Abschlußbericht Projekt HOTEL (SI 372)

Christoph Benzmüller, Michael Kohlhase, Jörg Siekmann

30. Juni 2004

1 Allgemeine Angaben

1.1 DFG-Geschäftszeichen

HOTEL, SI 372

1.2 Antragsteller

1. Jörg H. Siekmann, Prof. Dr. (Ph.D.), grad. Ing.
2. Michael Kohlhase, Prof. Dr.
   (jetzt International University of Bremen)

1.3 Institut/Lehrstuhl

1. AG Siekmann, FR 6.2 Informatik
   Universität des Saarlandes
   Postfach 1150
   66041 Saarbrücken
   Tel: 0681/302 4754
   Fax: 0681/302 2235

2. (seit September 2003)
   School of Engineering & Science
   International University Bremen
   Campus Ring 12
   28759 Bremen
   Tel: 0421/200 3140
   Fax: 0421/200 3103

1.4 Thema

Gleichheitsbeweisen höherer Stufe
HOTEL (**H**igher-**O**rder **T**heorem Proving with **E**qua**L**ity)

1.5 Berichtszeitraum, Förderungszeitraum insgesamt

1/1996–5/1999 (1. Förderungsperiode), 6/1999–2/2004 (2. Förderungsperiode)

1.6 Liste der Publikationen

Wir geben hier eine kurze kommentierte Übersicht zu den wichtigsten Publikationen die im Projekt HOTEL entstanden sind oder wesentlich durch die Forschungsleistungen im Projekt HOTEL begünstigt wurden.

Bücher und Buchbeiträge

- [39] Eine Entwicklung einer zusammenhängenden Theorie der Deduktion höherer Stufe. Dieses Buchmanuskript entwickelt einen Resolutions- und Tableau-Kalküle höhere Stufe, und deren Semantische Theorie (aufbauend of [15]). Weiterhin wird ein sortierter λ-Kalkül entwickelt und das Unifikationsproblem gelöst.
- [42] Entwicklung des Tableau-Kalküls höherer Stufe mit Semantik und Vollständigkeitsbeweis für Henkin-Semantik.

Internationale Fachzeitschriften

- [15]: Landschaft von Semantiken für Logik höherer Stufe die motiviert ist durch unterschiedliche Rollen von Extensionalität und Gleichheit; Erweiterung des Beweisprinzips der abstrakten Konsistenz für Logik höherer Stufe; für jede Semantik wird ein vollständiger und korrekter ND-Kalkül angegeben
- [12]: Übersichtsartikel zum Resolutionsbeweisen in Logik höherer Stufe; Vergleich des im Projekt entwickelten extensionalen Resolutionsansatzes mit früheren Ansätzen.
- [14]: Integration des an der CMU entwickelten höherstufigen Beweisers TPS in das ΩMEGA-System; Entwicklung eines Ansatzes zur taktikbasierten Beweistransformation; diese Zusammenarbeit mit der CMU wurde durch das HOTEL Projekt ermöglicht.

Internationale Konferenzen

- [18]: Modellierung und Anwendung von LEO als ein Ω-ANTS Beweisagent; Zusammenspiel mit anderen Beweisern und Koordination durch das Ω-ANTS System im mathematischen Assistenzsystem ΩMEGA.
- [10]: Erweiterung der extensionalen Resolution höherer Stufe zur extensionalen Paramodulation höherer Stufe und extensionalen RUE-Resolution höherer Stufe.
- [23]: Beweiser LEO.
- [22]: Extensionale Resolution höherer Stufe.

Internationale Workshops

- [19]: Untersuchung des Phänomens der Komprehension in Logik höherer Stufe und des Verlusts von handhabbaren Beiweisen in Logik erster und höherer Stufe.

Habilitationen, Dissertation, Diplomarbeiten

- [8]: Dissertation von Christoph Benzmüller.

Technische Berichte

- [13]: Projektantrag im Aktionsplan Informatik von Christoph Benzmüller; erreichte die finale Evaluationsstufe wurde aber nicht zur Förderung angenommen.
- [17]: Erweiterung des Beweisprinzips der Abstrakten Konsistenz für Schnitt-freie Anwendungskontexte in der Logik höherer Stufe.
- [16]: Vorversion des Zeitschriftenartikels [15].
- [11]: RUE-Resolution in Logik erster Stufe ist nicht vollständig; in diesem Artikel wird gezeigt das die verwendete Gegenbeispiele zur Vollständigkeit nicht für den im Projekt entwickelten extensionalen RUE-Resolutionsansatz höherer Stufe gelten.

- [9]: Dissertation von Christoph Benzmüller als Technischer Bericht.
- [7]: Technischer Bericht zur extensionalen Paramodulation und RUE-Resolution höherer Stufe.
- [20]: Technischer Bericht zur extensionalen Resolution höherer Stufe.
- [21]: Erste Vorversion zum Zeitschriftenartikels [15].
- [6]: Erster technischer Bericht zur extensionalen Resolution höherer Stufe.

2 Arbeits- und Ergebnisbericht

2.1 Ausgangsfragen und Zielsetzung des Projekts

Im Projekt **HOTEL** sollten Verfahren zur Gleichheitsbehandlung für eine Logik höherer Stufe untersucht werden. Dabei sollten grundlegende Techniken zur Mechanisierung der Gleichheit, wie sie aus der Logik erster Stufe bekannt sind, so verallgemeinert werden, daß sie für die klassische Typtheorie (Logik höherer Stufe basierend auf dem einfach getypten λ-Kalkül) einsetzbar werden. Diese Verfahren sollten dann im automatischen Beweisen höherer Stufe eingesetzt werden und es sollte eine realistische Verwendbarkeit des automatischen Beweisens höherer Stufe in der Mathematik und in der Programmverifikation untersucht werden. Das Leo-System sollte als ein mathematischer Dienst zur Verfügung gestellt werden und Fallstudien sollten weitere Erweiterungen des Leo-Systems begleiten.

2.2 Darstellung und Einordnung der erreichten Ergebnisse

Das wichtigste Forschungsresultate des Projekts sind grundlegender, theoretischer Natur: (a) die Entwicklung einer Landschaft von Semantiken (die durch die verschiedene Rollen der Gleichheitsrelation motiviert sind) für die Logik höherer Stufe, (b) die Adaption des Beweisprinzips der Abstrakten Konsistenz für die Semantiken, und (c) die Bereitstellung von Interaktions-orientierten sowie Maschinen-orientierte Kalkülen und deren Korrektheits- und Vollständigkeitsanalyse mithilfe der eingeführten Methodologie unter (a) und (b).

Die grundlegende Bedeutung der Resultate wird u.a. durch den *Journal of Symbolic Logic* Gutachter der wichtigsten Publikation des Projekts [15] untermauert: “This is a very significant paper which provides much needed foundations for further work in this area, . . . ”

Erschwert wurde die Kontinuität und Effizienz der Projektarbeit in HOTEL in der zweiten Förderungsphase durch den Wechsel des Projektleiters Michael Kohlhase an die Carnegie Mellon University (2000–2003, Heisenberg Stipendium) und seinen Ruf an die International University Bremen, das einjährige Postdoktorandenstudium von Christoph Benzmüller an der University of Birmingham und der University of Edinburgh zu einem anderen Forschungsthema (Agenten-orientiertes Beweisen) und seiner Übernahme der Leitung des OMEGA Projektes nach seiner Rückkehr in 2001 an die Universität des Saarlandes. Durch gegenseitige Forschungsaufenthalte an der Carnegie Mellon University bzw. der Universität des Saarlandes gelang es uns dennoch die Forschung im Projekt weiter zu führen. Die Zusammenarbeit mit dem TPS Projekt (speziell mit Chad Brown, dem derzeitigen Hauptentwickler) wurde sogar weiter verbessert (inklusive gemeinsamer Publikationen). Diese Kooperation soll noch weiter verbessert werden durch die geplanten Wechsel von Chad Brown an die Universität des Saarlandes und seine Mitarbeit im Projekt Alonzo (derzeit von Christoph Benzmüller bei der DFG beantragt; Kooperation mit Michael Kohlhase in Bremen ist vorgesehen).

2.2.1 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 [30] 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 [33, 34, 3] 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 [2] ist eine Adaption von Smullyan's abstrakter Konsistenzmethode [47] 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 Leistungen im Projekt [15, 13, 8, 21], und die Zusammenarbeit mit Chad Brown (CMU, Pittsburgh, USA) [26] 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 Semantiken sind nicht nur als 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 [15]).

2.2.2 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 [44, 45] einen der ersten Ansätze zur Mechanisierung der Logik höherer Ordnung. Dieser basiert auf dem Tableauxverfahren und greift Ideen von Schütte [46] und Takeuti [48] auf. Wichtige Pionierarbeiten folgten dann kurze Zeit später mit Peter Andrews' Vorschlag zur Resolution höherer Ordnung [2], Jensen und Pietrzykowski's Ansatz [38] und speziell Gerard Huet's *Constrained Resolution* [35, 36].

Eine der größten Herausforderungen für die Mechanisierung von Logiken höherer Ordnung ist die Unentscheidbarkeit und Komplexität der Unifikation höherer Stufe [43, 37, 31]. Erste nichtvollständige Verfahren zur vollen Unifikation höherer Ordnung werden in [32, 27, 28] präsentiert und die ersten vollständigen Verfahren gehen zurück auf [35] und auf [38]. 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[1] und bietet dennoch ein vollständiges und semi-entscheidbares Verfahren zur Beantwortung der Lösbarkeit der aufgesammelten Unifikationsconstraints.

[1]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.

In den 80er Jahren hat Peter Andrews die Konnektionsmethode [4, 5, 24] 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 [1] belegt.

Der Resolutionsansatz wurde erst Anfang der 90er wieder aufgegriffen. Michael Kohlhase integriert in seiner Dissertation [40] Sorten in die Logik höherer Ordnung und entwickelt eine sortierte Unifikation höherer Ordnung, die er dann als Motor seines sortierten Resolutionskalküls einsetzt. Die Dissertation von Kohlhase (sowie [41]) lieferte auch erste Anstöße zur Verbesserung der Extensionalitäts- und Gleichheitsbehandlung, sowie zur Adaption des Beweisprinzips der abstrakten Konsistenz für Henkin Semantik.

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 Instantiierung 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.

Maschinenorientierte Kalküle Hinsichtlich Problem (I) wurden durch die Forschungsleistungen im Projekt HOTEL wesentliche Fortschritte erreicht [12, 8, 10, 22]: statt die Extensionalitätsaxiome blind dem Suchraum hinzuzufügen, erfolgen gegenseitige und zielgerichtete Aufrufe zwischen Unifikation und Beweissuche. Der im Projekt entwickelte Kalkül $\mathcal{ER}$ ist der erste Kalkül für klassische Logik höherer Ordnung, der ohne zusätzliche Axiome Henkin-Vollständigkeit garantiert.

Stichwortartig lassen sich die Beiträge des Projekts aus Kalkülsicht wie folgt zusammenfassen:

- Analyse und Klärung der Rolle der Gleichheit im automatischen Beweisen in klassischer Logik höherer Ordnung.
- Entwicklung eines Henkin-vollständigen Resolutionsansatzes, der zusätzliche Extensionalitätsaxiome im Suchraum vermeidet.
- Entwicklung Henkin-vollständiger primitiver Gleichheitsverfahren (RUE-Resolution und Paramodulation), die ebenfalls ohne zusätzliche Axiome auskommen.

Problem (II) wird, begünstigt auch durch die Zusammenarbeit mit dem Projekt HOTEL und die gemeinsame Arbeit [15], aktuell in der Dissertation von Chad Brown adressiert [25]: statt eine blinde Aufzählung von Formelschemata für Mengenvariablen durchzuführen sammelt ein Constraintsystem Bedingungen über die zu instantiierenden Mengenvariablen auf, generiert daraus zielgerichtet Instanziierungsvorschläge und propagiert diese dann an das Beweissystem.

Durch diese Forschungsbeiträge wurde gezeigt, daß der Anteil blinder Suche in der Mechanisierung von Logiken höherer Ordnung drastisch reduziert werden kann.

Interaktionsorientierte Kalküle Zusätzlich zu diesen Fortschritten bei maschinenorientierten Kalkülen wurden als Nebenprodukte im Projekt auch verschiedene höherstufige Varianten von Gentzen's Kalküls des natürlichen Schließens entwickelt. Konkret wurde in [15] für jeden der dort eingeführten Semantikbegriffe ein solcher Kalkül angegeben und als korrekt und vollständig bewiesen. Diese Arbeit ist relevant nicht zuletzt als theoretische Fundierung des Basiskaküls des mathematischen Assistenzsystems OMEGA, das im Sonderforschungsbereich SFB 378 in Saarbrücken entwickelt wird.

2.2.3 Beweiser LEO

Interessanterweise ergeben sich insbesondere bei vielen einfachen Problemen positive Effekte für den Suchraum durch den expressiveren Repräsentationsformalismus in Logik höherer Ordnung. 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 im Projekt entwickelten Beweiser LEO [23] wurde dies durch eine große Anzahl von Theoremen der Mengenlehre (der TPTP Datenbank des CASC-Beweiserwettbewerb[2] für Beweiser erster Ordnung) unter Verwendung höherstufiger Kodierungen demonstriert; siehe Abbildung 1 für eine Illustration der Beweissuche in LEO.

Das Beweiser LEO wurde im Rahmen des HOTEL Projekts vollständig in das Beweissystem ΩMEGA integriert. In Abbildung 2 ist ein (in Kapitel 4.4 in [12] diskutiertes) nicht-triviales höherstufiges Beweisproblem zu sehen: $(p = \lambda X_o.\, X_o) \vee (p = \lambda X_o.\, \neg X_o) \vee (p = \lambda X_o.\, \bot) \vee (p = \lambda X_o.\, \top)$. Es wurde im ΩMEGA System geladen und aktiviert und nun wird der Beweiser LEO per Menüauswahl darauf angewendet.

In Abbildung 3 erkennbar ist das LEO's Beweissuche erfolgreich war und das der Resolutionsbaum in LOUI, dem graphischen Interfaces von ΩMEGA, repräsentiert und analysiert werden kann.

Im Kontext von Vorlesungen an der Universität des Saarlandes wird das LEO System in der Lehre eingesetzt[3]. Besonders hilfreich ist dabei, das LEO nicht nur als vollautomatischer Beweiser aufgerufen werden kann, sondern auch menüunterstützt in ΩMEGA als interaktiver Beweiser; siehe Abbildung 4. Die Studenten können also interaktiv mit dem Kalkül arbeiten und sich dadurch besser mit den Problemen der höherstufigen (aber auch der erststufigen und propositionalen) Resolution vertraut machen.

2.2.4 Weitere Ergebnisse die durch das Projekt HOTEL gefördert wurden

Induktionsbeweisen Seit einem Drittel Jahrhundert ist Induktionsbeweisen eine Standarderweiterung des Theorembeweisens erster Stufe; typischerweise in Form expliziter Induktion. Wie man leicht sieht, sind Begriffe der induktiven Definition und der Wohlfundiertheit, die Peano-Axiome und ähnliche Axiome der strukturellen Induktion und das Theorem der Nöther'schen Induktion ausschließlich Formeln zweiter Stufe. Es wäre jedoch ein Irrtum, hieraus zu schließen, daß Theorembeweisen höherer Stufe spezieller Mechanismen der Induktion nicht bedürfe. Denn die reine Möglichkeit, einen mathematischen Beweis in irgendeiner Weise zu kodieren, ist nicht ausreichend. Um die tatsächliche Konstruktion mit Computer-Unterstützung bewerkstelligen zu können, müssen diese Beweise in ihrer speziell intendierten Form darstellbar sein. Da dies ja bereits der Hauptexistenzgrund des Theorembeweisens höherer Stufe neben dem der ersten Stufe ist und da Induktionsbeweise praktisch überall in der Mathematik häufig auftreten, kann die Notwendigkeit auch ihrer speziellen Unterstützung im Beweisen höherer Stufe nicht bestritten werden. In [50] ist es uns gelungen, zum ersten Mal aufzuzeigen, wie die *Descente Infinie* — in der Art und Weise wie sie von Mathematikern nun seit 24 Jahrhunderten als allgemeine Form der mathematischen Induktion verwendet wird — in dem Stand der Kunst entsprechende Sequenzen- und Tableau-Kalküle mit freien Variablen integriert werden kann; und zwar derart, daß eine effizientes Wechselspiel zwischen menschlicher Interaktion und Automatisierung möglich wird. Das wesentliche neue Konzept hierbei ist die unbeschränkte Anwendung von Induktionshypothesen. Die semantischen Anforderung werden von einer Reihe von zweiwertigen Logiken erfüllt, neben Klausellogik und klassischer Logik erster Stufe vor allem auch Logiken höherer Stufe (inkl. Modallogik).

2.3 Kooperationspartner im In- und Ausland

Carnegie Mellon University, USA

- TPS-Projekt von Prof. Peter Andrews
  Eine sehr enge Kooperation wurde im Projekt aufgebaut zur Arbeitsgruppe von Prof. Peter Andrews. In 1997 wurde Christoph Benzmüller von Peter Andrews als Gaststudent an die

[2] `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.

[3] Siehe z.B. `http://www.ags.uni-sb.de/~chris/lectures/fol-hol-tp/index.html`

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

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

In [29] 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 überträgt (wie man leicht erkennt) Ideen des Paramodulationsansatzes aus [10, 8] 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 $\cup = \lambda A_{\alpha\to o}, B_{\alpha\to o}, x_\alpha\text{.}\ (A\ x) \vee (B\ x)$ und $\cap = \lambda A_{\alpha\to o}, B_{\alpha\to o}, x_\alpha\text{.}\ (A\ x) \wedge (B\ x)$ 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:

$$C_1 : [\lambda X_\alpha\text{.}\ (b\ X) \vee ((c\ X) \wedge (d\ X)) = \lambda X_\alpha\text{.}\ ((b\ X) \vee (c\ X)) \wedge ((b\ X) \vee (d\ X))]^F$$

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

$$C_2 : [(b\ x) \vee ((c\ x) \wedge (d\ x)) \Leftrightarrow ((b\ x) \vee (c\ x)) \wedge ((b\ x) \vee (d\ x))]^F$$

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 Widerlegungsbeweis 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 [29] 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.

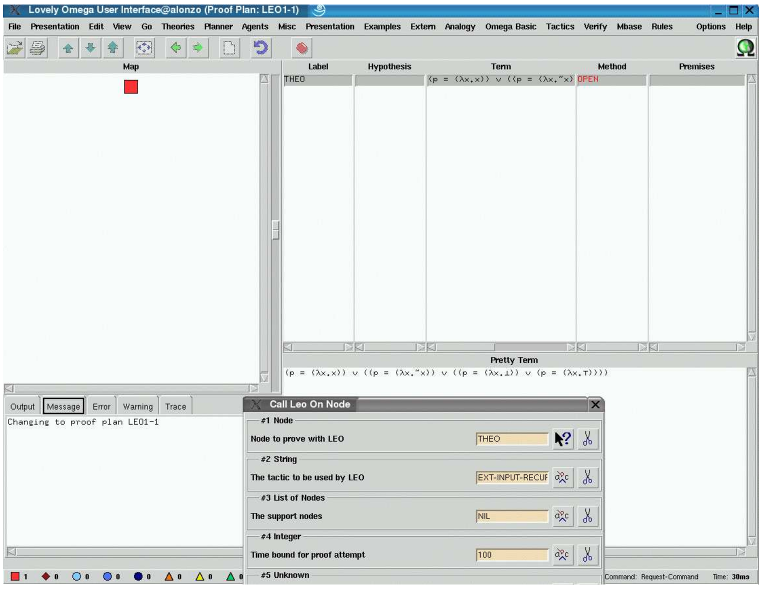

Abbildung 2: Menügesteuerter Aufruf von Leo aus der ΩMEGA Umgebung. Das Beweisproblem wurde zuvor geladen und wird in der graphischen Benutzeroberfläche $\mathcal{L}\Omega\mathcal{UI}$ angezeigt.

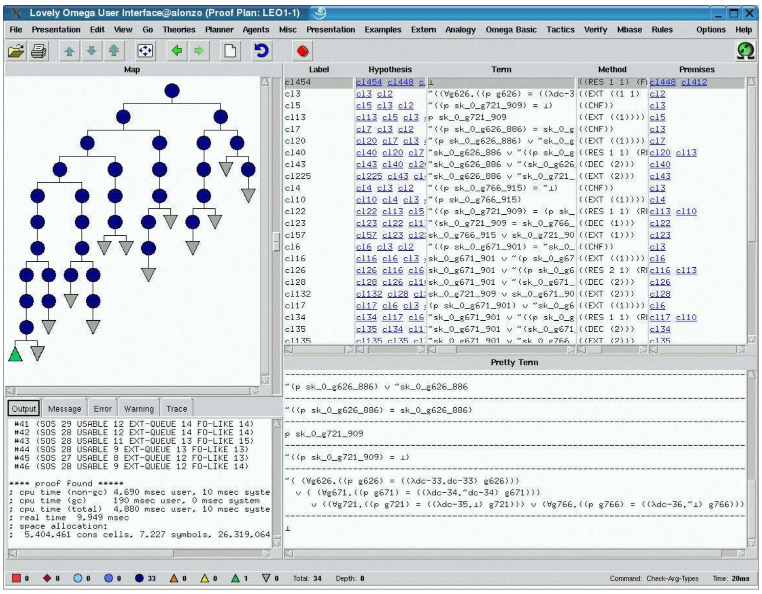

Abbildung 3: Illustration des Leo Beweisbaumes in ΩMEGA's graphischer Benutzeroberfläche $\mathcal{L}\Omega\mathcal{UI}$. Links wird der Widerlegungsbaum angezeigt, rechts oben der linearisierte Widerlegungsbeweis und rechts unten ein 'pretty print' des fokussierten Knotens. Alle Fenster sind per Hyperlinkmechanismus miteinander verbunden.

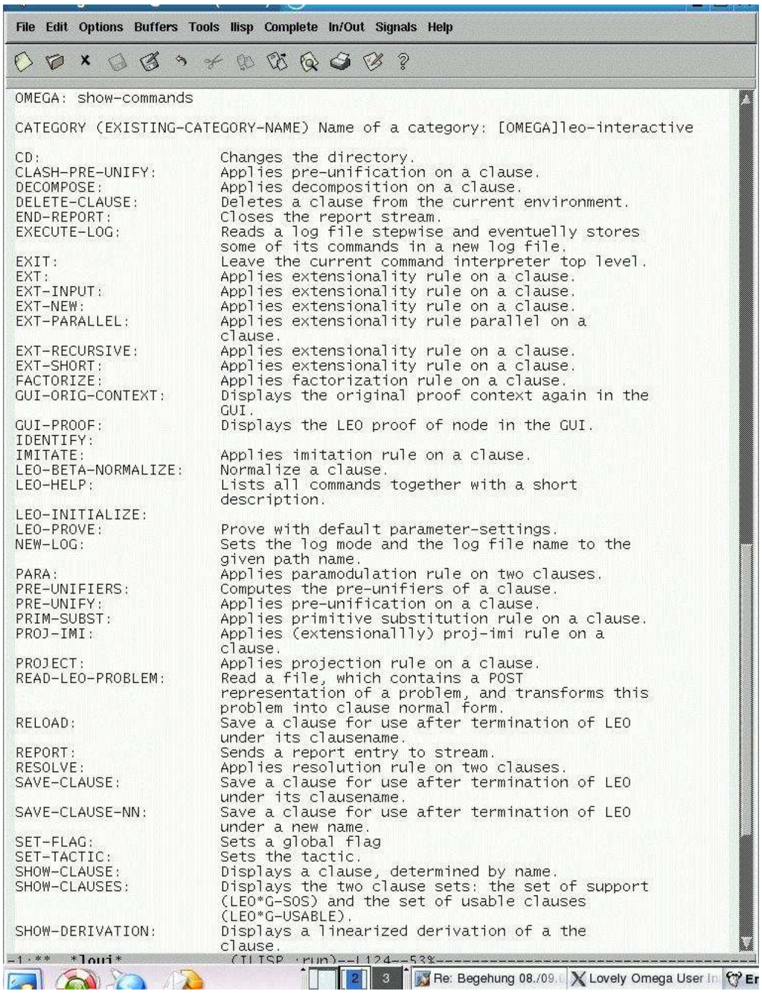

Abbildung 4: Leo wird als interaktiver und automatischer Beweiser in der Lehre eingesetzt. Die verfügbaren Kommandos zur interaktiven Beweiskonstruktion können durch das Kommando 'show-commands' angefragt werden.

CMU eingeladen. Er nahm dort an den *Introduction to Type Theory* Vorlesungen von Peter Andrews Teil, arbeitete sich in dessen höherstufigen Beweiser TPS ein, und baute den Kontakt zu Peter Andrews's Promotionsstudenten Matthew Bishop und Chad Brown auf. Mit beiden entstanden gemeinsame Publikationen [15, 13, 14].

- Prof. Frank Pfenning, Carnegie Mellon University, USA
  Während seines Aufenthalts an der CMU wurde Christoph Benzmüller auch von Frank Pfenning, dem früheren Schüler von Peter Andrews, mitbetreut. Frank Pfenning setzte diese Betreuung dann als Zweitgutachter der Dissertation von Christoph Benzmüller fort.

Mit der CMU fand ein regelmäßiger Austausch statt: Matt Bishop besuchte Saarbrücken in 1998, Chad Brown in 2001 und 2002. Umgekehrt besuchte Christoph Benzmüller die CMU in 2001 und 2002. Michael Kohlhase war Heisenberg-Stipendiat und adjungierter Associate Professor an der CMU von 2000 bis 2003.

The University of Birmingham, Birmingham, England Im Kontext eines Auslandsaufenthalts an der University of Birmingham (als Research Fellow in einem Projekt zum agentenbasierten Theorembeweisen in 2000) vertiefte Christoph Benzmüller auch die Kooperation zu Dr. Manfred Kerber auf dem Gebiet der Logik höherer Stufe: In praktischer Hinsicht wurde die Modellierung und Anwendbarkeit des Beweisers LEO als kooperativer Beweisagent [18] untersucht und auf der Theorieseite das Problem der Beschränkbarkeit von Mengenvariablensubstitutionen und Aspekte der Mengenkomprehension [19],

2.4 Qualifikation des wissenschaftlichen Nachwuchses

Die folgenden wissenschaftlichen Qualifikation von Nachwuchswissenschaftlern stehen in direktem Zusammenhang mit dem Projekt HOTEL bzw. wurden durch die Forschung im Projekt HOTEL begünstigt:

- Anerkennung der Habilitations-äquivalenten Leistungen von Christoph Benzmüller zur Einstellung als Hochschuldozent (C2) im Januar 2004 an der Universität des Saarlandes.
- Professur Michael Kohlhase an der International University Bremen, 2003.
- Heisenberg-Stipendium der Deutschen Forschungsgemeinschaft für Michael Kohlhase an der Carnegie Mellon University, USA. 2000–2003.
- Habilitation Michael Kohlhase. Universität des Saarlandes 1999.
- Promotion Christoph Benzmüller. Titel der Dissertation: Equality and Extensionality in Higher-Order Theorem Proving. 1999.
- Mehrere Diplomarbeiten die durch das Projekt HOTEL gefördert wurden: u.a. Ahmet Bozkurt, Gerald Klein, Lars Klein, Malte Hübner, Andreas Franke.

Literatur

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

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

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

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

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

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

[7] Christoph Benzmüller. An adaptation of paramodulation and rue-resolution to higher-order logic. Seki-Report SR-98-07, Department of Computer Science, Saarland University, 1998.

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

[9] Christoph Benzmüller. Equality and extensionality in higher-order theorem proving. Seki-Report SR-99-08, Department of Computer Science, Saarland University, 1999.

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

[11] Christoph Benzmüller. A remark on higher order RUE-resolution with EXTRUE. SEKI Technical Report SR-02-05, Fachbereich Informatik, Universität des Saarlandes, Saarbrücken, Germany, 2002.

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

[13] Christoph Benzmüller. Alonzo: Deduktionsagenten höherer ordnung für mathematische assistenzsysteme. Project proposal in the DFG Aktionsplan Informatik, 2003.

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

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

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

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

[18] Christoph Benzmüller, Mateja Jamnik, Manfred Kerber, and Volker Sorge. Experiments with an agent-oriented reasoning system. In Franz Baader, Gerhard Brewka, and Thomas Eiter, editors, *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, 2001.

[19] Christoph Benzmüller and Manfred Kerber. A lost proof. In *Proceedings of the IJCAR 2001 Workshop: Future Directions in Automated Reasoning*, pages 13–24, Siena, Italy, 2001.

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

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

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

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

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

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

[26] Chad E. Brown. *Set Comprehension in Higher-Order Logic.* PhD thesis, Carnegie Mellon University, 2003. Draft Version. To appear.

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

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

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

[30] Kurt Gödel. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte der Mathematischen Physik*, 38:173–198, 1931. English translation in [49].

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

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

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

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

[35] Gérard P. Huet. *Constrained Resolution: A Complete Method for Higher Order Logic*. PhD thesis, Case Western Reserve University, 1972.

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

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

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

[39] Michael Kohlhase. Higher-order automated theorem proving.

[40] Michael Kohlhase. *A Mechanization of Sorted Higher-Order Logic Based on the Resolution Principle*. PhD thesis, Fachbereich Informatik, Universität des Saarlandes, 1994.

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

[42] Michael Kohlhase. Higher-order automated theorem proving. In Wolfgang Bibel and Peter Schmitt, editors, *Automated Deduction – A Basis for Applications*, volume 1, pages 431–462. Kluwer, 1998.

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

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

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

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

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

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

[49] Jean van Heijenoort. *From Frege to Gödel : 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, 1967.

[50] Claus-Peter Wirth. Descente infinie + Deduction. *Logic Journal of the IGPL*, 12(1):1–96, 2004. `www.ags.uni-sb.de/~cp/p/d/welcome.html`.