

Teilprojekt MI 4:

**OMEGA:
Ressourcenadaptive Beweisplanung**

3.1 Allgemeine Angaben zum Teilprojekt MI 4

3.1.1 Thema

Ressourcenadaptive Beweisplanung

3.1.2 Fachgebiet und Arbeitsrichtung

Informatik / Künstliche Intelligenz

3.1.3 Leiter/in

Name, Geburtsdatum:	Prof.Dr. Siekmann, Jörg, 5.8.1941 Dr. Benzmüller, Christoph, 8.9.1968 PD.Dr. Melis, Erica, 28.2.1949
Dienstanschrift:	Fachrichtung Informatik Universität des Saarlandes D-66041 Saarbrücken
Telefon:	0681-302-5275 4574 4629
Telefax:	0681-302-5076
E-mail:	siekmann chris melis@ags.uni-sb.de

Ist die Stelle des Leiters/der Leitern des Projektes befristet?

Siekmann:	<input checked="" type="checkbox"/> Nein	<input type="checkbox"/> Ja, befristet bis zum
Benzmüller:	<input type="checkbox"/> Nein	<input checked="" type="checkbox"/> Ja, (zunächst) befristet bis zum 31.12.2003
Melis:	<input checked="" type="checkbox"/> Nein	<input type="checkbox"/> Ja, befristet bis zum

3.1.4 Überführung

entfällt

3.1.5 Vorgesehene Untersuchungen

In dem Teilprojekt sind vorgesehen:

Untersuchungen am Menschen ☐ ja ☒ neinklinische Studien im Bereich der
somatischen Zell- oder Gentherapie ☐ ja ☒ neinTierversuche ☐ ja ☒ neingentechnologische Untersuchungen ☐ ja ☒ nein**3.1.6 Bisherige und beantragte Förderung des Teilprojektes im Rahmen des
Sonderforschungsbereichs (Ergänzungsausstattung)**

Haushalts- jahr	Personalkosten	Sächliche Verw.- ausgaben	Investitionen	gesamt
bis 1998	715.2			715.2
1999	240			240
2000	240			240
2001	254.4			254.4
Zwischen- summe	734.4			734.4
2002	274.8	8	47	329.8
2003	274.8	8		282.8
2004	274.8	8		282.8

(Beträge in TDM)

3.2 Zusammenfassung

Im Projekt Ω MEGA wird ein mathematisches Assistenzsystem entwickelt, in dessen Mittelpunkt die wissensbasierte Beweisplanung steht. Beweisplanung abstrahiert von der Beweissuche auf Kalkülebene des traditionellen automatischen Beweisens. Dazu werden häufig wiederkehrende und mathematisch motivierte Muster einzelner Beweisschritte zu sogenannten Beweismethoden zusammengefaßt. Interpretiert als Planoperatoren werden diese dann in einem deliberativen Suchprozess zu Beweisplänen verkettet; domänenspezifisches, mathematisches Vorgehenswissen kann dabei zur Steuerung der Plankonstruktion eingesetzt werden.

Der Erfolg des Beweisplanens hängt entscheidend von der Güte des Methoden- und Strategiewissens ab und ist daher besonders in solchen Anwendungsgebieten erfolgreich, in denen lange Erfahrungen und damit ein reichhaltiges Domänenwissen vorliegen. Zum Beispiel also in etablierten Gebieten der Mathematik. In neuen – oder wenig strukturierten – Anwendungsgebieten aber, wie beispielsweise der Programmverifikation, ist dieses Wissen im allgemeinen nicht verfügbar und generische Vorgehensweisen und Heuristiken sind daher meist erfolgreicher.

In der Praxis liegt natürlich fast immer eine Mischform der beiden oben genannten Situationen vor, und unsere zentrale Arbeitshypothese für diesen Förderzeitraum ist eine Synthese der deliberativen, wissensbasierten Beweisplanung einerseits und einer agentenbasierten, reaktiven Beweissuche andererseits. Insbesondere soll unsere Systemarchitektur, die diese Synthese unterstützt, auch eine flexible Kooperation zwischen Beweisplanung und externen Problemlösesystemen ermöglichen.

Eine wichtige Technik zur Einschränkung des Suchraumes ist *Meta-Reasoning* über die Ziele des ressourcenadaptiven Beweisplaners und dazu sollen neue Techniken entwickelt werden.

Unsere bisherigen Erfahrungen mit der Beweisplanung zeigen deutlich, daß eine mathematisch motivierte, abstrakte Beweisplanungsebene nötig ist, die nicht unbedingt wie die bisherigen Beweismethoden direkt und lokal zu einem Beweis auf Kalkülebene expandierbar sein muß.

Das erweiterte Ω MEGA-System soll in Fallstudien evaluiert werden und es sollen Ansätze zum Lernen von Vorgehenswissen in der Beweisplanung untersucht werden.

3.3 Stand der Forschung

Der aktuelle Stand der internationalen Forschung, insbesondere zu den Themen Beweisplanung, Integration von Deduktionssystemen, mathematische Wissensbanken und agentenbasiertes Theorembeweisen wurde durch Vorarbeiten der Antragsteller mit geprägt. In diesem Abschnitt wird zunächst der Schwerpunkt auf die externe Forschung gelegt, während die eigenen Arbeiten in Abschnitt 3.4 skizziert werden. Wir verweisen ferner auf den SFB Ergebnisbericht zum Projekt OMEGA (B1).

Theorembeweisen mit Hilfe mehrerer Systeme

Die Integration von spezialisierten Beweissystemen erfolgt zum einen zur Leistungssteigerung durch Kooperation und zum anderen als Anwendung von Deduktionssystemen innerhalb größerer Systemumgebungen. So wurde der in Edinburgh entwickelte und auf Induktion spezialisierte Beweisplaner λ CIAM (Richardson, Smaill & Green, 1998) beispielsweise mit dem Beweissystem HOL gekoppelt (Boulton, Slind & Gordon, 1998). Eine Einbettung von Deduktionssystemen in formale Softwareentwicklungsumgebungen wurde u.a. dadurch realisiert, daß der Induktionsbeweiser INKA (Autexier, Hutter, Mantel & Schairer, 1999) in das VSE-tool (Autexier et al., 2000) integriert wurde. In einer Zusammenarbeit mehrerer europäischer Universitäten wird 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.

Zur Integration heterogener externer mathematischer Services, wie Deduktionssysteme, Computeralgebrasysteme, Modellgenerierer, etc., in eine mathematische Assistenzumgebung wurde in unserer Gruppe das MATHWEB (Franke & Kohlhase, 1999; Franke, Hess, Jung, Kohlhase & Sorge, 1999) System entwickelt. Dieses System wird unter anderem von mehreren Projektpartnern im EU-Netzwerk CALCULEMUS verwendet.

Beweisplanung

Im Forschungsgebiet KI-Planen werden zunehmend manuell akquiriertes oder automatisch extrahiertes Kontrollwissen und domänenspezifische Constraints eingesetzt (siehe den Überblick (Fox, 2000)). Typischerweise ist jedoch dieses Wissen sehr einfach im Vergleich zu dem Kontrollwissen im Beweisplanen. Eine Spezialrichtung der Steuerung in manchen Planungsdomänen basiert auf sogenannten *heavy-tail Verteilungen* von Lösungsversuchen. Mit solchen

Untersuchungen läßt sich eine Steuerung begründen, die einen statistisch motivierten Neustart von randomisierten Strategien vorschlägt (Gomes, Selman, McAloon & Tretkoff, 1998).

Ein anderer deutlicher Trend – mindestens in Gruppen, die sich mit realistischen, komplexen Planungsproblemen befassen – ist die Behandlung verschiedener Constraints zur Einschränkung des Suchraumes sowie die Implementation von Systemen, die mehrere Strategien zur Lösung eines Problems verwenden können (Wilkins & Myers, 1998; Fink, 1998; Howe, Dahlman, Hansen, Mayrhauser & Scheetz, 1999). Auch in der Deduktion gibt es inzwischen mehrere Ansätze, einen Beweiser mit verschiedenen Heuristiken parallel oder nacheinander laufen zu lassen (Schumann, 1996; Wolf, 1998) bzw. verschiedene Beweissysteme bei der Lösung eines Problems einzusetzen (Denzinger & Fuchs, 1999).

Bundy analysiert die Erfahrungen mit dem gegenwärtigen Beweisplanen (Bundy, 1999), speziell mit den Beweisplanern CIAM und λ CIAM. Die dort genannten Probleme sind teilweise allgemein und auch für den Ω MEGA-Beweisplaner relevant und wir werden auf einige in unserer weiteren Arbeit eingehen, insbesondere die geringe Robustheit, sowie die mangelnde Abstraktion im Beweisplanen.

Agentenbasierte und nebenläufige Architekturen

Multiagentensysteme und agentenbasierte Architekturen sind seit einigen Jahren sehr verbreitet, denn die Möglichkeit, komplexe Aufgaben auf mehrere spezialisierte Agenten zu verteilen, ist für viele Anwendungsgebiete interessant. Einen Überblick über das Gebiet bietet das Lehrbuch (Weiss, 1999). Weitere Monographien zu diesem Thema sind (Jennings & Wooldridge, 1998; Wooldridge & Veloso, 1999; Wooldridge, 2000). (Wilkins & Myers, 1998) stellt einen Multiagentenplanungsansatz der KI-Planung vor, der in der Praxis erfolgreich angewendet wurde.

Eine Klassifikation paralleler Methoden und Strategien speziell für das traditionelle automatische Theorembeweisen gibt (Bonacina, 2000). Die dort vorgestellte Taxonomie paralleler Beweisstrategien unterscheidet zwischen Parallelisierung von Berechnungen auf Termebene, auf Klausel- bzw. Inferenzebene und auf der Ebene der Beweissuche. Die Studie beschränkt sich auf das klassische automatische Theorembeweisen, und befaßt sich nicht mit taktischen Theorembeweisern, Beweisplanern, Modellgenerierern und der Beweissuche mit Hilfe integrierter Computeralgebrasysteme. Der agentenbasierter Ansatz in TECHS (Denzinger & Fuchs, 1998, 1999) realisiert eine parallelisierte und kooperative Beweissuche heterogener traditioneller Beweissysteme für die Logik erster Stufe. Dabei werden partielle Resultate

zwischen verschiedenen Beweisern ausgetauscht und die Kommunikation der Systeme wird durch *Referees* auf Empfänger- und Senderseite gefiltert. TECHS erweitert den TEAMWORK Ansatz (Denzinger & Fuchs, 1994), der sich auf kooperatives agentenbasiertes Beweisen mit einem einzigen Kernbeweiser konzentrierte. Weitere Architekturen für agentenbasiertes Theorembeweisen werden in (Fisher, 1997; Wolf, 1998) diskutiert.

Lerntechniken im automatischen Beweisen

Techniken des Maschinellen Lernens wurden im automatischen Beweisen (Fuchs & Fuchs, 1998; Denzinger & Schulz, 1996; Schulz, 2000) naturgemäß zunächst auf Entscheidungen traditioneller Theorembeweiser angewendet, also beispielsweise die Auswahl von Klauseln und die Parametrisierung von Systemen.

An der Universität Birmingham wurde in Zusammenarbeit mit Saarbrücken ein erster Ansatz zum *Lernen von Methoden* untersucht (Jamnik, Kerber & Benzmüller, 2000, 2001). Dieser Ansatz hat eine Menge von Beispielbeweisen als Eingabe, versucht darin gemeinsame Muster zu erkennen, um dann entsprechende Beweismethoden zu generieren. Dazu wird zunächst von den Eingabebeweisen abstrahiert und dann werden die Sequenzen von Inferenzregeln analysiert. Aus diesen Ausgangssequenzen wird dann ein regulärer Ausdruck konstruiert, der möglichst alle Sequenzen erklären kann. Dieser Ausdruck über Inferenzregeln skizziert die zu erlernende Beweismethode und wird unter Verwendung der ‘Precondition Analysis’ (Silver, 1984; Desimone, 1987) zu einer anwendbaren Beweismethode verfeinert. Bisher ist dieser Gesamtansatz noch auf lineare Eingabebeispiele beschränkt.

Im KI-Planen wurden Lerntechniken mehrfach zum Lernen von Operatoren und Kontrollwissen eingesetzt (Korf, 1985; Etzoni, 1993; Minton, 1993; Borrajo & Veloso, 1996; Kambhampati, Katukam & Qu, 1996; Prez, 1996; Wang, 1996; Leckie & Zukerman, 1998; Khardon, 1999). Insbesondere im Prodigy Projekt an der Carnegie Mellon University gab es mehrere Ansätze in dieser Richtung, siehe (Veloso et al., 1995).

3.4 Eigene Vorarbeiten

Die Vorarbeiten sind bereits im Berichtband ausführlicher dargestellt und werden daher hier nur in aller Kürze und mit teilweise anderen Akzenten zusammengefaßt: *Beweisplanen*, *Integration externer Systeme*, *mathematische Wissensbank* und *Fallstudien*. Vergleichsweise ausführlicher wird hier über die neuen Arbeitsgebiete *agentenbasiertes Theorembeweisen* und

Lernen von Steuerungswissen berichtet.

Beweisplanen

Das Ω MEGA-Projekt ist im Berichtszeitraum zur international führenden Arbeitsgruppe im Beweisplanen geworden. Insbesondere haben wir in der letzten Förderperiode das *wissensbasierte* Beweisplanen entwickelt (Melis, 1998b; Melis & Siekmann, 1999a, 1999b) sowie in mehreren Fallstudien erprobt (Melis, 1998a; Cheikhrouhou & Siekmann, 1998; Meier & Sorge, 2000b, 2000a). Diese Technik macht neben der Verwendung allgemeinen mathematischen Wissens die Notwendigkeit domänenspezifischer mathematischer Methoden, Strategien und Kontrollwissens explizit. Dieses domänenspezifische Wissen schließt auch domänenspezifische externe Systeme, wie etwa Computeralgebrasysteme und Constraintlöser ein. Daher widmeten sich eine Reihe von Untersuchungen den Mechanismen zur Integration solcher Systeme (Melis & Sorge, 1999; Melis, Zimmer & Müller, 2000b). Darüberhinaus wurden Standard-Constraintlöser an die Erfordernisse des Beweisplanens angepaßt und ein neues System zur symbolischen *Constraintberechnung* entwickelt (Melis, Zimmer & Müller, 2000a).

In den Fallstudien haben wir uns neben der Erarbeitung von (wenigen) domänenspezifischen Kontrollregeln die Frage gestellt, welche allgemeinen Regeln und Mechanismen für die Steuerung der Suche entwickelt und eingesetzt werden können, siehe (Melis & Pollet, 2000) und (Meier, Gomes & Melis, 2001b). Das Multi-Strategie-Beweisplanen (Melis & Meier, 1999, 2000) erlaubt es, ein Problem mit mehreren Lösungsstrategien anzugehen. Die Notwendigkeit eines solchen Vorgehens steht inzwischen auch international außer Frage, es gilt nun den geschaffenen Systemrahmen mit weiteren mathematisch sinnvollen Strategien auszufüllen und insbesondere die strategische Kontrolle zur Steuerung der Strategieauswahl zu verbessern.

Neue wissenschaftliche Fragestellungen sind durch unsere Arbeit im Beweisplanen entstanden, von denen hier nur zwei kurz erwähnt werden sollen: (1) In den Fallstudien zeigte sich, daß der Grundsatz der direkten Expansion von Methoden in den ND-Kalkül zu starken Einschränkungen führt und darüberhinaus die derzeitig verwendete Plandatenstruktur zu überladen und ineffizient ist. (2) In Kooperation mit dem Teilprojekt KNAC (B3) wurden erste experimentelle Untersuchungen zur Verwendung von Beweisplanmethoden in Lernmaterial durchgeführt (Melis, Glasmacher, Ullrich & Gerjets, 2001). Diese Untersuchungen sollen in der dritten Förderperiode konsolidiert werden.

Agentenbasierte Architekturen für das Theorembeweisen

In einer Pilotstudie wurde der Kommando-Vorschlagsmechanismus Ω -ANTS (Benzmüller & Sorge, 1998; Benzmüller & Sorge, 1999a, 2000) zur Unterstützung des Benutzers im interaktiven Beweisen entwickelt. Idee dieser zweistufigen Agenten- und Blackboardarchitektur ist es, die zum Teil sehr berechnungsintensiven Suchaufgaben in der Generierung von Kommandovorschlägen voneinander zu entkoppeln und sie nebenläufig zu realisieren.

Während der Benutzer über die weitere Konstruktion eines partiellen Beweises nachdenken kann, nutzt dieser Vorschlagsmechanismus in reaktiver Weise die nun freien Systemressourcen aus, um nach erfolgversprechenden Taktiken zu suchen oder die Anwendbarkeit externer Beweissysteme zu überprüfen und dies an den Benutzer zu kommunizieren (Benzmüller & Sorge, 2000). Externe Systeme werden dabei von Ω -ANTS indirekt über den mathematischen Softwarebus MATHWEB angesprochen. Die ermittelten Kommandovorschläge werden heuristisch bewertet und dem Benutzer dynamisch angezeigt. Wegen der Komplexität externer Systeme ist der kritische Umgang mit verfügbaren Systemressourcen eine Herausforderung. Deshalb wurde die ursprünglich ressourcenadaptierte Ω -ANTS Architektur in (Benzmüller & Sorge, 1999a, 1999b) konzeptuell zu einer ressourcenadaptiven weiterentwickelt.

Der Ω -ANTS-Vorschlagsmechanismus wurde um einen Automatisierungsmechanismus ergänzt und in einer Kooperation mit der Universität Birmingham zu einer allgemeinen Architektur für das agentenbasierte Theorembeweisen erweitert (Benzmüller, Jamnik, Kerber & Sorge, 2000, 1999) und in einem Prototypen realisiert (Benzmüller, Jamnik, Kerber & Sorge, 2001).

Lernen von Steuerungswissen

Neben den Arbeiten zum *Lernen von Methodenwissen* (siehe *Stand der Forschung*) haben wir auch erste Untersuchungen zum *Lernen von Kontrollwissen* durchgeführt. Dieser Ansatz benutzt statistische Daten über Wahrscheinlichkeitsverteilungen von Problemlöseversuchen in einer Klasse von Problemen, um Steuerungswissen über die Weiterführung oder den Abbruch eines Versuchs zu lernen (Meier, Gomes & Melis, 2001a). In diesen Arbeiten sind Techniken, die aus der KI bekannt sind (Gomes et al., 1998), auf das Gebiet des automatisches Theorembeweisen angewendet worden. Im Weiteren werden wir mit anderen Randomisierungen und Untersuchungen der Wahrscheinlichkeitsverteilung die Möglichkeiten dieser Technik für Steuerungsentscheidungen, speziell für strategische Kontrolle ausschöpfen.

3.5 Arbeitsprogramm (Ziele, Methoden, Zeitplan)

Das langfristige Ziel unserer Forschungen ist die weitere Entwicklung der wissensbasierten Beweisplanung und im besonderen (i) die Integration des deliberativen, wissensbasierten Beweisplanens mit dem reaktiven agentenbasierten Theorembeweisen, (ii) die Trennung der Beweisplanebene von der Logikebene und (iii) das Lernen von Vorgehenswissen.

Die Integration (i) sowie die Trennung der Beweisplanrepräsentation von dem ND-Kalkül (ii) erfordert umfangreiche Veränderungen an der Architektur des Ω MEGA-Systems, und wird einen erheblichen Implementierungsaufwand nach sich ziehen.

3.5.1 Methoden und Arbeitspakete

Methodisch werden wir im Ω MEGA-Projekt wie folgt vorgehen: (i) Empirische Untersuchungen in Form von Fallstudien und Experimenten, (ii) die Modellierung der Ergebnisse und (iii) die Umsetzung dieser Modellierung in einem Redesign und prototypischer Implementierung im Ω MEGA-System.

Empirische Evaluierungen sind sowohl für den derzeitigen Multi-Strategie-Planer MULTI als auch für den agentenbasierte Beweiser Ω -ANTS geplant. Nach Abschluß der Versuchsreihen werden die Vor- und Nachteile von MULTI und Ω -ANTS explizit gemacht und in einer Modellierungsphase in einer konkreten integrierten Systemarchitektur umgesetzt. Der Modellierung des erweiterten Systems wird sich dann die Implementierung anschließen, bevor das erweiterte System anhand von Fallstudien erneut empirisch evaluiert wird. Ganz ähnlich werden wir vorgehen bei der Entkoppelung der Planungsebene von der logischen Ebene. Die Erfahrungen aus Fallstudien sollen genutzt werden, um Anforderungen an die (neue) Planungsebene und deren Sprache auszuarbeiten.

Experimentelle Methoden werden in Zusammenarbeit mit dem Lehrstuhl Prof. Tack zur Untersuchung der Wirksamkeit von Beweisplanmethoden genutzt.

AP1: Integration von Beweisplanung und agentenbasiertem Theorembeweisen

Durch die Kopplung von wissensbasiertem Beweisplanen mit agentenbasierten Techniken wird die zentrale Ressourcengröße *Steuerungswissen* ergänzt durch die Größen *Rechen- bzw. Planungszeit* und *Speicherplatz*. Dadurch ergeben sich neue Möglichkeiten zur Steuerung der Beweissuche, da bisherige Erfahrungen zeigen, daß sich die Stärken und Schwächen der deliberativen Beweisplanung und des agentenbasierten Theorembeweisens sinnvoll ergänzen. Eine Idee der Integration ist, daß die reaktiven Beweisagenten den deliberativen Planungspro-

zess dynamisch über die Optionen in der Beweissuche informieren und die Entscheidungen dann mit Hilfe des verfügbaren Vorgehenswissens getroffen werden. Es sind aber auch andere Arbeitsteilungen denkbar: zum Beispiel zunächst die Berechnung eines Inselplanes durch die deliberative Planung und dann das Ausfüllen der Lücken durch den Agentenmechanismus.

Konkret sollen also die Ansätze von MULTI und Ω -ANTS gekoppelt werden und ihre Integration soll in einem agentenbasierten System mit robuster Beweisplanung münden. Die Agentenebene soll in dem System eine (möglichst) vollständige Beweissuche im Kalkül des natürlichen Schließens realisieren und es ferner ermöglichen, die bisher sequentialisierten Methoden-Anwendbarkeitstests zu parallelisieren. Die deliberative Planungsphase soll dagegen weitgehend von dem zugrundeliegenden ND-Kalkül entkoppelt werden und eine abstrakte Planungsebene anbieten. Neben der zentralen Ressource *verfügbares mathematisches Vorgehenswissen* rücken nun wieder die traditionellen Größen *Rechenzeit* und *Speicherplatz* in den Mittelpunkt.

Die Integration der Beweisplanung mit dem agentenbasierten Beweisen verspricht folgende Fortschritte:

1. Eine Kombination von deliberativen und reaktiven Problemlösemechanismen
2. Eine robuste (*not brittle*) Beweisplanung
3. Eine Verbesserung der Ressourcenadaptivität des Beweisplanens

1.) Die deliberative Steuerung im Beweisplaner soll um ein reaktives ressourcengesteuertes Zusammenspiel heterogener Beweisagenten erweitert werden. Die Beweisagenten kapseln dabei in uniformer Weise Kalkülregeln, Taktiken, Methoden und externe Systeme. Ein Ziel ist es beispielweise, daß zur Planzeit generierte Unterprobleme stets mit adäquaten Mitteln und Ressourcen angegangen werden können. Das heißt, daß beispielweise ein Beweisagent, der für einen klassischen Beweiser erster Stufe arbeitet, autonom überprüft, ob er ein Unterproblemlösung selbst lösen kann. Bei Erfolg zeigt er dies dem Beweisplaner an. Interessant ist auch das Zusammenspiel mit einem Modellgenerier der reaktiv nach Gegenbeispielen sucht. Die Herausforderung besteht also darin, ein geeignetes und erfolgreiches Zusammenspiel zwischen einem eng gesteuerten Vorgehen mittels domänenspezifischer Beweisstrategien und einem reaktiven Spiel mit Beweismethoden und externen Systemen zu modellieren und zu realisieren. Das Suchverhalten soll sich an der Qualität und Quantität des verfügbaren Steuerungswissen orientieren, das heißt, je weniger Steuerungswissen verfügbar ist, je stärker soll das Gesamtsystem auf die reaktiven Beweisagenten zurückgreifen, im schlimmsten Fall bis zu einer reaktiven uninformierten Beweissuche auf ND-Ebene.

Argumente für eine derartige Verbindung deliberativer und reaktiver Vorgehensmuster liefert auch ein Blick auf das Problemlöseverhalten eines Mathematikers (siehe z.B. (Hadamard, 1945)). Ein Problem in einer vertrauten Domäne löst ein Mathematiker typischerweise dadurch, daß er zunächst abwägt, welche der ihm bekannten Beweistechniken für diese Domäne im gegebenen Fall am geeignetsten erscheint, bevor er diese dann in routinierter Weise anwendet. Scheitert er dabei, greift er auf alternative Strategien zurück oder weicht leicht von der routinemäßigen Vorgehensweise ab, indem er eine geeignete Adaption bzw. Reparatur im Hinblick auf die konkrete Problemstellung durchführt. Ein solches Verhalten wird durch das deliberative Beweisplanen widergespiegelt.

Anders sieht es aus, wenn er eine ihm bis dahin fremde Problemdomäne exploriert. Es bleibt ihm dann nichts anderes übrig, als mit den ihm bekannten Methoden und Strategien *zu spielen*. Möglicherweise entdeckt er dabei Lösungsstrategien, die er erfolgreich auf die neue Domäne übertragen kann, oder er entwickelt allmählich ein Gespür für die Strukturen in der neuen Domäne mithilfe dessen er dann bekannte Strategien adaptieren kann.

2.) Die Beweisplanung setzt adäquates Methoden- und Steuerungswissen zum Lösen von Problemen in einer Anwendungsdomäne voraus. Fehlt solches Wissen, z.B. bei der Exploration neuer Domänen, bricht die Planung typischerweise zusammen. Wir sehen zwei Möglichkeiten, um das Problem der *Zerbrechlichkeit* (*brittleness*, A. Bundy) im Beweisplanen anzugehen. Zum einen soll eine reaktive Integration externer Systeme ermöglichen, lösbare Unterprobleme als solche zu erkennen und dann das passende externe System auch einzusetzen. Diese reaktive Integration externer Systeme könnte sich insbesondere im Zusammenspiel mit dem Inselplanen als fruchtbar erweisen. Inselplanen strukturiert den Suchraum, indem es Unterziele zur Lösung des Gesamtproblems vorgibt, die jeweiligen Ableitungen und Beziehungen der Unterziele aber offen läßt. Das heißt, die Hoffnung ist, daß die übrigbleibenden Unterprobleme von den Agenten in reaktiver Weise geschlossen werden können.

Fehlt es an Methoden und Steuerungswissen, greift der Mechanismus auf anwendbare externe Systeme und Taktiken zurück. Dieser stufenweise Abstieg soll im schlimmsten Fall bis zu einer vollständigen Beweissuche im ND-Kalkül reichen. Werden dann auf dieser Ebene Unterziele generiert, die wiederum auf informierter Ebene angegangen werden können, so soll der Mechanismus in adäquater Weise reagieren.

3.) Durch die Kombination von traditionellem Beweisplanen mit agentenbasiertem Theorembeweisen ergeben sich neue Möglichkeiten zur Realisierung von ressourcenadaptivem Beweisplanen. Für das rein agentenbasierte Beweisplanen konnte ein ressourcenadaptives Ver-

halten realisiert werden (siehe AP3).

AP2: Beweisplanung

Dieses Arbeitspaket besteht aus mehreren zusammenhängenden Teilaufgaben: (i) die verbesserte Steuerung der Beweissuche durch Meta-Reasoning, (ii) die weitere Entwicklung von MULTI insbesondere der Strategien und des strategischen Kontrollwissens, (iii) die prinzipielle Abstraktion der Beweisplanebene von der Logikebene, (iv) die stärkere Hierarchisierung der Beweisplanung und (v) Untersuchungen zur kognitiven Adäquatheit von Beweisplanmethoden. Neben diesen Forschungsaufgaben werden Arbeiten zur Akquisition und Repräsentation von Beweisplanmethoden und Forschungen zum Constraintlösen im Beweisplanen für verschiedene mathematische Domänen fortgeführt, letztere in Zusammenarbeit mit dem Projekt NEP (MI 6).

Steuerung der Beweissuche durch Meta-Reasoning. Für die Beweisplanung haben wir bisher hauptsächlich Vorgehenswissen (in Form von Kontrollregeln) für bestimmte mathematische Domänen akquiriert und implementiert. Für die untersuchten Klassen von Theoremen haben wir nur wenige Kontrollregeln benötigt – pro Unterkapitel eines Standardlehrbuches etwa vier bis fünf. Das heißt, wir haben versucht, nicht immer speziellere Kontrollregeln und Methoden zu entwerfen, mit denen immer mehr Einzelfälle abgedeckt werden können, sondern wenige allgemeinere Kontrollregeln zu finden. Einige von uns entwickelte Methoden, etwa zu indirekten Beweisen und zur Diagonalisierung, gehen bereits über mathematische Domänengrenzen hinweg. In der kommenden Förderperiode wollen wir neben der Kodierung speziellen Domänenwissens auch weiteres generelles Kontrollwissen finden: Kontrollwissen, das sich auf den Zustand des Constraintlösers bezieht (beispielsweise: solche Beweisschritte werden bevorzugt, die die Instantiierung von Variablen stärker einschränken und so bestimmte Suchzweige einschränken), das sich auf die Theorie bezieht, zu der das Problem gehört (beispielsweise: bestimmte Hauptsätze oder Techniken werden in einer Theorie oft angewendet), das sich auf die Behandlung von Zusatzinformation bezieht (beispielsweise: die Dimension mathematischer Größen (kg , km , km/h), die an bestimmten Stellen einer Problemlösung eine Rolle spielen kann) und auf bisherige erfolgreiche bzw. erfolglose Beweisversuche (beispielsweise: um Fortfahren zu können, muß man an der Stelle X eine Fallunterscheidung einführen). Für letzteres ist insbesondere die Entdeckung, Analyse und Modellierung von mathematischer Reflexion über Fehler in Beweisversuchen geplant, vergleichbar den in der Gruppe von Alan Bundy entwickelten *Critics* (Ireland, 1992; Ireland & Bundy, 1996).

Multi-Strategie-Planung. Neben der Untersuchung und Entwicklung neuer Strategien in MULTI, ist die Entdeckung und Entwicklung *strategischen* Kontrollwissens unabdingbar. Daher soll die Steuerung der Multi-Strategie-Planung weiterentwickelt werden und um ein explizites Ressourcenreasoning ergänzt werden. Hierfür werden wir einerseits untersuchen, ob sich *heavy-tail* Techniken auch für das Wechseln zwischen Strategien anwenden lassen. Andererseits soll heuristisches Vorgehenswissen für Klassen von Problemen akquiriert und formalisiert werden.

Abstraktere Beweisplanebene. Als wesentliche und prinzipielle Schlussfolgerung aus Problemen, die in den bisherigen Fallstudien aufgetreten sind, soll die Beweisplanung noch stärker als bisher von der Ebene des zugrunde liegenden Logikkalküls getrennt werden. Der bisherige Ansatz, der vorsieht, daß jede Beweisplanmethode lokal expandierbar sein soll, hat sich als zu einschränkend für das Beweisplanen mathematischer Probleme erwiesen. Die Gründe dafür sind:

- Die Reihenfolge der Beweisplanungsschritte wird bereits im Prozeß des Planens festgelegt und das macht das nachträgliche Einfügen von Schritten oder die Umordnung von Beweisschritten, die aus logischen Gründen erforderlich sein kann (z.B. zur Einhaltung von Eigenvariablenbedingungen oder um geeignete Hypothesenmengen zu erhalten), unmöglich oder ineffizient.
- Die derzeitige Behandlung von Typen und Sorten ist vergleichsweise ineffizient und tritt auch an Stellen auf, an denen sie mathematisch irrelevant ist. Eine Behandlung von Typ- und Sorteninformationen auf einer niedrigeren Abstraktionsebene könnte sinnvoll sein, was im mathematischen Vorgehen auch Sinn macht, denn diese Informationen werden in erster Linie benutzt, um den Suchraum einzuschränken (z.B. in Algebraaufgaben beim Einsetzen von Größen mit einer bestimmten Dimension). Diese Behandlung kann erleichtert werden durch die Einführung einer weniger logikorientierten Planungsebene.
- Manche mathematischen Objekte sind nur mit großem Aufwand logisch korrekt zu repräsentieren (beispielsweise punktweise definierte Funktionen). Oft sind aber die Details der konkreten logischen Repräsentation für das Beweisplanen irrelevant, ja sogar störend, wenn beispielsweise durch die logische Repräsentation einfacher mathematischer Konzepte riesige Ausdrücke entstehen. Es wäre oft natürlicher und effizienter eine detaillierte, logisch korrekte Repräsentation erst anschließend an den Beweisplanungsprozess einzuführen.
- Die derzeitig verwendete Plandatenstruktur ist zu überladen, einerseits mit logischen Informationen und andererseits mit Informationen über den Planungsverlauf (Methoden-

matchings, Variablenbindungen, erfolglose Alternativen bis hin zur Strategieinformationen) und dadurch ineffizient.

Es gibt erste Vorstellungen dazu, welche Eigenschaften die neue, weniger logikorientierte Beweisplanebene haben sollte: Beweispläne sollen hierarchisch und halbgeordnet sein. Zusätzliche Information könnte etwa in Form von Constraints (der Reihenfolge, der Instantiierung etc.) gespeichert sein. Methoden, die mathematisch wenig relevante logische Inferenzen ausführen, sollen zunächst zurückgestellt werden. Es ist denkbar, daß die Beweisplanung künftig nur noch “Inseln” eines logischen Beweises finden wird und andere Mechanismen (z.B. das reaktive, agentenbasierte Beweisen) die verbliebenen Lücken füllen. Das wird aber nicht heißen, daß Suche völlig aus der Beweisplanung verschwinden wird.

Wir werden zur Lösung dieser Forschungsfrage zunächst eine mathematisch motivierte Sprache für (Zwischen-)ziele und Beweisplanschritte entwickeln und die Beweispläne mit Constraints – etwa für die Reihenfolge von Schritten in den endgültigen Beweisen – anreichern. Wir wollen zwar prinzipiell die globale Übersetzbarkeit von Beweisplänen in logische Beweise erhalten, damit deren logische Korrektheit überprüft werden kann, aber das wird hier nicht mehr im Vordergrund stehen. Die theoretischen Untersuchungen auf der Grundlage und als Erweiterung von *logical frameworks* sollen im Schwesterprojekt λ -PLAN (MI 5) durchgeführt werden.

Insgesamt ist die Definition und Realisierung dieser Ebene des Beweisplanens eine völlig offene Forschungsfrage und das Ergebnis noch nicht abzuschätzen. Es ist aber beispielsweise vorherzusehen, daß die rekursive *lokale* Expandierbarkeit von Beweisplänen zu ND-Beweisen verlorengehen wird.

Kognitive Adäquatheit von Beweisplanmethoden. Zur Konsolidierung und Vertiefung der vorläufigen Ergebnisse der vergangenen Förderperiode (Melis et al., 2001), planen wir gemeinsam mit Dr. Gerjets (SFB-Projekt STAR-LIKE, EM 3) weitere Experimente, in denen die kognitive Effektivität eines Instruktionsformats, das auf Methoden der Beweisplanung in Ω MEGA beruht, gezeigt werden soll. Offene Fragen sind etwa: Was sind die Gründe für die bessere Performanz der Probanden? Wie wirken sich diese Instruktionen bei schwierigen Testproblemen und bei Testproblemen, die oberflächlich unähnlich zu dem Instruktionsbeispiel sind aus? Darüberhinaus soll das Selektionsverhalten von Probanden unter Ressourcenbeschränkungen untersucht werden. Versuchspersonen mit niedrigem oder hohem Vorwissen erhalten dabei in einer Lernphase die drei bereits evaluierten Instruktionsformate zur Auswahl. Das Selektionsverhalten der Versuchspersonen und dessen Zusammenhang mit der Performanz in nachfolgenden Problemlöseaufgaben sollen erfasst werden.

AP3: Agentenbasiertes Theorembeweisens

Die bisherigen Pilotstudien zum agentenbasierten taktischen Theorembeweisen werfen folgende Fragen auf:

Reaktive generische Inferenzmaschine. Obwohl der Ω -ANTS Mechanismus im Kontext von Ω MEGA entstanden ist, bietet er einen einheitlichen und logikunabhängigen Rahmen zur agentenbasierten Modellierung von Kalkülregeln, Taktiken, Methoden und externen Systemen. Die Anwendbarkeit von Regeln, Taktiken, Methoden und externen Systemen wird dann reaktiv von den modellierten Agenten analysiert.

Parallelisierung von Berechnungen auf Term-, Inferenz-, und Suche Ebene. Der Ω -ANTS Mechanismus kann im Zusammenspiel von Ω MEGA und MATHWEB im Sinne von (Bonaccina, 2000) parallelisiert werden, um Berechnungen auf Termebene, Inferenzebene und auf der Ebene der Beweissuche nebenläufig zu berechnen. Eine weitere technische Frage betrifft die Kommunikation von Ressourcen an Beweisagenten, die externe Beweissysteme kapseln und mittels MATHWEB über das Internet distribuieren. Ziel ist, die MATHWEB Agenten die Ressourcenvorgaben der Ω -ANTS Agenten übernehmen zu lassen und anhand dieser die Berechnungsprozesse der Beweiser auf der lokalen Maschine zu kontrollieren.

Dynamisierung der Menge aktiver Inferenzregeln. Die Menge der jeweils aktiven Inferenzregeln kann im Ω -ANTS Ansatz grundsätzlich zur Laufzeit verändert werden. Dies erfolgt durch das Addieren, Löschen und Modifizieren der Beweisagenten, die eine Inferenzregel ein-kapseln. Dies wird durch eine deklarative Agentenspezifikationssprache unterstützt.

Flexible Integration und Kooperation von heterogenen externen Systemen. Die genannte Dynamisierung kann auch eine flexible Integration externer Systeme unterstützen, die nun wie Inferenzregeln zur Laufzeit in den Beweisprozess integriert werden. Es muß untersucht werden, inwieweit eine flexible Kooperation zwischen den integrierten externen Systemen unterstützt werden kann; z.B. zwischen Beweisern für die Logik höherer und erster Stufe, Modellgenerierern und Computeralgebrasystemen.

Mechanismen zum Ressourcenadaptiven Beweisen. Die Performanz der Ω -ANTS Beweisagenten wird durch Selbstevaluationen ermittelt und im System kommuniziert. Dieses Performanzwissen kann nun verwendet werden, um eine ressourcenadaptive Steuerung zu erreichen.

Integration von Interaktion und Automatisierung. Die Vorschläge von Ω -ANTS werden dynamisch und zeitgleich an Benutzer und Beweiser geliefert und ein automatischer Beweisprozess kann an jeder Stelle unterbrochen und interaktiv fortgesetzt werden, bzw. umgekehrt. Ω -ANTS qualifiziert sich damit auch als ein potentiell interessanter Mechanismus für

das *Mixed-Initiative Beweisplanen*.

In den laufenden Untersuchungen wird mithilfe von Ω -ANTS ein Zusammenspiel von taktischem Theorembeweisen, externen Beweissystemen, externen Computeralgebrasystemen, Modellgenerieren und Beweisen im ND-Kalkül realisiert. Weil eine automatisierte Beweissuche ND-Kalkül nur sehr schwer zu realisieren ist, setzen wir dazu den an der Carnegie Mellon University entwickelten NIC-Kalkül (Byrnes, 1999; Prawitz, 1965) ein. Die derzeitige Abbildung der Steuerungsmechanismen dieses Kalküls auf kooperierende Ω -ANTS Agenten gilt es weiter zu verbessern und theoretisch zu fundieren.

AP4: Evaluierung anhand von Fallstudien

Fallstudien werden, wie schon in unserer bisherigen Forschung, einerseits der Entdeckung neuer Phänomene im mathematischen Problemlösen dienen, die dann in der Beweisplanung simuliert werden, und andererseits der Erprobung der Repräsentation, der Mechanismen und der Architektur, die neu entwickelt werden.

Geplant ist eine weiterführende Studie aus dem Gebiet der Analysis (das Lehrbuch (Bartle & Sherbert, 1982)), die auch eine Ausweitung unserer Erfahrungen aus der früheren Studie zu epsilon-delta-Beweisen bringen soll. Eine weitere Studie wird sich mit Problemen aus der elementaren Mengentheorie befassen und soll in enger Kooperation mit dem Projekt DIALOG (MI 3) durchgeführt werden. Dieses Projekt untersucht den tutorieller Dialog mit einem mathematischen Assistenzsystem und greift auf Ω MEGA als dynamische Wissensquelle zurück. Als konkrete mathematische Untersuchungsdomäne wurde in diesem Projekt zunächst die elementare Mengentheorie ausgewählt, die deshalb in einem ausreichenden Maße automatisiert werden soll. Darüber hinaus werden wir wahrscheinlich Problemklassen aus dem Gebiet der Algebra untersuchen, da sich deren Repräsentation, Problematik und Techniken deutlich von denen in der Analysis unterscheiden und dies für eine Evaluation der neuen Architektur, die ja nicht nur in einem Gebiet Vorteile bringen soll, sinnvoll sein wird.

In den Fallstudien sollen, wo dies möglich ist, die Performanz von MULTI, Ω -ANTS und des integrierten Ansatzes verglichen werden. Die Analyse des Problemlöseverhaltens wird variierende Ressourcenbeschränkungen einbeziehen.

AP5: Grundlagen für das Lernen von Steuerungswissen

In der Beweisplanung ist die Wissensakquisition, wie auch aus anderen Gebieten der Künstlichen Intelligenz bekannt, schwierig und aufwendig. Daher wollen wir, sofern unsere Kapazitäten dafür ausreichen, beginnen, Teile des benötigten Wissens, speziell des Kontrollwissens, automatisch aus (guten) Beispielen zu lernen. Für die Forschungen zum automatischen Lernen von Steuerungswissen ist zu beachten, daß die meisten Techniken des Maschinellen Lernens sehr viele Lernbeispiele (und Testbeispiele) benötigen. Wir werden daher zunächst an einfachen Beispielen erproben, ob und wie die gewählte Technik erfolgreich anwendbar ist, um dann das Kontrollwissen in größeren Klassen von Theoremen zu lernen.

Zunächst muß geklärt werden, was gute Beispiele bzw. Beispielmengen auszeichnet, sowie für welches Kontrollwissen und für welche Methoden es möglich sein wird, erfolgreich Techniken des Maschinellen Lernens anzuwenden. Außerdem muß die Frage geklärt werden, ob das Lernen von Methodenwissen erfolgversprechender ist, als das Lernen von Kontrollwissen. Die verschiedenen bekannten Techniken (reinforcement learning, rule learning, decision trees, concept learning, genetic algorithms, inductive logic programming) sollen hinsichtlich dieser Lernaufgaben evaluiert werden.

Weitere Lernmechanismen sollen die Selbstevaluation in Ω -ANTS ausnutzen, um Performanzinformationen zu Beweismethoden in einzelnen Problemdomänen zu akquirieren und beim Lernen von Steuerungswissen entsprechend zu berücksichtigen. Auf diese Art sollte sich zum Beispiel ermitteln lassen, wie erfolgversprechend der Einsatz externer Systeme unter gegebenen Ressourcenbeschränkungen in einzelnen Problemdomänen ist.

AP6: Verbesserung der Infrastruktur

Das Ω MEGA-System integriert inzwischen mehr als 20 externe Beweissysteme, Computer Algebra Systeme und Constraint Löser über den mathematischen Software Bus MATHWEB. Wichtig ist hier die Verbesserung der Ressourcenverwaltung der MATHWEB Agenten und die konsequente Umsetzung zu einer ONE-MATHWEB Architektur, in der sich einzelne Broker des MATHWEB Systems dynamisch beim Netzwerk an- und wieder abmelden können. Weiterhin sollen Systeme, die noch nicht über MATHWEB integriert sind (z.B. Induktionsbeweiser), angebunden werden.

Zur Zeit entsteht in einer Kooperation mit Michael Kohlhase an der Carnegie Mellon University die verteilte mathematische Wissensbank MBASE. An der Entwicklung dieser Wissensbank, die eine mathematische Ontologie unterstützen soll, auf die verschiedene Systeme über

das WWW zugreifen können, werden wir als Erstanwender mitarbeiten. D.h. wir wollen auf die Entwicklung von Funktionalitäten dieses Systems, die aus der Anwendungsperspektive des Ω MEGA-System wichtig erscheinen, Einfluß nehmen. Dies betrifft z.B. die Verwaltung von Methoden- und Kontrollwissen.

Das graphische Benutzerinterface $\mathcal{L}\Omega\mathcal{U}\mathcal{I}$ muß den Weiterentwicklungen von Ω MEGA angepaßt werden. Darüberhinaus wäre es sinnvoll, $\mathcal{L}\Omega\mathcal{U}\mathcal{I}$ auf Benutzbarkeitseigenschaften zu untersuchen und anschließend die Schwachstellen zu beseitigen. Letztere Arbeiten werden gemeinsam mit der Gruppe von Prof. Tack/Dr. Gerjets (EM 3 STAR-LIKE) angestrebt, benötigen allerdings zusätzliche Hilfskräfte.

3.5.2 Zeitplan

Zuerst sollen die laufenden Arbeiten und Untersuchungen zu AP2 (Beweisplanung) und AP3 (agentenbasiertes Theorembeweisen) abgeschlossen werden. Auf den Resultaten dieser Studien wird AP1 (Integration beider Ansätze) aufsetzen. Die anderen Arbeiten aus AP1 zum Meta-Reasoning und zur Erarbeitung einer neuen, weniger logikabhängigen Beweisplanebene werden über den gesamten Zeitraum laufen. AP6 (Verbesserung der Infrastruktur) und AP5 (Lernen) sind weitgehend unabhängig von AP1, AP2 und AP3. Die Fallstudien (AP4) werden über den gesamten Zeitraum laufen, teilweise aber abhängig sein von der Fertigstellung einer integrierten Architektur (AP1).

3.6 Stellung innerhalb des Programms des Sonderforschungsbereichs

Zur Zeit sind folgende Kooperationen innerhalb des SFB 378 geplant:

- Mit dem Projekt λ -PLAN (MI 5) ist eine Zusammenarbeit zur theoretischen Fundierung des Beweisplanens geplant. Dies ist für das Ω MEGA-Projekt besonders wichtig, da wir eine mathematisch und kognitiv motivierte Beweisplanebene entwickeln wollen, deren logischer Status unklar ist.
- Das Projekt DIALOG (MI 3) untersucht den multi-modalen Dialog mit einem tutoriellen Assistenzsystem für die Mathematik und das Ω MEGA System übernimmt dabei die Rolle der dynamischen mathematischen Wissensquelle. Im DIALOG Projekt sollen einfache Mengenprobleme als Ausgangspunkt für die Untersuchungen herangezogen werden. Deshalb soll in Kooperation zwischen den Projekten DIALOG und Ω MEGA untersucht werden, inwieweit diese Domäne durch die agentenbasierte Beweisplanung ab-

gedeckt werden kann; dabei liegt der Aspekt nicht nur im Finden von Beweisen bzw. Beweisplänen, sondern ebenso im Entdecken von Gegenbeispielen und der Generierung von Gegenbeispielen.

- Mit NEP (MI 6) werden wir zusammenarbeiten, um das Constraintlösen in der Beweisplanung für neue Domänen anzuwenden und um Informationen über einen Constraintspeicher für die Steuerung von Beweisen auszunutzen.
- Mit STAR-LIKE (EM 3) soll zur Untersuchung der kognitiven Adäquatheit von Beweisplanmethoden kooperiert werden, zum Beispiel in empirisch-psychologischen Studien.

Literatur

- Autexier, S., Hutter, D., Langenstein, B., Mantel, H., Rock, G., Schairer, A., Stephan, W., Vogt, R. & Wolpers, A. (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, S., Hutter, D., Mantel, H. & Schairer, A. (1999). System description: Inka 5.0 – a logic voyager. In *Proc. of 16th International Conference on Automated Deduction* (S. 207–211). Springer-Verlag.
- Bartle, R. & Sherbert, D. (1982). *Introduction to real analysis* (2 Aufl.). Wiley.
- Benzmüller, C., Jamnik, M., Kerber, M. & Sorge, V. (1999). Agent based mathematical reasoning. In *Proc. of the Calculemus Workshop: Systems for Integrated Computation and Deduction*.
- Benzmüller, C., Jamnik, M., Kerber, M. & Sorge, V. (2000). Resource guided concurrent deduction. In *Proc. of the AISB'2000 Symposium 'How to design a functioning mind'*.
- Benzmüller, C., Jamnik, M., Kerber, M. & Sorge, V. (2001). *An agent based approach to reasoning*. (Submitted to International Joint Conference on Artificial Intelligence 2001).
- Benzmüller, C. & Sorge, V. (1998). A blackboard architecture for guiding interactive proofs. In *Proc. of 8th International Conference on Artificial Intelligence: Methodology, Systems, Applications*.
- Benzmüller, C. & Sorge, V. (1999a). Critical agents supporting interactive theorem proving. In *Proc. of the 9th Portuguese Conference on Artificial Intelligence* (S. 208–221). Springer-Verlag.
- Benzmüller, C. & Sorge, V. (1999b). Towards fine-grained proof planning with critical agents. In *Proc. of the 6th Workshop on Automated Reasoning* (S. 19–21).
- Benzmüller, C. & Sorge, V. (2000). Ω -ANTS – an open approach at combining interactive and automated theorem proving. In *Proc. of the Calculemus Symposium 2000*. A.K.Peters.
- Bonacina, M. P. (2000). A taxonomy of parallel strategies for deduction. *Annals of Mathematics and Artificial Intelligence*.
- Borrajo, D. & Veloso, M. (1996). Lazy incremental learning of control knowledge for efficiently obtaining quality plans. *AI Review Journal. Special Issue on Lazy Learning*, 10, 1–34.

- Boulton, R., Slind, K. & Gordon, A. B. M. (1998). An interface between clam and hol. In *Proc. of the 11th International Conference on Theorem Proving in Higher Order Logics* (S. 87-104). Springer-Verlag.
- Bundy, A. (1999, August). *A critique of proof planning*. (This was prepared as a book chapter for Bob Kowalski's festschrift)
- Byrnes, J. (1999). *Proof search and normal forms in natural deduction*. Dissertation, Dept. of Philosophy, Carnegie Mellon University, Pittsburgh, PA 15213, USA.
- Cheikhrouhou, L. & Siekmann, J. (1998). Planning Diagonalization Proofs. In *Proc. of 8th International Conference on Artificial Intelligence: Methodology, Systems, Applications* (S. 167–180). Springer-Verlag.
- Dennis, L. A., Collins, G., Norrish, M., Boulton, R., Slind, K., Robinson, G., Gordon, M. & Melham, T. (2000). The prosper toolkit. In *Proc. of the 6th International Conference on Tools and Algorithms for the Construction and Analysis of Systems*. Springer-Verlag.
- Denzinger, J. & Fuchs, D. (1999). Cooperation of Heterogeneous Provers. In *Proc. of the 16th International Joint Conference on Artificial Intelligence* (S. 10–15). Morgan Kaufmann.
- Denzinger, J. & Fuchs, M. (1994). Goal oriented equational theorem proving using team work. In *KI-94: Advances in Artificial Intelligence – Proc. of KI-94, 18th German Annual Conference on Artificial Intelligence* (S. 343–354). Springer-Verlag.
- Denzinger, J. & Fuchs, M. (1998). A comparison of equality reasoning heuristics. In *Automated deduction. a basis for applications* (Bd. II: Systems and Implementation Techniques, S. 361-382). Kluwer Academic Publishers.
- Denzinger, J. & Schulz, S. (1996). Learning domain knowledge to improve theorem proving. In *Proc. of 13th International Conference on Automated Deduction* (S. 62-76). Springer-Verlag.
- Desimone, R. (1987). Learning control knowledge within an explanation-based framework. In *Progress in Machine Learning — Proc. of 2nd European Working Session on Learning, EWSL-87*. Bled, Yugoslavia: Sigma Press.
- Etzoni, O. (1993). Acquiring search control knowledge via static analysis. *Artificial Intelligence*, 60.
- Fink, E. (1998). How to solve it automatically: Selection among problem-solving methods. In *Proc. of the Fourth International Conference on AI Planning and Scheduling* (S. 128-136). AAAI Press.
- Fisher, M. (1997). An Open Approach to Concurrent Theorem Proving. In *Parallel Processing for Artificial Intelligence* (Bd. 3). Elsevier.
- Fox, M. (Hrsg.). (2000). *Workshop on analysing and exploiting domain knowledge for efficient planning*. AAAI Press.
- Franke, A., Hess, S. M., Jung, C. G., Kohlhase, M. & Sorge, V. (1999). Agent-oriented integration of distributed mathematical services. *Journal of Universal Computer Science*, 5, 156–187.
- Franke, A. & Kohlhase, M. (1999). System description: MATHWEB, an agent-based communication layer for distributed automated theorem proving. In *Proc. of the 16th Conference on Automated Deduction* (S. 217–221). Springer-Verlag.
- Fuchs, M. & Fuchs, M. (1998). Feature-based learning of search-guiding heuristics for theorem proving. *AI Communications*, 11, 175–189.
- Gomes, C. P., Selman, B., McAloon, K. & Tretkoff, C. (1998). Randomization in backtrack

search: Exploiting heavy-tailed profiles for solving hard scheduling problems. In *Proc. of the Fourth International Conference on AI Planning and Scheduling* (S. 208–213). AAAI Press.

- Hadamard, J. (1945). *The psychology of invention in the mathematical field*. Princeton: Princeton Univ. Press.
- Howe, A., Dahlman, E., Hansen, C., Mayrhauser, A. von & Scheetz, M. (1999). Exploiting competitive planner performance. In *Proc. of Fifth European Conference on Planning*. Springer-Verlag.
- Ireland, A. (1992). The use of planning critics in mechanizing inductive proofs. In *International Conference on Logic Programming and Automated Reasoning* (S. 178-189). Springer-Verlag.
- Ireland, A. & Bundy, A. (1996). Extensions to a generalization critic for inductive proof. In *Proc. of 13th International Conference on Automated Deduction* (S. 47-61). Springer-Verlag.
- Jamnik, M., Kerber, M. & Benzmüller, C. (2000). Towards learning new methods in proof planning. In *Proc. of the Calculemus Symposium 2000*. A.K.Peters.
- Jamnik, M., Kerber, M. & Benzmüller, C. (2001). *Learning method outlines in proof planning*. (Submitted to International Joint Conference on Artificial Intelligence 2001).
- Jennings, N. R. & Wooldridge, M. J. (Hrsg.). (1998). *Agent technology: Foundations, applications, and markets*. Springer-Verlag.
- Kambhampati, S., Katukam, S. & Qu, Y. (1996). Failure driven search control for partial order planners: An explanation based approach. *Artificial Intelligence*, 88.
- Khardon, R. (1999). Learning action strategies for planning domains. *Artificial Intelligence*, 113(1-2), 125-148.
- Korf, R. (1985). Macro-operators: A weak method for learning. *Artificial Intelligence*, 26, 35-77.
- Leckie, C. & Zukerman, I. (1998). Inductive learning of search control rules for planning. *Artificial Intelligence*, 101(1-2), 63-98.
- Meier, A., Gomes, C. & Melis, E. (2001a). Extending the reach of proof planning. (Submitted to International Joint Conference on Artificial Intelligence 2001).
- Meier, A., Gomes, C. & Melis, E. (2001b). Heavy-tailed behavior and randomization in proof planning. In *Workshop on Model-Based Validation of Intelligence at AAAI Spring Symposium*.
- Meier, A. & Sorge, V. (2000a). Exploring Properties of Residue Classes. In *Proceedings of the Calculemus Symposium 2000*. A.K. Peters.
- Meier, A. & Sorge, V. (2000b). Exploring the Domain of Residue Classes. In *Workshop on The Role of Automated Deduction in Mathematics at CADE-17* (S. 50–54).
- Melis, E. (1998a). The “limit” domain. In *Proc. of the Fourth International Conference on AI Planning and Scheduling* (S. 199–206) AAAI Press.
- Melis, E. (1998b). AI-techniques in proof planning. In *European Conference on Artificial Intelligence* (S. 494-498). Kluwer.
- Melis, E., Glasmacher, C., Ullrich, C. & Gerjets, P. (2001). Automated proof planning for instructional design. (Submitted to COGSCI 2001).
- Melis, E. & Meier, A. (1999). Proof planning with multiple strategies ii. In *FLoC’99 workshop*

on Strategies in Automated Deduction (S. 61-72).

- Melis, E. & Meier, A. (2000). Proof planning with multiple strategies. In *First International Conference on Computational Logic* (S. 644-659). Springer-Verlag.
- Melis, E. & Pollet, M. (2000). Domain knowledge for search heuristics in proof planning. In *AIPS 2000 workshop: Analyzing and Exploiting Domain Knowledge* (S. 12-15).
- Melis, E. & Siekmann, J. (1999a). Concepts in proof planning. In *Intellectics and computational logic. papers in honor of Wolfgang Bibel* (S. 249-264). Kluwer.
- Melis, E. & Siekmann, J. (1999b). Knowledge-based proof planning. *Artificial Intelligence*, 115(1), 65-105.
- Melis, E. & Sorge, V. (1999). Employing external reasoners in proof planning. In *Proc. of the Calcuemus Workshop: Systems for Integrated Computation and Deduction* (S. 123-134).
- Melis, E., Zimmer, J. & Müller, T. (2000a). Extensions of constraint solving for proof planning. In *European Conference on Artificial Intelligence* (S. 229-233). IOS Press.
- Melis, E., Zimmer, J. & Müller, T. (2000b). Integrating constraint solving into proof planning. In *Proc. of Conference on Frontiers of Combining Systems*. Springer-Verlag.
- Minton, S. (Hrsg.). (1993). *Machine learning methods for planning*. Morgan Kaufmann.
- Prawitz, D. (1965). *Natural deduction: A proof-theoretical study*. Almqvist & Wiskell.
- Prez, M. (1996). Representing and learning quality-improving search control knowledge. In *Machine Learning: Proc. of the Thirteenth International Conference*. Morgan Kaufmann.
- Richardson, J., Smaill, A. & Green, I. (1998). Proof planning in higher-order logic with λ clam. In *Proc. of the 15th International Conference on Automated Deduction* (S. 129-133). Springer-Verlag.
- Schulz, S. (2000). *Learning search control knowledge for equational deduction*. Dissertation, Fakultät für Informatik, Technische Universität München.
- Schumann, J. (1996). Sicotheo: simple competitive parallel theorem provers. In *Proc. of 13th International Conference on Automated Deduction* (S. 240-244). Springer-Verlag.
- Silver, B. (1984). Precondition analysis: Learning control information. In R. Michalski, J. Carbonelle & T. Mitchell (Hrsg.), *Machine Learning 2*. Tioga Publishing Company.
- Veloso, M., Carbonell, J., Pérez, M. A., Borrajo, D., Fink, E. & Blythe, J. (1995). Integrating planning and learning: The PRODIGY architecture. *Journal of Experimental and Theoretical Artificial Intelligence*, 81-120.
- Wang, X. (1996). A multistrategy learning system for planning operator acquisition. In *The Third International Workshop on Multistrategy Learning*. Harpers Ferry.
- Weiss, G. (Hrsg.). (1999). *Multiagent systems : a modern approach to distributed artificial intelligence*. MIT Press.
- Wilkins, D. & Myers, K. (1998). A multiagent planning architecture. In *Proc. of the Fourth International Conference on AI Planning and Scheduling* (S. 154-162). AAAI Press.
- Wilkins, D. E. & Myers, K. L. (1998). A Multiagent Planning Architecture. In *Proc. of the Fourth International Conference on AI Planning and Scheduling* (S. 154-162). AAAI Press.
- Wolf, A. (1998). Strategy selection for automated theorem proving. In *Proc. of 8th International Conference on Artificial Intelligence: Methodology, Systems, Applications*. (S.

452-465).

Wooldridge, M. (2000). *Reasoning about rational agents*. Cambridge, Massachusetts - London, England: MIT Press.

Wooldridge, M. J. & Veloso, M. (Hrsg.). (1999). *Artificial intelligence today : recent trends and developments* (Bd. 1600). Springer-Verlag.

3.7 Ergänzungsausstattung für das Teilprojekt MI 4

PK: Personalbedarf und -kosten (Begründung vgl. 3.7.1)

SV: Sächliche Verwaltungsausgaben (Begründung vgl. 3.7.2)

I: Investitionen (Geräte über DM 20.000 brutto; Begründung vgl. 3.7.3)

Bewilligung 2001				2002			2003			2004		
PK	Verg.- Gr.	Anz.	Betrag DM	Verg.- Gr.	Anz.	Betrag DM	Verg.- Gr.	Anz.	Betrag DM	Verg.- Gr.	Anz.	Betrag DM
	Bat IIa SHK	2	213.600	Bat IIa SHK	2	213.600	Bat IIa SHK	2	213.600	Bat IIa SHK	2	213.600
		2	40.800		3	40.800		3	40.800		3	40.800
	zus.:	4	254.400	zus.:	5	274.800	zus.:	5	274.800	zus.:	5	274.800
SV				Kosten- kategorie oder Kennziff.		Betrag DM	Kosten- kategorie oder Kennziff.		Betrag DM	Kosten- kategorie oder Kennziff.		Betrag DM
				547		8.000	547		8.000	547		8.000
				zus.		8.000	zus.		8.000	zus.		8.000
I				Investitionsmittel insges.			Investitionsmittel insges.			Investitionsmittel insges.		
				47.000								

3.7.1 Begründung des Personalbedarfs

	Name, akad. Grad, Dienststellung	engeres Fach des Mitarbeiters	Institut der Hochschule oder der außeruniv. Ein- richtung	Mitarbeit im Teilprojekt in Std./Woche (beratend: B)	auf dieser Stelle im SFB tätig seit	beantragte Einstufung in BAT
Grundausrüstung						
3.7.1.1 wissenschaftl. Mitarbeiter (einschl. Hilfskräfte)	1. Jörg Siekmann, Prof. Dr. Universitätsprofessor 2. Christoph Benzmüller, Dr., Hochschulassistent 3. Erica Melis, Dr., Privatdozentin 4. Helmut Horacek, Dr., Privatdozent 5. Aljoscha Buschardt, SHK 6. Walter Stephan, SHK	Informatik Informatik Informatik Informatik Computerlinguistik Computerlinguistik	FR Informatik FR Informatik FR Informatik FR Informatik FR Computerlinguistik FR Computerlinguistik	5 10 10 10 8 8	1.1.1996 1.1.2001 1.1.1996 1.1.1996	
3.7.1.2 nichtwissenschaftl. Mitarbeiter	7. Irmtraud Stein, Sekretärin			5	1.1.1996	
Ergänzungsausrüstung						
3.7.1.3 wissenschaftliche Mitarbeiter (einschl. Hilfskräfte)	8. Andreas Meier, Dipl. Inform., Wiss. Mitarb. 9. Andreas Franke, Dipl. Inform., Wiss. Mitarb. 10. Achim Bergmeister, SHK 11. Malte Hübner, SHK 12. Siegfried Scholl, SHK 13. Christian Ludt, SHK	Informatik Informatik Informatik Informatik Informatik Informatik	FR Informatik FR Informatik FR Informatik FR Informatik FR Informatik FR Informatik	38,5 38,5 19 19 19 19	1.1.2000	BAT IIa BAT IIa SHK SHK SHK SHK
3.7.1.4 nichtwissenschaftl. Mitarbeiter						

(Stellen, für die Mittel *neu* beantragt werden, sind mit X gekennzeichnet.)

Mitarbeiter durchnummerieren und Aufgabenbeschreibung nachfolgend erläutern.

Bitte Verfahrensgrundsätze der Deutschen Forschungsgemeinschaft zur Bezahlung wissenschaftlicher Mitarbeiter beachten.

Aufgabenbeschreibung von Mitarbeitern der Grundausrüstung

Die Antragsteller, Jörg Siekmann (MA 1.), Christoph Benzmüller (MA 2.) und Erica Melis (MA 3.) leiten das Projekt und koordinieren die Arbeitspakete. Erica Melis wird sich schwerpunktmäßig auf die Weiterentwicklung des Beweisplanens (AP2) konzentrieren und Christoph Benzmüller auf die Weiterentwicklung des agentenbasierten Beweisen (AP3). Die Integration beider Ansätze sowie die Verbesserung der Infrastruktur und die Durchführung der Fallstudien wird von den Antragstellern gemeinsam betreut. Helmut Horacek (MA 4.) wird bei Arbeiten zur Verbesserung der Präsentation von Beweisen in Ω MEGA beratend zur Verfügung stehen.

Die studentischen Hilfskräfte Aljoscha Buschard (SHK 5.) und Walter Stephan (SHK 6.) werden die mathematische Datenbank MBASE mit dem CASL Tool verbinden.

Aufgabenbeschreibung von Mitarbeitern der Ergänzungsausrüstung

Andreas Meier (MA 8.): Integration von Beweisplanung und agentenbasiertem Beweisen (AP1), Multi-Strategie-Beweisplanung (AP2), Fallstudien (AP4), Lernen (AP5)

Andreas Franke (MA 9.): Abstraktion im Beweisplanen (AP2), Fallstudien (AP4) und Infrastruktur (AP6)

Achim Bergmeister (SHK 10.) Agentenbasiertes Theorembeweisen (AP3)

Malte Hübner (SHK 11.) Agentenbasiertes Theorembeweisen (AP3)

Siegfried Scholl (SHK 12.) Infrastruktur (AP6) insbesondere MATHWEB und MBASE

Christian Ludt (SHK 13.) Infrastruktur (AP6), insbesondere Weiterentwicklung des graphischen Benutzerinterfaces $\mathcal{L}\Omega\mathcal{U}\mathcal{I}$.

3.7.2 Aufgliederung und Begründung der Sächlichen Verwaltungsausgaben (nach Haushaltsjahren)

	2002	2003	2004
Für Sächliche Verwaltungsausgaben stehen als <i>Grundausstattung</i> voraussichtlich zur Verfügung:	2000	2000	2000
Für Sächliche Verwaltungsausgaben werden als <i>Ergänzungsausstattung</i> beantragt (entspricht den Gesamtsummen „Sächliche Verwaltungsausgaben“ in Übersicht 3.7)	8.000	8.000	8.000

(Alle Angaben in DM.)

Begründung zur *Ergänzungsausstattung* der Sächlichen Verwaltungsausgaben

- **(547) Sonstiges**

Die jährlichen Sachmittelzuwendung aus der Grundausstattung des Lehrstuhls von Prof. Siekmann durch die Universität des Saarlandes beträgt derzeit DM 5.800 jährlich; während sie an anderen Lehrstühlen in der Informatik zum Teil über DM 20.000 betragen. Allein durch die DFG gefördert werden derzeit die Projekte OMEGA und LISA innerhalb des SFB, sowie vier weitere DFG Einzelprojekte. Für jedes dieser Projekte geht die DFG von einer Summe von DM 1000 – 2000 aus, die aus der Grundausstattung zur Verfügung stehen sollten. Für den Lehrstuhl von Prof. Siekmann ergibt sich damit ein Mißverhältnis zwischen den ca. DM 8000 die von der DFG als Leistungen aus der Grundausstattung für die Einzelprojekte erwartet werden und den tatsächlich Zuwendungen von nur DM 5.800 durch die Universität des Saarlandes. In den vergangenen Jahren wurde dieses Mißverhältnis immer deutlicher; ein Resultat der Situation ist beispielsweise, daß wir vor einem Monat die Datensicherung für unsere DFG Projekte durch das Rechenzentrum der Universität reduzieren mußten; für eine umfassende Sicherung der kompletten Mitarbeiter Daten müßten jährlich ca. 3000 DM aufgebracht werden. Insgesamt entstanden in den vergangenen Jahren immer wieder kritische Engpässe in der Sachmittelversorgung, die nur durch den räumlichen Zusammenschluß der Ω MEGA Projektgruppe und den DF-KI Projektgruppen von Prof. Siekmann überwunden werden konnten. Durch den Wegzug der DF-KI Projektgruppen in das neue DF-KI Gebäude fallen selbst diese (nicht ganz regelkonformen) Synergiemöglichkeiten nun vollständig weg.

Wir beantragen zum Ausgleich des akuten Mißverhältnisses zwischen den tatsächlichen jährlichen Sachmittelzuwendungen aus der Grundausstattung und dem Bedarf in den SFB

Projekten zur Gewährleistung eines reibungslosen Forschungsablaufs eine Erhöhung der Sachmittelzuwendungen um jährlich DM 8.000

3.7.3 Investitionen (Geräte über DM 20.000,— brutto und Fahrzeuge)

	Beantragt für das Haushaltsjahr		
	2002	2003	2004
Summe:	47.000		

(Alle Preisangaben in DM *einschl.* MwSt., Transportkosten etc.)

Begründung zur Ergänzungsausstattung der Investitionen

- **(515) Kleingeräte (Anschaffungspreis bis 20.000)**

Für die Weiterentwicklung des Ω MEGA Systems, zur Vorführung bei unseren Kooperationspartnern und für Demonstrationen auf Konferenzen beantragen wir zwei leistungsfähige Notebooks. Das Ω MEGA System kann mittels der MATHWEB Architektur leicht über das Internet verteilt werden und davon machen wir auch regen Gebrauch. Für Arbeiten am System (z.B. bei kurzfristigen Aufenthalten bei Projektpartnern) ist aber eine Installation des Ω MEGA Kerns notwendig. Dieses setzt ein installiertes und *lizenzpflichtiges* Allegro Common Lisp 6.0 voraussetzt, was oft nicht gegeben ist. Auch bei Systemdemonstrationen auf Konferenzen stellt sich das Problem, daß oft kein direkter Zugang zum Internet vom Rednerpult aus möglich ist, bzw. die Browser nicht entsprechend konfiguriert sind und beispielsweise die Mozart Oz-Applets unseres graphischen Benutzerinterfaces nicht ausgeführt werden können. Ein besonderes, aus Mozart Oz bedingtes, Problem besteht hinsichtlich des Startens des Ω MEGA-Systems hinter einem Firewall. Diese Probleme machen Ω MEGA Demonstrationen auf Konferenzen ohne eigene Notebooks derzeit nur bedingt möglich. Deshalb beantragen wir zwei leistungsfähige Notebooks (je DM 8.000) auf denen aktuelle Vorführversionen des Systems installiert werden sollen. DM 16.000

Wir beantragen eine Neuausstattung mit Arbeitsplatzrechnern (je DM 7.000) für beide Mitarbeiter auf der Ergänzungsstelle und zwei weitere Hilfskraftarbeitsplätze (je DM 6.000). Die Bildschirme sind mittlerweile ca. 5 Jahre alt und könnten sich bei einer Weiterverwendung über die gesamte kommende Förderperiode zu einem Problem entwickeln. Auch die Ausstattung der Rechner entspricht nicht mehr den Mindestanforde-

rungen im Hinblick auf die arbeitsspeicher- und rechenintensiven Anwendungen im Rahmen des Ω MEGA Projekts. In der kommenden Förderperiode werden diese Leistungsanforderungen enorm ansteigen, da sich durch die Integration ressourcengesteuerter Prozesse (Multi-Strategie-Planung und agentenbasiertes Theorembeweisen) mehr Möglichkeiten für große Testreihen ergeben. Auch die Bereitstellung unserer Beweissysteme über MATHWEB an andere Gruppen (z.B. im Kontext des DORIS Systems in der Computerlinguistik) setzt schnelle und leistungsfähige Rechner voraus. DM 26.000

Durch eine Investition der ACTIVEMATH Projektgruppe am DFKI, mit der wir eng kooperieren, konnten wir die Netzanbindung einiger Mitarbeiter am Lehrstuhl Siekmann bereits auf 100 Megabit erhöhen. Um auch die SFB Mitarbeiter und Hilfskräfte an dieser verbesserten Infrastruktur teilhaben zu lassen und gleichstarke Entwicklungsumgebungen zu garantieren, sind zwei 100 Megabit Switches notwendig (zu je DM 2500). Insbesondere im Hinblick auf die kommenden Fallstudien mit dem Multi-Strategie-Planer sowie dem agentenbasierten Beweisen mit externen Systemen, die über das Netz verteilt sind und deren Zusammenspiel durch MATHWEB koordiniert wird, wird die Netzbelastung stark ansteigen. DM 5.000