

Teilprojekt MI 3:

DIALOG:

Tutorieller Dialog mit einem Mathematik-Assistenzsystem

3.1 Allgemeine Angaben zum Teilprojekt MI 3

3.1.1 Thema

Tutorieller Dialog mit einem Mathematik-Assistenzsystem

3.1.2 Fachgebiet und Arbeitsrichtung

Künstliche Intelligenz und Computerlinguistik

3.1.3 Leiter/in

Name, Geburtsdatum:	Pinkal, Manfred, 24.08.1949 Siekmann, Jörg, 05.08.1941 Benzmüller, Christoph, 08.09.1968
Dienstanschrift:	FR Computerlinguistik / FR Informatik / FR Informatik Universität des Saarlandes D-66041 Saarbrücken
Telefon:	0681-302-4343 / 0681-302-5275 / 0681-302-4574,
Telefax:	0681-302-4351 / 0681-302-5076 / 0681-302-4351
E-mail:	pinkal@coli.uni-sb.de siekmann@ags.uni-sb.de chris@ags.uni-sb.de

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

Pinkal	<input checked="" type="checkbox"/> Nein	<input type="checkbox"/> Ja, befristet bis zum
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

3.1.4 Überführung

entfällt

3.1.5 Vorgesehene Untersuchungen

In dem Teilprojekt sind vorgesehen:

Untersuchungen am Menschen	<input checked="" type="checkbox"/> ja	<input type="checkbox"/> nein
klinische Studien im Bereich der somatischen Zell- oder Gentherapie	<input type="checkbox"/> ja	<input checked="" type="checkbox"/> nein
Tierversuche	<input type="checkbox"/> ja	<input checked="" type="checkbox"/> nein
gentechnologische Untersuchungen	<input type="checkbox"/> ja	<input checked="" type="checkbox"/> 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	1449.6			1449.6
2002	254.4	2.4		256.8
2003	254.4	1.2		255.6
2004	254.4			254.4

(Beträge in TDM)

3.2 Zusammenfassung

Ziel des Projekts ist die empirische Untersuchung, die Modellierung und die Implementierung natürlich-sprachlichen Dialogs in einer tutoriellen Anwendung für ein mathematisches Teilgebiet. Übergeordnete Motivation ist die Untersuchung und Realisierung von Dialogtechniken in einer inhaltlich und pragmatisch anspruchsvollen Anwendung, die einerseits echtes

Sprachverstehen mit einer Tiefensemantik und Dialogstrategien und andererseits echte Wissensverarbeitung voraussetzt.

Im Projekt soll das Konzept eines dialogbasierten tutoriellen Systems realisiert werden, das dem Benutzer mathematische Begriffe aus dem Bereich der Mengentheorie erläutert und interaktiv bei der Durchführung und beim Verstehen von Beweisen Hilfestellung leistet. Das Fachwissen des Tutorsystems soll in dynamischer Form durch Interaktion mit vorhandenen Werkzeugen der Beweisplanung und Mathematik-Assistenz beigesteuert werden.

Das Projekt ist zwar implementierungsorientiert, besitzt aber einen starken empirischen Schwerpunkt. Im Einzelnen sind mehrere aufeinander aufbauende Simulationsexperimente (als *Wizard-of-Oz*-Studien), die konzeptuelle Modellierung und softwaremäßige Realisierung der zentralen Dialogkomponente, sowie die Implementierung eines Demonstratorsystems zur tutoriellen Dialogführung geplant.

3.3 Stand der Forschung

Beide am Projekt beteiligte Gruppen haben seit längerem in den verschiedenen für das DIA-LOG-Projekt relevanten Teilbereichen geforscht und den internationalen Forschungsstand durch ihre Vorarbeiten mitgeprägt. Um Redundanz zu vermeiden, werden wir Aspekte des Forschungsstandes, die auf eigenen Arbeiten beruhen, schwerpunktmäßig unter Abschnitt 3.4 darstellen, und verweisen im Übrigen auf die SFB Ergebnisberichte zu den Projekten OMEGA, LISA, CHORUS, und AGINT.

Mathematische Assistenzsysteme

In den vergangenen drei Jahrzehnten wurden zahlreiche automatische und semi-automatische Deduktionssysteme entwickelt. Neben der Motivation Unterstützungssysteme für die Mathematik zu schaffen, galt das Interesse zunächst auch der Verbesserung des *Commonsense Reasoning* in der KI. Im Hinblick auf das letztgenannte Ziel ist der große Durchbruch bis heute aus mehreren Gründen ausgeblieben. Im Bereich der Mathematik hingegen konnten Teilerfolge erreicht werden, wie z.B. die Lösung des bis dato von menschlichen Mathematikern ungelösten *Robbins Problems* mit dem Beweiser EQP (McCune, 1997). Es ist allerdings auch auf längere Sicht nicht zu erwarten, dass solche Beweissysteme menschlichen Mathematikern vergleichbare Fähigkeiten entwickeln. Ein grundsätzliches Problem bei der Benutzung von Deduktionssystemen liegt in der Kommunikation, weil die logischen Beschreibungsebenen und

die mathematischen Ausdrucksmittel sich fundamental unterscheiden. Optimalerweise sollte ein natürlich-sprachlicher Dialog mit dem System auf einer den Bedürfnissen und der Sprache des Mathematikers bzw. des Lernenden angepassten Abstraktionsebene möglich sein. Traditionelle vollautomatisierte Beweissysteme wie z.B. die Beweiser erste Stufe OTTER (McCune, 1994; McCune & Wos, 1997), oder SPASS (Weidenbach, 1997) sind dafür ungeeignet, weil sie maschinennahe Kodierungen mathematischer Konzepte voraussetzen.

In dieser Hinsicht erscheinen Systeme erfolversprechender, die a) eine expressivere Sprache verwenden, b) Abstraktionsmechanismen bereitstellen, c) Kommunikation auf der Ebene mathematischer Konzepte in einer geeigneten Form unterstützen, d) Teil-Automatisierungen von Routineaufgaben erlauben (evtl. auch durch integrierte traditionelle Beweissysteme) und e) umfangreiche Wissens-Management Aufgaben übernehmen. Beispiele von Systemen, die in diese Richtung zielen, sind die Beweiser für Logiken höherer Stufe TPS (Andrews et al., 1996), PVS (Owre, Rushby & Shankar, 1992), HOL (Gordon, 1985), λ CLAM (Richardson, Smaill & Green,) und Ω MEGA (Benz Müller et al., 1997). Ein Beispiel für einen Abstraktionsmechanismus ist die Faktenebene (Huang, 1994), die Anwendungen von Definitionen, Sätzen, etc. auf einer von der reinen Kalkülebene abstrahierten Form beschreibt. Einige Beweiser, wie z.B. HOL und ISABELLE (Paulson, 1994), unterstützen interaktives Beweisen mit Beweistaktiken (Gordon, Milner & Wadsworth, 1979) und abstrahieren damit von der reinen Kalkülebene. Beweisplanungssysteme wie λ CLAM oder Ω MEGA gehen noch einen Schritt weiter und erlauben die Planung von Beweisen mit Hilfe von Beweismethoden. Beweismethoden sind um zusätzliches Wissen angereicherte Beweistaktiken; sie werden in der Beweisplanung als Planoperatoren verwendet. Wesentliches Charakteristikum der genannten Systeme ist, dass sie ein explizites Beweisobjekt aufbauen und verwalten. Textuelle oder graphische Benutzerschnittstellen ermöglichen die Kommunikation dieser Beweisobjekte an den Benutzer, sie erfordern aber Kenntnisse der logikzentrierten Repräsentationsmechanismen der einzelnen Systeme. Einige Systeme, z.B. Ω MEGA, HOL, und λ CLAM, integrieren externe Unterstützungssysteme in die automatisierte Beweissuche oder bieten sie dem Benutzer interaktiv als Spezialwerkzeuge an. Insbesondere bei Ω MEGA umfassen die integrierten Systeme neben verschiedenen traditionellen Beweissystemen auch Modellgenerierer, Constraintsolver, und Computeralgebrensysteme wie MAPLE (Char et al., 1992) oder GAP (GAP, 1998); insgesamt etwa 20 verschiedene Systeme.

Ein interessantes Anwendungsgebiet für mathematische Assistenzsysteme sind Mathematik-Lernumgebungen. In der Arbeitsgruppe Siekmann entsteht derzeit im Projekt ACTIVE MATH (Melis, 2000b, 2000a) eine solche Lernumgebung, die das mathematische Assistenzsystem Ω MEGA integriert. Die Anwendung von Deduktionssystemen in einem solchen Kontext ist

deshalb nicht unrealistisch, da die Komplexität und Schwierigkeit der Aufgaben in vielen konkreten mathematischen Domänen bereits im Bereich des Leistungsvermögens heutiger Deduktionssysteme liegt. In Lernumgebungen spielen adäquate Repräsentationsformen zur Übermittlung mathematischer Sachverhalte eine wichtige Rolle. Eine Möglichkeit ist die Verwendung von Diagrammen.

Tutorsysteme

Tutorielle Systeme besitzen, ebenso wie natürlich-sprachliche Dialogsysteme, zwar eine lange Forschungstradition, aber erst in den letzten Jahren hat sich in beiden Bereichen ein Stand herausgebildet, der es interessant macht, sie im Zusammenhang *eines* Projektes zu betrachten. Tutorielle Systeme der neusten Generation versuchen den Studenten mittels eines Dialogs über den Gegenstand einer tutoriellen Sitzung dazu anzuhalten, sich dieses tutorielle Ziel gedanklich zu erarbeiten. „Dialog“ ist hier allerdings noch nicht unbedingt im Sinne eines natürlich-sprachlichen Dialogs gemeint. In Ansätzen wurden natürlich-sprachliche Fähigkeiten mit einer eingeschränkten Dialogführung in tutoriellen Systemen umgesetzt. Dazu gehören unter anderem a) Autotutor (Person, Graesser, Harter & Mathews, 2000), ein System das Hinweise zur Beantwortung von Benutzerfragen in gesprochener Sprache erzeugt und durch Mimik eines Tutoragenten unterstützt, b) Geometrietutor (Alevan & Koedinger, 2000), der von dem Studenten eine exakte natürlich-sprachliche Beschreibung eines geometrischen Konzepts erfordert und Hinweise bei fehlerhafter Ausdrucksweise von seiten der Studenten gibt, und c) Ms.Lindquist (Heffernan & Koedinger, 2000), ein Algebratutor, der den Studenten bei der Erarbeitung einer Formel zur Beschreibung eines mäßig komplexen Zusammenhangs durch Zerlegung in Teilaufgaben und Exploration von Beispielen unterstützt. Tatsächlich handelt es sich bei diesen Systemen also um Systeme mit dialogischen Elementen, ohne dass aber das eigentliche Potenzial, das im natürlich-sprachlichen Dialog steckt, ausgenutzt wird. Grundsätzlich fehlt die Möglichkeit der flexiblen natürlich-sprachlichen Eingabe (stattdessen finden menübasierte Verfahren Verwendung oder es wird, wie im Geometrie-Tutor, der exakte Wortlaut verlangt), ebenso übrigens die flexible Behandlung von Formeleingaben (unerwartete Gleichungen können vom System nicht überprüft werden). Die Inhalte der tutoriellen Sitzungen sind zudem in all diesen Systemen sehr eng gehalten. Die zentrale Rolle von Dialogfähigkeiten für den Erfolg tutorieller Sitzungen konnte empirisch nachgewiesen werden (Moore, 2000).

Ein generelles Problem existierender Tutorsysteme auf der Architekturebene ist die fehlende Trennung von Domänenrepräsentation, tutorieller Strategie und natürlich-sprachlichem Dia-

log. Das erschwert die Identifikation domänenspezifischer Aspekte und domänenunabhängiger Anteile und ein modulares Vorgehen bei der Systementwicklung. Perspektiven für eine bessere Lösung werden bei den eigenen Vorarbeiten erwähnt.

Dialogsysteme

Dialogsysteme gehören zu den ersten Anwendungen, mit denen sich die Sprachverarbeitungs-forschung überhaupt beschäftigt hat. Ein Charakteristikum war bis in die Neunziger Jahre hin-ein der geschriebene Interaktionsmodus (in Ermangelung von brauchbaren Spracherkennungs- und Sprachsynthesewerkzeugen), und damit eine gewisse Vernachlässigung von Fragen der Benutzerakzeptanz (die bei geschriebener Interaktion ohnehin nur in geringem Ausmaß vor-handen ist). Eine zweite durchgängige Eigenschaft war der konsequente wissensbasierte An-satz, bei dem das Verhalten des Dialogsystems durch elementare Schlußtechniken auf der Grundlage unterschiedlicher Wissensquellen in ganz wesentlichem Ausmaß bestimmt wurde. Dies betrifft die linguistische Verarbeitung (Analyse und Generierung), die Domänenmodel-lierung, die Dialogplanung sowie die Intentions- und Planerkennung.

Planbasierte Methoden zur Beschreibung aufgabenorientierter, kooperativer Dialoge folgen oft den grundlegenden Arbeiten zur Theorie von Aktion und Interaktion in (Cohen & Le-vesque, 1990), mit Erweiterungen auf gemeinsame Aktionen von im Diskurs involvierten Agenten (Cohen & Levesque, 1991). Beispiele für umfangreiche wissens- und planbasierte Dialogsysteme aus den frühen Neunziger Jahren sind das TRAINS-System (Allen, Miller, Ringger & Sikorski, 1996) und TACITUS (Hobbs, Stickel, Appelt & Martin, 1990).

Die Forschungstradition hat wichtige Einsichten über Voraussetzungen und Strukturen von natürlich-sprachlichen Dialogen erbracht, ist aber im Blick auf praktische Anwendbarkeit nicht erfolgreich gewesen. Die Technik, Dialogverhalten durchgängig als Resultat eines Infe-renzprozesses auf der Grundlage von Wissen unterschiedlichster Art zu berechnen, hat zu ein-drucksvollen Demonstratorsystemen geführt. Sie sind aber generell zu fragil und zu langsam, und eine Erweiterung auf eine realistische Abdeckung wäre zu teuer, um sie zu ernsthaften Kandidaten für Anwendungssysteme zu machen.

In den letzten zehn Jahren haben vor allem rapide Fortschritte in der Spracherkennungstechno-logie zu einem stark wachsenden Markt für natürlich-sprachliche Dialogschnittstellen geführt. Dies hat eine starke Fokussierung auf extrem einfache Dialogtechniken bewirkt, die im Blick auf Benutzerakzeptanz, d.h. hier zunächst auf Effektivität, Robustheit und Echtzeitverhal-ten optimiert werden. Der Regelfall sind Dialogmodelle in Form endlicher Automaten (mit

unterschiedlichen, aber nicht sehr weitgehenden Flexibilisierungsvarianten), Sprachverstehen auf der Ebene von Schlüsselworterkennung, Sprachausgabe durch (parametrisierte) vorgefertigte Ausgaben, und Einbezug von Domänenwissen auf einer recht trivialen Ebene (z.B. Datenbank-Abfragen für Fahrplanauskunft). Beispiele für Dialogsysteme auf dieser Ebene sind (Aust, Oerder, Seide & Steinbiss, 1995; Carpenter & Chu-Carroll, 1998). Ein weiterer Aspekt anwendungsorientierter Dialogforschung ist Multimodalität. Beispielsweise werden multimodale Dialoge im Kontext des MALIN-Systems analysiert, das auf der LINLIN-Architektur beruht (Ahrenberg, Dahlbäck, Jönsson & Thur'ee, 1996; Dahlbäck & Jönsson, 1997; Heeman, Johnston, Denney & Kaiser, 1998).

Seit einigen Jahren gibt es unterschiedliche interessante Versuche, flexible, inhaltsorientierte Dialoggestaltung zu realisieren, ohne sich wie die traditionelle Dialogverarbeitung auf die KI-vollständigen Probleme der allgemeinen Wissensrepräsentation und -verarbeitung einzulassen. Dies kann man erreichen, indem man die jeweils domänen- und anwendungsspezifisch beschränkten und spezialisierten Funktionalitäten des zu modellierenden Dialogtyps ausnutzt, die Mächtigkeit der verwendeten Verfahren genau darauf abstellt, und insbesondere unaufwändige flache Dialogmodellierung zur Präkompilierung uniformer Dialogbestandteile mit sorgfältig ausgewählten Aspekten wissensbasierter Dialogverarbeitung kombiniert. Ein Beispiel ist die sog. ARTEMIS-Technologie, mit der es gelungen ist, ein Dialogsystem zu realisieren, das Angaben in gesprochener Sprache vor dem Hintergrund eines Auskunftssystems interpretiert und mit dem Benutzer einen Klärungsdialog führt, bis die gewünschte Information oder ein angemessener Ersatz dazu gefunden und dem Benutzer mitgeteilt wird. Eine andere vielversprechende Richtung baut auf dem Konzept der sog. Dialogspiele auf (Carlson, 1985; Carletta et al., 1997), die nur bestimmte Folgen von Dialogzügen zulassen oder auf der Modellierung komplexer dynamischer Informationszustände (Traum et al., 1999), die regelbasiert Dialogzüge generieren. Wir kommen auf diese in den EU-Projekten TRINDI und SIRIDUS vorangetriebene Forschungsrichtung unter „Eigene Vorarbeiten“ zurück.

Natürlich-sprachliche Analyse und Generierung

Das DIALOG-Projekt wird in seinen empirischen Simulationsteilen automatische Analyse und Generierung nicht oder nur sehr eingeschränkt verwenden. Im Demonstrator wollen wir die in den Arbeitsgruppen reichlich vorhandenen Techniken pragmatisch einsetzen, um exemplarische tutorielle Dialogsequenzen zu realisieren. Wir fassen uns deshalb bei der Beschreibung des Forschungsstandes kurz und verweisen auf den Abschnitt 3.4.

Zu Spracherkennung wollen wir nur anmerken, dass es zur Zeit kommerzielle „demonstratorfähige“ Systeme gibt (z.B. das ASR1600/M-System der Firma Lernout & Hauspie); in der Synthese sind für das Englische akzeptable Vollsynthesysteme erhältlich (Sproat, 1997). Seit einiger Zeit ist wegen der ständig wachsenden Arbeitsspeicherkapazitäten die wesentlich verständlichere und gefälligere Wortkonkatenation ein interessantes alternatives Verfahren.

Zur linguistischen Analyse: Flache syntaktisch-semantische Analyse, wie sie in allen Anwendungssystemen zur natürlich-sprachlichen Dialogführung eingesetzt wird, reicht nicht aus, um für die Dialogführung in einem mathematischen Tutorsystem verlässliches und hinreichend detailliertes Sprachverstehen zu ermöglichen. Dies gilt nicht nur für Ansätze, die auf Schlüsselwortsuche basieren, sondern auch für statistische Parser und die Technik des Chunk-Parsing (Abney, 1996; Skut & Brants, 1998). Tiefe linguistische Analyse, etwa im HPSG oder TAG-Format, mit diskurssemantischen Zielrepräsentationen (Pollard & Sag, 1994; Joshi & Schabes, 1997; Kamp & Reyle, 1993), wird auf absehbare Zeit generell noch nicht in der nötigen Abdeckung und Robustheit zur Verfügung stehen. Für mathematische Lehrbuchtexte ist sie vermutlich erreichbar, wie unsere Vorarbeiten im LIMA-Projekt gezeigt haben (s. 3.4), für Schüler-Dialogbeiträge in einem Tutorsystem, die gesprochen und mindestens moderat spontan-sprachlich sind, dagegen illusorisch. Vielversprechend sind Techniken, die flache und tiefe Verfahren in geeigneter Weise kombinieren (s. ebenfalls 3.4).

In der Generierung besteht eine vergleichbare Situation: Anspruchsvolle inhaltsorientierte Anwendungen wie das beabsichtigte Tutor-System kommen nicht nur mit vorgeformten Ausgabemustern aus. Auf der anderen Seite liefert eine vollständig wissensbasierte Textplanung mit konsequenter Concept-to-Speech-Generierung zur Zeit im Blick auf Benutzerakzeptanz noch keine wirklich optimalen Resultate. Deshalb wurden mittlerweile verschiedene Kompromisse eingegangen. So konnten anhand der als Konsensus-Architektur bezeichneten Aufteilung des Generierungsprozesses in die drei Phasen Textplanung, Satzplanung und linguistische Realisierung und insbesondere deren sequentieller Verarbeitung (Reiter, 1994) vielfach akzeptable Texte generiert werden. Trotz bekannter theoretischer Nachteile verwenden praktisch alle anwendungsorientierten Systeme die vereinfachende pipeline-artige Architektur. Zusätzlich besteht wegen des hohen Modellierungsaufwands für Gegenstandsbereich und sprachliche Kompetenzen die Bestrebung, die Modellierung der beteiligten Phänomene nur soweit zu betreiben, wie es der gewünschten Funktionalität eines Anwendungssystems entspricht (Reiter & Mellish, 1993).

3.4 Eigene Vorarbeiten

Die Arbeitsgruppe von Jörg Siekmann entwickelt seit über einem Jahrzehnt das mathematische Assistenzsystem Ω MEGA. Insbesondere im Hinblick auf die Beweisplanung sowie die Integration heterogener Spezialsysteme nimmt die Gruppe in der internationalen Forschung eine Spitzenstellung ein. Es bestehen zahlreiche Kooperationen mit den wichtigsten Universitäten und die Gruppe koordiniert das EU-Netzwerk CALCULEMUS zur Integration von Deduktion und symbolischer Berechnung. Die Teilgruppe um Helmut Horacek und Armin Fiedler beschäftigt sich seit mehreren Jahren mit der natürlich-sprachlichen Präsentation von Beweisen. Zum DFKI besteht eine enge Kooperation zur Einbindung von Deduktionssystemen in webbasierte Lernumgebungen für die Mathematik.

Nach seiner Promotion in der Ω MEGA-Gruppe hat Christoph Benzmüller in England erfolgreich ein Projekt zum ressourcengesteuerten agentenbasierten Theorembeweisen mit heterogenen externen Systemen abgeschlossen. Seit seiner Rückkehr leitet er die Ω MEGA-Gruppe.

Die Arbeitsgruppe von Manfred Pinkal hat ihren langjährigen Schwerpunkt in der linguistischen, insbesondere semantischen Verarbeitung. Die SFB-Projekte CHORUS und LISA haben unter verschiedenen Perspektiven die Realisierung des Konzepts einer „realistischen semantischen Verarbeitung“ verfolgt. Im EU-Projekt FraCaS wirkte Pinkal an der Etablierung eines allgemeinen Frameworks für den Bereich Computational Semantics mit. Im Dialogübersetzungsprojekt Verbmobil war die Gruppe von Pinkal unter anderem im Bereich der Speech-Language-Schnittstelle tätig und hatte außerdem die Verantwortung für das Arbeitspaket „Robuste semantische Verarbeitung“. Das Dialogthema wurde zunächst indirekt ebenfalls im Verbmobil-Projekt thematisiert, in dem die Arbeitsgruppe von Pinkal auch für den Bereich der Diskurs- und Dialogsemantik zuständig war. Seit mehreren Jahren beteiligt sich die Gruppe intensiv an der Grundlagenforschung zur Dialogmodellierung als Partner in den EU-Projekten TRINDI und SIRIDUS. Neben diesen eher grundlagenorientierten Projekten gibt es kleinere Anwendungsprojekte zur Dialogsteuerung von Fahrstühlen und natürlich-sprachlichen Bedienoberflächen im Automobil.

Mathematische Assistenzsysteme

In einer Vorarbeit wurde die Eignung des ressourcenadaptiven, agentenbasierten Theorembeweislers Ω -ANTS (Benzmüller & Sorge, 1999, 1998) für einfache Probleme aus der elementaren Mengentheorie nachgewiesen (Benzmüller, Jamnik, Kerber & Sorge, 2001); insge-

samt wurden 20.000 automatisch generierte Mengengleichungen erfolgreich durch Ω -ANTS verifiziert bzw. falsifiziert. Der Ω -ANTS-Ansatz integriert dazu auch leistungsstarke externe Unterstützungssysteme (Benzmüller, Jamnik, Kerber & Sorge, 1999; Benzmüller & Sorge, 2000). Dazu gehören der im DFG-Projekt HOTEL entwickelte Beweiser LEO, dessen Eignung für die Mengentheorie bereits in (Benzmüller & Kohlhase, 1998) diskutiert wurde, und der Modellgenerierer SATCHMO, der wie in (Benzmüller et al., 2001) skizziert Gegenbeispiele für Nicht-Theoreme generieren kann.

Ω MEGAS Beweisplaner wurde in der vergangenen Förderperiode zu einem Multistrategie-Planer (Melis & Meier, 1999, 2000) weiterentwickelt und die Mechanismen zum Meta-Reasoning wurden weiter verbessert. In der kommenden Förderperiode soll der reaktive, agentenbasierte Ansatz Ω -ANTS im Rahmen des Projekts OMEGA mit der traditionellen, deliberativen Beweisplanung im Ω MEGA-System gekoppelt werden. Die Vorteile und das Potenzial dieser Integration, insbesondere auch für unser Projekt, wurde durch die Arbeiten (Meier & Sorge, 2000; Meier, Pollet & Sorge, 2000) gezeigt, in denen etwa 8000 Theoreme der Gruppentheorie und eine reiche Auswahl an Theoremen eines Analysis-Lehrbuchs bewiesen werden konnte.

Tutorsysteme

In der Arbeitsgruppe Siekmann wurde in einer ersten Ausbaustufe das Mathematik-Lernsystem ACTIVEMATH entwickelt, das benutzeradaptiv Lerndokumente generiert und webbasiert mehrere Servicesysteme einbindet (Melis, 2000b; Melis et al., 2001). ACTIVEMATH hat eine verteilte, offene Architektur, die Erweiterungen des Systems erleichtert und die Speicherung und Verwaltung verschiedener Wissensarten in modularer Weise unterstützt. Beispielsweise werden das mathematische Wissen, die pädagogischen Regeln und das Wissen über den Benutzer getrennt gehalten. Die Kommunikation der Komponenten erfolgt in der Standardsprache XML-RPC. Die Einbindung externer Servicesysteme erweitert die Möglichkeiten eines Lernsystems vor allem in Richtung exploratives Lernen und Metakognition (Melis & Fiedler, 2000). In (Melis & Horacek, 2000) wird ein tutorieller Dialog erläutert, der sich an der Problemlösungsstrategie des Beweisplaners orientiert.

Dialogsysteme

Im EU-Projekt TRINDI (Traum et al., 1999) wird das Konzept der Dialogspiele auf komplexe Informationszustände übertragen. Jeder „Dialog-Zug“ eines Sprechers wird dabei vor

dem Hintergrund eines Informationszustandes interpretiert und (regelbasiert) auf einen Folgezustand abgebildet. Durch die Wahl entsprechender Elemente in den Informationszuständen lässt sich dieser generelle Rahmen der Dialogmodellierung flexibel an unterschiedliche Dialogszenarien anpassen; so können z.B. verschiedene Sprachschnittstellen zum Benutzer verwendet oder auch Planungskomponenten integriert werden. Mit dem TRINDI-KIT (Larsson, Bohlin, Bos & Traum, 1999) wurde eine Software-Umgebung implementiert, die die Entwicklung und Evaluierung von Dialogsystemen auf der Basis dynamischer Informationszustände unterstützt und einfache Standardmodule für Sprachein- und -ausgabe zur Verfügung stellt. Diese Arbeiten werden derzeit im Projekt SIRIDUS fortgeführt.

In (Horacek, 2000) wird das Konzept gemeinsamer Pläne (Grosz & Kraus, 1996, 1998) um die Behandlung semantisch reichhaltiger kommunikativer Aktionen, wie das Erklären eines mathematischen Beweises, erweitert. Dabei werden vorrangig Kompromisse zwischen Dialogbeitragslänge und inhaltlicher Vollständigkeit auf der Basis Grice'scher Maxime modelliert. Diese Techniken sind für Beweiszusammenfassungen in einer interaktiven Exploration eines Beweises für das Projekt einsetzbar.

Mit *P.rex* (Fiedler, 2001b) wurde ein Beweiserklärungssystem entwickelt, das dem Benutzer Beweise interaktiv erläutert. Das System ist in der Lage Dialoge in eingeschränkter Form zu führen, wie Unterbrechungen durch den Benutzer bei unzureichenden Erklärungen oder Klärungsdialoge bei Mehrdeutigkeiten (Fiedler, 2001a).

Natürlich-sprachliche Analyse und Generierung

Im Rahmen des Projekts Verbmobil wurde für die Übersetzung gesprochener Dialoge eine Komponente zur Diskurs- und Dialogsemantik entwickelt (Bos & Heine, 2000). Diese Komponente beinhaltet ein Dialoggedächtnis, das zur Anaphernresolution und Interpretation von einfachen Ellipsen genutzt wird.

Im SFB-Projekt LISA wurden Techniken zur Auflösung von Ellipsen und deakzentuierten Phrasen basierend auf Unifikation höherer Stufe und deren Interaktion mit Anaphernresolution untersucht (Gardent, Kohlhase & Konrad, 1998; Gardent, 1999). Im Bereich Generierung wurde gezeigt, dass Baumbeschreibungsgrammatiken sowohl für die Generierung als auch für das Parsen verwendet werden können (Duchier & Thater, 1999; Thater & Gardent, sub.).

Im Arbeitspaket „Robuste semantische Verarbeitung“ des Verbmobil-Projekts wurde erstmals ein Verfahren für die linguistische Verarbeitung spontan-sprachlicher Dialogbeiträge reali-

siert, das partielle Analyseresultate unterschiedlicher Parser, insbesondere eines HPSG-basierten „tiefen“ Analysemoduls, eines Chunk-Parsers und eines statistischen Parsers unter Verwendung heuristischer Regeln zu approximativen Interpretationen der Gesamtäußerung kombiniert (Worm, 2000; Pinkal, Rupp & Worm, 2000).

In der Arbeitsgruppe Siekmann beschäftigt man sich schon seit zwei Jahrzehnten mit der Beweispräsentation, in neuerer Zeit sind von uns Systeme entwickelt worden, die dem Benutzer Beweise in natürlicher Sprache darbieten. Das System *PROVERB* (Huang & Fiedler, 1996, 1997) ist in der Lage, Beweise auf einer vorgegebenen Abstraktionsebene in einem lehrbuchähnlichen Stil wiederzugeben. Darauf aufbauend wurde das Nachfolgesystem *P.rer* (Fiedler, 2001b) entwickelt, das darüber hinaus in der Lage ist, Beweise interaktiv und benutzerangepasst zu erklären (Fiedler, 1999).

In (Horacek, 1999) werden psychologisch motivierte Abstraktionen zur Beweispräsentation verfeinert. Darüber hinaus wird die Inferenzfähigkeit des Adressaten modelliert und für mathematische Beweise ausgearbeitet. In (Horacek, 1998) wird auch die Anwendbarkeit der Techniken zur Berücksichtigung der Inferenzfähigkeit des Adressaten für die Textplanung bei alltäglichen Dialogen gezeigt. Dabei ist insbesondere das Design der Schnittstelle zwischen Modellierung der Inferenzen und den Domänenkonzepten von Bedeutung.

Spezifische Vorarbeiten

In den Projekten AGINT, OMEGA, und LISA wurden bzw. werden in der laufenden Förderungsperiode folgende spezifische Vorarbeiten für das Projekt erbracht:

1. In der SFB-internen Kooperation zwischen den Projekten LISA, PERFORM, OMEGA, und AGINT wurde die Struktur und Bedeutung mathematischer Texte untersucht. Diese Arbeiten sind auf einer konkreteren Stufe im Projekt LIMA (Landesmittelprojekt) wiederaufgenommen worden zur tiefen syntaktisch-semantischen Analyse mathematischer Texte (Baur, 1999). Das LIMA-System besteht aus einer HPSG-Grammatik, die Zielrepräsentationen in Lambda-DRT erzeugt (Kohlhase, Kuschert & Pinkal, 1996; Bos, Mastenbroek, McGlashan, Millies & Pinkal, 1994). Das System ist in der Lage Definitionen, Sätze und Beweise eines einführenden Buches über Analysis einzulesen, und in eine semantische Repräsentation zu überführen. Im Rahmen des LIMA-Projektes wurde die textstrukturelle Analyse mathematischer Dokumente (Definitionen, Beweise) geleistet, Mechanismen für textsortenspezifische diskurssemantische Referenz (Anapher mit Hilfe von Meta-Variablen, Definitions- und Beweisabschnittsnummerierungen) bereitgestellt

und die Grammatik auf Mischkonstruktionen aus Formeln und natürlich-sprachlichen Elementen erweitert. Alle Resultate sind für den Einsatz im geplanten Projekt unmittelbar relevant.

2. Als Infrastruktur zur Integration verschiedener heterogener Module wurde der mathematische Softwarebus MATHWEB entwickelt. Im Kontext des Ω MEGA Systems unterstützt MATHWEB die Integration und Kommunikation mathematischer Servicesysteme über das Internet. In der Lernumgebung ACTIVEMATH (DFKI) wird MATHWEB bereits in einem erweiterten Kontext erfolgreich verwendet. In der Arbeitsgruppe Pinkal wurde MATHWEB bereits direkt in der Dialogmodellierung eingesetzt: DORIS ist ein System, das verteiltes Theorembeweisen via MATHWEB benutzt, um Präsuppositionen in Dialogen zu überprüfen. In diesem Projekt soll MATHWEB ebenfalls zur Integration der jeweiligen Einzelkomponenten des Systems eingesetzt werden. Die Komponenten können dann auch bei Bedarf über das Internet verteilt werden.
3. In Kooperation mit der Universität Birmingham wurden 20.000 einfache Gleichungen über Mengen durch den agentenbasierten taktischen Theorembeweiser Ω -ANTS verifiziert bzw. falsifiziert. Beispielsweise liefert Ω -ANTS einen mengentheoretisch adäquaten Beweis für das Theorem $\forall x, y, z. (x \cup y) \cap z = (x \cap z) \cup (y \cap z)$. Im Falle von Nicht-Theoremen wie $\forall a, b, c. (a \cap b) \cup c = (a \cap b) \cap c$ werden im letzten Schritt durch einen Modellgenerierer Gegenbeispiele synthetisiert und im System textuell angezeigt.
4. Zum tutoriellen Dialog über mengentheoretische Fragen wurde eine erste Vorstudie durchgeführt. Eine Schülerin (11. Klasse) und ein Informatikstudent (1. Studienabschnitt) hatten die Aufgabe, mit Hilfe des Tutorsystems den Beweis zu einem Satz aus der Mengenlehre zu entwickeln. Sprachein- und -ausgabe erfolgten über Tastatur bzw. Bildschirm (ohne Beschränkungen auf der Benutzer- oder Systemseite), der Ablauf des tutoriellen Dialogs war vom Versuchsleiter (Wizard) partiell und informell vorstrukturiert. Diese methodisch sehr anspruchslose Studie hat bereits eine Reihe wichtiger Resultate erbracht, unter anderem:
 - Schülerin und Student hatten Schwierigkeiten, sehr feinkörnige Schritte auf der Logikebene explizit zu formulieren. Während die Schülerin der Konfrontation mit detaillierten logischen Fragestellungen auswich, weil ihr explizites logisches Schließen scheinbar fremd war, behandelte der Student diese Ebene bewusst implizit, weil er sie scheinbar für zu trivial hielt.
 - Es bestand vor allem bei der Schülerin eine starke Tendenz zu Meta-Dialogen (ver-

mutlich, da ihr die methodischen Voraussetzungen nicht klar waren). Ein wesentliches Problem wird die geeignete Benutzerführung darstellen, da der Dialog bei freier Benutzerinitiative leicht auf eine Meta-Ebene gerät, die mit heutiger Technologie nicht zu beherrschen ist.

- Die Notwendigkeit zur weiteren Abstraktion der Beweise auf höhere Ebenen wurde deutlich.

3.5 Arbeitsprogramm (Ziele, Methoden, Zeitplan)

3.5.1 Ziele

Konkretes Ziel des Projekts ist die empirische Untersuchung, die Modellierung und die Implementierung natürlich-sprachlichen Dialogs in einer tutoriellen Anwendung für ein mathematisches Teilgebiet. Übergeordnete Motivation ist die Untersuchung und Realisierung von Dialogtechniken in einer inhaltlich und pragmatisch anspruchsvollen Anwendung, die einerseits echtes Sprachverstehen und andererseits echte Wissensverarbeitung voraussetzt. Dies ist keineswegs ein Vorstoß in Neuland. Wissensbasierte Dialogverarbeitung ist ein altes, in der Vergangenheit intensiv und im Endeffekt ohne durchschlagenden Erfolg bearbeitetes Thema in KI- und Sprachverarbeitungsforschung. Wie im Folgenden ausgeführt, besteht auf Grund der aktuellen Voraussetzungen (generell und insbesondere am Ort) und auf Grund einer veränderten methodischen Herangehensweise aus unserer Sicht die begründete Aussicht auf substantielle und anwendungsrelevante Resultate.

Ausgangspunkt des Projekts ist das Konzept eines tutoriellen Systems, das dem Benutzer mathematische Begriffe aus dem Bereich der Mengentheorie erläutert und interaktiv beim Entwickeln und Verstehen von Beweisen Hilfestellung leistet. Für die Interaktion mit dem Benutzer soll natürliche Sprache verwendet werden, und zwar nicht nur in Form vorgefertigter Versatzstücke, sondern in einem Dialog, der syntaktische Flexibilität, semantische Orientierung, und pragmatische Situiertheit als Grundvoraussetzungen effizienter Sprachverwendung ausnutzt. Das Fachwissen des Tutorsystems soll ebenfalls nicht auf der Grundlage vorgefertigter Wissens-elemente eingebracht werden, sondern in Form einer semantisch durchmodellierten Domäne mit starken und effizienten Inferenzwerkzeugen.

Im Einzelnen wollen wir Erkenntnisse gewinnen über

- die Komplexität und die notwendigen Freiheitsgrade der Dialogführung im tutoriellen Bereich (z.B. benutzerinitiierte Klärungsdialoge zu verwendeten mathematischen Kon-

zepten und Gründen für Entscheidungen bei der Beweisführung),

- die angemessene Abstraktionsebene und Granularität in den Beweisen, ggf. in Abhängigkeit von unterschiedlichen Ausgangsvoraussetzungen bei den Nutzern,
- die Interaktion der Dialogstruktur mit der Struktur des Tutorsystems und den Strukturen, die durch das Beweisplanungssystem eingebracht werden,
- die sprachzentrierte, aber multimodale Gestaltung der Dialogbeiträge in Analyse und Generierung (mit Anteilen aus freier Sprache, Formeln, Grafik und Multiple-Choice-Elementen).

Der methodische Schwerpunkt in der beantragten Projektphase wird in der empirischen Erarbeitung und Modellierung grundlegender Ergebnisse aus diesen Bereichen liegen. Im Einzelnen sind vorgesehen:

- Mehrere aufeinander aufbauende *Simulationsexperimente* (als *Wizard-of-Oz*-Studien), die in zunehmendem Grad realistische Interaktionsbedingungen für den Benutzer bieten (Übergang von geschriebener zu gesprochener und multimodaler Interaktion) und zunehmend standardisierte und teilautomatisierte Komponenten (als Resultate der konzeptuellen Modellierung) verwenden.
- Die *konzeptuelle und softwaremäßige Modellierung* der zentralen Dialogkomponente, die Konzeption von Schnittstellen zu Tutor- und Beweisassistenzsystem sowie zu natürlich-sprachlicher Analyse und Generierung
- Die Implementierung eines *Demonstrators*, der exemplarische tutorielle Sitzungen mit Elementen der natürlich-sprachlichen Dialogführung durchführen kann.

Die Bereiche Beweisplanung und Mathematik-Assistenz, Tutorielle Systeme sowie natürlich-sprachliche Analyse und Generierung (inklusive Spracherkennung und -synthese) liefern wichtige Voraussetzungen, sowohl für Design und Realisierung der fortgeschrittenen *Wizard-of-Oz*-Studien als auch für die Implementierung des Demonstratorsystems. Da die zentralen Projektziele bereits sehr ambitiös sind, werden wir hier äußerst pragmatisch vorgehen, d.h., vorhandene Werkzeuge verwenden und ggf. an die Bedürfnisse des Projekts anpassen; notwendige Entwicklungen in Kooperation mit Partnern im SFB und anderen Partnerprojekten am Ort betreiben; mächtige Werkzeuge konsequent nur bei Bedarf einsetzen und soweit möglich, mit „flachen“ und ggf. trivialen Verfahren (z.B. Chunk-Parsing in der Spracheingabe, pre-recorded speech in der Sprachausgabe, Automaten für Subdialoge) kombinieren.

3.5.2 Methoden und Arbeitspakete

Methodisch werden empirische Untersuchungen, die Modellierung einer komplexen Systemarchitektur und die Implementierung und Validierung einzelner Teilkomponenten beziehungsweise eines Demonstrationssystems eingesetzt. Das methodische Vorgehen werden wir im Folgenden bei der Beschreibung der einzelnen Arbeitspakete näher präzisieren.

Das grundlegende Arbeitspaket, dessen Resultate in Modellierung und Implementierung eingehen werden, ist das AP1: „Experimentelle Testumgebungen und empirische Untersuchungen“. Darauf aufbauen sollen die Modellierungspakete AP2: „Dialogmodellierung“ und AP3: „Architektur und Schnittstellen“. Die Arbeitspakete AP4: „Bereitstellung und Aufbereitung des mathematischen Wissens“ und AP5: „Sprachschnittstellen“ liefern einerseits dem AP1 zu (Strukturierung und Teilautomatisierung der Simulationsexperimente), und bilden andererseits mit den Resultaten von AP2 und AP3 die Basis für die Realisierung des Demonstrators (AP6).

AP1: Experimentelle Testumgebung und empirische Untersuchungen

Im Laufe des Projektes soll eine Reihe von sogenannten *Wizard-of-Oz*-Studien (Fraser & Gilbert, 1991) durchgeführt werden, die Aufschluss über die grundlegenden Phänomene von tutoriellen Dialogen für einen begrenzten Teilbereich der Mathematik geben sollen.

Im Mittelpunkt der Experimente steht die Frage nach der geeigneten Dialogstruktur. Bei der Analyse soll auf für die tutorielle Anwendung relevante Dialogakte, deren Verknüpfung im Dialogmodell und mögliche Ansätze zur Modularisierung (stark standardisierte Subdialoge) geachtet werden, außerdem auf Arten der Korrelationen zwischen Dialogstruktur und Beweisplan sowie unterschiedlichen Dialogebenen (z.B. Metakommunikation zur Änderung der Beweisstrategie oder explizite Übergabe der Dialoginitiative an das System). Andererseits sollen Missverständnisse in der Kommunikation zwischen Benutzer und System untersucht werden, da diese speziell in Tutordialogen sehr schnell zum Verfehlen des Dialogziels führen können. Ein wichtiger Punkt ist auch die Analyse des Erinnerungsvermögens der Benutzer an zuvor im Dialog aufgetretene Fakten und Aspekte; diese Ressourcenbeschränkung des Benutzers soll im System später angemessen reflektiert werden und stehen damit in engem Zusammenhang mit den psychologischen Projekten zur Ressourcenbeschränkung des menschlichen Arbeitsgedächtnisses.

In den Vorstudien war erkennbar (siehe Punkt 4 in den spezifischen Vorarbeiten), dass eine

weitere Abstraktion von der zu stark logikorientierten Ebene im Dialog nötig ist, detaillierte Argumentationen auf der Ebene rein logischer Inferenzen wurden generell vermieden.

Als konkrete mathematische Domäne werden in den empirischen Untersuchungen zunächst einfache Beispiele aus der *naiven Mengenlehre* betrachtet. Diese Domäne ist elementar in der Mathematik, setzt bei den Probanden nur wenig Vorkenntnisse voraus und ist (siehe spezifische Vorarbeiten) durch die agentenbasierte Beweisplanung in Ω MEGA beherrschbar. Die *Wizard-of-Oz*-Studien sollen in drei Stufen vorgenommen werden. In einer ersten Stufe werden wir die Vorstudie insoweit replizieren, als wir von linguistisch unrestringierter Spracheingabe über Terminal und nur teilweise vorstrukturierten Systemreaktionen ausgehen. Wir werden allerdings mit einer größeren Zahl von Versuchspersonen arbeiten und eine (etwas) größere Zahl repräsentativer Anwendungsfälle durchvariieren. Die tutorielle Sitzung selbst werden wir etwas stärker durchstrukturieren, auf Systeminitiative abstellen, Grenzen für sinnvoll zu behandelnde Dialogbeiträge ziehen und bei Überschreitung der Grenzen durch die Testperson Reparaturmechanismen vorsehen. Die Resultate der ersten Studie gehen in die erste explizite Realisierung eines Dialogmodells ein (AP 2), das wiederum eine Grundlage für die zweite Studie bildet.

Die zweite und dritte Studie werden gesprochene Sprache, z.T. in Kombination mit über die Tastatur eingegebenen Formeln, verwenden. Zur Protokollierung (Verschriftung) werden wir kommerzielle Spracherkennungswerkzeuge mitverwenden. In beiden Studien soll neben der zentralen Dialogmaschine die Systeme Ω MEGA, *P.rex*, MBASE und MATHWEB in die experimentelle Umgebung eingebunden werden. Dabei sollen in einer Variante der losen Anbindung die Beschränkungen dieser Einzelsysteme durch den Versuchsleiter kompensiert werden. In einer anderen Variante sollen die gegenwärtigen und auch mittelfristig abzusehenden Beschränkungen voll übernommen werden, um konzeptuelle Beschränkungen für das intendierte Gesamtsystem zu ermitteln. In der dritten Studie werden wir zusätzlich grafische Elemente (z.B. Euler-Diagramme) in der Präsentation verwenden und deiktische Handlungen bei Benutzeräußerungen zulassen. Die Anwendung grafischer Mittel ist in Abb. 1 illustriert, einer Visualisierung eines Gegenbeweises des im Abschnitt 3.4 angeführten Nicht-Theorems.

Grundsätzlich sollen für die Experimente als Versuchspersonen Probanden mit unterschiedlichen mathematischen Vorkenntnissen gewählt werden, um Indikatoren für ein erstes Benutzermodell aus den Dialogen abzuleiten. Ob wir das in der Vorstudie repräsentierte Spektrum an Vorwissen einhalten können oder etwas einschränken müssen, um die Parallelentwicklung zu unterschiedlicher tutorieller und Dialogmodelle zu vermeiden, soll sich nach dem ersten Experiment herausstellen.

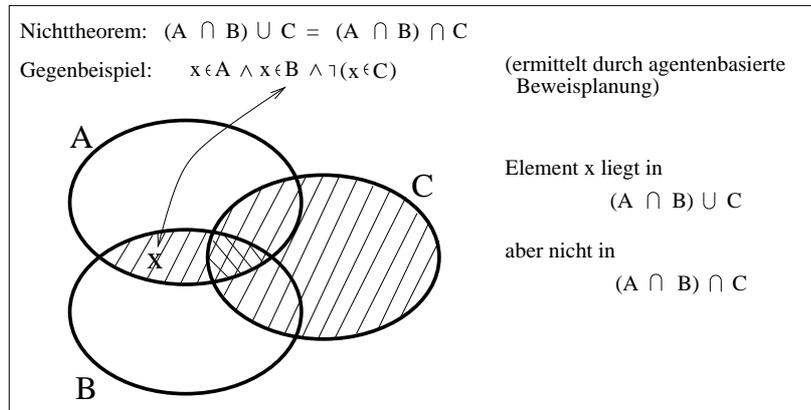


Abbildung 1. Gegenseitige Überführbarkeit logischer und graphischer Repräsentationen in der Mengenlehre.

AP2: Dialogmodellierung

Bei der Realisierung der Dialogkomponente wollen wir uns an die Methodik anlehnen, die in den EU-Projekten TRINDI und SIRIDUS unter Beteiligung der Arbeitsgruppe Pinkal entwickelt worden ist (Traum et al., 1999; Ericsson, Lewin & Rupp, 2000). Der linguistisch orientierte Anteil der Dialogmodellierung soll sich am Framework der Dialogspieltheorie orientieren. Über die geeignete Definition von Informationszuständen und Dialog-Zügen, die über Update-Regeln Informationszustände auf Folgezustände abbilden, lassen sich für bestimmte Dialogformen und Modi geeignete Formen der Dialogführung durchspielen und in ihrer Komplexität relativ einfach auf den jeweiligen Zweck adaptieren. Ein weiterer methodischer Grundzug des Dialogspiel-Konzeptes ist die Trennung von linguistisch motivierten und begründeten Dialogstrukturen, beispielsweise das „Grounding“, explizite oder implizite dialogische Mechanismen, die die Konsistenz des Dialogverständnisses bei den Partnern sicher stellen (Traum & Hinkelman, 1992; Traum, 1994), und die Schachtelung von Dialogen durch (eventuell rekursive) Klärungsprozesse, die durch das Konzept der „Question-under-Discussion“ (Ginzburg, 1995) modelliert wird, auf der einen Seite, und durch externe Planungsprozesse bedingte Dialogsteuerung (aus der tutoriellen Komponente und der Beweisplanung angestoßen) andererseits.

Wir werden überprüfen, wie weit wir Grundgerüst und Elemente der TRINDI-KIT Software verwenden können, die bereits eine abstrakte Dialog Move Engine (DME) implementiert, Formate und Datenstrukturen zur Definition von Informationszuständen und Update-Regeln zur Verfügung stellt und einfache Schnittstellen zur Integration externer Module vorsieht (Larsson et al., 1999).

Die inhaltliche Ausgestaltung dieser Schnittstellen ist ein weiteres wesentliches Ziel dieses Arbeitspaketes. Im Wesentlichen handelt es sich um die Schnittstelle zum Tutorsystem und zum mathematischen Beweissystem. Wichtige Teilaufgabe wird die Spezifikation einer internen Beschreibungssprache sein, die die Kommunikation zwischen der Dialogkomponente, den Tutorstrategien und dem mathematischen Beweissystem erlaubt.

AP3: Architektur und Schnittstellen

Ausgehend von den Resultaten der empirischen Untersuchungen in AP1 soll die Architektur des Gesamtkomplexes „tutorieller Dialog“ modelliert und die Schnittstellen der Einzelkomponenten definiert werden. Die Architektur soll erstens in die Demonstrator-Implementierung eingehen. Zweitens soll eine Grundlage für die fortgeschrittenen Experimente aus AP1 gelegt werden (für die allerdings nicht alle Komponenten gleichzeitig realisiert sein müssen). Einige Teile sind bereits grundsätzlich verfügbar: Beispielsweise soll das Ω MEGA System die Funktion der dynamischen mathematischen Wissensquelle übernehmen. Der Dynamikaspekt besteht darin, dass zur Laufzeit mathematische Inhalte, wie z.B. vom Benutzer geäußerte mathematische Sachverhalte, von Ω MEGA unter Zuhilfenahme der integrierten externen Beweiskomponenten auf ihren Wahrheitsgehalt analysiert werden können. Statisches mathematisches Wissen, wie Definitionen, Sätze, etc. werden über die mathematische Datenbank MBASE bereitgestellt. Weitere Voraussetzungen sind vorgegeben durch den Beweisleser LIMA (Analyse) und das Beweispräsentationswerkzeug *P.rex* (Generierung). Die Integration der Einzelkomponenten soll auf der Basis von MATHWEB erfolgen. Kommunikationsprotokolle können auf dem OMDOC-Format oder auf XML-RPC aufbauen.

Während einige Einzelkomponenten also bereits zur Verfügung stehen, müssen andere erst modelliert und realisiert werden. Dies gilt insbesondere für die Dialogmaschine und die tutorielle Einheit. Die Dialogmaschine ist das wesentliche Resultat aus AP2. Die tutorielle Einheit bildet als zentrale Steuereinheit den Kern des intendierten Dialogsystems. Unser Forschungsinteresse gilt aber in erster Linie der multimodalen natürlich-sprachlichen Verarbeitung. Die tutorielle Steuerung soll nur soweit verfolgt werden, wie es für den Demonstrator zwingend erforderlich ist. Insbesondere soll auf existierende Systeme wie ACTIVE MATH des DFKI und den Geometrietutor der CMU zurückgegriffen werden.

Ein erster hypothetischer Modellentwurf für das Gesamtsystem ist in Abbildung 2 abgebildet. Dieser Entwurf soll nach Ausarbeitung der Kernfunktionalitäten des Systems entsprechend ergänzt und verfeinert werden.

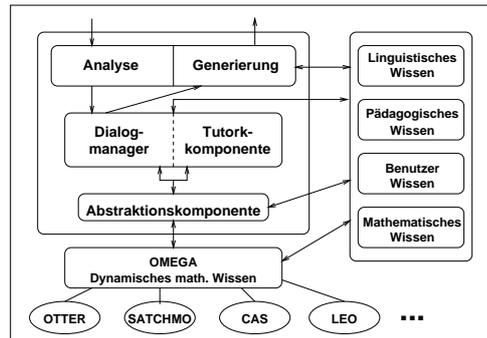


Abbildung 2. Ein erster Vorschlag für die Gesamtarchitektur.

AP4: Bereitstellung und Aufbereitung des mathematischen Wissens

Mathematisches Wissen wird im intendierten Dialogsystem auf verschiedene Arten zur Verfügung gestellt: Erstens durch den Zugriff auf die Definitionen, Sätze, etc. in der mathematischen Datenbank MBASE, zweitens durch das explizit repräsentierte Meta-Wissen in Ω MEGAs Beweisplanung, und drittens durch Ω MEGAs Potenzial mathematische Probleme dynamisch zur Laufzeit zu analysieren, möglicherweise in Kooperation mit den integrierten externen Unterstützungssystemen. Insbesondere die Möglichkeit der dynamischen Verifikation oder Falsifikation von Äußerungen des Benutzers durch die angeschlossenen Beweissysteme ist ein spezielles Merkmal des intendierten Systems.

Die Vorversuche haben allerdings aufgezeigt, dass die in Ω MEGA verwendeten Repräsentationsformate für Beweise bzw. Beweispläne aus der Perspektive eines tutoriellen Dialogs noch zu stark kalkülorientiert sind und sich in der jetzigen Form nur bedingt als mathematisches Hintergrundwissen eignen. Beispielweise ist in vielen Fällen eine Abstraktion von der Reihenfolge vertauschbarer Kalkülschritte erforderlich. Einfache aussagenlogische Inferenzen werden zudem vom Benutzer bewusst oder unbewusst implizit behandelt, wie die Vorstudie zeigte. Es müssen daher ein geeignetes Abstraktionsniveau für die Resultate, die von Ω MEGA geliefert werden, definiert und eine entsprechende Abstraktionskomponente realisiert werden. Zur Verbesserung der Beweisplanung in Ω MEGA wurde ebenfalls die Notwendigkeit einer höheren Abstraktion als der derzeit noch zu logikzentrierten Plandatenstruktur erkannt. Es soll deshalb untersucht werden, inwieweit sich die Anforderungen an diesen Abstraktionsschritt aus der Perspektive des Dialogsystems und aus der Perspektive der Beweisplanung (MI 4) miteinander verbinden lassen.

In den Vorstudien wurden zwar zahlreiche Beispiele aus der Mengentheorie untersucht, doch

handelt es sich dabei um relativ einfache Mengengleichungen. Die Bandbreite der beherrschbaren Theoreme und Nichttheoreme soll in Kooperation mit dem Ω MEGA Projekt weiter ausgedehnt werden. Sobald in Ω MEGA die traditionelle Beweisplanung und das agentenbasierte Theorembeweisen integriert sind, soll relevantes Steuerungswissen für die Mengentheorie akquiriert und kodiert werden, und dann sollen entsprechende Fallstudien durchgeführt werden. Unter anderem sollen sämtliche Theoreme aus der Sammlung (Trybulec & Swieczkowska, 1989) automatisiert werden. Ein ebenso wichtiger Aspekt gilt aber dem Erkennen und der Repräsentation — z.B. in Form von Euler-Diagrammen — von Gegenbeispielen. Die Aufgabenverteilung ist in dieser Kooperation: OMEGA (MI 4) ist für die Automatisierung der Beispiele verantwortlich und DIALOG für die abstrahierende Aufbereitung.

AP5: Sprachschnittstellen

Die Sprachschnittstellen haben im Projekt subsidiäre Funktion. In den empirischen Experimenten wird automatische Sprachverarbeitung nicht verwendet, oder doch nur in den Grenzfällen der patternbasierten Ein- und Ausgabe. Für den Demonstrator ist sehr breite Abdeckung weder erforderlich noch erreichbar. Allerdings wollen wir grundsätzliche domänen- und anwendungsspezifische Aufgabenstellungen in Analyse und Generierung untersuchen und in einer gewissen Breite behandeln können.

In der Analyse greifen wir allgemein auf die diskurssemantischen Arbeiten des LISA-Projektes und die robusten semantischen Analysetechniken aus Verbmobil zurück. Speziell und konkret setzen wir in verschiedener Hinsicht auf die oben beschriebenen Vorarbeiten im LIMA-Projekt auf. Dies betrifft unter anderem die Analyse von gemischten Benutzereingaben (natürliche Sprache und Formeln), die diskurssemantische Modellierung mathematischer Texte, mit textsortenspezifischen Besonderheiten wie der Referenz mit Hilfe von Meta-Variablen und die in LIMA erst partiell realisierte Übersetzung von natürlich-sprachlich orientierten semantischen Repräsentationen in die Plandatenstruktur von Ω MEGA. Die Arbeit wird im Wesentlichen in der Erweiterung der Abdeckung und im Einbezug von dialogspezifischen Phänomenen bestehen (LIMA wurde zur Analyse von Lehrbuchtexten entwickelt), z.B. einfache Formen von elliptischen Äußerungen und das Abfangen spontansprachlicher Phänomene durch die Kombination der tiefen Analyse mit flachen Verfahren.

Im Bereich der Generierung geht es vor allem darum, einige in der Arbeitsgruppe verfügbare Teilsysteme zusammenzuführen, aneinander anzupassen, und ihre Funktionalität gemäß den Projektzielen zu erweitern.

Zunächst wird die am DFKI in der Gruppe von Prof. Uszkoreit entwickelte Satzrealisierungskomponente TG/2 (Busemann, 1996), die in der aktuellen Phase des DFG-Projekts FABEON in das Beweispräsentationssystem der Arbeitsgruppe integriert wird, an die Äußerungsarten und Formulierungen des DIALOG Projekts angepasst. Dabei ist insbesondere der Kontextmechanismus für die motivierte Auswahl von Alternativen relativ zu dem Benutzermodell einzusetzen. Eine echte Erweiterung dieser Komponente besteht in der Mischung von Text- und Formelteilen in der Ausgabe, was ein Eingreifen in Interna des Grammatikformalismus bedingt. Diese Kombination von Darstellungsmöglichkeiten ist auch für andere Anwendungen der Textgenerierung von großem Interesse.

Im Bereich der Planung von Äußerungen ist das Beweispräsentationssystem *P.rex* (Fiedler, 1999, 2001b) um einige Funktionalitäten zu erweitern. Dazu gehören eine Ausweitung des Repertoires an Sprechakten und die Einbindung multimodaler Elemente, wie etwa Diagramme, in den Planungsprozess. Letzteres erfordert eine Verfeinerung der Referenzgenerierungsverfahren im System, was sich auch bis in die Satzrealisierungskomponente fortsetzt. Fehlende Flexibilität bei der satzübergreifenden Behandlung von Referenzen hat sich bei früheren Anwendungen von TG/2 (Busemann & Horacek, 1998; Horacek & Busemann, 1998) als Mangel erwiesen.

AP6: Realisierung eines Demonstrators

Zur Erprobung und Validierung der entwickelten Architektur und ihrer Komponenten soll ein Demonstrator entstehen, der in einer möglichen zweiten Förderphase zu einem prototypischen tutoriellen System erweitert werden kann. Wir werden bei der Realisierung des Demonstrators äußerst pragmatisch vorgehen müssen, da die Komplexität der Aufgabenstellung enorm ist. Das bedeutet, dass wir von dem empirisch fundiert und detailliert ausgearbeiteten Zentralbereich mit Dialogmaschine und tutorieller Komponente konsequent vorhandene Systeme und Entwicklungen einsetzen und evtl. anpassen werden. Dazu gehören MATHWEB zur Integration von Systemen sowie deren Distribuierung über das Internet, LIMA als Ausgangspunkt für die Sprachanalyse, *P.rex* als Ausgangspunkt für die Generierung und Präsentation von Beweisen, Ω MEGA als Beweissystem und die mathematische Wissensbank MBASE zur Verwaltung der mathematischen Objektontologien. Das OMDOC-Format und XML-RPC werden als Kommunikationsprotokoll genutzt und das DFKI-Projekt ACTIVEMATH liefert ein mathematisches Tutorsystem. Wir werden, wo dies ohne Akzeptanzverlust möglich ist, auf allen Ebenen auch mit vorgefertigten Mustern arbeiten. Einerseits betrifft das die Ebene der Sprachein- und -ausgabe, wo wir tiefe und flache Verfahren konsequent verknüpfen wollen,

andererseits die Gesamtstruktur des tutoriellen Systems: Wir streben für den Demonstrator eine Form der tutoriellen Interaktion an, die natürlich-sprachlichen Dialog als wesentlichen Bestandteil enthält, möglicherweise aber im Wechsel mit präsentationsbasierten Passagen.

Der Demonstrator soll zunächst in einer „Null-Version“ erstellt werden, in der Architektur und Schnittstellen im Prinzip realisiert sind; seine Endversion soll die zentralen Resultate aus den konzeptionellen Arbeitspaketen berücksichtigen. Ziel ist zunächst eine Version mit geschriebener Ein- und Ausgabe. Beabsichtigt ist eine Variante mit gesprochenen Dialogbeiträgen, ohne dass hier viel zusätzliche Arbeit investiert werden soll. Dies ist für die Synthese unproblematisch (Wortkonkatenation), für die Erkennung erwarten wir, dass die kommerzielle Entwicklung in wenigen Jahren auf dem Stand ist, dass eine einfache First-Best-Schnittstelle, die eine Wortkette liefert, ausreichende Erkennungssicherheit bietet.

3.5.3 Zeitplan

AP1 auf der einen Seite und AP2 und AP3 auf der anderen Seite werden zeitlich leicht versetzt (AP1 zuerst) und zyklisch bearbeitet. Mit der Bereitstellung des mathematischen Wissens kann sofort begonnen werden, hinsichtlich der Konzeption der Abstraktionskomponente müssen die empirischen Untersuchungen (AP1) aber abgewartet werden. Leicht zeitlich versetzt zu den empirischen Untersuchungen und der Modellierung des Gesamtsystems soll AP5 (Sprachschnittstellen) bearbeitet werden. Die partiell realisierten Teilkomponenten sollen vom 2. Projektjahr an zu einem Demonstrator verbunden werden (AP6).

3.6 Stellung innerhalb des Programms des Sonderforschungsbereichs

Das Projekt ist dem Bereich „Modellierung und Inferenz“ zugeordnet und dies beschreibt den Spannungsbogen der Nachbardisziplinen und Projekte.

Einerseits werden wir sehr eng mit dem mathematischen Assistenzsystem Ω MEGA (MI 4) zusammenarbeiten und die Zuarbeit der Projekte λ -PLAN (MI 5) und ACTIVEMATH benötigen.

Für Beiträge zur Sprachverarbeitung sind wir auf die Projekte CHORUS (MI 2) und PERFORM (MI 1) angewiesen und werden hier eng zusammenarbeiten und Systemkomponenten austauschen und gemeinsam wirken.

Mit dem Projekt NEP (MI 6) werden wir so wie in der Vergangenheit eng kooperieren und von dort die Expertise über Realisierungs- und Implementierungskonzepte nutzen.

Mit dem Projekt READY (EM 4) verbindet uns die gemeinsame Zielvorstellung eines multimodalen Dialogs, jedoch mit unterschiedlicher Ausprägung: während in READY ein Dialog unter realen Ressourcenbeschränkungen untersucht wird, soll in diesem Projekt gerade umgekehrt von einer tiefen semantischen Analyse, tutoriell gesteuerter Dialogführung und wissensintensiver Beispieldomäne ausgegangen werden. Eine solche Zielvorstellung unterliegt völlig anderen Ressourcenbeschränkungen.

Literatur

- Richardson, J., Smaill, A. & Green, I. Proof planning in higher-order logic with λ clam. In C. Kirchner & H. Kirchner (Hrsg.), *Proceedings of the 15th International Conference on Automated Deduction (CADE-98)* (Bd. 1421, S. 129–133). Springer.
- Abney, S. (1996). Partial Parsing via Finite-State Cascades. In J. Carroll (Hrsg.), *Workshop on Robust Parsing (Eighth European Summer School in Logic, Language, and Information)* (S. 8–15).
- Ahrenberg, L., Dahlbäck, N., Jönsson, A. & Thur'ee, A. (1996). *Costumizing interaction for natural language interfaces* (Tech. Rep. Nr. 1(1)). Linköping Electronic articles in Computer and Information Science. (<http://ww.ep.liu.se/ea/cis/1996/001/>)
- Aleven, V. & Koedinger, K. (2000). The need for tutorial dialog to support self-explanation. In *Papers from the 2000 AAAI Fall Symposium on Building Dialogue Systems for Tutorial Applications* (S. 14–19). AAAI Press.
- Allen, J. F., Miller, B. W., Ringger, E. K. & Sikorski, T. (1996). A Robust System for Natural Spoken Dialogue. In A. Joshi & M. Palmer (Hrsg.), *Proceedings of the 1996 Annual Meeting of the Association for Computational Linguistics (ACL'96)* (S. 62–70). San Francisco: Morgan Kaufmann Publishers.
- Anderson, M., Cheng, P. & Haarslev, V. (Hrsg.). (2000). *A proposal for automatic diagrammatic reasoning in continuous domains* (Nr. 1889). Springer Verlag.
- Andrews, P. B., Bishop, M., Issar, S., Nesmith, D., Pfenning, F. & Xi, H. (1996). TPS: A theorem proving system for classical type theory. *Journal of Automated Reasoning*, 16(3), 321–353.
- Aust, H., Oerder, M., Seide, F. & Steinbiss, V. (1995). The Philips automatic train timetable information system. *Speech Communication*, 17, 249–262.
- Baur, J. (1999). *Syntax und Semantik mathematischer Texte — ein Prototyp*. Unveröffentlichte Master's Thesis, Saarland University.
- Benz Müller, C., Cheikhrouhou, L., Fehrer, D., Fiedler, A., Huang, X., Kerber, M., Kohlhase, M., Konrad, K., Melis, E., Meier, A., Schaarschmidt, W., Siekmann, J. & Sorge, V. (1997). Omega: Towards a mathematical assistant. In *Proceedings of 14th International Conference on Automated Deduction (CADE)* (S. 252–255).
- Benz Müller, C., Jamnik, M., Kerber, M. & Sorge, V. (1999). Agent based mathematical reasoning. In *Proceedings of the Calculemus Workshop: Systems for Integrated Computation and Deduction*. Trento, July.
- Benz Müller, C., Jamnik, M., Kerber, M. & Sorge, V. (2001). *An agent based approach to reasoning*. (Submitted to IJCAI-01)

- Benzmüller, C. & Kohlhase, M. (1998). Leo – a higher-order theorem prover. In *Proceedings of the 15th International Conference on Automated Deduction (CADE-15)*. LINDAU, Germany.
- Benzmüller, C. & Sorge, V. (1998). A blackboard architecture for guiding interactive proofs. In *Proceedings of 8th International Conference on Artificial Intelligence: Methodology, Systems, Applications (AIMSA'98)*. Sozopol, Bulgaria.
- Benzmüller, C. & Sorge, V. (1999). Critical agents supporting interactive theorem proving. In P. Borahona & J. J. Alferes (Hrsg.), *Proceedings of the 9th Portuguese Conference on Artificial Intelligence (EPIA'99)* (S. 208–221). Evora, Portugal: Springer.
- Benzmüller, C. & Sorge, V. (2000). Oants – an open approach at combining interactive and automated theorem proving. In *Proceedings of the Calculemus Symposium 2000*. St. Andrews, Schottland: A.K.Peters.
- Bos, J. & Heine, J. (2000). Discourse and Dialogue Semantics for Translation. In W. Wahlster (Hrsg.), *Verbmobil: Foundations of Speech-to-Speech Translation* (S. 336–347). Berlin, Heidelberg, New York: Springer Verlag.
- Bos, J., Mastenbroek, E., McGlashan, S., Millies, S. & Pinkal, M. (1994). A Compositional DRS-based Formalism for NLP-Applications. In *Proceedings of the International Workshop on Computational Semantics* (S. 21–31). Tilburg.
- Busemann, S. (1996). Best-first surface realization. In *Eighth International Natural Language Generation Workshop*. Herstmonceaux, Univ. of Brighton, UK.
- Busemann, S. & Horacek, H. (1998). A flexible shallow approach to text generation. In *International Workshop on Natural Language Generation*. Niagara-on-the-Lake.
- Carletta, J., Isard, A., Isard, S., Kowtko, J. C., Doherty-Sneddon, G. & Anderson, A. H. (1997). The reliability of a dialogue structure coding scheme. *Computational Linguistics*, 23(1), 13–32.
- Carlson, L. (1985). *Dialogue games*. D. Reidel Publishing Company.
- Carpenter, B. & Chu-Carroll, J. (1998). Natural Language Call Routing: A Robust Self-Organizing Approach. In R. Mannell & J. Roberts-Ribes (Hrsg.), *Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP '98)*. Sydney, Australia.
- Char, B. W., Geddes, K. O., Gonnet, G. H., Leong, B. L., Monagan, M. B. & Watt, S. M. (1992). *First leaves: a tutorial introduction to maple v*. Berlin: Springer Verlag.
- Cohen, P. & Levesque, H. (1990). A theory of rational action and interaction. In *Proc. of the 28th Annual Meeting of the Association for Computational Linguistics (ACL-90)*.
- Cohen, P. & Levesque, H. (1991). Conformation and joint action. In *Proc. of the 12th International Joint Conference on Artificial Intelligence (IJCAI-91)* (S. 951-957).
- Constable, R. L., Allen, S., Bromly, H., Cleaveland, W., Cremer, J., Harper, R., Howe, D., Knoblock, T., Mendler, N., Panangaden, P., Sasaki, J. & Smith, S. (1986). *Implementing mathematics with the nuprl proof development system*. Englewood Cliffs, New Jersey: Prentice-Hall.
- Cornes, C., Courant, J., Fillitre, J.-C., Huet, G., Manoury, P., Muoz, C., Murthy, C., Parent, C., Paulin-Mohring, C., Saïbi, A. & Werner, B. (1995). *The Coq proof assistant reference manual, version 5.10* (rapport technique Nr. 177). INRIA.
- Dahlbäck, N. & Jönsson, A. (1997). Integrating domain specific focusing in dialog models. In *Proceedings of Eurospeech '97*. Rhodes, Greece.

- Duchier, D. & Thater, S. (1999). Parsing with tree descriptions: a constraint-based approach. In *Sixth International Workshop on Natural Language Understanding and Logic Programming (NLULP'99)* (S. 17–32). Las Cruces, New Mexico.
- Ericsson, S., Lewin, I. & Rupp, C. (2000). *Dialogue Moves in Negotiative Dialogues* (Tech. Rep.). SIRIDUS. (SIRIDUS Deliverable D1.2)
- Fiedler, A. (1999). Using a cognitive architecture to plan dialogs for the adaptive explanation of proofs. In T. Dean (Hrsg.), *Proceedings of the 16th International Joint Conference on Artificial Intelligence (IJCAI)* (S. 358–363). Stockholm, Schweden: Morgan Kaufmann.
- Fiedler, A. (2001a). Dialog-driven adaptation of explanations of proofs. (Submitted to IJCAI-01)
- Fiedler, A. (2001b). *P.r.ex*: An interactive proof explainer. (Submitted to IJCAR-01)
- Fraser, N. N. & Gilbert, G. N. (1991). Simulating Speech Systems. *Computer, Speech and Language*, 5(1), 81-99.
- GAP – Groups, Algorithms, and Programming, Version 4*. (1998). (<http://www-gap.dcs.st-and.ac.uk/~gap>)
- Gardent, C. (1999). Unifying Parallels. In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics*. Maryland, US.
- Gardent, C., Kohlhase, M. & Konrad, K. (1998). Higher-Order Colored Unification: a Linguistic Application. *Technique et Science Informatique*, 18(2), 181–209.
- Ginzburg, J. (1995). Resolving Questions, I. *Linguistics and Philosophy*, 18(5), 459–527.
- Gordon, M. (1985). *HOL: a machine oriented formulation of higher-order logic* (Tech. Rep. Nr. 68). University of Cambridge, Computer Laboratory.
- Gordon, M., Milner, R. & Wadsworth, C. (1979). *Edinburgh lcf: A mechanized logic of computation*. Springer Verlag.
- Grosz, B. & Kraus, S. (1996). Collaborative plans for complex group action. *Artificial Intelligence*, 86(2), 269-357.
- Grosz, B. & Kraus, S. (1998). The evolution of shared plans. In A. Rao & M. Wooldridge (Hrsg.), *Foundations and theories of rational agency*.
- Heeman, P., Johnston, M., Denney, J. & Kaiser, E. (1998). Beyond structured dialogues: Factoring out grounding. In *Proceedings of ICSLP'98*. Sydney, Australien.
- Heffernan, N. & Koedinger, K. (2000). Intelligent tutoring systems are missing the tutor: Building a more strategic dialog-based tutor. In *Papers from the 2000 AAAI Fall Symposium on Building Dialogue Systems for Tutorial Applications* (S. 14–19). AAAI Press.
- Hobbs, J., Stickel, M., Appelt, D. & Martin, P. (1990). *Interpretation as abduction* (Tech. Rep. Nr. SRI Technical Note 499). SRI International.
- Horacek, H. (1998). Generating inference-rich discourse through revisions of RST-trees. In *Proceedings of AAAI-98*. Morgan Kaufmann. (angenommen)
- Horacek, H. (1999). Presenting proofs in a human-oriented way. In *CADE-99* (S. 142-156).
- Horacek, H. (2000). Tailoring inference-rich descriptions through making compromises between conflicting principles. *International Journal on Human Computer Studies*, 53, 1117-1146.
- Horacek, H. & Busemann, S. (1998). Towards a methodology for developing application-oriented report generation. In *KI-98*. Bremen.

- Huang, X. (1994). Reconstructing proofs at the assertion level. In (S. 738–752).
- Huang, X. & Fiedler, A. (1996). Presenting machine-found proofs. In M. McRobbie & J. Slaney (Hrsg.), *Proceedings of the 13th Conference on Automated Deduction* (S. 221–225). New Brunswick, NJ, USA: Springer Verlag.
- Huang, X. & Fiedler, A. (1997). Proof verbalization as an application of NLG. In M. E. Pollack (Hrsg.), *Proceedings of the 15th International Joint Conference on Artificial Intelligence (IJCAI)* (S. 965–970). Nagoya, Japan: Morgan Kaufmann.
- Jamnik, M. (1999). *Automating diagrammatic proofs of arithmetic arguments*. Dissertation, Division of Informatics, University of Edinburgh.
- Joshi, A. K. & Schabes, Y. (1997). Tree Adjoining Grammars. In G. Rozenberg & A. Salomaa (Hrsg.), *Handbook of formal languages* (Bd. 3, S. 69–124). Berlin, Heidelberg, New York: Springer.
- Kamp, H. & Reyle, U. (1993). *From Discourse to Logic*. Dordrecht/Boston/London: Kluwer Academic Publishers.
- Kohlhase, M., Kuschert, S. & Pinkal, M. (1996). A type-theoretic semantics for λ -DRT. In P. Dekker & M. Stockhof (Hrsg.), *Proceedings of the 10th Amsterdam Colloquium* (S. 479–498).
- Larsson, S., Bohlin, P., Bos, J. & Traum, D. (1999). *TRINDIKIT 1.0 Manual*. (Deliverable D2.2 – Manual)
- McCune, W. (1997). Solution of the robbins problem. *Journal of Automated Reasoning*, 19(3), 263–276.
- McCune, W. & Wos, L. (1997). Otter CADE-13 competition incarnations. *Journal of Automated Reasoning*, 18(2), 211–220. (Special Issue on the CADE-13 Automated Theorem Proving System Competition)
- McCune, W. W. (1994). *Otter 3.0 reference manual and guide* (Technical Report Nr. ANL-94-6). Argonne, Illinois 60439, USA: Argonne National Laboratory.
- Meier, A., Pollet, M. & Sorge, V. (2000). *Exploring the domain of residue classes* (Seki Report Nr. SR-00-04). Saarbrücken, Germany: Fachbereich Informatik, Universität des Saarlandes.
- Meier, A. & Sorge, V. (2000). Exploring Properties of Residue Classes. In M. Kerber & M. Kohlhase (Hrsg.), *Proceedings of the Calculemus Symposium 2000*. St. Andrews, Großbritannien: AK Peters, New York, NY, USA.
- Melis, E. (2000a). Activemath: Web-based environment for the dynamic generation of mathematics documents employing external systems. In *10. Arbeitstreffen der GI-Fachgruppe 1.1.5 "Intelligente Lehr-/Lernsysteme"* (S. 110-117). debis.
- Melis, E. (2000b). The 'interactive textbook' project. In *Proceedings of CADE-17 workshop on deduction in Education* (S. 26-34).
- Melis, E., Andres, E., Franke, A., Gogvadse, G., Kohlhase, M., Libbrecht, P., Pollet, M. & Ullrich, C. (2001). ACTIVE MATH system description. In *Artificial Intelligence and Education*.
- Melis, E. & Fiedler, A. (2000). On the benefit of expert services in mathematics education systems. In B. D. Boulay (Hrsg.), *Proceedings of ITS-2000 workshop on Modeling Human Teaching Tactics and Strategies* (S. 84-85). Montreal.
- Melis, E. & Horacek, H. (2000). Dialog issues for a tutor system incorporating expert problem solvers. In R. Friedman & C. Rose (Hrsg.), *AAAI Fall Symposium on Building Dialog*

Systems for Tutorial Applications (S. 37-44). AAAI.

- Melis, E. & Meier, A. (1999). Proof planning with multiple strategies ii. In H. K. B. Gramlich & F. Pfenning (Hrsg.), *FLoC'99 workshop on Strategies in Automated Deduction* (S. 61-72).
- Melis, E. & Meier, A. (2000). Proof planning with multiple strategies. In J. Loyd, V. Dahl, U. Furbach, M. Kerber, K. Lau, C. Palamidessi, L. Pereira & Y. S. P. Stuckey (Hrsg.), *First International Conference on Computational Logic* (Bd. 1861, S. 644-659). Springer-Verlag.
- Moore, J. (2000). What makes human explanations effective? In *In Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society* (S. 131-136). Hillsdale, NJ. Earlbaum.
- Owre, S., Rushby, J. M. & Shankar, N. (1992). PVS: a prototype verification system. In (S. 748-752).
- Paulson, L. C. (1994). *Isabelle: A generic theorem prover*. Springer Verlag.
- Person, N., Graesser, A., Harter, D. & Mathews, E. (2000). Dialog move generation and conversation management in autotutor. In *Papers from the 2000 AAAI Fall Symposium on Building Dialogue Systems for Tutorial Applications* (S. 45-51). AAAI Press.
- Pinkal, M., Rupp, C. & Worm, K. (2000). Robust Semantic Processing of Spoken Language. In W. Wahlster (Hrsg.), *Verbmobil: Foundations of Speech-to-Speech Translation* (S. 321-335). Berlin, Heidelberg, New York: Springer Verlag.
- Pollard, C. & Sag, I. A. (1994). *Head-driven phrase structure grammar*. Chicago/London: University of Chicago Press.
- Reiter, E. (1994). Has a consensus nl generation architecture appeared, and is it psycholinguistically plausible? In *Proc. of the 7th International Workshop on Natural Language Generation (INLGW-94)* (S. 163-170).
- Reiter, E. & Mellish, C. (1993). Optimizing the costs and benefits of natural language generation. In *Proc. of the 13th International Joint Conference on Artificial Intelligence (IJCAI-93)* (S. 1164-1169). Chambery, France.
- Skut, W. & Brants, T. (1998). A Maximum-Entropy Partial Parser for Unrestricted Text. In *Proceedings of the Sixth Workshop on Very Large Corpora*. Montreal, Canada.
- Sproat, R. (Hrsg.). (1997). *Multilingual Text-to-Speech Synthesis: The Bell Labs Approach*. Boston: Kluwer Academic Publishers.
- Thater, S. & Gardent, C. (sub.). *Static Generation*. (Submitted)
- Traum, D. (1994). *A Computational Theory of Grounding in Natural Language Conversation*. Dissertation, University of Rochester, Department of Computer Science, Rochester, NY.
- Traum, D., Bos, J., Cooper, R., Larsson, S., Lewin, I., Matheson, C. & Poesio, M. (1999). *A Model of Dialogue Moves and Information State Revision* (Tech. Rep. Nr. Deliverable D2.1). Trindi.
- Traum, D. & Hinkelman, E. (1992). Conversation acts in task-oriented spoken dialogue. *Computational Intelligence*, 8(3). (Special Issue on Non-literal Language)
- Trybulec, Z. & Swieczkowska, H. (1989). Boolean properties of sets. *Journal of Formalized Mathematics*, 1.
- Weidenbach, C. (1997). SPASS: Version 0.49. *Journal of Automated Reasoning*, 18(2), 247-252. (Special Issue on the CADE-13 Automated Theorem Proving System Competition)

Worm, K. (2000). *Robust Semantic Processing for Spoken Language*. Dissertation, University of Saarbrücken.

3.7 Ergänzungsausstattung für das Teilprojekt MI 3

- 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	2	213.600	Bat IIa	2	213.600	Bat IIa	2	213.600	Bat IIa	2	213.600
	SHK	2	40.800	SHK	2	40.800	SHK	2	40.800	SHK	2	40.800
	zus.:	4	254.400	zus.:	4	254.400	zus.:	4	254.400	zus.:	4	254.400
SV												
				Kosten- kategorie oder Kennziff.	Betrag DM	Kosten- kategorie oder Kennziff.	Betrag DM	Kosten- kategorie oder Kennziff.	Betrag DM			
				547	2.400	547	1200	-	-			
				-	-	-	-	-	-			
				-	-	-	-	-	-			
				-	-	-	-	-	-			
				-	-	-	-	-	-			
				zus.	2.400	zus.	1.200	zus.	-			
I				Investitionsmittel insges.	Investitionsmittel insges.	Investitionsmittel insges.	Investitionsmittel insges.					
				-	-	-	-					

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
Grundaussstattung						
3.7.1.1 wissenschaftl. Mitarbeiter (einschl. Hilfskräfte)	1. Jörg Siekmann, Prof. Dr., Universitätsprofessor 2. Manfred Pinkal, Prof. Dr., Universitätsprofessor 3. Christoph Benz Müller, Dr., Hochschulassistent 5. Ivana Kruijff-Korbayova, Dr., wiss. Mitarb. 6. Erica Melis, Dr. habil., Privatdozent	Informatik Computerlinguistik Informatik Computerlinguistik Informatik	FR Informatik FR Computerlinguistik FR Informatik FR Computerlinguistik FR Informatik	5 5 10 8 5	1.1.96 1.1.96 1.1.01 1.1.98 1.1.98	
3.7.1.2 nichtwissenschaftl. Mitarbeiter	7. Irmtraud Stein, Verwaltungsangestellte 8. Helga Riedel	– –	FR Informatik FR Computerlinguistik	3 3	1.1.96 1.1.96	
Ergänzungsaussstattung						
3.7.1.3 wissenschaftliche Mitarbeiter (einschl. Hilfskräfte)	9. Armin Fiedler, Dipl. Inform., wiss. Mitarb. 10. N.N., wiss. Mitarb. 11. Abdel Hakim Freihab, SHK 12. Alexander Weber, SHK	Informatik Computerlinguistik Computerlinguistik Informatik	FR Informatik FR Computerlinguistik FR Computerlinguistik	38,5 38,5 19 19	1.1.96	BAT IIa BAT IIa SHK SHK
3.7.1.4 nichtwissenschaftl. Mitarbeiter						

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

Aufgabenbeschreibung von Mitarbeitern der Grundausrüstung

Jörg Siekmann (MA 1.) Projektleitung, Anleitung und Betreuung im Bereich der Künstlichen Intelligenz

Manfred Pinkal (MA 2.) Projektleitung, Anleitung und Betreuung im Bereich der Computerlinguistik

Christoph Benzmüller (MA 3.) Projektleitung, Koordination der Arbeitspakete, wissenschaftliche Mitarbeit aus Richtung des dynamischen mathematischen Assistenzsystems

Helmut Horacek (MA 4.) Anleitung und wissenschaftliche Mitarbeit zu Aspekten der Sprachverarbeitung, insbesondere der Generierung, dem Dialogmanagement und der tutoriellen Steuerung

Ivana Kruijff-Korbayova (MA 5.) Anleitung und wissenschaftliche Mitarbeit zum Aspekt Sprachverarbeitung

Erica Melis (MA 6.) Anleitung und wissenschaftliche Mitarbeit zum Aspekt tutorielle Systeme für die Mathematik

Irmtraud Stein (MA 7.) Sekretariat

Helga Riedel (MA 8.) Sekretariat

Aufgabenbeschreibung von Mitarbeitern der Ergänzungsausrüstung

(MA 9.) Für die Durchführung des Projekts ist einerseits Informatik-Expertise, insbesondere im Bereich der Verarbeitung mathematischen Wissens, aber auch in Architektur- und Implementierungsfragen für die Empirische Experimente und Demonstrator, nötig. Für diese Stelle ist Dipl.-Inform. Armin Fiedler vorgesehen.

(MA 10.) Der Mitarbeiter soll in der Dialogverarbeitung und bei den Schnittstellen des Dialogs zu anderen Komponenten mitwirken und insbesondere den experimentellen und den Modellierungsteil mit abdecken. Außerdem ist die Stelle für die Entwicklung der Sprachschnittstellen erforderlich.

(SHK 11.) Implementierungsarbeiten in verschiedenen Arbeitspaketen. Hierfür ist Hakim Freihat vorgesehen.

(SHK 12.) Hilfe bei der Vorbereitung, Durchführung und Auswertung der Versuche; Mitwirkung bei der Grammatikentwicklung. Hierfür ist Alexander Weber vorgesehen.

3.7.2 Aufgliederung und Begründung der Sächlichen Verwaltungsausgaben (nach Haushaltsjahren)

	2002	2003	2004
Für Sächliche Verwaltungsausgaben stehen als <i>Grundausrüstung</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)	2400	1200	-

(Alle Angaben in DM.)

Begründung zur *Ergänzungsausstattung* der Sächlichen Verwaltungsausgaben

Sonstiges (547) Zur Erarbeitung der Phänomene eines natürlich-sprachlichen Dialogs mit einem mathematischen Assistenzsystem sind in AP1 Experimentreihen vorgesehen. Wir gehen von drei Experimenten mit je 40 Versuchspersonen zu je DM 20 aus. Die Vorstudien haben gezeigt, das man pro Experiment etwa 1,5 Stunden ansetzen muss. Zwei Experimente sollen im ersten Jahr durchgeführt werden und ein drittes im zweiten Jahr. Es ergibt sich ein Gesamtbedarf von ca. DM 3600.

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

	Beantragt für das Haushaltsjahr		
	2002	2003	2004
Summe:			

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