer Algebra) zu entwickeln. Die möglichen Verfahren sind: (i) Kooperation von Deduktionssystem und Computeralgebrasystem auf gleicher Ebene, (ii) Einbettung eines Deduktionssystems in ein Computeralgebrasystem, bzw. umgekehrt, die (iii) Einbettung eines Computeralgebrasystems in ein Deduktionssystem. Um es in der Sprache von CALCULEMUS auszudrücken: (i) $DS \equiv CAS$, (ii) $DS \subseteq CAS$, oder (iii) $CAS \subseteq DS$.

Geht man von einem Beweissystem als Basissystem wie in (iii) aus, gibt es grundsätzlich zwei verschiedene Möglichkeiten zur Integration:

1. (Integration auf Kalkülebene) Direkte domänenspezifische Erweiterung des Kalküls eines Beweissystem durch Spezialregeln; d.h. durch den Einbau einer Theorie. Diese Methode wird vornehmlich für Systeme erster Ordnung angewendet und Beispiele sind die Erweiterung des Superpositionsansatzes durch den Einbau einer Shostak Theorie [Ganzinger *et al*, 2003], der *Theorem Proving Modulo* Ansatz [Dowek *et al*, 2001], oder die historische T-Resolution und die T-Unifikation [Baader & Siekmann, 1994, Bürckert *et al*, 1988, Stickel, 1981].

2. (Integration auf Systemebene) Die Alternative ist eine Integration domänenabhängiger Spezialsysteme. Ein Beispiel ist die am SRI entwickelte Verifikationsumgebung PVS [Owre *et al*, 1996], die viele domänenspezifische Entscheidungsprozeduren integriert. Auch im CALCULEMUS Netzwerk hat sich bisher der Integrationsansatz auf Systemebene als vielversprechender erwiesen. Beispiele sind das *Constraint Contextual Rewriting* CCR [Armando & Ranise, 2002], eine generalisierte Form von Rewriting, die eine effektive Plug & Play Integration von Entscheidungsprozeduren in die Formelsimplifikation ermöglicht, und das MathSat System [Giunchiglia *et al*, 2000, Audemard *et al*, 2002], eine verallgemeinerte Architektur zur Integration propositionaler Beweiser und Berechnungsalgorithmen zur effizienten Lösung Boole'scher und mathematischer Aussagen. Vorteile im Vergleich zu (1.) liegen vor allem in der Flexibilität des Ansatzes. Eine zusätzliche Herausforderung ist die Korrektheit der Integration, die z.B. durch eine Transformation der externen Repräsentation in den Repräsentationsformalismus des Basissystems erreicht werden kann, um dort eine Verifikation durchzuführen. Ein Beispiel ist das Elan System das solche Transformationen in COQ unterstützt [Nguyen *et al*, 2002]. In unserer eigenen Arbeitsgruppe wurden die Transformationssysteme TRAMP [Meier, 2000] und SAPPER [Sorge, 2000] entwickelt.

[21] `http://www.eurice.de/calculemus/`

Ein weiteres Beispiel für eine erfolgreiche Integration ist der in Edinburgh entwickelte und auf Induktion spezialisierte Beweisplaner $\lambda Clam$ [Richardson *et al*, 1998], der mit dem Beweissystem HOL gekoppelt wurde [Boulton *et al*, 1998]. Eine Einbettung von Deduktionssystemen in formale Softwareentwicklungsumgebungen wird mit dem Induktionsbeweiser INKA [Autexier *et al*, 1999] im VSE-tool [Autexier *et al*, 2000] realisiert. In einer Zusammenarbeit mehrerer europäischer Universitäten wurde im Kontext der formalen Methoden das PROSPER System [Dennis *et al*, 2000] entwickelt. Dabei handelt sich um eine offene Architektur zur Anbindung formaler Verifikationssysteme in industrielle CAV und CASE Werkzeuge.

Eine Übersicht über den derzeitigen Stand der Forschung in diesem Bereich liefern die Übersichtsdarstellungen zum CALCULEMUS-Projekt [Benzmüller, 2003c, Benzmüller, 2003a, Benzmüller, 2003b]. Um junge Wissenschaftler in dieses Forschungsgebiet einzuführen und zu begeistern wurde in 2002 erstmals die sehr erfolgreiche CALCULEMUS Herbstschule ausgerichtet [Benzmüller & (eds.), 2002a, Benzmüller & (eds.), 2002b, Benzmüller & (eds.), 2002c, Zimmer & (eds.), 2002].

Zur Unterstützung der Integration heterogener externer mathematischer Dienste, wie Deduktionssysteme, Computeralgebrasysteme, Modellgenerierer, und Spezialverfahren wurde in der eigenen Arbeitsgruppe der MATHWEB Software Bus [Franke & Kohlhase, 1999, Franke *et al*, 1999] entwickelt, an den heute mehr als 22 verschiedene mathematische Dienste gekoppelt sind und der bei Spitzenbelastung einige tausend Beweisprobleme pro Tag bearbeitet. Ein mit MATHWEB verwandtes System ist MetaPRL [22] [Hickey *et al*, 2003], das an der Cornell University entwickelt wird.

Die Integration heterogener Spezialsysteme wirft konsequenterweise die Frage nach automatischer Unterstützung für deren Koordination auf. Verschiedene internationale Forschungsprojekte stellen sich derzeit dieser Herausforderung: Das EU Projekt MONET[23] (Mathematics on the Net) konzentriert sich dabei auf die Koordination von Berechnungsverfahren in der Computer Algebra. Das ETI-Projekt[24] an der Universität Dortmund koordiniert Verifikationstools für Realzeitsysteme und Modelchecking. Das QSL Projekt von Claude und Hélène Kirchner am LORIA, Nancy, konzentriert sich ebenfalls auf den Verifikationskontext und strebt verbesserte Mechanismen für eine benutzerkontrollierte Koordination von Systemen an. Vom Grundgedanken her ähnlich ist auch die Forderung nach einem Semantic Grid [Roure, 2003] sowie das QPQ (QED Pro Quo) Projekt[25] am SRI, Stanford, USA. Letzteres plant den Aufbau einer großen internationalen Datenbank deduktiver Softwarekomponenten, u.a. mit dem Ziel eine verbesserte Wiederverwendbarkeit dieser Werkzeuge in mathematischen Assistenzsystemen zu erreichen.

2.1.4 Semantische Mediatoren

Während mathematische Assistenzsysteme bis Mitte der 90er Jahre noch als rein monolithische Systeme entwickelt wurden, die insbesondere auch ihr domänenspezifisches Wissen in systeminternen und von außen meist unzugänglichen Wissensbanken verwalteten, werden diese Wissensbanken in den vergangenen Jahren zunehmend ausgelagert und als eigenständige Module bereitgestellt. Ein Beispiel für diese Entwicklung ist die Auslagerung der Wissensbank des Beweissystems COQ und die Kooperation von COQ mit der Wissensbank HELM[26].

Auch MBASE, eine mathematische Wissensbank, die aus dem ΩMEGA-Projekt hervorging und nun durch die Arbeitsgruppe von Prof. Michael Kohlhase an der International University Bremen entwickelt wird, ist ein gutes Beispiel für diesen Trend. MBASE kooperiert zunehmend mit anderen Projekten und die weltweit größte mathematische Wissensbank MIZAR, sowie die Beweiser TPS und ΩMEGA können

[22] `http://cvs.metaprl.org`
[23] `http://monet.nag.co.uk/cocoon/monet/`
[24] `http://eti.cs.uni-dortmund.de/eti/servlet/ETISiteApplication`
[25] `http://www.qpq.org`
[26] `http://helm.cs.unibo.it/`

mittlerweile ihre Wisseneinheiten nach MBASE exportieren. Eine wichtige Bedeutung kommt solchen (mittelfristig im Internet verteilten und verbundenen) mathematischen Wissensbanken beispielsweise in zukunftsweisenden Forschungsinitiativen wie dem MKM Netzwerk zu [Buchberger *et al*, 2003]. Ein Ziel von MKM ist es, die Mathematik im 21. Jahrhundert insofern zu revolutionieren, als das derzeit noch vornehmlich in Papierform existierende mathematische Wissen zunächst in digitalisierter Form und später in strukturierter und semantisch zugänglicher Form in solchen mathematischen Wissensbanken bereitgestellt wird (das polnische MIZAR Projekt, das bereits seit Mitte der 80er existiert, kann als ein Vorreiter dieser Vision angesehen werden). Eine ähnliche Zukunftsvision unterliegt auch dem europäischen Euler-Projekt[27].

Für den Mathematiker ergibt sich damit die Perspektive leistungsfähige (semantische) Suchmaschinen und semantische Mediatoren beim Aufstöbern mathematischer Inhalte nutzen zu können und mathematische Assistenzsysteme werden dieses mathematische Wissen ebenso einsetzen können. Es sollte also zum Beispiel möglich sein, schwierige Beweisprobleme bzw. Unterprobleme durch Nachschlagen in einer mathematischen Wissensbank zu lösen anstatt sie immer wieder neu im Beweissystem direkt zu beweisen.

Damit stellt sich das bisher recht wenig untersuchte Problem, nämlich das mathematische Wissensbanken an möglichst redundanzfreier Verwaltung mathematischer Inhalte interessiert sind, während Probleme in Beweissystemen typischerweise in vielen verschiedenen aber semantisch äquivalenten Variationen auftreten können. Die Akquisition *passenden* mathematischen Wissens für einen konkreten Anwendungskontext kann (ähnlich zu deduktiven Softwarekomponenten) wiederum mit Hilfe kleiner spezieller Beweisverfahren über die logische Äquivalenz bzw. Implikation beider Aussagen gesucht werden.

Da sich die Mathematik in der Praxis typischerweise Repräsentationsformalismen höherer Ordnung bedient, liegt es also nahe semantische Mediatoren für die Logik höherer Ordnung zu entwickeln, die Mathematiker ebenso wie automatische Beweiser bei der Suche nach passenden Wisseneinheiten unterstützen.

Der Begriff eines Mediators wird in [Wiederhold, 1992] eingeführt und zwei konkrete Beispiele für Mediatoren sind TSIMMIS [Papakonstantinou *et al*, 1996] und KOMET [Calmet *et al*, 1997]. Eigene Vorarbeiten des Antragstellers sind [Benzmüller *et al*, 2003d, Benzmüller *et al*, 2001c]. Es ist zu erwarten, das sich die Forschungsaktivitäten in diesem jungen und bisher von europäischen Forschern dominierten Gebiet in den kommenden Jahren verstärken werden. Insbesondere werden e-Learning Systeme für die Mathematik entwickelt und in der Mathematikausbildung (*'Mathematikführerschein'*[28]) eingesetzt; solche Systeme können von reichen Wissensbanken mit formalisiertem mathematischen Wissen profitieren und sind oft auf eine direkte Kooperation mit diesen ausgelegt.

2.1.5 Praktische Anwendungen

Logik höherer Ordnung wird heute in den folgenden Gebieten genutzt und für die jeweils spezifischen Anwendungen entwickelt:

- Funktionale Programmiersprachen (wie ML[29], Haskell[30], oder OCAML[31]; siehe auch [Prehofer, 1998])
- Logik Programmierung höherer Ordnung (wie λ-Prolog[32]; siehe auch [Nadathur & Miller, 1994])
- Sprachverarbeitung (siehe z.B. [Kohlhase, 1998])
- Automatisches und Interaktives Theorem Beweisen (siehe Referenzen in diesem Antrag)

[27] `http://www.emis.de/projects/EULER/`

[28] `http://www.rp-online.de/news/wissenschaft/bildung/2003-2006/mathefuhrerschein.html`

[29] `http://www.smlnj.org/`

[30] `http://www.haskell.org/`

[31] `http://caml.inria.fr/`

[32] `http://www.cse.psu.edu/~dale/lProlog/`

- Formale Methoden, Verifikation (siehe Referenzen in diesem Antrag)
- Datenbanktheorien
- Komponentenbasierte Softwareentwicklung[33]

2.2 Eigene Vorarbeiten

Das hier beantragte Forschungsprojekt ist durch die aktuellen Entwicklungen und Forschungstendenzen im Bereich mathematischer Assistenzsysteme und interaktiver Beweisumgebungen und dem zunehmenden Interesse an Logik höherer Ordnung motiviert. Aus der Perspektive des Antragstellers bietet es die Chance, die verschiedenen persönlichen Forschungsinteressen und -aktivitäten der zurückliegenden Jahre in einem kohärenten und langfristig ausgelegten Gesamtansatz und -system zusammenzuführen. Das Projekt profitiert darüberhinaus von den nationalen und internationalen Forschungskonsortien, die der Antragsteller leitet bzw. an denen er beteiligt ist:

- CALCULEMUS[34] Research Training Network zur *Integration von Symbolischem Schließen und Symbolischer Berechnung* in mathematischen Assistenzsystemen.
  Funktion: Koordinator des Netzwerks und Leiter des Netzwerkknotens an der Universität des Saarlandes.
- MKM[35] (*Mathematical Knowledge Management*) Netzwerk; Ziel ist eine umfassende computerorientierte Neugestaltung der Verwaltung und Verarbeitung mathematischen Wissens im 21. Jahrhundert.
  Funktion: Leiter des Netzwerkknotens an der Universität des Saarlandes.
- Projekt ΩMEGA[36] zur *Ressourcenadaptiven Beweisplanung* im Sonderforschungsbereich SFB 378 *Ressourcenadaptive kognitive Prozesse* an der Universität des Saarlandes.
  Funktion: Antragsteller (gemeinsam mit Prof. Siekmann und Dr. Melis).
- Projekt DIALOG[37] zum Thema *Tutorieller Dialog mit einem mathematischen Assistenzsystem* im SFB 378.
  Funktion: Antragsteller (gemeinsam mit Prof. Pinkal und Prof. Siekmann).

Leistungsfähige automatische Beweisagenten für Logik höherer Ordnung werden positive Effekte auf diese Forschungsinitiativen und darüber hinaus haben. Deren Entwicklung und Integration wird aber in keinem dieser Projekte bisher direkt gefördert.

Im Folgenden werden die Vorarbeiten des Antragstellers weiter aufgeschlüsselt und nach den für das beantragte Projekte relevanten Forschungsbereichen skizziert.

2.2.1 Logik Höherer Ordnung

Die Dissertation [Benzmüller, 1999a] des Antragstellers untersucht Gleichheit und Extensionalität im automatischen Beweisen in Logik höherer Ordnung. Die betrachtete Sprache ist die klassische Typtheorie, d.h. wie auch für das Forschungsprojekt vorgesehen, eine Logik höherer Ordnung basierend auf Church's einfach getypten λ-Kalkül.

Zunächst werden in [Benzmüller, 1999a] unterschiedlich starke Semantikbegriffe für klassische Logiken höherer Ordnung erarbeitet, die durch die jeweils unterschiedlichen Rollen, die die Gleichheit in

[33] `http://clip.dia.fi.upm.es/COLOGNET-WS/`; area: *Componentbased Software Engineering*.
[34] `http://www.eurice.de/calculemus/`
[35] `http://monet.nag.co.uk/mkm/`
[36] `http://www.coli.uni-sb.de/sfb378/projects.phtml`
[37] `http://www.coli.uni-sb.de/sfb378/projects.phtml`

ihnen einnimmt, motiviert sind. Jedem der eingeführten Semantikbegriffe wird dann eine Menge abstrakter Konsistenzeigenschaften zugeordnet mit dem Ziel, die Analyse der Verbindung zwischen Syntax und Semantik von Beweiskalkülen für die klassische Logik höherer Ordnung zu erleichtern. Dieser Teil der Arbeit wurde in den vergangenen Jahren (und zuletzt in Kooperation mit Chad Brown, CMU) weiter verfeinert; siehe [Benzmüller *et al*, 2003b, Benzmüller *et al*, 2003a].

Der Hauptbeitrag der Dissertation sind drei neue Beweiskalküle: ER (extensionale Resolution höherer Ordnung), EP (extensionale Paramodulation höherer Ordnung) und ERUE (extensionale RUE-Resolution höherer Ordnung). Diskutiert werden diese Kalküle auch in [Benzmüller & Kohlhase, 1998a, Benzmüller, 1999b, Benzmüller, 1997] und mit früheren Ansätzen verglichen werden sie in [Benzmüller, 2002]. Idee dieser Kalküle ist es, die Mechanisierung definierter und primitiver Gleichheit in klassischer Logik höherer Ordnung zu verbessern. Im Gegensatz zu den zuvor in der Literatur diskutierten Widerlegungsansätzen erreichen diese Kalküle Vollständigkeit hinsichtlich der Henkin Semantik [Henkin, 1950], ohne dem Suchraum zusätzliche Extensionalitätsaxiome hinzufügen zu müssen.

Der Kalkül ER wurde im Beweissystem LEO implementiert und die Eignung dieses Systems zum Beweisen von Aussagen über Mengen wurde in Fallstudien gezeigt [Benzmüller & Kohlhase, 1998b, Benzmüller, 1999a].

Begünstigt wurden diese Ergebnisse durch einen Forschungsaufenthalt an der Carnegie Mellon University (CMU), Pittsburgh, USA, um dort von der ungebrochenen Tradition ürsprünglich von dem verstorbenen Alonzo Church, und heute von seinem Schüler Peter Andrews und dessen Schülern, insbesondere Frank Pfenning, zu profitieren. Diese Kooperation mit der CMU haben wir seither weiter ausgebaut (zuletzt mit Chad Brown) und sie führte zu den gemeinsamen Publikationen [Benzmüller *et al*, 2003a, Benzmüller *et al*, 2003b, Benzmüller *et al*, 2003c, Benzmüller *et al*, 1999a, Benzmüller *et al*, 1999c] und verschiedenen, bisher noch unveröffentlichten, Manuskripten.

Zu den eigenen Vorarbeiten gehören auch die Kalküle des natürlichen Schließens für die jeweiligen Semantiken [Benzmüller *et al*, 2003a], die zur Zeit durch Chad Brown um entsprechende Sequenzenkalküle ergänzt werden. Damit steht erstmals eine ganze Landschaft vollständiger Referenzkalküle für die verschiedenen Semantikbereiche der Logik höherer Ordnung bereit.

Hinsichtlich der Automatisierung von Logik höherer Ordnung gibt es jedoch noch zahlreiche offene Fragen. Ein sehr zentrales Problem, das durch die Dissertation des Antragstellers aufgeworfen wurde, betrifft das Zusammenspiel der *FlexFlex*-Regel [Snyder & Gallier, 1989] mit den neuen Extensionalitätsregeln in der Unifikation. Die starke Vermutung, daß die *FlexFlex*-Regel in diesem Kontext – wie in der Prä-Unifikation [Huet, 1975, Gallier & Snyder, 1989] – weiterhin vermieden werden kann, ist trotz großer internationaler Anstrengung noch nicht formal bewiesen worden. Eine weitere, noch nicht formal bewiesene Hypothese bezieht sich auf die Einschränkbarkeit der rekursiven Aufrufe aus der Unifikation an die Beweissuche auf Unifikationspaare mit Basistyp ι oder o. Auch die Reduzierbarkeit dieser Aufrufe auf Unifikationsconstraints bestehend aus einem einzigen Unifikationsliteral erscheint plausibel. Weitere solche den Suchraum einschränkende Hypothesen wurden zwar erarbeitet — ein formaler Nachweis der Vollständigkeit sowie deren Implementierung und Erprobung in der Praxis ist aber noch offen.

Weitere Fragen betreffen den gezielten Umgang mit der primitiven Substitution und das geschickte Zusammenspiel zwischen Kalkülen erster und höherer Ordnung; die Idee dabei ist Unterziele in der Beweissuche mit den jeweils besten und adäquaten Mitteln zu attackieren, also eine Art adaptierende Beweissuche zwischen erster und höherer Stufe zu ermöglichen. Das Problem der primitiven Substitution wird derzeit von Brown in Rahmen seiner Dissertation untersucht [Brown, 2002, Brown, 2003].

2.2.2 Integration von Beweissystemen

Der Antragsteller hat einen Ansatz zur Integration interaktiver und automatischer Beweissysteme auf Systemebene entwickelt, der sich die Konzepte des taktikbasierten Beweisens und der Beweisplanung, also einer abstrakten Kalkülebene, zunutze macht. Dabei werden die Kalkülregeln des externen Systems als Beweismethoden bzw. Beweistaktiken im zentralen System widergespiegelt und die eigentliche Beweistransformation wird durch die stufenweise Expansion der Planoperatoren bzw. ihrer Beweistaktiken realisiert.[38] Eine wesentliche Aufgabe der Integration besteht dann darin, die Expansionabbildung der einzelnen widergespiegelten Beweistaktiken zu definieren, aber umso mehr externe Systeme auf diese Weise in ein zentrales System integriert sind, desto einfacher wird diese Aufgabe. Denn sehr grob beschriebene oder sehr große Beweisschritte können im zentralen System bei Ihrer Expansion als offene Unterprobleme aufgefaßt und an ein geeignetes externes System weitergereicht werden, das im Erfolgsfall dann einen feinkörnigeren Beweisplan zurückliefert.

Dieser Ansatz erlaubt es also, in einer hierarchischen Vorgehensweise die Transformation eines externen Beweises in das interne Format eines zentralen Systems durch andere, bereits integrierte Systeme sinnvoll zu unterstützen. Ferner erfolgt die eigentliche Beweistransformation in transparenter Weise im zentralen System selbst.

Dieser Ansatz wurde angewendet und prototypisch erprobt im Rahmen der Integration des Beweisers höherer Ordnung TPS in das mathematische Assistenzsystem ΩMEGA und ist ausführlich beschrieben in [Benzmüller *et al*, 1999a].

2.2.3 Agentenbasiertes Theorembeweisen

Das Paradigma der Multi-Agenten Systeme hat die Künstliche Intelligenz und die moderne Informatik entscheidend beeinflußt und Multi-Agenten Systeme haben heute bereits viele Anwendungen in der industriellen Praxis gefunden [Weiss, 1999]. Intelligente Agenten sind pro-aktive Softwarepakete, die zur eigenständigen Lösung von Teilaufgaben geeignet sind und in der Lösungssuche kooperieren. Dazu müssen sie intelligent, kommunikativ, kooperativ und autonom sein.

Auf den Bereich der Deduktionssysteme übertragen, haben wir die Vision pro-aktiver Deduktionsagenten entwickelt, die heterogene Spezialsysteme kapseln und kooperativ ein mathematisches Theorem bearbeiten und zu beweisen versuchen. Zu dieser Forschungsrichtung hat der Antragsteller in Kooperation mit Volker Sorge durch die Entwicklung des Ω-ANTS-Systems beigetragen [Benzmüller & Sorge, 1998, Benzmüller & Sorge, 1999b, Benzmüller *et al*, 1999b, Benzmüller & Sorge, 1999a, Benzmüller *et al*, 2000, Benzmüller *et al*, 2001b, Benzmüller *et al*, 2001a, Benzmüller, 2001, Benzmüller & Sorge, 2002].

Zunächst entstand der Kommando-Vorschlagsmechanismus Ω-ANTS zur Unterstützung des Benutzers im interaktiven Beweisen [Benzmüller & Sorge, 1998, Benzmüller & Sorge, 1999a, Benzmüller & Sorge, 1999b]. Der Ansatz basiert auf einer zweistufigen Blackboard-Architektur und verfolgt das Ziel, die zum Teil sehr berechnungsintensiven (und potentiell unentscheidbaren) Suchaufgaben bei der Generierung von Kommando-Vorschlägen bis auf eine sehr feinkörnige Ebene voneinader zu entkoppeln und somit deren nebenläufige Realisierung zu ermöglichen. Dies wiederum begünstigt einen dynamischen Kommando-Vorschlagsmechanismus, der die gesamten Systemressourcen zwischen zwei aufeinander folgenden Benutzerinteraktionen ausnutzen kann, um nach geeigneten Vorschlägen (z.B. Anwendung einer Beweismethode oder -taktik, einer Kalkülregel oder eines externen Beweissystems) für das weitere Vorgehen zu suchen. Abstrakt gesehen ermöglicht dies dem mathematischen Assistenzsystem in eine Konkurrenz zum Benutzer zu treten und ihm dadurch besser zu assistieren.

[38] Zur Verständlichkeit muss bemerkt werden, daß hier nicht von Beweistaktiken im Edinburgh LCF-Stil die Rede ist [Gordon *et al*, 1979], die immer garantiert korrekte abstrakte Inferenzen beschreiben. Der Taktikbegriff umfasst hier auch solche abstrakten Inferenzen (Unterspezifikation), die teilweise auch inkorrekte abstrakte Beweisschritte zulassen. Solche inkorrekten abstrakten Inferenzen werden später bei der Expansion auf Basiskalkülebene entdeckt.

Ein wichtiges Charakteristikum dieses Ansatzes ist es, daß sämtliche Agenten (die *Knowledge Sources* der Blackboards) auch noch zur Laufzeit definiert und modifiziert werden können [Benzmüller & Sorge, 1999a]. Dies gilt auch für die Heuristiken, die die Priorisierung der generierten Vorschläge bestimmen. Insgesamt ergibt sich damit ein höchst flexibler, parametrisierbarer, reaktiver und logikunabhängiger Mechanismus.

Die ursprüngliche Intention des Vorschlagsmechanismus war die Unterstützung des Benutzers, das darüber hinausgehende Potential des Ansatzes – insbesondere im Hinblick auf die Automatisierung des taktikbasierten Theorembeweisens im Zusammenspiel mit externen Systemen – wurde dann aber schnell deutlich. Die Idee dazu ist recht einfach: Anstatt dem Benutzer die letztliche Auswahl aus der heuristisch geordneten Menge von vorgeschlagenen Inferenzen zu überlassen, kann dies auch die Maschine übernehmen, entweder völlig autonom oder in einem semi-automatischen Zusammenspiel von Maschine und Benutzer. Ein einfacher Kommando-Ausführungsmechanismus verbunden mit einem Backtrackingmechanismus bildet deshalb den Kern des agentenbasierten Beweisers Ω-ANTS, siehe auch [Benzmüller & Sorge, 2000, Benzmüller *et al*, 2001b].

Aus kognitionswissenschaftlicher Perspektive ergibt sich damit eine interessante Architektur: Sie unterstützt in einheitlicher Form benutzerorientierte abstrakte Inferenzen und eine theoretisch fundierte Beweissuche im Kalkül des natürlichen Schließens. Zudem können externe Systeme flexibel zur Laufzeit als abstrakte Inferenzen integriert werden. Interessant erscheint auch die Möglichkeit der Integration kritischer Beweisagenten, welche über die Beweiskonstruktion auf Objektebene reflektieren und die Steuerungsmechanismen des Gesamtsystems zur Laufzeit adaptieren könnten.

In einem einjährigen Forschungsprojekt an der Universität Birmingham hat der Antragsteller in Kooperation mit Manfred Kerber, Mateja Jamnik[39] (Birmingham) und Volker Sorge[40] (damals Saarbrücken) diese Ideen weiterverfolgen können und ein agentenbasiertes, reaktives und ressourcengesteuertes Zusammenspiel externer Unterstützungssysteme mit dem taktikbasierten Beweisen sowie dem Beweisen im Kalkül des natürlichen Schließens prototypisch realisiert, siehe auch [Benzmüller *et al*, 2001a, Benzmüller *et al*, 2001b, Benzmüller & Sorge, 2002, Benzmüller *et al*, 2000, Benzmüller *et al*, 1999b]

Zu den externen Unterstützungssystemen gehören verschiedene automatische Beweiser, ein Modellgenerierer und ein kommerzielles Computeralgebrasystem. Erste Fallstudien (beispielweise die automatische Klassifikation von mehr als 10.000 Mengengleichungen in Theoreme und Nichttheoreme) haben die Leistungsfähigkeit und das Potential dieses Ansatzes bestätigt [Benzmüller *et al*, 2001b]. Internationale Anerkennung fanden diese Forschungsleistungen u.a. durch einen eingeladenen Hauptvortrag (siehe [Benzmüller, 2001]) auf der britischen AISB Convention 'Agents and Cognition' in 2001 in York und in weiteren eingeladenen Vorträgen an der CMU und der Cornell University, USA.

3 Ziele und Arbeitsprogramm

3.1 Ziele

Das langfristige Ziel meiner Forschung ist die Entwicklung eines großen mathematischen Assistenzsystems und dessen Integration und internationale Verwendung im entstehenden mathematischen Semantischen Netz.[41]

Das unmittelbare Ziel des Projekts ist es Beweisagenten für die Logik höherer Ordnung zu entwickeln, deren Leistungsfähigkeit und Effizienz kontinuierlich zu verbessern, eine Koordinationsfähigkeit

[39] Mateja Jamnik ist inzwischen Lecturer in Cambridge.

[40] Volker Sorge ist inzwischen Lecturer in Birmingham.

[41] `http://www.win.tue.nl/dw/monet/`

mit anderen Beweiserkomponenten zu gewährleisten und diese exemplarisch für die Aλonzo/ΩMEGA-Systemumgebung zu realisieren. Mittelfristig soll dadurch dem übergeordneten Ziel *leistungsfähige und kommerziell verwertbare Assistenzsysteme für die Mathematik im entstehenden mathematischen semantischen Netz* zugearbeitet werden.

Die konkreten Einzelziele decken das Spektrum von der Grundlagenforschung bis hin zu ersten praktischen Anwendungen der daraus resultierenden Systeme ab.

Die Arbeitspakete I und II dienen der Verbesserung der Mechanisierung von Logik höherer Stufe in Theorie und Praxis. Die verbesserten Beweiser für HOL werden in den Arbeitspaketen III und IV in den anwendungsorientierten Kontext des entstehenden mathematischen semantischen Netzes eingebettet. Arbeitspaket V adressiert die Dissemination der Ergebnisse des Projekts.

AP I (Grundlagen) Semantik und Mechanisierung von HOL Einerseits besteht in verschiedenen Disziplinen (z.B. der Informatik, Mathematik, und Sprachverarbeitung) ein zunehmendes Interesses an Logik höherer Ordnung als intuitivem und ausdrucksmächtigem Repräsentationsformalismus. Andererseits gibt es jedoch noch immer viele offene Grundlagenprobleme hinsichtlich der Semantik und Mechanisierbarkeit. Im Arbeitspaket I dieses Projekts ist diese Diskrepanz Motivation und Forschungsinteresse und unter Ausnutzung bzw. Verstärkung bestehender internationaler Kooperationen (insbesondere durch Gewinnung von Chad Brown als Mitarbeiter in diesem Projekt) sollen diese Probleme bearbeitet werden.

AP II (Systementwicklung 1) Beweiser für die Logik höherer Stufe Implementierungen der verbesserten Mechanisierungsansätze sollen in den Systemen LEO (Resolutionsmethode) oder TPS (Konnektionsmethode) erfolgen. Effizienzsteigernde Techniken wie sie in Beweisern der ersten Stufe verwendet werden sollen auf Systeme für die Logik höherer Stufe übertragen werden. Die Transformation der maschinenorientierten Beweisformate dieser Beweiser in eine benutzerorientierte und intuitivere Repräsentation soll gewährleistet werden.

AP III (Systementwicklung 2) Agentenbasierte Integration und Koordination Ziel dieses Arbeitspakets ist es die in I entwickelten Verfahren in allgemeiner Form als leistungssteigernde Dienste auf Systemebene zur Verfügung zu stellen. In Anlehnung bzw. Kooperation mit aktuellen Arbeiten zum Aufbau eines mathematischen semantischen Netzes soll eine agentenbasierte Modellierung dieser Beweissysteme erfolgen, d.h. die Subsysteme sollen in der Lage sein als pro-aktive Agenten autonom und in verallgemeinerter Weise Beweisprobleme in unterschiedlichen System- und Problemkontexten zu lösen. Ferner soll die Koordinierbarkeit der Beweisagenten höherer Ordnung mit Beweisern erster Ordnung und mit Computeralgebrasystemen untersucht werden; dabei sind drei Steuerungsansätze interessant: (a) zentrale Steuerung durch das mathematische Assistenzsystem (z.B. durch Beweisplanung oder das Ω-ANTS-System), (b) dezentrale Steuerung in einem Netzwerk kooperierender Agenten (z.B. im Kontext der geplanten Erweiterung des MATHWEB-Systems [Zimmer, 2003]; dies erfordert eine enge Kooperation mit Edinburgh, Saarbrücken und Bremen, wo dieses System zur Zeit entwickelt wird), und (c) eine direkte, intern in den Beweisagenten höherer Ordnung gesteuerte Kooperation mit Spezialsystemen. Neben dem ΩMEGA-System sollen die Beweisagenten höherer Ordnung mittelfristig auch in anderen mathematischen Assistenzsystemen (z.B. in den Systemen des CALCULEMUS-Netzwerks) eingesetzt werden.

AP IV (Anwendung) Semantische Mediatoren (basierend auf HOL) Der aktuelle Aufbau großer mathematischer Wissensbanken soll durch die Bereitstellung semantischer Mediatoren für Logik höherer Ordnung unterstützt werden. Da die mathematische Praxis typischerweise expressive höherstufige Repräsentationen nutzt, werden auch die enstehenden Wissensbanken (wie MBASE es bereits tut) höherstufige Repräsentationen bereitstellen. Dazu soll eine unabhängige Architektur für semantische Mediatoren entwickelt werden, welche die Systeme aus Arbeitspaket I kapseln und mit dem MBASE-Projekt in der Praxis anwenden. Neben einem Einsatz in mathematischen Assistenzsystem und als semantische Suchmaschinen für den Mathematiker sollen die Mediatoren auch in entstehenden e-Learning Systemen eingesetzt werden (wie z.B. dem Saarbrücker ACTIVEMATH-System das MBASE als Wissensbank verwendet). Eine direkte Anwendung ergibt sich auch für das DIALOG Projekt im SFB 378.

AP V (Dissemination) Einführende Literatur und Soziologische Ziele Der Mangel an einführender Literatur zum Themengebiet soll durch zwei geplante Lehrbücher (Semantik und Kalküle zu HOL, OMEGA/AλONZO-System) verbessert werden.

Die Sichtbarkeit des Gebietes auf internationalen Konferenzen soll durch die Mitarbeit am geplanten Beweiserwettbewerb für Beweiser höherer Ordnung ähnlich dem existierenden jährlichen CASC-Wettbewerb für Beweiser erster Ordnung auf der CADE Konferenz verstärkt werden.

3.2 Konkretes Arbeitsprogramm

3.2.1 AP I (Grundlagen) Semantik und Mechanisierung von HOL

1. Semantik:

   (a) (Modellklassen) Aufbauend auf den eigenen Vorarbeiten sollen Modellklassen für HOL weiter untersucht und verfeinert werden. Die Semantiken in [Benzmüller, 1999a, Benzmüller *et al*, 2003a] sind motiviert durch unterschiedliche Einschränkungen hinsichtlich Gleichheit und Extensionalität; zusätzliche Parameter wie das Auswahlaxiom, der Descriptionoperator oder das Unendlichkeitsaxiom, die nicht ohne Bedeutung für die Mathematik sind, wurden bisher ausgeklammert.

   (b) (Abstrakte Konsistenz) Für jede der eingeführten Modellklassen sollen abstrakte Konsistenzbedingungen angegeben werden, die eine Vollständigkeitsanalyse von Kalkülen auf der Basis dieser syntaktischen Kriterien unterstützen. Durch die Vorarbeiten in [Benzmüller *et al*, 2003a] wurde dies bereits für verschiedene Modellklassen geleistet und exemplarisch auf Kalküle des natürlichen Schließens angewendet. Allerdings erfordern diese abstrakten Konsistenzbedingungen eine sehr starke Saturationseigenschaft, deren Nachweis Beweismethoden erfordert, die ebenso stark sind wie die der Cut-Elimination; siehe [Benzmüller *et al*, 2003c]. Deshalb sind diese Konsistenzbedingungen typischerweise auch nur für Kalküle mit Cut (z.B. dem Kalkül des natürlichen Schließens) problemlos anwendbar. Zur Unterstützung von Vollständigkeitsanalysen maschinenorientierter Kalküle muß deshalb die Saturationsbedingung wie in [Benzmüller *et al*, 2003c] vorgeschlagen durch entsprechend geeignetere Kriterien ersetzt werden. Hinweise hierzu können aus den Kalkülen zur primitiven Gleichheitsbehandlung EP und ERUE [Benzmüller, 1999a] gewonnen werden.

   (c) (Cut-Elimination) In diesem Kontext soll auch die Frage der Cut-Elimination in Logiken höherer Ordnung weiter untersucht werden. Insbesondere sollen unsere Resultate mit den Aussagen anderer Projekte (wie z.B. denjenigen des *Theorem Proving Modulo Ansatzes* [Dowek *et al*, 2001]) verglichen werden.

   (d) (Annotationen und Partialität) Es soll ein semantischer Bezugsrahmen geschaffen werden für annotierte Logiken in denen funktionale und Boole'sche Terme hinsichtlich der Gültigkeit/Nichtgültigkeit von Boole'scher und funktionaler Extensionalität gekennzeichnet werden können. In einer solchen Semantik könnten extensional ausgezeichnete Funktionen und nicht-extensional ausgezeichnete Funktionen koexistieren was in den bisherigen Modellklassen nicht möglich ist.

Ebenso soll versucht werden einen semantischen Bezugsrahmen zu entwickeln der partielle Funktionen unterstützt, etwa in Anlehnung an Arbeiten wie [Stump, 2003, Kerber & Kohlhase, 1997].

2. Kalküle, Extensionalität und Gleichheit

   (a) (Kalküle für Modellklassen) Für die im Projekt eingeführten Modellklassen sollen vollständige maschinenorientierte Kalküle (Resolution und Matrixverfahren) und interaktionsorientierte Kalküle (natürliches Schließen und Sequenzenkalkül) entwickelt werden.

   (b) (Verfeinerung, Extensionalität) Diese Kalküle (insbesondere die maschinenorientierten) sollen weiter verfeinert und durch Strategien angereichert werden. Verschiedene Hypothesen vollständigkeitserhaltender Einschränkungen (Vermeidbarkeit der *FlexFlex*-Regel, Rekursive Aufrufe aus Unifikation an Beweissuche nur für Basistypen, Rekursive Aufrufe an Beweissuche nur bei Unifikationsproblemen bestehend aus einem einzigem Unifikationsconstraint, etc.) liegen für die extensionale Resolution, wie im Abschnitt 2.2.1 angedeutet, bereits vor.

   (c) (Mechanisierung der Gleichheit) Ein wichtiger Aspekt ist die Verbesserung der Mechanisierung der Gleichheitsrelation. Dabei sollen parallel die Ansätze zur Paramodulation und Differenzreduzierung aus [Benzmüller, 1999a] und [Brown, 2003] weiter verbessert werden. Auch die geschickte Steuerung von Definitionsexpansionen soll weiter untersucht werden; beispielsweise ist die Arbeit zur *dualen Instanziierung* [Bishop & Andrews, 1998] noch nicht auf den extensionalen Kontext übertragen worden.

   (d) (Nicht-Normalform Kalküle) Der gegenwärtige Trend der Entwicklung maschinenorientierter Kalküle ohne Normalform ist auch für HOL ein konsequenter Schritt. In der extensionalen Resolution [Benzmüller, 1999a] sind Unifikation und Beweissuche durch die Möglichkeit gegenseitiger Aufrufe auf gleicher konzeptueller Ebene vereint und dadurch wurde in gewisser Weise eine extensionale Unifikation höherer Ordnung eingeführt, die sich bei der Unifikation von Formeln der Mächtigkeit des Beweisers bedienen kann. Die Angleichbarkeit der Formeln zweier zu unifizierender, nicht-normaler Literale liegt deshalb in deren Leistungsspektrum, während die Unifikation erster Ordnung die Mittel hierzu nicht bereitstellt. Dies ist auch eine Anforderung an die semantischen Mediatoren in Arbeitspaket IV: diese sollen ebenfalls die semantische Angleichbarkeit von Formeln analysieren.

   (e) (Parametrisierte Kalküle) Unter der Voraussetzung einer erfolgreichen Semantikkonstruktion in (1d) sollen Kalküle für annotierte Logiken entwickelt werden, in denen funktionale und Boole'sche Terme hinsichtlich der Gültigkeit/Ungültigkeit von Extensionalitätseigenschaften gekennzeichnet werden können. Idee ist, daß die Denotate zweier funktionaler Terme nur dann identifiziert werden müssen, wenn sie punktweise gleich sind *und* beide Terme als funktional extensional ausgezeichnet sind. In einer solchen Logik könnte man also beispielsweise gleichzeitig über *mathematische Funktionen* (in denen von ihrer konkreten Realisierung abstrahiert wird; d.h. nur ihr Ein-/Ausgabeverhalten ist interessant) und über *funktionale Programme* reden und schließen.

       Unter gleichen Voraussetzungen sollen Kalküle für Logiken mit partiellen Funktionen untersucht werden.

3. Mengenvariablen

   (a) (Mengenvariablen und Induktion) Aktuelle Arbeiten von Brown [Brown, 2002, Brown, 2003] definieren und untersuchen erstmals einen zielgerichteten Ansatz für die Instanziierung von Mengenvariablen, die für mathematische Anwendungen besonders wichtig und zudem fundamental für HOL sind. Dieses Problem ist auch als Problem der *Primitiven Substitution* bekannt. Bisherige Ansätze verfolgten eine Ratestrategie (*choose-and-check*) um diese Variablen a priori mit Formelschemata zu instanziieren; siehe z.B. [Andrews *et al*, 1996]. Diese a priori Herangehensweise ist in folgender Weise verwandt mit dem bekannten Ansatz der expliziten Induktion [Boyer & Moore, 1979, Walther, 1993]: In Logik höherer Ordnung kann das Induktionsprinzip axiomatisiert und dem Suchraum hinzugefügt werden. Das Finden und Anwenden eines geeigneten Induktionschemas (Induktionsvariablen und Induktionsordnung) beispielsweise entpuppt sich bei genauerer Betrachtung dann als ein Mengenvariablen-Instanziierungsproblem. Das Problem der Primitiven Substitution ist also äquivalent dazu, eine Induktionsordnung a priori zu raten. A posteriori Methoden, wie sie nun von Brown untersucht werden und die auf der Sammlung von Constraints beruhen, verallgemeinern in analoger Weise den Ansatz der impliziten Induktion [Wirth, 2003]. Diese Zusammenhänge sollen untersucht und domänenspezifische Strategien, die für die implizite oder explizite Induktion in Systemen der ersten Ordnung (wie INKA) entwickelt wurden, sollen nach Möglichkeit in den Projektkontext übertragen werden.

   (b) (Domänenspezifische Strategien) Es soll (analog zur Induktion) versucht werden weitere Anwendungsbereiche für domänen- und kontextabhängige Strategien zu identifizieren. Das mögliche/notwendige Zusammenspiel zwischen domänenspezifischen und domänenunabhängigen Vorgehensweisen und Strategien zur Instanziierung von Mengenvariablen soll untersucht und entsprechende Steuerungsmechanismen sollen entwickelt werden.

   (c) (Beispiel: Theorie der Gröbner Basen [Buchberger, 1985, Buchberger, 1992]) Die Hypothese von Bruno Buchberger, RISC und Uni Linz (Kooperationspartner in CALCULEMUS und ΩMEGA), daß es nur wenige kreative mathematische Prinzipien in der Mathematik gibt, die es jedoch fruchtbar miteinander zu kombinieren gilt, soll im Zusammenhang zur Mengenvariablen-Instanziierungs-Problematik untersucht werden. Als Entwickler der Theorie der Gröbner Basen hat Bruno Buchberger bei der Modellierung mathematischer Theorien große Erfahrungen gesammelt. Im Kontext der Logik höherer Ordnung betrachtet, bilden sich die kreativen Aspekte dieser Modellierungsaufgabe gerade wieder ab auf das Problem der geschickten Instanziierung von Mengenvariablen. Die Frage ist, ob sich die von Bruno Buchberger vermuteten mathematischen Prinzipien identifizieren lassen und entsprechend durch domänenspezifische oder domänenunabhängige Vorgehensweisen und Strategien bei der Mengenvariablen-Instanziierung modellieren lassen.

   (d) (Übertragung auf andere Kalküle) Die zielgerichteten Mechanisierungsansätze zur Mengenvariablen-Instanziierung sollen auch auf andere Kalküle übertragen werden. Die Arbeiten von Brown entstehen im Kontext von Sequenzenkalkül und Konnektionsmethode. Sie sollen zumindest auf den Resolutionskalkül ER, den Paramodulationskalkül EP, den Differenzreduzierungskalkül ERUE, bzw. deren beabsichtigte Verfeinerungen übertragen werden.

3.2.2 AP II (Systementwicklung 1) Beweiser für HOL

1. (System Implementierungen) Die entwickelten maschinenorientierten Kalküle sollen in den Beweissystemen LEO (Resolutionsverfahren) und TPS (Matrixverfahren) realisiert und getestet werden. Eine Studie soll die beiden Ansätze vergleichen und mit der Situation in Logik erster Ordnung

in Beziehung setzen. In der Logik erster Ordnung sind zur Zeit Resolutionsverfahren dominierend; allerdings lassen sich deren essentielle Techniken wie Termindexing und Literal-Selektion nicht ohne weiteres auf die Logik höherer Ordnung übertragen.

2. (Effizienzsteigernde Techniken) Erfolgreiche, effizienzsteigernde Techniken für Beweiser erster Ordnung — Termindexing [Ramakrishnan *et al*, 1998], starke Literal-Selektionsfunktionen wie im Superpositionsansatz, usw. — sollen nach Möglichkeit auf Logiken höherer Ordnung übertragen werden. Allerdings sind dabei erhebliche Probleme zu erwarten: Beispielsweise ist Termindexing höherer Ordnung modulo extensionaler Gleichheit schwer realisierbar, weil dies i.A. volles Theorembeweisen innerhalb des Indexing erfordert. Auch die Adaption starker Literal-Selektionsfunktionen für Logik höherer Ordnung erscheint schwierig, da Wohlordnungskriterien, wie $\forall x.F > \{x \leftarrow T\}F$ für alle Terme T der betrachteten Logik (Instanzen quantifizierter Formeln sind immer kleiner als die Originalformel) in HOL Probleme bereiten: Die HOL Formel $\forall x_o.x_o$ hat als Instanz $\{x_o \leftarrow \forall x_o.x_o\}x_o$, welche zur Ausgangsformel $\forall x_o.x_o$ reduziert.

   Ansätze zum Termindexing höherer Ordnung in einem nicht-extensionalen Kontext werden zur Zeit durch Brigitte Pientka, CMU, im Rahmen ihrer Dissertation untersucht [Pientka, 2003]; es soll untersucht werden, ob sich Anwendungen dieser Arbeit für den extensionalen Kontext ergeben.

3. (Beweistransformationen) Wichtig für die Integration der Beweiser höherer Ordnung in mathematische Assistenzsysteme ist die Transformierbarkeit der maschinenorientierten Beweisrepräsentation (Resolution oder Matrix) in die typischen benutzerorientierten Beweisrepräsentationen im mathematischen Assistenzsystem (z.B. natürliches Schließen oder Sequenzenkalkül und deren Hochtransformation auf den Assertion-Level [Huang, 1994] und in Natürliche Sprache [Fiedler, 2001b]). Die Motivationen ist: (a) Ein Basissystem wie etwa ΩMEGA erfordert die Übersetzung der Beweise höherer Ordnung in das eigene Beweisrepräsentationsformat, um sie dort zu verifizieren und dem Gesamtbeweis hinzuzufügen, (b) Beweiserklärungssysteme wie P.rex [Fiedler, 2001a], die in der Lage sind Beweise in natürlichsprachliche Repräsentationen zu überführen, erfordern typischerweise eine vorgeschaltete Transformation in eine benutzerorientierte Beweisrepräsentation.

   Es soll deshalb versucht werden die Transformationsmechanismen von TPS (Konnektionsbeweise höherer Ordnung in natürliches Schließen) und von TRAMP (Resolutionsbeweise erster Ordnung in natürliches Schließen) für die im Projekt entwickelten Kalküle zu erweitern. Alternativ kann eine Erweiterung des in den Vorarbeiten (Abschnitt 2.2.2) skizzierten taktikbasierten Transformationsansatzes gewählt werden.

   Bessere Repräsentationen sind vor allem deshalb von Bedeutung, weil der getypte λ-Kalkül höherer Ordnung eine ungewohnte und für den Alltagsgebrauch zu umständliche und aufwendige Syntax hat (dies betrifft konkret den λ-Binder und die Typannotationen). Dies wird von vielen Kritikern gern als Argument gegen Church's Formalismus benutzt, obwohl die prinzipiellen Vorteile anerkannt werden.

3.2.3 AP III (Systementwicklung 2) Agentenbasierte Integration und Koordination

1. (Modellierung als Agenten) Die Systeme LEO und TPS sollen als pro-aktive Agenten modelliert exemplarisch in das mathematische Assistenzsystem ΩMEGA integriert werden. Dort sollen sie autonom Unterziele die möglicherweise in Ihrem Leistungspektrum liegen erkennen und attackieren. Diese Integration erfordert eine enge Kooperation mit dem ΩMEGA-Projekt. Als Infrastruktur für die agentenbasierte Modellierung sollen die aktuellen Arbeiten [Zimmer, 2003] am MATHWEB-System (Saarbrücken, Edinburgh, Bremen) aufgegriffen werden.

2. (Koordination) Möglichkeiten zur Kooperation der Beweisagenten höherer Ordnung mit Beweisern der ersten Ordnung und Computeralgebrasystemen sollen untersucht werden. Drei Steuerungsansätze zur Koordination kommen in Betracht: (a) zentrale Steuerung durch ein mathematisches Assistenzsystem (z.B. durch die Beweisplanung und das Ω-ANTS-System), (b) dezentrale Steuerung in einem Netzwerk kooperierender Agenten (z.B. im Kontext der geplanten Erweiterung des MATHWEB-Systems [Zimmer, 2003]), und (c) eine direkte, intern in den Beweisagenten höherer Ordnung gesteuerte Koordination von Spezialsystemen.

   Hinsichtlich Möglichkeit (a) kann auf den eigenen Vorarbeiten aus Abschnitt 2.2.3 oder auf Arbeiten unserer Gruppe wie [Meier *et al*, 2002, Cohen *et al*, 2003] aufgebaut werden.

   Hinsichtlich Möglichkeit (b) bietet sich eine Kooperation mit den aktuellen Arbeiten zur Koordination von MATHWEB-Diensten an [Zimmer, 2003].

   Möglichkeit (c) wurde auf der Basis verteilter bzw. nebenläufiger Architekturen für Resolutionsbasierte Verfahren in Beweisern der ersten Ordnung bereits erfolgreich angewendet; eine Taxonomie verteilter Beweiser erster Ordnung wird beispielsweise in [Bonacina, 2000] diskutiert und agentenbasierte Architekturen in [Denzinger & Fuchs, 1999, Dahn & Denzinger, 1998, Fisher & Ireland, 1998, Wolf, 1998, Fisher, 1997]. Im Bereich Logiken höherer Ordnung ist ein wesentlich größerer Erfolg solcher Architekturen zu erwarten, weil diese weitaus mehr Aspekte aufweisen die nebenläufige Realisierung nahelegen: Logik höherer Ordnung ist von vielen zusätzlichen Verzweigungspunkten in der Beweissuche geprägt bzw. geplagt — z.B. hinsichtlich der Instanziierung von Mengenvariablen oder der grundsätzlich infinitären Unifikation.

3. (White-Box Integration) Unter Ausnutzung der Arbeit in APII(3) soll eine Transformation der Beweisresultate der Beweisagenten höherer Ordnung in den Beweisrepräsentationsformalismus des Basissystems unterstützt werden. Die konkrete Durchführung der Transformation ist aber erfahrungsgemäß sehr zeitaufwendig (evtl. aufwendiger als die Beweissuche selbst) und soll deshalb von den Beweisagenten nur auf Nachfrage und vom eigentlichen Beweisprozess entkoppelt erfolgen.

4. (Benutzerunterstützung) Die Benutzer des mathematischen Assistenzsystems sollen durch die Beweisagenten höherer Ordnung besser unterstützt werden. Dabei sollen die Beweisagenten nicht in direkte Kommunikation mit dem Mathematiker treten, sondern im Hintergrund versuchen partielle Beweise zu erzeugen, die (nach vorausgegangener Beweistransformation; APIII(3)) in eine natürliche Repräsentation für den Benutzer überführt werden.

3.2.4 AP IV (Anwendung) Semantische Mediatoren (basierend auf HOL)

1. (Problemanalyse) Retrieval von Information mit Hilfe von Semantik und Ontologien (ähnlich dem *Semantic Web*) soll aus einer strukturierten mathematischen Wissensbank erfolgen. Dabei geht es nicht nur darum, eine einzige Formel (z.B. ein fokussiertes Ziel in einem mathematischen Assistenzsystem) mit einem Theorem in der Wissensbank *syntaktisch* zu vergleichen, sondern wie das folgende Beispiel zeigt, eine *semantische* Beziehung mit Hilfe von Systemen höherer Ordnung herzustellen. Der semantische Vergleich umfaßt dabei folgende Aspekte: (i) Möglicherweise sind logische Umformungen notwendig um den Zusammenhang zwischen fokussiertem Beweisziel und einem Theorem in der Datenbank zu erkennen; ein Theorem der abstrakten Form $\forall\overline{X}.\, A \Leftrightarrow B$ kann eben nicht syntaktisch auf einen äquivalenten Beweiskontext wie $\forall\overline{X}.\, A \Rightarrow B \wedge (\neg A \Rightarrow \neg B)$ abgebildet werden. (ii) Die Variablen des Theorems in der Datenbank können dabei mit Untertermen aus dem fokussierten Beweisziel instanziiert werden und umgekehrt können (iii) freie Variablen (Meta-

variablen) aus dem fokussierten Ziel mit Untertermen aus dem Theorem belegt werden. Diese drei Aspekte müssen in Kombination und mit Hilfe von Deduktionswerkzeugen gelöst werden.

Zur Illustration aller drei Aspekte nehmen wir ein einfaches Theorem der naiven Mengenlehre in der Datenbank an wie

$$\forall M_{\alpha\to o}, N_{\alpha\to o}.M \setminus N = \emptyset \Leftrightarrow M \subseteq N$$

wobei die Mengensubtraktion $\setminus$ und die leere Menge $\emptyset$ definiert sind als $\lambda U_{\alpha\to o}, V_{\beta\to o}, Z_{\alpha}.(U\ Z) \wedge \neg(V\ Z)$ und $\lambda Z_{\alpha}.\bot$, und α eine Typvariable ist.

Als Beispiel für ein fokussiertes Beweisziel betrachten wir

$$\begin{array}{c}
(\{X_{nat}|X\ Quadratzahl\} \subseteq \{X_{nat}|X\ gerade\} \Rightarrow \\
\qquad A = \{X_{nat}|X\ Quadratzahl\} \setminus \{X_{nat}|X\ gerade\}) \\
\wedge \\
(\{X_{nat}|X\ Quadratzahl\} \setminus \{X_{nat}|X\ gerade\}) = A \Rightarrow \\
\qquad \{X_{nat}|X\ Quadratzahl\} \subseteq \{X_{nat}|X\ gerade\})
\end{array}$$

wobei die Mengen $\{X_{nat}|X\ Quadratzahl\}$ und $\{X_{nat}|X\ gerade\}$ kodiert sind als $\lambda X_{nat}.(Quadratzahl\ X)$ und $\lambda X_{nat}.(gerade\ X)$, und A eine im automatischen Assistenzsystem zuvor eingeführte Metavariable vom Typ $nat \to o$ sein soll. Dieses Ziel kann einerseits von einem Beweiser der Logik höherer Stufe direkt gelöst werden (insbesondere der zweite Konjunkt ist trivial); dabei müßte die Instanziierung $\emptyset$ (eingeschränkt auf Typ nat) für die Metavariable A synthetisiert werden.

Andererseits kann dieses Ziel aber auch durch das (zuvor unabhängig bewiesene) Theorem in der Datenbank geschlossen werden. Dazu müssen die Variablen M und N mit den konkreten Mengen $\{X_{nat}|X\ Quadratzahl\}$ und $\{X_{nat}|Xgerade\}$ instanziiert werden (siehe (ii)) und die Metavariable A mit $\emptyset$ und die Typvariable α mit nat (siehe (iii)). Zusätzlich muß die logische Äquivalenz (bzw. die Implikation von Theorem auf das Ziel) unter dieser Instanziierung analysiert werden (siehe (i)). Es ist leicht einzusehen, daß diese drei Probleme nicht unabhängig gelöst werden können und den Einsatz eines Beweisers wie LEO bedingen. Ein Vorteil bei diesem Vorgehen ist, das die Instanziierung der Metavariable A nicht synthetisiert werden muss, sondern per Matching/Unifikation gegen das Theorem ermittelt werden kann. Für die Mechanisierung bedeutet dies eine (partielle) Vermeidbarkeit des Mengenvariablen-Instanziierungsproblems.

Das fokussierte Beweisziel befindet sich ebenso wie die Theoreme der Datenbank in einem Theoriekontext. Wenn die Übereinstimmung/Verträglichkeit dieser Kontexte aber nicht bereits a priori durch einfache Mechanismen gewährleistet ist, ergeben sich möglicherweise zusätzlich Beweisprobleme für den semantischen Mediator. Diese Problematik soll näher untersucht und ein entsprechendes Verfahren gefunden werden, das solche Theoriebezüge in das semantische Retrieval von mathematischem Wissen einbeziehen kann und ein Auffinden der mathematischen Sachverhalte ermöglicht, so wie es heute von den Suchmaschinen im Netz bzw. den semantikorientierten Suchmaschinen des zukünftigen *Semantic Web* bereitgestellt wird.

2. (Architektur) Auf der Grundlage der obigen Analyse soll eine Architektur für semantische Mediatoren für das Szenario entworfen werden. Diese Architektur soll nach Möglichkeit die effizienten syntaktischen und beschränkt semantischen Filtermechanismen die von solchen mathematischen Wissensbanken zum Retrieval von Information typischerweise zur Verfügung gestellt werden, berücksichtigen. Hierzu bietet sich eine zweistufige Architektur an wie in [Benzmüller *et al*, 2003d]

vorgeschlagen. Eine konkrete Zusammenarbeit ist dabei mit dem MBASE Projekt von Michael Kohlhase an der International University Bremen, der CMU, und dem DFKI (Semantic Web) geplant. Eine Kooperation mit dem DFKI ist insbesondere auch hinsichtlich des MAYA-Projekts[42] [Autexier *et al*, 2002]) interessant. Als Vorarbeit steht eine Mediatoren-Modellierung basierend auf dem eigenen Ω-ANTS-System zur Verfügung.

3. (Nebenläufigkeit) Mathematische Wissensbanken enthalten heute bereits einige tausend Theoreme und für die Zukunft kann ein enormes Wachstum angenommen werden. Auch bei starken vorgeschalteten Filtern werden sich für die Mediatoren noch viele Theorem-Kandidaten ergeben. Für jeden dieser Kandidaten ist also eine i.A. unentscheidbare Frage hinsichtlich ihrer Anwendbarkeit auf das fokussierte Problem im mathematischen Assistenzsystem zu untersuchen. Dies motiviert eine nebenläufige Suche nach anwendbaren Theoremen. Optimalerweise kann wie im Ω-ANTS-Vorschlagsmechanismus ein Anytime-Verhalten erreicht werden, bei dem sich das vorgeschlagene anwendbare mathematische Wissen heuristisch ständig verbessert und das bei Unterbrechung und späterer Wiederaufnahme der Suche nicht zwangsläufig alle Informationen neu berechnet.

4. (Ressourcenadaptive Steuerung) Es sollen ressourcenadaptive Steuerungsmechanismen für die nebenläufigen Mediatoren entwickelt werden, welche die verfügbaren Systemressourcen, Informationen über die im Kontext maximal erwünschte Bearbeitungszeit, sowie verfügbares Kontextwissen aus dem mathematischen Assistenzsystem berücksichtigen. Insbesondere sollen Techniken wie sie heute in Suchmaschinen (z.B. Google) verwendet werden entsprechend übertragen werden.

5. (Erweiterte Verwendung) Einsatzmöglichkeiten im entstehenden mathematischen semantischen Netz sollen über MBASE und ΩMEGA hinausgehend aufgezeigt und unterstützt werden. Kooperationen sind beispielsweise geplant im MKM- und dem CALCULEMUS-Netzwerk.

6. (Suchmaschine für die Mathematik) Die Mediatoren wurden bisher aus der Perspektive eines mathematischen Assistenzsystems als Klienten diskutiert. In analoger Weise sollen direkte Problemanfragen durch Mathematiker im Sinne einer semantischen Suchmaschine für die Mathematik unterstützt werden. In diesem Kontext erhält die Problematik der notwendigen Klärung des Theoriebezugs der Anfrage eine zusätzliche Qualität.

3.2.5 AP V (Dissemination) Einführende Literatur und Soziologische Ziele

- (Literatur) Zwei einführende Lehrbücher (über die Semantik und Kalküle zu HOL und über das AλONZO/ΩMEGA-System) sollen die Ergebnisse des Projekts zusammenfassen und das Heranführen junger Wissenschaftler in das Arbeitsgebiet unterstützen.

- (Tutorien) Geplant ist auch die Durchführung von Tutorien und Kursen auf Sommerakademien, Sommerschulen, und Konferenzen.

- (Beweiserwettbewerb) Die entwickelten Beweissysteme höherer Ordnung sollen in dem geplanten internationalen Beweiserwettbewerb eingesetzt und mit anderen Beweisern höherer Ordnung und erster Ordnung verglichen werden. Dazu soll beim Aufbau eines Beweiserwettbewerbs ähnlich dem jährlichen CASC-Wettbewerb für Beweiser erster Ordnung mitgearbeitet werden. Eine vorgelagerte Aufgabe besteht in der Zusammenstellung einer Datenbank von Beweisproblemen in Logik höherer Ordnung; dies soll in Kooperation mit dem MBASE-Projekt von Michael Kohlhase erfolgen. Um auch Vergleichsanalysen mit Beweisern der ersten Ordnung zu unterstützen sollen die Einträge

[42] `http://www.dfki.de/~inka/maya.html`

in dieser Datenbank möglichst auf die ihnen entsprechenden Problemkodierungen der im CASC-Wettbewerb verwendeten TPTP Datenbank verweisen.[43]

3.2.6 Vorgehen und Zeitplan

Das skizzierte Arbeitsprogramm ist — insbesondere in Bezug auf Arbeitspaket I — sehr ambitioniert und eine vollständige Lösung all der skizzierten Forschungsfragen kann natürlich nicht mit Gewißheit prognostiziert werden. In diesen Forschungsfragen geht der Antrag von einer weiterhin funktionierenden bzw. noch intensiveren Kooperation mit anderen Experten des Gebietes aus (z.B. mit dem TPS-Projekt von Peter Andrews, CMU oder mit Michael Kohlhase, IUB). Insbesondere die personelle Besetzung der Mitarbeiterstelle für Arbeitspaket I wird einen entscheidenden Einfluß auf die zu erwartenden Resultate haben, weil bereits die Einarbeitung eines begabten, im Gebiet aber wenig erfahrenen Mitarbeiters, einen sehr langen Zeitraum beanspruchen kann.

Die Arbeitspakete III, IV, und V sind dagegen weitgehend unabhängig von dieser Problematik: AP III und IV verwenden zwar die Systeme LEO und/oder TPS diese stehen in ihren derzeitigen Versionen aber bereits zur Verfügung und sollen durch AP I und AP II kontinuierlich verbessert werden. Unter Einhaltung geeigneter Kommunikationsschnittstellen sollte deshalb ein periodischer Austausch der verbesserten Systeme in den Arbeitspaketen III und IV möglich sein.

Die geplante zeitliche Abfolge der Forschungsleistungen wird in Abbildung 2 illustriert.

3.2.7 Anmerkung

Ausgehend von der langfristigen Vision und der mittelfristigen Zielsetzung des Antragstellers, wurden in diesem Antrag die Projektziele zkizziert und die geplanten Arbeitspakete allgemein beschrieben. Auf einer weiteren Konkretisierungsebene müssten nun eigentlich die Arbeitspakete für den Fachexperten technisch spezifiziert werden; dies würde jedoch über den von der DFG vorgegebenen Seitenrahmen hinausgehen. Ist eine weitere Detaillierung dennoch erwünscht, werde ich dem auf Anfrage gerne Folge leisten.

[43] Erste Gespräche dazu wurden bereits mit Geoff Sutcliff (dem Organisator der CASC Wettbewerbe), Michael Kohlhase (zur Bereitstellung der Problemformulierungen mittels MBASE), Chad Brown und Peter Andrews geführt.

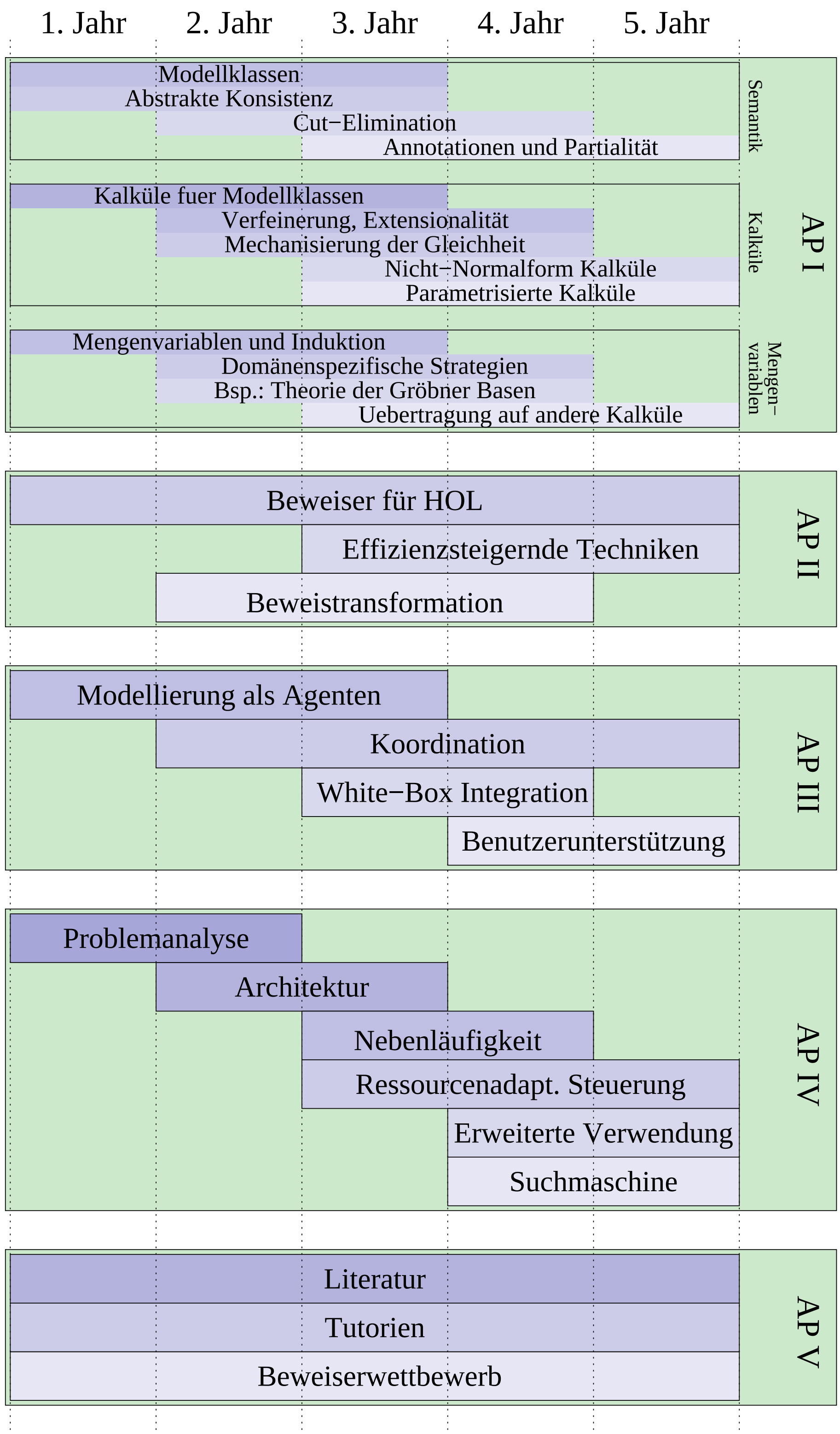

Abbildung 2: Projektzeitplan.

4 Beantragte Mittel

4.1 Personalbedarf

1. eine Projektleiterstelle (BaT Ia oder Juniorprofessur[44]) für die gesamte Dauer des Vorhabens.

2. drei wissenschaftliche Mitarbeiterstellen BAT IIa für die gesamte Dauer des Vorhabens.

   Begründung:
   Mitarbeiter A (Chad Brown der sein Interesse signalisiert hat; siehe Anlage 4) soll in enger Zusammenarbeit mit dem Antragsteller an den Arbeitspaketen I, II und IV arbeiten. Die Resultate des Projekts in AP I sind jedoch schwer prognostizierbar. Ein Mitarbeiter der mit der Thematik weniger vertraut ist, würde eine lange Einarbeitungszeit beanspruchen, was zu Verzögerungen im Projekt führen kann. In diesem Fall müssen die Erwartungen entsprechend angepaßt werden.

   Mitarbeiter B (N.N.) soll unter der Leitung des Antragsteller das Arbeitspaket III bearbeiten und bei II mitarbeiten.

   Mitarbeiter C: (N.N.) soll unter der Leitung des Antragsteller das Arbeitspaket III bearbeiten und bei II mitarbeiten.

3. drei wissenschaftliche Hilfskräfte
   19 Stunden pro Woche für die gesamte Dauer des Vorhabens

   Begründung: In den Arbeitspaketen II, III, und IV soll eine wissenschaftliche Hilfskraft bei der Implementierung von Systemkomponenten und der Durchführung von Experimenten mitarbeiten. Für den anspruchsvollen Bereich der Logik höherer Ordnung lassen sich meist nur geeignete Studenten mit Abschluß (Bachelor oder Diplom) rekrutieren, deshalb die Beantragung *wissenschaftlicher* Hilfskräfte.

4. eine studentische Hilfskraft,
   19 Stunden pro Woche für die gesamte Dauer des Vorhabens

   Begründung: Ein reibungsloser Forschungsbetrieb der Arbeitsgruppe und die umfangreicher Implementierungsarbeiten setzten eine funktionierende Systemadministration für die Gruppe voraus. Dies soll durch eine studentische Hilfskraft unterstützt werden.

4.2 Verbrauchsmaterial

Gewährleistung der Grundausstattung durch die Universität des Saarlandes.
Darüberhinausgehend werden 3 Notebooks für den gesamten Förderungszeitraum beantragt. Für Projektpräsentationen auf Konferenzen und bei Projektpartnern stellen diese mittlerweile ein *unverzichtbares Arbeitsmittel* dar.

Bedarf[45]: 3×3000 Euro

[44] Die Frage ob BaT Ia oder Juniorprofessur soll nach evtl. Bewilligung mit der Universität des Saarlandes verhandelt werden.

[45] Ein Angebot liegt dem Antrag bei; Anlage 3

4.3 Reisen

Durch das Reisekostenbudget sollen abgedeckt werden:

1. Publikation von Projektresultaten auf internationalen Konferenzen; zu diesen gehören

   (a) IJCAI: International Joint Conference on Artificial Intelligence

   (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) verschiedene kleinere oder nationale KI Konferenzen und Workshops: KI, AIMSA, EPIA, etc.

   Die Statistiken der vergangenen Jahre innerhalb der ΩMEGA-Gruppe und die Präsentationen des Antragstellers (siehe Liste der Vorträge unter `http://www.ags.uni-sb.de/~chris/cv/Chris-Academic-Experience.html`), zeigen daß für die Mitarbeiter ca. 2 Konferenzteilnahmen pro Jahr und für den Projektleiter (Antragsteller) ca. 3 Konferenzen angesetzt werden sollten. Die Konferenzgebühren auf internationalen Konferenzen sind in den vergangenen Jahren enorm gestiegen und betragen oft zwischen 500 und 800 US Dollar. Beantragt werden deshalb:

Projektleiter:	5000 Euro/Jahr
Pro Mitarbeiter (aktive Teilnahme an ca. 2 Konferenzen pro Jahr):	$3 \times$ 3000 Euro/Jahr

2. Forschungsaufenthalte von Mitarbeitern/Studenten bei Projektpartnern: 3000 Euro/Jahr

3. Einladung von Gastwissenschaftlern 1500 Euro/Jahr

4.4 Publikationskosten

Keine.

4.5 Sonstige Kosten

Keine.

5 Voraussetzungen für die Durchführung des Vorhabens

5.1 Zusammensetzung der Arbeitsgruppe

- Dr. Christoph Benzmüller (Projektleiter, Universität des Saarlandes)
- Chad Brown (CMU) — (AP I, II, V; in Kooperation mit Antragsteller; siehe Anlage 4)
- Weitere potentielle Mitarbeiter sollen über die Job-Börse im Internet und die Studienstiftung rekrutiert werden; es gibt zur Zeit wieder eine wesentlich bessere Bewerbersituation.

5.2 Zusammenarbeit mit anderen Wissenschaftlern

Mit folgenden Wissenschaftlern soll auf nationaler Ebene eine konkrete Kooperation stattfinden:

- in AP I: mit Michael Kohlhase an der International University Bremen.
- In AP III: mit dem ΩMEGA-Projekt an der Universität des Saarlandes (Jörg Siekmann, Serge Autexier, Claus-Peter Wirth), dem Ω-ANTS-Projekt an der Universität des Saarlandes, dem MATHWEB-Projekt an der Universität des Saarlandes (Jürgen Zimmer, z.Zt. in Besuch der Gruppe von Alan Bundy in Edinburgh) und dem ACTIVEMATH-Projekt am DFKI (Erica Melis, Carsten Ullrich, Andreas Meier).
- In AP IV und AP V: mit dem MBASE-Projekt (Michael Kohlhase) an der International University Bremen, sowie dem Maya-Projekt am DFKI.

Eine Kooperation zur Thematik *kooperative Deduktionswerkzeuge* ist mit Christoph Kreitz in Potsdam geplant.

Ausbildungsaufenthalte von Mitarbeitern und Studenten bei den Kooperationspartnern bzw. gegenseitige Tutorien sind sehr wichtig und gemeinsame Arbeitstreffen sollen regelmäßig stattfinden.

5.3 Arbeiten im Ausland und Kooperation mit ausländischen Partnern

Mit folgenden Wissenschaftlern ist eine konkrete Kooperation auf internationaler Ebene geplant:

- in AP I: mit dem TPS-Projekt (Peter Andrews) und Frank Pfenning an der Carnegie Mellon University, USA und mit Bruno Buchberger in Linz, Österreich.
- In AP III: mit dem Ω-ANTS-Projekt (Volker Sorge, Manfred Kerber) an der University of Birmingham und mit der Gruppe von Alan Bundy in Edinburgh. Zudem ist eine Anlehnung an die Standards internationaler Initiativen (z.B. MONET-Netzwerk oder CALCULEMUS-Netzwerk) für Beschreibungen von mathematischen Diensten wichtig.
- In AP V: mit dem TPS-Projekt (Peter Andrews) und Geoff Sutcliffe in Miami zum Aufbau eines Beweiserwettbewerbs für Beweiser höherer Stufe.

Ausbildungsaufenthalte von Mitarbeitern und Studenten bei den Kooperationspartnern und gemeinsame Tutorien sind sehr wichtig; gemeinsame Arbeitstreffen sollen regelmäßig stattfinden (um Reisemittel zu sparen können dazu auch andere Mittel, wie beispielsweise CALCULEMUS-Netzwerktreffen genutzt werden).

5.4 Apparative Ausstattung

Die Arbeitsplatzrechner für die Arbeitsgruppe werden aus der Grundausstattung bereitgestellt.

5.5 Laufende Mittel für Sachausgaben

Sachmittel, Telefonkosten usw. für die Arbeitsgruppe werden aus der Grundausstattung bereitgestellt.

5.6 Sonstige Voraussetzungen

Keine.

6 Erklärungen

6.1 Erklärungen zur Stellung zu anderen Projekten

Der Antragsteller beabsichtigt im Falle einer Förderung die derzeitige Projektleitung der ΩMEGA-Gruppe von Prof. Siekmann abzugeben. Eine weitere Nutzung des ΩMEGA-Rahmensystems und eine inhaltliche Kooperation zur Bereitstellung und Integration leistungsfähiger Beweisagenten höherer Ordnung in die ΩMEGA-Umgebung ist jedoch sehr wichtig, da eine vollständige Systemneuentwicklung mit den hier beantragten Mitteln und in diesem Zeitrahmen völlig unrealistisch wäre. Große Deduktionssysteme haben heute Entwicklungszeiten über zehn Jahren und erfordern einen hohen Personal- und Entwicklungsaufwand in dieser Zeitspanne.

6.2

Ein Antrag auf Finanzierung dieses Vorhabens wurde bei keiner anderen Stelle eingereicht. Wenn wir einen solchen Antrag stellen, werden wir die Deutsche Forschungsgemeinschaft unverzüglich benachrichtigen.

6.3

Der Vertrauensdozent des Univ. des Saarlandes, Prof. Dr. Janocha wurde von der Antragstellung unterrichtet und eine Unterstützungszusage der Universität des Saarlandes ist diesem Antrag beigefügt.

6.4

Die Hochschule an der das Forschungsprojekt angesiedelt werden soll ist die:

Universität des Saarlandes

Begründung: Die Universität des Saarlandes zählt mit den renommierten Forschungsinstitutionen MPII und DFKI im Bereich der Informatik zu den national und international herausragenden Forschungsstandorten.[46] Im Bereich der Formalen Methoden, der Logik und der Deduktionssysteme ist die Saarbrücker Informatik zusammen mit dem Max Planck Institut für Informatik (MPII) und dem Deutschen Forschungszentrum für künstliche Intelligenz (DFKI) unbestritten weltweit führend, sowohl was die Bilanz und das Rating der wissenschaftlichen Arbeiten angeht[47] als auch hinsichtlich der Größe und Anzahl der angesiedelten Forschergruppen in diesem Gebiet.

In Kooperation mit dem MPII und dem DFKI wird das gesamte Spektrum von der Grundlagenforschung in der Logik, über anwendungsorientierte Forschung zur Automatisierbarkeit des logischen Schließens bis hin zu Anwendungen von Beweissystemen in den Bereichen Formale Methoden und e-Learning abgedeckt. Hinzu kommt, daß der von mir gewählte *Bottom-Up* Forschungsansatz ausgehend vom derzeitigen ΩMEGA-System eine enge Kooperation mit der bestehenden ΩMEGA-Gruppe von Prof. Siekmann

[46] Bei verschiedenen Evaluationen hat die Saarbrücker Informatik im deutschen Hochschulbereich den ersten Platz belegt und weltweit wird sie unter die ersten fünf Universitäten (Carnegie Mellon, Stanford, Berkeley, MIT, Edinburgh) eingereiht.

[47] Hierzu werden der Citation Index, die Anzahl der Saarbrücker Publikationen auf internationalen Tagungen, sowie Bücher und Zeitschriftenartikel Saarbrücker Autoren sowie das Drittmittelaufkommen als Maßstab genommen und evaluiert.

voraussetzt. Die ΩMEGA-Arbeitsgruppe von Prof. Siekmann ist derzeit die weltweit größte Forschergruppe im Bereich der mathematischen Assistenzsysteme und bietet durch ihre heterogene Projektlandschaft in diesem Arbeitsbereich eine hervorragende Ausgangsbasis für das Projekt. In der Grundlagenforschung zur Automatisierung der Logik nimmt die Arbeitsgruppe von Prof. Ganzinger am MPII eine internationale Spitzenstellung ein und bildet einen sehr guten lokalen Kooperationspartner für die Automatisierung der Logik höherer Ordnung. Hinsichtlich der Anwendungen mathematischer Assistenzsysteme ist das DFKI idealer Kooperationspartner mit seinen Arbeitsgruppen in den Bereichen e-Learning und den Anwendungen von Formalen Methoden/Verifikation in zahlreichen sicherheitsrelevanten Industrieprojekten.

Die wissenschaftlichen Interaktionsmöglichkeiten für das konkrete Forschungsvorhaben sind deshalb an jeder anderen Hochschule auch im internationalen Vergleich nicht so optimal.

Zudem leitet der Antragsteller seit 2001 von Saarbrücken aus das *EU Research Training Network* (RTN) CALCULEMUS. Das CALCULEMUS RTN wird im 5. Rahmenprogramm der EU gefördert und endet im September 2004. Eine Verlängerung bzw. ein Neuantrag im 6. Rahmenprogramm ist vom CALCULEMUS Konsortium beschlossen und der Antragsteller wurde erneut als Leiter und Koordinator unter Ausnutzung der erfolgreich aufgebauten Infrastruktur in Saarbrücken gewählt.

7 Unterschriften

Dr. Christoph Benzmüller

8 Anlagen

(1) Unterstützungszusage der Universität des Saarlandes

(2) Dissertation und fünf weitere ausgewählte Publikationen des Antragstellers in 2-facher Ausfertigung

(3) Angebot für Notebook

(4) Interessensbekundung von Chad Brown an dem Projekt mitzuarbeiten

Literatur

[Allen *et al*, 2000] Allen, S., Constable, R., Eaton, R., Kreitz, C. and Lorigo, L. (2000). The Nuprl open logical environment. In McAllester [McAllester, 2000].

[Andrews, 1971] Andrews, Peter B. (1971). Resolution in type theory. *Journal of Symbolic Logic*, 36(3):414–432.

[Andrews, 1972] Andrews, Peter B. (1972). General models and extensionality. *Journal of Symbolic Logic*, 37(2):395–397.

[Andrews, 1976] Andrews, Peter B. (1976). Refutations by matings. *IEEE Trans. Comp.*, C-25(8):801–807.

[Andrews, 1981] Andrews, Peter B. (April 1981). Theorem proving via general matings. *Journal of the Association for Computing Machinery*, 28(2):193–214.

[Andrews, 1986] Andrews, Peter B. (1986). *An Introduction to Mathematical Logic and Type Theory: To Truth Through Proof.* Academic Press.

[Andrews, 2002] Andrews, Peter B. (2002). *An Introduction to Mathematical Logic and Type Theory: To Truth Through Proof.* Applied Logic Series. Kluwer, Boston;Dordrecht;London.

[Andrews *et al*, 1996] Andrews, P., Bishop, M., Issar, S., Nesmith, D., Pfenning, F. and Xi, H. (1996). TPS: A theorem proving system for classical type theory. *Journal of Automated Reasoning*, 16(3):321–353.

[Andrews *et al*, 2000] Andrews, Peter B., Bishop, Matthew and Brown, Chad E. (June 2000). System description: Tps: A theorem proving system for type theory. In McAllester, David A., (ed.), *Proceedings of 17th International Conference on Automated Deduction*, volume 1831 of *LNCS*, pages 164–169. Springer.

[Armando & Ranise, 2002] Armando, Alessandro and Ranise, Silvio. (2002). Constraint Contextual Rewriting. *Journal of Symbolic Computation. Special issue on First Order Theorem Proving, P. Baumgartner and H. Zhang editors*.

[Audemard *et al*, 2002] Audemard, Gilles, Bertoli, Piergiorgio, Cimatti, Alessandro, Korniłowicz, Artur and Sebastiani, Roberto. (2002). Integrating Boolean and Mathematical Solving: Foundations, Basic Algorithms and Requirements. In Calmet, Jacques, Benhamou, Belaid, Caprotti, Olga, Henocque, Laurent and Sorge, Volker, (eds.), *CALCULEMUS-2002: Symposium on the Integration of Symbolic Computation and Mechanized Reasoning*, volume 2385 of *LNAI*. Springer.

[Autexier *et al*, 1999] Autexier, Serge, Hutter, Dieter, Mantel, Heiko and Schairer, Axel. (1999). System description: Inka 5.0 – a logic voyager. In *Proc. of 16th International Conference on Automated Deduction CADE-16*, number 1632 in LNAI, pages 207–211. Springer.

[Autexier *et al*, 2000] Autexier, Serge, Hutter, Dieter, Langenstein, Bruno, Mantel, Heiko, Rock, Georg, Schairer, Axel, Stephan, Werner, Vogt, Roland and Wolpers, Andreas. (2000). VSE: Formal Methods Meet Industrial Needs. *Special Issue on Mechanized Theorem Proving for Technology Transfer of the STTT-Springer International Journal on Software Tools for Technology Transfer*, 3(1):66–77.

[Autexier *et al*, 2002] Autexier, S., Hutter, D., Mossakowski, T. and Schairer, A., (2002). The development graph manager maya, In Helene Kirchner, editor, Proceedings of 9th International Conference on Algebraic Methodology And Software Technology (AMAST'02). Springer Verlag, 2002.

[Baader & Siekmann, 1994] Baader, Franz and Siekmann, Jörg. (1994). Unification theory. In Gabbay, Dov, (ed.), *Logic in Artificial Intelligence and Logic Programming*. Oxford University Press.

[Barendregt & Geuvers, 2001] Barendregt, H. and Geuvers, H. (2001). Proof assistants using dependent type systems. In Robinson, J. and Voronkov, A., (eds.), *Handbook of Automated Reasoning*, volume 2. Elsevier Science.

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

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

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

[Benzmüller & Kohlhase, 1998a] Benzmüller, Christoph and Kohlhase, Michael. (1998a). Extensional higher-order resolution. In Kirchner, Claude and Kirchner, Hélène, (eds.), *Proceedings of the 15th International Conference on Automated Deduction (CADE-15)*, number 1421 in LNAI, pages 56–71, Lindau, Germany. Springer.

[Benzmüller & Kohlhase, 1998b] Benzmüller, Christoph and Kohlhase, Michael. (1998b). LEO – a higher-order theorem prover. In Kirchner, Claude and Kirchner, Hélène, (eds.), *Proceedings of the 15th International Conference on Automated Deduction (CADE-15)*, number 1421 in LNAI, pages 139–143, Lindau, Germany. Springer.

[Benzmüller & Sorge, 1998] Benzmüller, Christoph and Sorge, Volker. (1998). A blackboard architecture for guiding interactive proofs. In Giunchiglia, Fausto, (ed.), *Proceedings of 8th International Conference on Artificial Intelligence: Methodology, Systems, Applications (AIMSA'98)*, number 1480 in LNAI, pages 102–114, Sozopol, Bulgaria. Springer.

[Benzmüller & Sorge, 1999a] Benzmüller, Christoph and Sorge, Volker. (1999a). Critical agents supporting interactive theorem proving. In Borahona, Pedro and Alferes, Jose J., (eds.), *Proceedings of the 9th Portuguese Conference on Artificial Intelligence (EPIA'99)*, number 1695 in LNAI, pages 208–221, Evora, Portugal. Springer.

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

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

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

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

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

[Benzmüller, 1999b] Benzmüller, Christoph. (1999b). Extensional higher-order paramodulation and RUE-resolution. In Ganzinger, Harald, (ed.), *Proceedings of the 16th International Conference on Automated Deduction (CADE-16)*, number 1632 in LNAI, pages 399–413, Trento, Italy. Springer.

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

[Benzmüller, 2002] Benzmüller, Christoph. (2002). 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.

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

[Benzmüller, 2003b] Benzmüller, Christoph. (2003b). 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)*, Rome, Italy.

[Benzmüller, 2003c] Benzmüller, Christoph. (2003c). 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,.

[Benzmüller *et al*, 1997] Benzmüller, Christoph, Cheikhrouhou, Lassaad, Fehrer, Detlef, Fiedler, Armin, Huang, Xiaorong, Kerber, Manfred, Kohlhase, Michael, Konrad, Karsten, Melis, Erica, Meier, Andreas, Schaarschmidt, Wolf, Siekmann, Jörg and Sorge, Volker. (1997). OMEGA: Towards a mathematical assistant. In McCune, William, (ed.), *Proceedings of 14th International Conference on Automated Deduction (CADE-14)*, number 1249 in LNAI, pages 252–255, Townsville, Australia. Springer.

[Benzmüller *et al*, 1999a] Benzmüller, Christoph, Bishop, Matt and Sorge, Volker. (1999a). Integrating TPS and OMEGA. *Journal of Universal Computer Science*, 5:188–207.

[Benzmüller *et al*, 1999b] Benzmüller, Christoph, Jamnik, Mateja, Kerber, Manfred and Sorge, Volker. (1999b). Agent based mathematical reasoning. *Electronic Notes in Theoretical Computer Science, Elsevier*, 23(3):21–33.

[Benzmüller *et al*, 1999c] Benzmüller, Christoph, Sorge, Volker and Byrnes, John. (1999c). OANTS for interactive ATP. Draft, AG Siekmann, Saarland University.

[Benzmüller *et al*, 2000] Benzmüller, Christoph, Jamnik, Mateja, Kerber, Manfred and Sorge, Volker. (2000). Resource guided concurrent deduction. In Kerber, Manfred and Kohlhase, Michael, (eds.), *Symbolic Computation and Automated Reasoning*, pages 245–246. A.K.Peters.

[Benzmüller *et al*, 2001a] Benzmüller, Christoph, Jamnik, Mateja, Kerber, Manfred and Sorge, Volker. (2001a). An agent-oriented approach to reasoning. In *Proceedings of the Calculemus Workshop 2001*, pages 48–63, Siena, Italy.

[Benzmüller *et al*, 2001b] Benzmüller, Christoph, Jamnik, Mateja, Kerber, Manfred and Sorge, Volker. (2001b). Experiments with an agent-oriented reasoning system. In Baader, Franz, Brewka, Gerhard and Eiter, Thomas, (eds.), *KI 2001: Advances in Artificial Intelligence, Joint German/Austrian Conference on AI, Vienna, Austria, September 19-21, 2001, Proceedings*, number 2174 in LNAI, pages 409–424. Springer.

[Benzmüller *et al*, 2001c] Benzmüller, Christoph, Meier, Andreas and Sorge, Volker. (2001c). Distributed assertion retrieval. In *First International Workshop on Mathematical Knowledge Management RISC-Linz*, pages 1–7, Schloss Hagenberg.

[Benzmüller *et al*, 2003a] Benzmüller, Christoph, Brown, Chad and Kohlhase, Michael. (2003a). Higher order semantics and extensionality. *Submitted to Journal of Symbolic Logic; accepted final revision*.

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

[Benzmüller *et al*, 2003c] Benzmüller, Christoph, Brown, Chad E. and Kohlhase, Michael. (2003c). Semantic techniques for cut-elimination in higher order logic. Technical Report Draft Version, Carnegie Mellon University, Pittsburgh, PA.

[Benzmüller *et al*, 2003d] Benzmüller, Christoph, Meier, Andreas and Sorge, Volker. (2003d). Bridging theorem proving and mathematical knowledge retrieval. In *Festschrift in Honour of Jörg Siekmann*, LNAI. To appear.

[Bibel, 1983] Bibel, Wolfgang. (1983). Matings in matrices. *Communications of the ACM*, 26:844–852.

[Bishop & Andrews, 1998] Bishop, M. and Andrews, P. (1998). Selectively instantiating definitions. In Kirchner and Kirchner [Kirchner & Kirchner, 1998].

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

[Boulton *et al*, 1998] Boulton, Richard, Slind, Konrad, Bundy, Alan and Gordon, Mike. (1998). An interface between CLAM and HOL. In Grundy, J. and Newey, M., (eds.), *Proceedings of the 11th International Conference on Theorem Proving in Higher Order Logics*, number 1479 in LNCS, pages 87–104, Canberra. Springer.

[Boyer & Moore, 1979] Boyer, Robert S. and Moore, J Strother. (1979). *A Computational Logic*. Academic Press, New York, USA.

[Brown, 2002] Brown, Chad E. (2002). Solving for set variables in higher-order theorem proving. In Voronkov, Andrei, (ed.), *Proceedings of the 18th International Conference on Automated Deduction (CADE-19)*, number 2392 in LNAI, pages 144–149, Copenhagen, Denmark. Springer.

[Brown, 2003] Brown, Chad E. (2003). *Set Comprehension in Higher-Order Logic*. Unpublished Ph.D. thesis, Carnegie Mellon University, Draft Version. To appear.

[Buchberger, 1985] Buchberger, Bruno. (1985). *Gr öbner-Bases: An Algorithmic Method in Polynomial Ideal Theory, In N.K. Bose (ed.), Multidimensional Systems Theory - Progress, Directions and Open Problems in Multidimensional Systems Theory, chapter 6*, pages 184–232. Reidel Publishing Company, Dodrecht - Boston - Lancaster, second edition to appear 2003.

[Buchberger, 1992] Buchberger, Bruno. (1992). History and basic features of the critical-pair/completion procedure. *Journal of Symbolic Computation*, 3(1/2):272–279. (Earlier version appeared in: Proceedings of the Conference on Rewrite Technique and Applications, Dijon, May 1985, Lecture Notes in Computer Science, Vol. 202, Springer, 1985, pp. 1-45).

[Buchberger, 1999] Buchberger, Bruno, (July 1999). Theory exploration versus theorem proving. Invited talk at the 1999 Calculemus Symposium, IRST, Trento, Italy.

[Buchberger, 2000] Buchberger, Bruno. (October 2000). Theory exploration with theorema. In Jebelean, T., Negru, V. and Popovici, A., (eds.), *Analele Universitatii Din Timisoara, Ser. Matematica-Informatica, Vol. XXXVIII, Fasc.2, 2000, (Proceedings of SYNASC 2000, 2nd International Workshop on Symbolic and Numeric Algorithms in Scientific Computing*, pages 9–32, Timisoara, Rumania. ISSN 1124-970X.

[Buchberger, 2001a] Buchberger, B. (2001a). Theorema: A short introduction. *Mathematica Journal*, 8(2):247–252.

[Buchberger, 2001b] Buchberger, B. (September 2001). Theorema: Extending mathematica by automated proving. In Ungar, D., (ed.), *Proceedings of PrimMath 2001 (The Programming System Mathematica in Science, Technology, and Education)*, pages 10–11, University of Zagreb, Electrotechnical and Computer Science Faculty.

[Buchberger, 2003] Buchberger, Bruno, (July 2003). Verified algorithm development by lazy thinking. Invited talk at IMS 2003 (International Mathematica Symposium 2003, Imperial College, London, Imerial College).

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

[Bundy, 1988] Bundy, A. (1988). The use of explicit plans to guide inductive proofs. In Lusk, E. and Overbeek, R., (eds.), *Proceedings of the 9th Conference on Automated Deduction*, number 310 in LNCS, pages 111–120, Argonne, Illinois, USA. Springer.

[Bürckert *et al*, 1988] Bürckert, Hans-Jürgen, Herold, Alexander, Kapur, Deepak, Siekmann, Jörg H., Stickel, Mark E., Tepp, Michael and Zhang, Hantao. (1988). Opening the AC-unification race. *Journal of Automated Reasoning*, 4(4):465–474.

[Calmet *et al*, 1997] Calmet, J., Jekutsch, S., Kullmann, P. and Schü, J. (1997). KOMET: A system for the integration of heterogenous information sources. In *Proceedings of the 10th International Symposium on Methodologies for Intelligent Systems (ISMIS)*.

[Church, 1940] Church, Alonzo. (1940). A formulation of the simple theory of types. *Journal of Symbolic Logic*, 5:56–68.

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

[COQ, 2003] INRIA. (2003). *The Coq Proof Assistant Reference Manual*, See `http://coq.inria.fr/doc/main.html`.

[Coquand & Huet, 1985] Coquand, Thierry and Huet, Gérard. (1985). A theory of constructions. In *Semantics of Data Types*. Springer.

[Dahn & Denzinger, 1998] Dahn, I. and Denzinger, J. (1998). Cooperating theorem provers. In Bibel, W. and Schmitt, P.H., (eds.), *Automated Deduction—A Basis for Applications*, volume II, pages 383–416. Kluwer.

[Darlington, 1971] Darlington, J. L. (1971). Deductive plan formation in higher-order logic. *Machine Intelligence*, 7:129–137.

[de Bruijn, 1968] de Bruijn, N. G. (1968). The mathematical language automath, its usage and some of its extensions. In *Proc. Symp. Automated Deduction*, number 125 in LNCS. Springer.

[de Bruijn, 1980] de Bruijn, Nicolaas Govert. (1980). A survey of the project AUTOMATH. In Hindley, James R. and Seldin, Jonathan P., (eds.), *To H.B. Curry: Essays in Combinator Logic, Lambda Calculus and Formalisms*, pages 579–606. Academic Press.

[Dennis *et al*, 2000] Dennis, Louise A., Collins, Graham, Norrish, Michael, Boulton, Richard, Slind, Konrad, Robinson, Graham, Gordon, Mike and Melham, Tom. (2000). The prosper toolkit. In *Proceedings of the 6th International Conference on Tools and Algorithms for the Construction and Analysis of Systems*, volume 1785 of *LNCS*, Berlin, Germany. Springer Verlag.

[Denzinger & Fuchs, 1999] Denzinger, J. and Fuchs, D. (1999). Cooperation of Heterogeneous Provers. In Dean, Thomas, (ed.), *Proceedings of the 16th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 10–15, San Francisco, CA, USA. Morgan Kaufmann Publishers.

[Dowek *et al*, 2001] Dowek, Gilles, Hardin, Thérèse and Kirchner, Claude. (2001). HOL-$\lambda\sigma$ an intentional first-order expression of higher-order logic. *Mathematical Structures in Computer Science*, 11(1):1–25.

[Ernst, 1971] Ernst, G. W. (1971). A matching procedure for type theory. Technical report, Case Western Reserve University.

[Fiedler, 2001a] Fiedler, A. (2001a). *P.rex*: An interactive proof explainer. In Goré et al. [Goré *et al*, 2001].

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

[Fisher & Ireland, 1998] Fisher, Michael and Ireland, Andrew. (1998). Multi-agent proof-planning. to appear at the CADE-15 Workshop "Using AI methods in Deduction".

[Fisher, 1997] Fisher, Michael. (1997). An Open Approach to Concurrent Theorem Proving. In Geller, J., Kitano, H. and Suttner, C., (eds.), *Parallel Processing for Artificial Intelligence*, volume 3. Elsevier/North Holland.

[Franke & Kohlhase, 1999] Franke, Andreas and Kohlhase, Michael. (1999). System description: MATHWEB, an agent-based communication layer for distributed automated theorem proving. In Ganzinger [Ganzinger, 1999], pages 217–221.

[Franke *et al*, 1999] Franke, Andreas, Hess, Stephan M., Jung, Christoph G., Kohlhase, Michael and Sorge, Volker. (1999). Agent-oriented integration of distributed mathematical services. *Journal of Universal Computer Science*, 5:156–187.

[Gallier & Snyder, 1989] Gallier, Jean H. and Snyder, Wayne. (1989). Complete sets of transformations for general E-unification. *Theoretical Computer Science*, 1(67):203–260.

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

[Ganzinger, 1999] Ganzinger, H., (ed.). (1999). *Proceedings of the 16th Conference on Automated Deduction*, number 1632 in LNAI. Springer.

[Ganzinger *et al*, 2003] Ganzinger, H., Hillenbrand, T. and Waldmann, U. (2003). Superposition modulo a Shostak theory. In Baader, F., (ed.), *Automated Deduction — CADE-19*, volume 2741, pages 182–196.

[Geuvers, 2003] Geuvers, H. (2003). The calculus of constructions and higher order logic. In de Groote, Ph., (ed.), *The Curry-Howard isomorphism*, number 8 in Cahiers du Centre de logique, pages 139–191. Academia, Universit'e catholique de Louvain, Louvain-la-Neuve, Belgium. To appear.

[Giunchiglia *et al*, 2000] Giunchiglia, Fausto, Sebastiani, Roberto and Traverso, Paolo. (2000). Integrating SAT solvers with domain-specific reasoners. In Kerber, Manfred and Kohlhase, Michael, (eds.), *CALCULEMUS-2000, Systems for Integrated Computation and Deduction*, St. Andrews, Scotland. AKPeters. forthcoming.

[Gödel, 1931] Gödel, Kurt. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte der Mathematischen Physik*, 38:173–198. English translation in [van Heijenoort, 1967].

[Goldfarb, 1981] Goldfarb, Warren D. (1981). The undecidability of the second-order unification problem. *Theoretical Computer Science*, 13:225–230.

[Gordon & Melham, 1993] Gordon, M. and Melham, T. (1993). *Introduction to HOL – A theorem proving environment for higher order logic*. Cambridge University Press.

[Gordon *et al*, 1979] Gordon, M., Milner, R. and Wadsworth, C. (1979). *Edinburgh LCF: A Mechanized Logic of Computation*. Number 78 in LNCS. Springer.

[Goré *et al*, 2001] Goré, R., Leitsch, A. and Nipkow, T., (eds.). (2001). *Automated Reasoning — 1st International Joint Conference, IJCAR 2001*, number 2083 in LNAI. Springer.

[Gould, 1966] Gould, William Eben. (1966). A matching procedure for ω-order logic. Technical report, Applied Logic Corporation, One Palmer Square, Princeton, NJ.

[Harrison, 2000] Harrison, John. (2000). Formal verification of IA-64 division algorithms. In Aagaard, M. and Harrison, J., (eds.), *Theorem Proving in Higher Order Logics: 13th International Conference, TPHOLs 2000*, volume 1869 of *Lecture Notes in Computer Science*, pages 234–251. Springer-Verlag.

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

[Henkin, 1996] Henkin, Leon. (1996). The discovery of my completeness proofs. *The Bulletin of Symbolic Logic*, 2(2):127–157.

[Herbrand, 1930] Herbrand, Jaques. (1930). *Recherches sur la th éorie de la démonstration*. Unpublished Ph.D. thesis, Université de Paris, Englisch translation in [van Heijenoort, 1967].

[Hickey *et al*, 2003] Hickey, Jason, Nogin, Aleksey, Constable, Robert L., Aydemir, Brian E., Barzilay, Eli, Bryukhov, Yegor, Eaton, Richard, Granicz, Adam, Kopylov, Alexei, Kreitz, Christoph, Krupski, Vladimir N., Lorigo, Lori, Schmitt, Stephan, Witty, Carl and Yu, Xin. (2003). Metaprl — a modular logical environment. In *Proceedings of TPHOLs 2003*, LNCS. Springer. In Print.

[Hillenbrand & Löchner, 2002] Hillenbrand, Th. and Löchner, B. (2002). The next waldmeister loop. In Voronkov, (ed.), *Proceedings of the 18th International Conference on Automated Deduction*, number 2392 in LNCS, pages 486–500, Copenhagen, Denmark. Springer.

[Hillenbrand *et al*, 1999] Hillenbrand, Th., Jaeger, A. and Löchner, B. (1999). System description: Waldmeister — improvements in performance and ease of use. In Ganzinger [Ganzinger, 1999], pages 232–236.

[Holmes & Foss, 2001] Holmes, R. and Foss, A. (2001). The watson theorem prover. *Journal of Automated Reasoning*, 26(4):357–408.

[Holmes, 1999] Holmes, R. (1999). Subsystems of quine's "new foundations" with predicativity restrictions. *Notre Dame Journal of Formal Logic*, 40(2):183–196.

[Huang, 1994] Huang, Xiaorong. (1994). Reconstructing proofs at the assertion level. In Bundy, Alan, (ed.), *Proceedings of the 12th Conference on Automated Deduction*, number 814 in LNAI, pages 738–752, Nancy, France. Springer.

[Huet, 1972] Huet, Gérard P. (1972). *Constrained Resolution: A Complete Method for Higher Order Logic*. Unpublished Ph.D. thesis, Case Western Reserve University.

[Huet, 1973a] Huet, Gérard P. (1973a). A mechanization of type theory. In Walker, Donald E. and Norton, Lewis, (eds.), *Proceedings of the 3rd International Joint Conference on Artificial Intelligence (IJCAI73)*, pages 139–146.

[Huet, 1973b] Huet, Gérard P. (1973b). The undecidability of unification in third order logic. *Information and Control*, 22(3):257–267.

[Huet, 1975] Huet, Gérard P. (1975). An unification algorithm for typed λ-calculus. *Theoretical Computer Science*, 1:27–57.

[Jensen & Pietrzykowski, 1972] Jensen, D. C. and Pietrzykowski, T. (1972). A complete mechanization of ω-order type theory. In *Proceedings of the ACM annual Conference*, volume 1, pages 82–92.

[Kaufmann, 1992] Kaufmann, Stefan. (1992). *Mathematica als Werkzeug: eine Einführung*. Birkhäuser.

[Kerber & Kohlhase, 1997] Kerber, Manfred and Kohlhase, Michael. (1997). Mechanising partiality without re-implementation. In Brewka, Gerhard, Habel, Christopher and Nebel, Bernhard, (eds.), *Proceedings of the 21st Annual German Conference on Artificial Intelligence, KI'97*, number 1303 in LNAI, pages 123–134, Freiburg, Germany. Springer Verlag, Berlin, Germany.

[Kirchner & Kirchner, 1998] Kirchner, C. and Kirchner, H., (eds.). (1998). *Proceedings of the 15th Conference on Automated Deduction*, number 1421 in LNAI. Springer.

[Kohlhase, 1994] Kohlhase, Michael. (1994). *A Mechanization of Sorted Higher-Order Logic Based on the Resolution Principle*. Unpublished Ph.D. thesis, Fachbereich Informatik, Universität des Saarlandes.

[Kohlhase, 1995] Kohlhase, Michael. (1995). Higher-Order Tableaux. In Baumgartner, Peter, Hähnle, Rainer and Posegga, Joachim, (eds.), *Theorem Proving with Analytic Tableaux and Related Methods*, number 918 in LNAI, pages 294–309.

[Kohlhase, 1998] Kohlhase, M. (1998). Higher-order automated theorem proving. In Bibel, Wolfgang and Schmidt, Peter H., (eds.), *Automated Deduction: A Basis for Applications. Volume I, Foundations: Calculi and Methods*. Kluwer Academic Publishers, Dordrecht.

[Lucchesi, 1972] Lucchesi, Claudio. L. (1972). The undecidability of the unification problem for third order languages. Report CSRR 2059, University of Waterloo, Waterloo, Canada.

[Martin-Löf, 1994] Martin-Löf, Per. (1994). *Intuitionistic Type Theory*. Bibliopolis.

[McAllester, 2000] McAllester, D., (ed.). (2000). *Proceedings of the 17th Conference on Automated Deduction*, number 1831 in LNAI. Springer.

[McCune & Wos, 1997] McCune, William and Wos, Larry. (1997). Otter CADE-13 competition incarnations. *Journal of Automated Reasoning*, 18(2):211–220. Speacial Issue on the CADE-13 Automated Theorem Proving System Competition.

[McCune, 1994] McCune, W. (1994). Otter 3.0 reference manual and guide. Technical Report ANL-94-6, Argonne National Laboratory, Argonne, Illinois 60439, USA.

[McCune, 1997] McCune, W. (1997). Solution of the Robbins problem. *Journal of Automated Reasoning*, 19(3):263–276.

[Meier, 2000] Meier, A. (2000). TRAMP: Transformation of Machine-Found Proofs into Natural Deduction Proofs at the Assertion Level. In McAllester [McAllester, 2000].

[Meier *et al*, 2002] Meier, A., Pollet, M. and Sorge, V. (October 2002). 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. Steve Linton and Roberto Sebastiani, eds.

[Melis & Siekmann, 1999] Melis, E. and Siekmann, J. (1999). Knowledge-based proof planning. *Artificial Intelligence*, 115(1):65–105.

[Nadathur & Miller, 1994] Nadathur, Gopalan and Miller, Dale. (1994). Higher-order logic programming. Technical Report CS-1994-38, Department of Computer Science, Duke University.

[Nguyen *et al*, 2002] Nguyen, Quang Huy, Kirchner, Claude and Kirchner, Helene. (2002). External rewriting for skeptical proof assistants. *Journal of Automated Reasoning*, 29(3-4):309–336.

[Nipkow *et al*, 2002] Nipkow, Tobias, Paulson, Lawrence C. and Wenzel, Markus. (2002). *Isabelle/HOL: A Proof Assistant for Higher-Order Logic*. Number 2283 in LNCS. Springer.

[Owre *et al*, 1996] Owre, S., Rajan, S., Rushby, J.M., Shankar, N. and Srivas, M. (1996). PVS: Combining specification, proof checking, and model checking. In Alur, R. and Henzinger, T., (eds.), *Computer-Aided Verification, CAV '96*, number 1102 in LNCS, pages 411–414, New Brunswick, NJ. Springer.

[Papakonstantinou *et al*, 1996] Papakonstantinou, Y., Garcia-Molina, H. and Ullman, J. (1996). Medmaker: A mediation system based on declarative specifications. In *Proceedings of the 12th International Conference on Data Engineering*, pages 132–141. IEEE Computer Society.

[Paulson, 1994] Paulson, L. (1994). *Isabelle: A Generic Theorem Prover*. Number 828 in LNCS. Springer.

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

[Prehofer, 1998] Prehofer, Christian. (1998). *Solving Higher-Order Equations: From Logic to Programming*. Progress in theoretical computer science. Birkhäuser.

[Ramakrishnan *et al*, 1998] Ramakrishnan, I.V., Sekar, R. and Voronkov, A. (1998). Term indexing. In Robinson, J. and Voronkov, A., (eds.), *Handbook of Automated Reasoning*, volume 1. Elsevier Science.

[Redfern, 1996] Redfern, Darren. (1996). *The Maple Handbook: Maple V Release 4*. Springer.

[Riazanov & Voronkov, 2001] Riazanov, A. and Voronkov, A. (2001). Vampire 1.1 (system description). In Goré et al. [Goré *et al*, 2001].

[Richardson *et al*, 1998] Richardson, Julian, Smaill, Alan and Green, Ian. (1998). Proof planning in higher-order logic with λClam. In Kirchner and Kirchner [Kirchner & Kirchner, 1998], pages 129–133.

[Rizanow & Voronkov, 2002] Rizanow, A. and Voronkov, A. (2002). The design and implementation of vampire. *AI Communications*, 15(2-3):91–110.

[Robinson, 1968] Robinson, John A. (1968). New directions in theorem proving. In *Proceedings of IFIP Congress in Information Processing*, volume 68, pages 63–67. North Holland.

[Robinson, 1969] Robinson, John A. (1969). Mechanizing higher order logic. *Machine Intelligence*, 4:151–170.

[Roure, 2003] Roure, David De. (July/August 2003). On self-organization and the semantic grid. *IEEE Intelligent Systems*, 18(4):77–79.

[Schulz, 2002] Schulz, S. (2002). E: A brainiac theorem prover. *AI Communications*, 15(2-3):111–126.

[Schütte, 1960] Schütte, Kurt. (1960). Semantical and syntactical properties of simple type theory. *Journal of Symbolic Logic*, 25:305–326.

[Siekmann *et al*, 2002a] Siekmann, Jörg, Benzmüller, Christoph, Brezhnev, Vladimir, Cheikhrouhou, Lassaad, Fiedler, Armin, Franke, Andreas, Horacek, Helmut, Kohlhase, Michael, Meier, Andreas, Melis, Erica, Moschner, Markus, Normann, Immanuel, Pollet, Martin, Sorge, Volker, Ullrich, Carsten, Wirth, Claus-Peter and Zimmer, Jürgen. (2002a). Proof development with OMEGA. In Voronkov, Andrei, (ed.), *Proceedings of the 18th International Conference on Automated Deduction (CADE-19)*, number 2392 in LNAI, pages 144–149, Copenhagen, Denmark. Springer.

[Siekmann *et al*, 2002b] Siekmann, Jörg, Benzmüller, Christoph, Fiedler, Armin, Meier, A and Pollet, M. (2002b). Proof development with OMEGA: Sqrt(2) is irrational. In Baaz, Matthias and Voronkov, Andrei, (eds.), *Logic for Programming, Artificial Intelligence, and Reasoning, 9th International Conference, LPAR 2002*, number 2514 in LNAI, pages 367–387. Springer.

[Siekmann *et al*, 2003] Siekmann, Jörg, Benzmüller, Christoph, Fiedler, Armin, Meier, Andreas, Normann, Immanuel and Pollet, Martin. (2003). Proof development in OMEGA: The irrationality of square root of 2. In Kamareddine, Fairouz, (ed.), *Thirty Five Years of Automating Mathematics*, Kluwer Applied Logic series. Kluwer Academic Publishers. In Print.

[Smullyan, 1963] Smullyan, Raymond M. (1963). A unifying principle for quantification theory. *Proc. Nat. Acad Sciences*, 49:828–832.

[Snyder & Gallier, 1989] Snyder, Wayne and Gallier, Jean. (1989). Higher-Order Unification Revisited: Complete Sets of Transformations. *Journal of Symbolic Computation*, 8:101–140.

[Sorge, 2000] Sorge, V. (2000). Non-Trivial Computations in Proof Planning. In Kirchner, H. and Ringeissen, C., (eds.), *Frontiers of combining systems: Third International Workshop, FroCoS 2000*, volume 1794 of *LNAI*. Springer.

[Stenz & Wolf, 1999] Stenz, Gernot and Wolf, Andreas. (1999). E-setheo: Design, configuration and use of a parallel automated theorem prover. In Foo, Norman, (ed.), *Proceedings of AI'99: The 12th Australian Joint Conference on Artificial Intelligence*, number 1747 in LNCS, pages 231–243, Sydney, Australia. Springer.

[Stickel, 1981] Stickel, Mark E. (1981). A unification algorithm for associative-commutative functions. *JACM*, 28(3):423–434.

[Stump, 2003] Stump, Aaron. (2003). Subset types and partial functions. In Baader, Franz, (ed.), *Automated Deduction — CADE'19*, volume 2741 of *LNCS*, pages 151–165, Miami Beach, Fl, USA. Springer-Verlag.

[Takeuti, 1953] Takeuti, Gaisi. (1953). On a generalized logic calculus. *Japan Journal of Mathematics*, 23:39 f.

[van Heijenoort, 1967] van Heijenoort, Jean. (1967). *From Frege to G ödel : a source book in mathematical logic 1879-1931*. Source books in the history of the sciences series. Harvard University Press, Cambridge, MA, USA, 3rd printing, 1997 edition.

[Van Quine, 1973] Van Quine, Williard. (1973). *Mengenlehre und ihre Logik*. Vieweg, Braunschweig.

[Walther, 1993] Walther, Christoph. (1993). Mathematical induction. In Gabbay, Dov M., Hogger, C. J. and Robinson, J. Alan, (eds.), *Handbook of Logic in Artificial Intelligence*, volume 2, pages 127–228. Clarendon Press.

[Weidenbach *et al*, 1999] Weidenbach, C., Afshordel, B., Brahm, U., Cohrs, C., Engel, Th., Keen, E., Theobalt, C. and Topic, D. (1999). System description: SPASS version 1.0.0. In Ganzinger [Ganzinger, 1999], pages 378–382.

[Weiss, 1999] Weiss, G., (ed.). (1999). *Multiagent systems : a modern approach to distributed artificial intelligence*. MIT Press.

[Wiederhold, 1992] Wiederhold, Gio. (1992). Mediators in the architecture of future information systems. *IEEE Computer*, 25(3):38–49.

[Wirth, 2003] Wirth, Claus-Peter. (2003). History and future of implicit and inductionless induction: Beware the old jade and the zombie! In *Festschrift in Honour of J örg Siekmann*, LNAI. To appear.

[Wolf, 1998] Wolf, Andreas. (1998). P-SETHEO: Strategy Parallelism in Automated Theorem Proving. In de Swaart, Harry, (ed.), *Automated Reasoning with Analytic Tableaux and Related Methods*, volume 1397 of *LNAI*, page 320. Springer.

[Wolfram, 1988] Wolfram, Stephen. (1988). *Mathematica. A System for Doing Mathematics by Computer*. Addison-Wesley.

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

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