

Humboldt-Universität zu Berlin

Institut für Informatik

www.informatik.hu-berlin.de

Jahresbericht 2003



©Humboldt-Universität zu Berlin, 2003
Institut für Informatik
Unter den Linden 6
10099 Berlin

Besuchsanschrift: Rudower Chaussee 25, 12489 Berlin
Druck: Offset-Druckerei Gerhard Weinert GmbH, Saalburgstr. 3, 12099 Berlin
Redaktion: Prof. Wolfgang Reisg, Christine Henze
Redaktionsschluß: 31.12.2003

Vorwort

Seit Jahrzehnten beeinflusst Informatik die Gesellschaft in vielen Bereichen und sie ändert dabei laufend ihr eigenes Profil. In schnellen Schüben entstehen neue Konzepte und Produkte. Innovation in Forschung und Aktualisierung der Lehre ist der Schlüssel des Erfolgs.

Vor diesem Hintergrund freut sich das Institut, dass Prof. Dr. Martin Grohe als herausragender Vertreter des Gebietes "Logik in der Informatik" den Ruf des Senators zum 1. August 2003 angenommen hat. Prof. Grohe stärkt die Theoretische Informatik und schlägt mit seinen Arbeiten Brücken hin zu Anwendungen logik-basierter Konzepte in vielen Bereichen der Informatik.

Im Rahmen des Aktionsplans Informatik der Deutschen Forschungsgemeinschaft hat das Institut zwei Nachwuchs-Forschergruppen eingeworben und als Juniorprofessuren ausgestaltet. Damit konnte das Institut auch in Zeiten der Kürzung die Zahl seiner Professoren erhöhen! Seit April 2003 leitet Professor Dr. Felix Naumann die Gruppe "Informationsintegration", und Professor Dr. Tobias Scheffer die Gruppe "Wissensmanagement".

Zahlreiche Mitglieder des Instituts sind für ihre wissenschaftliche Arbeit im Jahr 2003 ausgezeichnet worden:

- Prof. Dr. Freytag hat zum vierten Mal einen "Faculty Award" von IBM erhalten
- Dr. Deryk Osthus gewann den European Prize in Combinatorics" (gemeinsam mit Daniela Kühn, FU Berlin), verliehen auf der Eurocomb 2003;
- Sören Sonnenburg bekam 1. Hugo – Geiger – Preis für seine Diplomarbeit zur Erkennung von Spleißstellen im Genom
- Dr. Stephan Kreutzer und Dr. Nicole Schweikardt, die beide seit Sommer 2003 am Institut als wissenschaftliche Mitarbeiter tätig sind, wurde gemeinsam der Dissertationspreis der Gesellschaft für Informatik für ihre Dissertationen im Bereich der Logik in der Informatik verliehen.

Im Dezember 2002 hat eine auswärtige Kommission unter Leitung von Prof. K. Mehlhorn vom Max Plank Institut für Informatik, Saarbrücken, die Forschung des Institutes evaluiert und ihm ausgezeichnete Leistungen bescheinigt. Das Institut bedankt sich bei der Kommission für ihre Empfehlungen zu weiteren Verbesserungen.

Im Sommersemester hat das Institut aus seinen drei Schwerpunkten

Algorithmen und Logik,

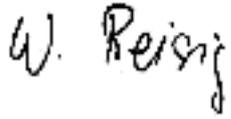
Modellbasierte Systementwicklung und

Große Datenräume in Web-basierten Umgebungen

heraus mit seiner Ringvorlesung "Schwerpunkte der Informatik" begonnen. Die Veranstaltungen waren insgesamt sehr gut besucht; das Institut wird die Ringvorlesung in Zukunft regelmäßig anbieten.

Mit dem Erwin Schrödinger-Zentrum stehen dem Institut nun sehr funktionale Räume zur Verfügung, so dass insbesondere die großen Lehrveranstaltungen des Grundstudiums nicht mehr in Mitte statt finden müssen.

Wolfgang Reisig
Geschäftsführender Direktor Informatik

Handwritten signature of Wolfgang Reisig in black ink.

Berlin, Januar 2004

Preface

For many decades Computer Science had an impact on various areas of society and has been continuously updating its own profile. New concepts and products are created at an increasing pace. Innovation in research and curricula updates are the keys to success.

The institute was very pleased that Professor Dr. Martin Grohe, an outstanding scientist in the areas of logic, algorithms and complexity, has accepted his appointment and joined our faculty on August 1, 2003. Professor Grohe strengthens Theoretical Computer Science and his work builds bridges to applications of logic-based concepts in numerous areas of computer science.

The institute established two young research groups and appointed two Junior Professors who are funded by the project “Aktionsplan Informatik” of Deutsche Forschungsgemeinschaft. Thus the institute was able to increase its number of professors even in this difficult period of financial cuts and employment freeze.

In April 2003 Professor Dr. Felix Naumann became the head of the group “Information Integration” and Professor Dr. Tobias Scheffer became the head of the group “Knowledge Management.”

A number of colleagues and staff of the institute have been awarded prizes for their scientific work in 2003:

- Prof. Dr. Freytag received the IBM Faculty Award for the 4th time
- Dr. Deryk Osthus (together with Daniela Kühn, FU Berlin) was awarded the European Prize in Combinatorics at Eurocomb 2003
- Sören Sonnenburg was awarded the 1st Hugo-Geier-Award for his Diploma Thesis on the detection of splice signals in genomes
- Dr. Stephan Kreutzer and Dr. Nikole Schweikardt who have been at the institute as research assistants since summer 2003 received the “Dissertationspreis” (Best Dissertation Award) of the “Gesellschaft für Informatik” (German Computer Society) for their dissertations in the area “Logic in Computer Science.”

An external commission lead by Professor K. Mehlhorn of the Max Planck Institute in Saarbrücken evaluated the institute in December 2002 and certified excellent research work. The Institute thanks the commission for advice on further improvements.

The Institute decided to focus and accelerate its research in:

Algorithms and Logic

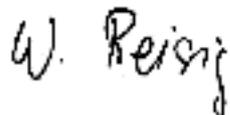
Model-driven Software Development and

Large Data Spaces in Web-based Environments

Based on these three research clusters the institute started a series of lectures under the title “Schwerpunkte der Informatik” in the Summer Semester. The lectures found large acceptance and will be held regularly in the future.

The newly-built Erwin Schrödinger-Zentrum provides functional rooms for our teaching activities and that we do not need to offer our undergraduate education in the Main Building in Berlin-Mitte any longer.

Wolfgang Reisig
Chairman

A handwritten signature in black ink, appearing to read "W. Reisig". The signature is written in a cursive, slightly slanted style.

Berlin, January 2004

Inhaltsverzeichnis

I.	Institutsorganisation	9
II.	Professoren am Institut für Informatik	11
III.	Lehr- und Forschungseinheiten	21
	<i>Theoretische Informatik</i>	
	Logik in der Informatik	
	<i>Leiter: Prof. Dr. Martin Grohe</i>	21
	Algorithmen und Komplexität	
	<i>Leiter: Prof. Dr. Hans Jürgen Prömel</i>	27
	Zufällige Strukturen und Algorithmen	
	<i>Leiter: Prof. Dr. Anusch Taraz</i>	27
	Komplexität und Kryptographie	
	<i>Leiter: Prof. Dr. Johannes Köbler</i>	40
	<i>Praktische Informatik</i>	
	Systemanalyse, Modellierung und Computersimulation „SAM“	
	<i>Leiter: Prof. Dr. Joachim Fischer</i>	45
	Theorie der Programmierung	
	<i>Leiter: Prof. Dr. Wolfgang Reisig</i>	60
	Softwaretechnik	
	<i>Leiter: Prof. Dr. Klaus Bothe</i>	70
	Datenbanken und Informationssysteme	
	<i>Leiter: Prof. Johann-Christoph Freytag, Ph. D.</i>	76
	Wissenschaftsmanagement in der Bioinformatik	
	<i>Leiter: Prof. Dr. Ulf Leser</i>	99
	Künstliche Intelligenz	
	<i>Leiter: Prof. Dr. Hans-Dieter Burkhard</i>	109
	Wissenschaftsmanagement	
	<i>Leiter: Prof. Dr. Tobias Scheffer</i>	134
	Datenanalyse	
	<i>Leiter: Prof. Dr. Egmar Rödel</i>	139
	Informatik in Bildung und Gesellschaft	
	<i>Leiter: Prof. Dr. Wolfgang Coy</i>	146
	Parallele und Verteilte Systeme	
	<i>Leiter: Prof. Dr. Alexander Reinefeld</i>	160

	Spezifikation, Verifikation und Testtheorie	
	<i>Leiter: Prof. Dr. Bernd-Holger Schlingloff</i>	175
	Informationsintegration	
	<i>Leiter: Prof. Dr. Felix Naumann</i>	179
	<i>Technische Informatik</i>	
	Rechnerorganisation und Kommunikation	
	<i>Leiter: Prof. Dr. Mirosław Malek</i>	183
	Signalverarbeitung/ Mustererkennung	
	<i>Leiterin: Prof. Dr. Beate Meffert</i>	198
IV.	Lehre	206
V.	Informationstechnik des Instituts für Informatik	224
	Rechnerbetriebsgruppe	224
	Netzwerkbetriebsgruppe	229
VI.	Ideen - Werkstatt und Studienberatung für Studentinnen und Schülerinnen	232
	<i>Leiterin: Dr. Mårta Gutsche</i>	

I. Institutsorganisation

Postadresse: Humboldt-Universität zu Berlin, Institut für Informatik
Unter den Linden 6, 10099 Berlin

Sitz: Rudower Chaussee 25 / Ecke Magnusstraße,
12489 Berlin

Geschäftsführender Direktor: PROF. DR. WOLFGANG REISIG
Sekretariat: BIRGIT HEENE, Tel.: 2093 3066, Raum IV 417

Prüfungsausschuß

Vorsitzender: PROF. DR. KLAUS BOTHE, Tel.: 2093 3008, Raum IV 201
Sprechzeit: dienstags, 13.00 – 14.00 Uhr in Raum II 323

Mitarbeiterin für Studium, Lehre und Prüfung

HEIDI NEUGEBAUER, Tel.: 2093 3000, Raum II 323
Sprechzeiten: montags 13.00 – 16.00 Uhr
dienstags 14.00 – 16.00 Uhr
mittwochs, freitags 9.00 – 12.00 Uhr

Studienfachberatung

Informatik (Diplom, Magister, Magisternebenfach): PROF. DR. BODO HOHBERG
Tel.: 2093 3068, Raum IV 411
Sprechzeit: donnerstags 14.00 – 15.00 Uhr

Informatik (Lehramt): PROF. DR. EGMAR RÖDEL, Tel.: 2093 3079, Raum IV 401
Sprechzeit: dienstags 15.30 – 16.30 Uhr

Studentische Studienberatung

Fachschaft: CHRISTIAN BECKER, Tel.: 2093 3923, Raum II 321
Sprechzeiten: dienstags 15.00 – 16.00 Uhr und per Mail

Studienberatung für Studentinnen und Schülerinnen

Leiterin: DR. MÄRTA GUTSCHE, Tel: 2093 5468, Raum IV 106
Sprechzeiten: mittwochs 10.00 – 17.00 Uhr und nach Vereinbarung

Ideenwerkstatt „Frauen in die Informatik“

Leiterin: DR. MÄRTA GUTSCHE, Tel: 2093 5468, Raum IV 106

Institutsrat

Direktor: PROF. DR. WOLFGANG REISIG, Tel: 2093 3065, Raum IV 416

Kommission Lehre und Studium

Vorsitzender: PROF. JOHANN-CHRISTOPH FREYTAG, PH. D., Tel.: 2093 3009, Raum IV 203

Bibliothekskommission

Vorsitzender: PROF. DR. KLAUS BOTHE, Tel.: 2093 3007, Raum IV 201

Haushaltskommission

Vorsitzender: PROF. DR. JOACHIM FISCHER, Tel. 2093 3109, Raum III 303

Institutsbibliothek

Leiterin: ANJA HERWIG, Tel.: 2093 2357

Verwaltung

Haushalt und Personal: RITA FALCK, Tel.: 2093 3002, Raum II 316

II. Professoren am Institut für Informatik



PROF. DR. MARTIN GROHE

Studierte von 1987 bis 1992 Mathematik mit Nebenfach Informatik an der Albert-Ludwigs-Universität Freiburg und promovierte dort im Jahre 1994 bei Heinz-Dieter Ebbinghaus in der Mathematischen Logik. Die Jahre 1995-96 verbrachte er als Postdoktorand an der Stanford University und der University of California in Santa Cruz. Anschließend kehrte er nach Freiburg zurück und habilitierte dort im Jahre 1998 an der mathematischen Fakultät. Im akademischen Jahr 2000-2001 hatte er eine Assistenzprofessur an der University of Illinois in Chicago inne, von dort wechselte er 2001 als Reader an die University of Edinburgh. Seit August 2003 ist er Professor am Institut für Informatik der Humboldt-Universität zu Berlin.

Die Forschungsinteressen von Professor Grohe liegen in den Bereichen Logik, Algorithmen, Komplexitätstheorie, Graphentheorie und Datenbanktheorie. 1999 wurde er für seine Arbeiten mit dem Heinz Maier-Leibnitz-Preis der Deutschen Forschungsgemeinschaft und des Bundesministeriums für Bildung und Forschung ausgezeichnet, und im Jahre 2001 wurde er von der amerikanischen Sloan Foundation zum „Alfred P. Sloan Fellow“ ausgewählt. Er ist Mitherausgeber der Zeitschriften „Journal of Symbolic Logic“ und „Journal of Discrete Algorithms“.



PROF. DR. HANS JÜRGEN PRÖMEL

Studierte Mathematik und Wirtschaftswissenschaften an der Universität Bielefeld (1974-1979) und promovierte 1982 dort im Fach Mathematik. In den Jahren 1984-1985 war er Gastprofessor an der University of California, Los Angeles. Von 1988-1994 hatte er eine Professur (C4) für Diskrete Mathematik an der Universität Bonn inne. Seit 1994 ist Prof. Prömel Inhaber des Lehrstuhls für Algorithmen und Komplexität an der Humboldt-Universität zu Berlin. Derzeit ist er als Vizepräsident für Forschung der Humboldt-Universität von seiner Professur beurlaubt.

Im Mittelpunkt der Forschungsinteressen von Prof. Prömel steht die Untersuchung diskreter mathematischer Strukturen wie Graphen und Netzwerke und ihre algorithmische Behandlung. Ein Schwerpunkt dabei ist die Ausnutzung des Zufalls, zum einen um strukturelle Einsichten zu gewinnen, zum anderen als Hilfsmittel beim Entwurf randomisierter und approximativer Algorithmen. Prof. Prömel ist Koautor des Buches „The Steiner Tree Problem. A Tour through Graphs, Algorithms and Complexity“ und Mitherausgeber mehrerer Sammelbände. Er gehört zudem derzeit den Editorial Boards der Zeitschriften „Random Structures and Algorithms“, „Combinatorics, Probability and Computing“, „Electronic Journal of Combinatorial Number Theory“ sowie „SIAM Journal on Discrete Mathematics“ an. Prof. Prömel ist unter anderem Sprecher der von der Deutschen Forschungsgemeinschaft geförderten Forschergruppe „Struktur, Algorithmen, Zufall“, Mitglied der Bundesjury beim Wettbewerb „Jugend forscht“ und Mitglied des Präsidiums der Deutschen Mathematiker-Vereinigung.



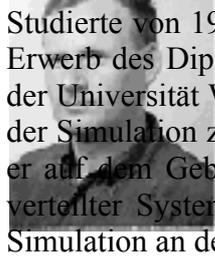
PROF. DR. ANUSCH TARAZ

Nach seinem Abitur in Großbritannien studierte Anusch Taraz von 1989 bis 1995 Mathematik, Informatik und Operations Research an der Universität Bonn. Im Anschluss an sein Diplom kam er an das Institut für Informatik der Humboldt-Universität und schloss 1999 seine Promotion mit einer Dissertation über Phasenübergänge in der Evolution partieller Ordnungen ab. Danach war er hier zunächst als wissenschaftlicher Assistent tätig, vertrat dann im Wintersemester 2001/02 den Lehrstuhl für Algorithmen und Komplexität und wurde im Mai 2002 auf die Juniorprofessur für Zufällige Strukturen und Algorithmen am Institut für Informatik berufen. Sein Arbeitsgebiet ist an der Schnittstelle von Theoretischer Informatik und Diskreter Mathematik angesiedelt und beschäftigt sich schwerpunktmäßig mit diskreten stochastischen Modellen, die Wahrscheinlichkeitsverteilungen auf dem Raum aller möglichen Inputs eines Algorithmus darstellen. Sie ermöglichen es einerseits, ein Verfahren nach seiner durchschnittlichen Leistung zu beurteilen und vermeiden auf diese Weise den übertriebenen Pessimismus von worst-case Paradigmen. Andererseits erfordern sie die genaue Kenntnis von Struktureigenschaften, die Eingaben mit hoher Wahrscheinlichkeit besitzen. Diese Kenntnis ist darüber hinaus eine wesentliche Voraussetzung für die Anpassung der Modelle an reale Fragestellungen, wie beispielsweise bei der Simulation von komplexen Netzwerken in der Informationstechnologie oder großen Datenbanken in der Bioinformatik.



PROF. DR. JOHANNES KÖBLER

Studierte von 1978 bis 1985 Informatik mit Nebenfach Mathematik an der Universität Stuttgart. Nach seiner Promotion im Jahr 1989 wechselte er an die Universität Ulm und habilitierte dort 1995 im Fach Theoretische Informatik. Seit Oktober 1999 ist er Professor für Algorithmen und Komplexität an der Humboldt-Universität zu Berlin. Die Forschungsinteressen von Prof. Köbler liegen auf den Gebieten Komplexitätstheorie, Algorithmisches Lernen und Kryptografie. Sein Hauptinteresse gilt der Komplexität konkreter algorithmischer Problemstellungen wie etwa der des Graphisomorphieproblems und Fragestellungen wie „Lässt sich die Effizienz von Algorithmen durch Zuhilfenahme von Zufallsentscheidungen oder von Interaktion steigern?“ Daneben vertritt Prof. Köbler die Lehrgebiete (probabilistische und approximative) Algorithmen, Automaten-theorie und formale Sprachen, Berechenbarkeitstheorie und Logik.

PROF. DR. JOACHIM FISCHER

Studierte von 1973 bis 1978 Mathematik an der Humboldt-Universität zu Berlin. Nach dem Erwerb des Diploms absolvierte er 1979 ein Ergänzungsstudium am Institut für Informatik der Universität Warschau. 1982 promovierte er an der Humboldt-Universität auf dem Gebiet der Simulation zeitdiskreter und zeitkontinuierlicher Prozesse. Sechs Jahre später habilitierte er auf dem Gebiet „Mathematische Informatik“ mit einer Arbeit zum „Rapid Prototyping verteilter Systeme“. 1994 wurde er zum Professor für Systemanalyse, Modellierung und Simulation an der Humboldt-Universität zu Berlin berufen.

Im Mittelpunkt des aktuellen Forschungsinteresses von Prof. Fischer steht die Entwicklung werkzeuggestützter Modellierungs- und Simulationsmethoden verteilter Systeme und deren Anwendung im Telekommunikationsbereich bei Einsatz verteilter Objekttechnologien. Einen Schwerpunkt bildet dabei die konzeptionelle Weiterentwicklung der genormten Spezifikationstechnik „*Specification and Description Language*“ (SDL) in ihrer Kombination mit weiteren praxisrelevanten *Computational-* und *Engineering-*Beschreibungstechniken wie OMG-UML, ITU-ODL und *OMG-Component IDL*.

Ein weiterer Schwerpunkt ist die Erforschung CORBA-basierter Plattformarchitekturen für Applikationen mit sowohl operationalen als auch *Stream*-basierten, multimedialen Interaktionen im Telekommunikationsbereich. Ein Großteil der an seinem Lehrstuhl betriebenen Forschungen wird aus Drittmitteln im Rahmen internationaler Projekte finanziert. Bedeutende industrielle Kooperationspartner der letzten Jahre waren T-Nova, Siemens-AG, NTT (Japan), EURESCOM GmbH und gecco.net AG.

Seine Mitarbeiter sind in verschiedenen internationalen Standardisierungsgremien wie der OMG, der ITU und der C++ *ABI Working Group* aktiv. Prof. Fischer selbst leitet als Rapporteur in der Studiengruppe 17 der ITU-T (Sprachen und allgemeine Software-Aspekte für Telekommunikationssysteme) derzeit zwei unterschiedliche Projekte. Er ist Mitglied des DIN-Ausschusses 21.1 und der Arbeitsgemeinschaft „Simulation“ in der Gesellschaft für Informatik (ASIM).

Prof. Fischer ist Mitautor mehrerer Fachbücher: „Digitale Simulation: Konzepte-Werkzeuge-Anwendungen“ (Akademie-Verlag Berlin 1990), „Objektorientierte Programmierung“ (Verlag Technik Berlin/München 1992) und „Objektorientierte Prozeßsimulation“ (Addison-Wesley-Verlag 1996).

Von 1997 bis 1998 leitete Prof. Fischer als Geschäftsführender Direktor die Verlagerung des Instituts von Berlin-Mitte nach Berlin-Adlershof.

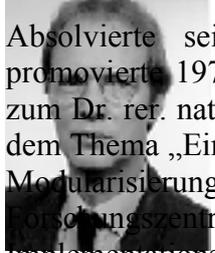
PROF. DR. WOLFGANG REISIG

Studierte in Karlsruhe und Bonn Physik und Informatik. Er war wissenschaftlicher Mitarbeiter und Assistent von 1974 bis 1983 an der Universität Bonn und der RWTH Aachen. Dort promovierte er 1979 bei Prof. Indermark zur Analyse kooperierender sequentieller Prozesse. 1983 vertrat er eine Professur an der Universität Hamburg und leitete anschließend bei der Gesellschaft für Mathematik und Datenverarbeitung ein Projekt zur Systemanalyse und -Modellierung. 1987 habilitierte er an der Universität Bonn und wurde danach zum Professor für Theoretische Informatik an die TU München berufen. Seit 1993 ist er Professor am Institut für Informatik der Humboldt-Universität zu Berlin.

Die Forschungsinteressen Prof. Reisigs umfassen Verteilte Algorithmen, insbesondere Methoden und formale Konzepte für ihren systematischen Entwurf, ihre Modellierung und den Nachweis ihrer Korrektheit, die Verwendung solcher Algorithmen und Modelle in der Praktischen Informatik (Rechnerarchitektur, verteilte Datenbanken) und der Angewandten

Informatik (rechnerintegrierte Steuerung und Organisation). In der Theorie der Programmierung interessieren ihn insbesondere Petrinetze, Prozessalgebren, Temporale Logik, algebraische Softwarespezifikation, Abstract-State Machines und unkonventionelle Rechner- und Algorithmenstrukturen. Seine Bücher über Petrinetze wurden in mehrere Sprachen übersetzt.

PROF. DR. KLAUS BOTHE



Absolvierte sein Mathematikstudium an der Humboldt-Universität zu Berlin und promovierte 1979 mit dem Thema „Spezifikation und Verifikation abstrakter Datentypen“ zum Dr. rer. nat. an gleicher Stätte. Dort habilitierte er dann auch 1986 zum Dr. sc. nat mit dem Thema „Ein algorithmisches Interface für Pascal-Compiler: Compiler-Portabilität durch Modularisierung“. Vom September 1986 bis Juli 1987 arbeitete er am ungarischen Forschungszentrum SZKI in Budapest zu den Themen Logische Programmierung, Implementationstechniken von Prolog und Expertensystemen. Von September 1991 bis Februar 1992 erhielt er ein Sonderforschungsstipendium der Humboldt-Stiftung, das er zu einem Aufenthalt in Erlangen bei Prof. Stoyan nutzte. Seit Dezember 1993 ist er Professor für Softwaretechnik und Theorie der Programmierung an der Humboldt-Universität zu Berlin. Die bisherigen wissenschaftlichen Arbeitsgebiete waren: Theorie der Programmierung, Compilerbau (hier wurden Projekte zu Problemen der Quelltexttransformation, zu Portierungstechniken sowie zur Einbeziehung modularer Softwarearchitekturen in den Compilerbau realisiert), Logische Programmierung sowie Expertensysteme (in Zusammenarbeit mit der Charité wurde an einem Expertensystem zur Nierendiagnostik gearbeitet). 1991 erschien unter Mitwirkung von S. Stojanow das Buch „Praktische Prolog-Programmierung“ im Verlag Technik Berlin, München.



PROF. JOHANN-CHRISTOPH FREYTAG, PH.D.

Begann sein Studium 1975 in Hamburg und setzte es an der Harvard Universität, MA, USA, fort, wo er 1985 seine universitäre Ausbildung mit dem Ph.D. in Applied Mathematics/ Computer Science abschloss. Danach arbeitete er zwei Jahre am IBM Almaden Research Center (ARC), CA, USA, am Starburst Datenbankprojekt mit, dessen Technologie im heutigen IBM-Datenbankprodukt DB2/UDB wiederzufinden sind. 1987 kehrte er nach Europa zurück und war für 2 Jahre am ECRC (European Computer Industry Research Centre) im Bereich der Anfragebearbeitung und Transaktionsverwaltung in deduktiven Datenbanken und objektorientierten Datenbanksystemen tätig. 1990 übernahm er den Aufbau der Database System Research Gruppe und des Database Technology Centers für Digital Equipment Inc., USA, in München, als deren Leiter er für fast vier Jahre Forschung und Technologietransfer im Bereich Datenbankoptimierung und Anwendung von Datenbanktechnologie im CIM-Bereich koordinierte und selbst forschend tätig war. Dabei entstanden innovative Arbeiten für DEC's Datenbanksystem Rdb/VMS und für das Produkt Database Integrator (DBI) als Teil des Technologietransfers. Im Oktober 1993 wechselte Prof. Freytag an die TU München, ehe er im Februar 1994 seine Tätigkeit als Professor für Datenbanken und Informationssysteme an der Humboldt-Universität aufnahm. Parallel zu diesen Tätigkeiten war er von 1986 bis 1993 für die Firma Codd & Date Inc., CA, weltweit in Industrieseminaren tätig. Seine wesentlichen Arbeitsgebiete umfassen Anfragebearbeitung in Datenbanksystemen, Optimierungstechniken für zentrale und parallele Datenbanksysteme, aktive Datenbanken, Workflow und Datenbanken, die

Entwicklung geeigneter Schnittstellen zu komplexen Anwendungen sowie alle Aspekte der Datenmodellierung. Seit mehr als drei Jahren widmet er sich im Bwsonderen dem Bereich Bioinformatik/Life Science.

Für seine Arbeiten erhielt Prof. Freytag den IBM-Faculty-Award in den Jahren 1999, 2001 und 2002, sowie den „IBM Shared University Research Grant“ (SUR-Grant) im Jahre 2001. Als „Technical Program Chair“ wird er im Jahr 2003 die „Very Large Database“ (VLDB-) Konferenz, die weltweit wichtigste Konferenz im Bereich Datenbanken, an der Humboldt-Universität zu Berlin organisieren.



PROF. DR. ULF LESER

Studierte Informatik an der Technischen Universität München und arbeitete danach 2 Jahre am Max-Planck-Institut für molekulare Genetik in Berlin an der Entwicklung von integrierten Datenbanken im Rahmen des Human Genome Projekt. Von 1997 bis 2000 promovierte er am Graduiertenkolleg „Verteilte Informationssysteme“ über Anfragealgorithmen in heterogenen Informationssystemen. Daneben forschte er an Annotationsworkflows für Hochdurchsatzexperimente in der molekular-biologischen Forschung, der Standardisierung von Datenmodellen und Zugriffsmethoden in Life Sciences auf Basis von CORBA und der qualitätsbasierten Anfrageoptimierung. Nach der Promotion ging er in die Industrie und leitete bei der UBIS AG Softwareentwicklungsprojekte im Bereich Data Warehousing, Beratungstechnologien im eCommerce und Wissensmanagement. Seit 2002 ist er Professor für Wissensmanagement in der Bioinformatik an der Humboldt Universität.



PROF. DR. HANS-DIETER BURKHARD

Studierte von 1962-68 Mathematik in Jena und Berlin. Zwischenzeitlich arbeitete er 1965/66 als Programmierer im Rechenzentrum der Deutschen Reichsbahn in Berlin. Er erwarb 1974 die Promotion A (Gebiet Automaten-theorie) und 1985 die Promotion B (Gebiet Verteilte Systeme). Seit 1972 war er Wissenschaftlicher Mitarbeiter an der Humboldt-Universität in den Bereichen Mathematik und Informationsverarbeitung. 1989/90 war er Mitglied des Runden Tisches an der Humboldt-Universität, und von 1991-98 war er Vorsitzender des Konzils. Im Herbst 1990 wurde er zum Dozenten berufen, 1992 erfolgte die Berufung zum Professor für Künstliche Intelligenz. In seiner wissenschaftlichen Tätigkeit entstanden theoretische und anwendungsorientierte Arbeiten auf den Gebieten Automaten-theorie, Schaltkreis-Diagnose, Petrinetze, Verteilte Systeme und Künstliche Intelligenz. Die aktuellen Interessengebiete sind Verteilte Künstliche Intelligenz, Agenten-orientierte Techniken, Fallbasiertes Schließen, Knowledge Management, Kognitive Robotik, Sozionik und Anwendungen der Künstlichen Intelligenz im Bereich der Medizin. Er ist Mitglied im Board of Trustees der internationalen RoboCup Federation und ECCAI-Fellow.

PROF. DR. TOBIAS SCHEFFER



Studierte von 1990 bis 1995 Informatik an der Technischen Universität Berlin. Seine anschließende Promotion wurde mit einem Ernst-von-Siemens-Stipendium gefördert. Nach Aufhalten bei Siemens Corporate Research in Princeton und an der University of New South Wales in Sydney promovierte er 1999 an der Technischen Universität Berlin über die statistischen Grundlagen des maschinellen Lernens. Er war als wissenschaftlicher Assistent an der Otto-von-Guericke-Universität Magdeburg beschäftigt, hat an zwei Startup-Unternehmen mitgearbeitet und wurde

2003 im Aktionsplan Informatik in das Emmy-Noether-Programm der Deutschen Forschungsgemeinschaft aufgenommen und zum Juniorprofessor für Wissensmanagement an der Humboldt-Universität zu Berlin berufen.

Die Interessen Prof. Scheffers umfassen Verfahren, die Wissen in großen Datenbanken – oder großen Textsammlungen – entdecken und nutzbar machen. Dazu gehören statistische, maschinelle Lernverfahren, Sampling- und aktive Lernalgorithmen, Assoziationsregel- und Subgruppenverfahren. Untersuchte Anwendungen stammen vor allem aus dem Dokumentenmanagement und der Bioinformatik.



PROF. DR. EGMAR RÖDEL

Studierte von 1960 bis 1965 Mathematik an der Humboldt-Universität zu Berlin. 1969 promovierte er zum Dr. rer. nat. mit dem Thema „Abhängigkeitsmaße für Zufallsgrößen mit Werten in einem separablen Hilbertraum“. 1971 bis 1974 übte er verschiedene Tätigkeiten in Industrie und Wirtschaft aus, ehe er seine Lehrtätigkeit an der Sektion Mathematik der Humboldt-Universität fortsetzte. 1974 bis 1990 war er als wissenschaftlicher Mitarbeiter am Organisations- und Rechenzentrum der

Humboldt-Universität in den Arbeitsgebieten Angewandte Statistik und Numerik, Optimierung, Entwicklung und Anwendung mathematisch-statistischer Software tätig und hielt Vorlesungen an der Sektion Mathematik zum Thema „Mathematische Statistik und Ökonometrie“. 1981 bis 1990 war Prof. Rödel Leiter der Arbeitsgruppe „Statistische Informationsverarbeitung“ der Hoch- und Fachschulen der DDR. 1983/84 absolvierte er einen 6-monatigen Studienaufenthalt an der Karls-Universität Prag, Institut für Mathematische Statistik, in der Gruppe von Frau Prof. Dr. J. Jureckova und habilitierte 1985 zum Dr. sc. nat. an der Sektion Mathematik der Universität Rostock mit dem Thema „Unabhängigkeitstests für zweidimensionale Verteilungen mit endlicher Kontingenz und ihre Adaption“. 1992 wurde er zum Professor für Datenanalyse an das Institut für Informatik berufen.



PROF. DR. WOLFGANG COY

Studium der Elektrotechnik, Mathematik und Philosophie an der TH Darmstadt mit dem Abschluss Diplomingenieur der Mathematik im Jahr 1972 und einer anschließenden Promotion in Informatik „Zur Komplexität von Hardwaretests“ im Jahr 1975. Es folgten wissenschaftliche Tätigkeiten an der TH Darmstadt, den Universitäten Dortmund, Kaiserslautern und Paris VI. 1979 Professur für Informatik an der Universität Bremen. Seit 1996 vertritt er das Gebiet Informatik in Bildung und Gesellschaft an der Humboldt-Universität zu Berlin. Seine Interessen in Lehre und Forschung liegen in den Bereichen Digitale Medien, Theorie der Informatik, Informatik und Gesellschaft sowie Sozial- und Kulturgeschichte der Informatik.

Buchveröffentlichungen: „Industrieroboter - Zur Archäologie der Zweiten Schöpfung“ (Berlin: Rotbuch 1985), Aufbau und Arbeitsweise von Rechenanlagen (Braunschweig/Wiesbaden: Vieweg, 2. Auflage 1991) und zusammen mit Lena Bonsiepen „Erfahrung und Berechnung - Zur Kritik der Expertensystemtechnik“ (Berlin et al.: Springer, 1989). Mitherausgeber von vier weiteren Büchern, zuletzt zusammen mit Martin Warnke und Christoph Tholen, „HyperKult-Theorie, Geschichte und Kontext Digitaler Medien“ (Basel/Frankfurt am Main: Stroemfeld 1997). Prof. Coy ist im Herausgebergremium der Zeitschrift „Informatik-Spektrum“ und Sprecher der DFG-Forschergruppe „Bild-Schrift-Zahl“.



PROF. DR. ALEXANDER REINEFELD

Alexander Reinefeld studierte zunächst Physik an der TU Braunschweig und anschließend Informatik an der Universität Hamburg und an der University of Alberta (Edmonton, Kanada). 1982 schloss er das Studium mit dem Diplom in Informatik ab und im Jahr 1987 promovierte er zum Dr. rer. nat., beides an der Universität Hamburg. Während seiner beiden einjährigen Forschungsaufenthalte in Edmonton als DAAD-Stipendiat bzw. als Sir Izaak Walton Killam Memorial Post-Doctoral Fellow widmete er sich in den Jahren 1984/ 85 und 1987/88 der Entwicklung effizienter Baum-Suchalgorithmen, die in der Künstlichen Intelligenz zum Fällen von Entscheidungen in komplexen Situationen eingesetzt werden.

Von 1983 bis 87 arbeitete Herr Reinefeld als wissenschaftlicher Mitarbeiter und von 1989 bis 92 als Hochschulassistent an der Universität Hamburg. In den dazwischenliegenden Jahren (1987-90) sammelte er Industrie-Erfahrung als Unternehmensberater in den Bereichen Systemanalyse, Datenbanken und Compilerbau.

1992 wechselte Herr Reinefeld als geschäftsführender Leiter an das Paderborn Center for Parallel Computing, das er maßgeblich als überregionales wissenschaftliches Institut der Universität Paderborn mit aufgebaut hat.

Seit 1998 leitet Herr Reinefeld den Bereich Computer Science am Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB). Diese Aufgabe ist verbunden mit einer Professur für Parallele und Verteilte Systeme am Institut für Informatik der Humboldt-Universität zu Berlin.



Prof. Dr. Holger Schlingloff

Studierte von 1978-1984 Informatik und Logik an der TU München und promovierte dort 1990 mit einer Arbeit zur temporalen Logik von Bäumen. Im Jahr 1991 war Prof. Schlingloff Gastwissenschaftler an der Carnegie-Mellon-Universität in Pittsburgh, PA. Von 1992 bis 1996 war er wissenschaftlicher Assistent am Institut für Informatik der TU München und danach bis 2001 Geschäftsführer des Bremer Instituts für Sichere Systeme (BISS) am Technologie-Zentrum Informatik (TZi) der Universität Bremen. In seiner Habilitation (2001) beschäftigte er sich mit partiellen Zustandsraumanalyseverfahren für sicherheitskritische Systeme. Seit 2002 ist Prof. Schlingloff Professor für Spezifikation, Verifikation und Testtheorie am Institut für Informatik der Humboldt-Universität, und gleichzeitig Leiter des Bereichs Softwaresynthese, Validierung und Test am Fraunhofer Institut für Rechnerarchitektur und Softwaretechnik FIRST. Seine Arbeitsgebiete sind die Software-Qualitätssicherung mit formalen Methoden, temporale Logik und Modellprüfung, sowie spezifikationsbasiertes Testen von eingebetteten Steuergeräten.



PROF. DR. FELIX NAUMANN

Seit Mai 2003 leitet Felix Naumann die DFG Nachwuchsgruppe für Informationsintegration am Institut für Informatik. Die Gruppe erforscht grundlegende Methoden zur Integration von Informationen verschiedener, autonomer Quellen. Felix Naumann studierte von 1990 an Wirtschaftsmathematik an der Technischen Universität Berlin und schloss 1997 das Studium mit einem Diplom ab. Als Mitglied des Berlin-Brandenburger Graduiertenkollegs „Verteilte Informationssysteme“ wechselte Naumann in die Informatik und forschte von 1997 bis 2000 am Lehrstuhl für Datenbanken an der Humboldt-Universität zu Berlin. In seiner Forschung verheiratete Naumann die Forschungsgebiete Datenbanken und Informationsqualität. Er konnte zeigen, dass die Berücksichtigung von Qualitätsmerkmalen wie Vollständigkeit und Zuverlässigkeit in Informationssystemen zu stark verbesserten und zugleich schnelleren Ergebnissen führt. Naumann promovierte in 2000 und erhielt für seine Arbeit den Dissertationspreis der Gesellschaft für Informatik. In 2001 und 2002 war er als Forscher am IBM Almaden Research Center in San Jose, Kalifornien beschäftigt. Dort entwickelte er mit Kollegen den Clio Prototyp, eine Software zur semi-automatischen Transformation von Daten verschiedener Strukturen. Seine Ergebnisse vollzogen erfolgreich den Technologietransfer in IBMs Produktabteilung, honoriert durch dem IBM Research Division Award.



PROF. DR. MIROSLAW MALEK

Erhielt 1970 das Diplom für Elektronik und promovierte 1975 auf dem Gebiet der Technischen Informatik an der Technischen Universität Wroclaw (Breslau), Polen. 1977 war er als Gastwissenschaftler der Universität zu Waterloo in Waterloo, Ontario, Canada. Danach folgten Assistent Professor, Associate Professor und Professor an der Universität zu Texas in Austin, wo er außerdem Inhaber der Bettie-Margaret-Smith- und Southwestern-Bell-Professur war. Im Juli 1994 wurde er zum Professor und Inhaber des Lehrstuhls für Rechnerorganisation und Kommunikation an der Humboldt-Universität zu Berlin berufen. Prof. Maleks Interessen richten sich auf High-Performance und responsives (fehlertolerantes, echtzeitfähiges) Rechnen für parallele und verteilte Rechnersysteme. Er nahm an zwei Pionierprojekten zum Parallelrechnen teil, die entscheidend zu Theorie und Praxis des Parallelen Netzwerkdesigns beitrugen. Er entwickelte die auf Vergleich basierende Methode für Systemdiagnose und hatte teil an der Entwicklung von WSI und Diagnosetechniken in Netzwerken, unterbreitete Vorschläge für den Entwurf konsensbasierter responsiver Rechnersysteme und veröffentlichte mehr als 120 Fachbeiträge und mit G. J. Lipovski das Buch „Parallel Computing: Theory and Comparisons“. Er war außerdem Herausgeber zweier Bücher über responsives Rechnen. Weiterhin führte er den Vorsitz, organisierte und war Programm-Komitee-Mitglied zahlreicher internationaler IEEE- und ACM-Konferenzen und Workshops. Er gehört zu den Herausgebergremien der Zeitschriften „Journal of Interconnection Networks“ und „Real-Time Systems Journal“. Während der Sommer 1984 und 1985 arbeitete er am IBM T. J. Watson Research Center, Yorktown Heights, N.Y. Er war Wissenschaftler am Office of Naval Research in London, Inhaber des IBM- Lehrstuhls an der Keio Universität in Japan in der Zeit von Juni 1990 bis August 1992 und Gastprofessor an der Stanford Universität in Kalifornien (1997/98), an der New York University (2001) und an der CNR/Universita di Pisa (2002).



PROF. DR. BEATE MEFFERT

Nach dem Abitur und einer gleichzeitigen Ausbildung als Funkmechanikerin studierte sie Theoretische Elektrotechnik an der Technischen Hochschule Ilmenau. Während der anschließenden Tätigkeit als wissenschaftliche Assistentin an der Sektion Elektronik der Humboldt-Universität zu Berlin 1976 Promotion (A) über Walshfunktionen und Anwendungen der Walshtransformation, 1983 Promotion (B) zur Theorie und Applikation der Sequenztechnik. 1984 Hochschuldozentin und fünf Jahre später ordentliche Professorin an der Sektion Elektronik der Humboldt-Universität. Seit 1993 Professorin für das Fachgebiet Signalverarbeitung/Mustererkennung am Institut für Informatik der Humboldt-Universität zu Berlin. Zu den bisherigen und gegenwärtigen Arbeitsgebieten gehören: Theorie und Applikation orthogonaler Transformationen; Grundlagen der Signalverarbeitung; Sequenztechnik; Erfassung, Verarbeitung und Klassifikation von Biosignalen zur Unterstützung der Diagnostik und zur Therapiekontrolle; Bildverarbeitung, speziell Datenreduktion im Spektralbereich, Klassifikationsverfahren; Lehrgebiete: Grundlagen der Signalverarbeitung, Mustererkennung, Bildverarbeitung.



PROF. DR. BODO HOHBERG

Studierte von 1959 bis 1964 Mathematik, Spezialisierungsrichtung Algebra, an der Humboldt-Universität zu Berlin und promovierte dort fünf Jahre später zum Thema „Zusammenhang struktureller Eigenschaften von Chomsky-Grammatiken und optimaler Sprachanalyse“. 1978 habilitierte er an der gleichen Universität mit dem Thema „Probleme der Dialog-compilation - Arbeitsweise und Realisierung des MS-Algol Dialog-compilers“.

1987 wurde die Fakultas docendi erteilt. Bis zu seiner Berufung zum außerplanmäßigen Professor 1994 lagen die Forschungsschwerpunkte im Bereich der Softwaretechnik, des systematischen Programmtests und des Compilerbaus. Ab 1995 erfolgte eine Neuorientierung auf Probleme der Modellierung und Verifikation verteilter Systeme und Algorithmen mit Schwerpunkt: Methoden zum systematischen Test verteilter Systeme. Darüber hinaus beschäftigt sich Prof. Hohberg mit der Weiterentwicklung von Methoden zum modularen Beweisen von Eigenschaften der durch Petrinetze modellierten verteilten Systeme.

III. Lehr- und Forschungseinheiten

Lehr- und Forschungseinheit

Logik in der Informatik

<http://www.informatik.hu-berlin.de/logik>

(bis August 2003:Automaten und Systemtheorie)

Leiter

PROF. DR. MARTIN GROHE

Tel.: (030) 2093 3078

E-Mail: grohe@informatik.hu-berlin.de

DR. KARSTEN SCHMIDT (GASTDOZENT ZUR VERTRETUNG DES FACHS BIS AUGUST 2003)

Tel.: (030) 2093 3083

E-Mail: kschmidt@informatik.hu-berlin.de

Sekretariat

BIRGIT EISENMANN

Tel.: (030) 2093 3080

Fax: (030) 2093 3081

E-Mail: eisenman@informatik.hu-berlin.de

Wissenschaftliche Mitarbeiter

DR. STEPHAN KREUTZER

DR. LOUCHKA POPOVA-ZEUGMANN

DR. NICOLE SCHWEIKARDT

Technikerin

DIPL.-ING. PETRA KÄMPFER

Tutoren

HIEP HAN (bis 30.09.2003)

DZIFA AMETOWOBLA (ab 01.10.2003)

FLORIAN KÜHNERT

MARK THURLEY

Logik spielt in der Informatik eine grundlegende Rolle, vergleichbar mit der Rolle der Analysis in der Physik und den traditionellen Ingenieursdisziplinen. Anwendungen der Logik findet man in so unterschiedlichen Bereichen der Informatik wie Rechnerarchitektur, Softwaretechnik, Programmiersprachen, Datenbanken, künstliche Intelligenz, Komplexitäts- und Berechenbarkeitstheorie.

Im Zentrum der Lehre und Forschung am Lehrstuhl für Logik in der Informatik steht das Grenzgebiet zwischen Logik, Algorithmik und Komplexitätstheorie. Wichtige algorithmische Probleme aus Anwendungen wie Datenbanksystemen oder künstlicher Intelligenz lassen sich auf natürliche Weise durch „logische“ Probleme modellieren. Effiziente

Algorithmen für die Probleme und ihre Komplexität lassen sich oft am besten anhand dieser abstrakten logischen Modelle untersuchen. Ebenso lassen sich grundlegenden Fragen der Komplexitätstheorie oft mittels logischer Methoden untersuchen, ein nützliches Paradigma besagt, dass die algorithmische Komplexität eines Problems sehr eng mit der sprachlich-logischen Komplexität zusammenhängt – sprich: Probleme, die schwer lösbar sind auch schwer beschreibbar und umgekehrt.

Dass diese abstrakten theoretischen Untersuchungen dann auch ganz konkrete Anwendungen haben, zeigen etwa Arbeiten der Arbeitsgruppe im Bereich der Datenbanksysteme.

Bis August 2003 wurde der Lehrstuhl von Dr. Karsten Schmidt vertreten. Die Neubesetzung durch Prof. Grohe brachte eine Umorientierung der Lehr- und Forschungsinhalte mit sich. Die neue Arbeitsgruppe befindet sich zurzeit noch im Aufbau.

Lehre

Vorlesungen

- Theoretische Informatik 1 (M. Grohe, WS 2003/2004)
- Lineare Optimierung (L. Popova-Zeugmann, SS 2003)
- Computergestützte Verifikation (K. SCHMIDT, SS 2003)

Seminare und Proseminare

- Logik und Komplexitätstheorie (M. Grohe, WS 2003/2004)
- Funktionale Programmierung (M. Grohe, WS 2003/2004)
- Formale Sprachen (L. Popova-Zeugmann, SS 2003)
- Programmverifikation (K. SCHMIDT, SS 2003)
- Prozessalgebra (K. SCHMIDT, SS 2003)
- Petrinetze (S. ROCH, SS 2003)
- Temporale Logik (S. ROCH, SS 2003)

Übungen

- Theoretische Informatik 1 (S. Kreutzer, L. Popova-Zeugmann, N. Schweikardt, WS 2003/2004)

Forschung

Projekt: Die Struktur Parametrischer Komplexitätsklassen

Ansprechpartner: PROF. DR. MARTIN GROHE

Zusammenarbeit: Albert-Ludwigs-Universität Freiburg, PROF. DR. JÖRG FLUM

Forschungsförderung: Deutsche Forschungsgemeinschaft

Die Komplexitätstheorie macht Aussagen über die zur Lösung von algorithmischen Problemen erforderlichen Ressourcen, wie etwa Rechenzeit. Dabei wird die Komplexität eines Problems üblicherweise als Funktion der Eingabegröße gemessen. Dieses einfache Modell führt zu einer klaren Einteilung in Klassen von leicht und schwer lösbaren algorithmischen Problemen, hat aber den Nachteil, dass gewisse feinere Strukturen der Eingabe nicht berücksichtigt und unter Umständen Probleme als „schwer“ klassifiziert werden, obwohl nur gewisse für die Praxis irrelevante Fälle schwer lösbar sind. Häufig besteht die Eingabe eines Problems aus mehreren Teilen. Als Beispiel betrachte man das Problem, eine Datenbankanfrage auszuwerten. Die Eingabe besteht hier aus der Anfrage und der Datenbank. Normalerweise ist die Datenbank um ein Vielfaches größer als die Anfrage. Die

parametrische Komplexitätstheorie berücksichtigt dies und ermöglicht eine verfeinerte Komplexitätsanalyse.

Ziel des Projektes ist es, ein klareres Bild der noch sehr unübersichtlichen Struktur der parametrischen Komplexitätsklassen und ihres Verhältnisses zu klassischen Klassen zu erlangen. Eine systematische Untersuchung der „Parameterabhängigkeit“ von Problemen soll eine realistischere Einschätzung ihrer Komplexität ermöglichen, als dies bisher möglich ist.

Projekt: Die Komplexität von Constraint-Satisfaction Problemen

Ansprechpartner: PROF. DR. MARTIN GROHE

Beteiligte Mitarbeiter: DR. STEPHAN KREUTZER, DR. NICOLE SCHWEIKARDT

Constraint-Satisfaction-Probleme (CSP) bilden eine natürliche Klasse von algorithmischen Problemen, die wichtige Anwendungen in ganz verschiedenen Bereichen wie künstliche Intelligenz, Datenbanken, automatische Verifikation und statistische Physik haben. Prominentestes Beispiel eines CSP, das auch in diesem Projekt eine wichtige Rolle spielen soll, ist das aussagenlogische Erfüllbarkeitsproblem.

Es ist seit langem bekannt, dass CSP im Allgemeinen NP-vollständig und damit, zumindest theoretisch, nicht effizient lösbar sind. In der Praxis hat es in den letzten Jahren jedoch enorme Fortschritte bei der Lösung insbesondere des aussagenlogischen Erfüllbarkeitsproblems gegeben. Inzwischen werden in industriellen Anwendungen Instanzen mit mehr als 10.000 Variablen routinemäßig gelöst.

Es liegt hier also eine deutliche Diskrepanz zwischen den theoretischen „worst-case“ Vorhersagen und der Praxis vor. Als Grund für diese Diskrepanz wird oft genannt, dass in der Praxis auftretende Instanzen „strukturiert“ sind. Allerdings ist es völlig unklar, welche strukturellen Eigenschaften hier relevant sind und wie diese von den üblicherweise eingesetzten Algorithmen ausgenutzt werden. Diese Fragen sollen im Mittelpunkt des Projekts stehen. Neben CSP und SAT als zentralem Beispiel soll hier auch eine Reihe verwandter Probleme, etwa Zählprobleme, untersucht werden.

Projekt: Zeitabhängige Systeme

Ansprechpartner: Dr. Louchka Popova-Zeugmann

Zusammenarbeit: HUB, IfI, LFG Rechnerorganisation und Kommunikation: DIPL.-INF. JAN RICHLING und TU Berlin, WE0432, Arbeitsgruppe Kommunikations- und Betriebssysteme: Dr.-ING. MATTHIAS WERNER

Das Ziel dieses Projekts ist die Entwicklung einer gemischt ereignis- und zeitgesteuerten Architektur (Message Scheduled System - MSS) für eingebettete Systeme und deren Verifizierung. Um das Verhalten solcher Systeme zu analysieren, sind adäquate Mittel zur Modellierung und Analyse notwendig.

Im Rahmen dieses Projekts wurden bereits existierende Werkzeuge an das konkrete Projekt adaptiert sowie neue entwickelt. Zur Modellierung und Analyse in speziellen endlichen Fällen wurden Timed Petri Netze mit Prioritäten eingesetzt. Eine Beschreibung mit prioritätenberücksichtigenden Time Petri Netzen als natürliche Modellierungsvorgehensweise wurde auch betrachtet.

Im Allgemeinen aber ist die Größe des Problems beliebig und deshalb durch Parameter definiert. Damit ist es nicht möglich, die Nichterreichbarkeit von unerwünschten Zuständen mittels Tools zu verifizieren, die darauf basieren, den Erreichbarkeitsraum eines Netzes zu berechnen. Um andere Tools zu erstellen, muss natürlich als Grundlage ein Algorithmus

entwickelt werden. Uns ist gelungen eine Methode zu entwerfen, die erstmals eine Zustandsgleichung für zeitabhängige Petri Netze aufstellt, die auch die zeitlichen Abhängigkeiten des Netzes berücksichtigt. Die ersten Resultate bezogen sich auf Timed Petri Netze ohne Prioritäten. Die Nichterreichbarkeit von Zuständen wird nun zurückgeführt auf die Lösungsmenge eines Ungleichungssystems. Die Hinzunahme von Prioritäten erhöht die Effektivität der Modellierung aber auch die Vielschichtigkeit des Analyseproblems. Die in der Vergangenheit entwickelte Methode zur Verifizierung von Nichterreichbarkeit gewisser Zustände in prioritätenlosen Netzen wurde verallgemeinert auf prioritätsbehaftete Timed Petri Netze. Die erweiterte Beweistechnik wurde nicht nur auf konkrete endliche Netze angewandt, sondern auch auf eine Klasse beliebig großer Timed Petri Netze, die durch Parameter beschrieben wird und Teilaspekte der Architektur MSS allgemein beschreibt.

Diese Beweistechnik soll auf die gesamte Architektur MSS angewandt werden und damit die Verifikation von MSS auf rein formalem Weg gezeigt werden.

Veröffentlichungen

M. BENEDIKT, M. GROHE, L. LIBKIN UND L. SEGOUFIN: *Reachability and connectivity queries in constraint databases*. Journal of Computer System Sciences 66: 169-206, 2003 (PODS'00 Special Issue).

P. BUNEMAN, M. GROHE, CH. KOCH: *Path Queries on Compressed XML*. In Proceedings of the 29th Conference on Very Large Data Bases (VLDB'03), S. 141-152, 2003.

Y. CHEN, M. GROHE, J. FLUM: *Bounded Nondeterminism and Alternation in Parameterized Complexity Theory*. In Proceedings of the 18th IEEE Conference on Computational Complexity (CCC'03), S. 13-29, 2003.

J. FLUM, M. GROHE: *Describing Parameterized Complexity Classes*. Information and Computation, 187: 291-319, 2003.

M. FRICK, M. GROHE UND CH. KOCH: *Query Evaluation on Compressed Trees*. In Proceedings of the 18th IEEE Symposium on Logic in Computer Science (LICS'03), S. 188-197, 2003.

M. GROHE: *The complexity of homomorphism and constraint satisfaction problems seen from the other side*. In Proceedings of the 44th IEEE Symposium on Foundations of Computer Science (FOCS'03), S. 552-561, 2003.

M. GROHE, N. SCHWEIKARDT: *Comparing the succinctness of monadic query languages over finite trees*. In Proceedings of the 17th International Workshop on Computer Science Logic (CSL'03), Lecture Notes in Computer Science 2803, S. 226-240, Springer-Verlag, 2003.

D. BERWANGER, E. GRÄDEL, S. KREUTZER: *Once upon a time in the west*. Determinacy, complexity and definability of path games. Proceedings of the 10th International Conference on Logic for Programming and Automated Reasoning (LPAR), vol. 2850 of LNCS, pp. 226-240, Springer-Verlag, 2003.

E. GRÄDEL, S. KREUTZER: *Will Deflation Lead to Depletion? On Non-Monotone Fixed Point Inductions*. IEEE Symp. on Logic in Computer Science (LICS), 2003.

S. KREUTZER: *Pure and Applied Fixed-Point Logics*, (in deutsch), in "Ausgezeichnete Informatik Dissertationen 2003" (D. Wagner et al., ed.), vol. D-3 of Lecture Notes in Informatics - Dissertations, pp. 59-68, Gesellschaft für Informatik (GI), 2003.

L. POPOVA-ZEUGMANN, M. WERNER, J. RICHLING: *Using State Equation to Prove Non-Reachability in Timed Petrinets*, Fundamenta Informaticae (FI), 187-202, IOS-Press, 2003.

M. WERNER, L. POPOVA-ZEUGMANN, J. RICHLING: *Non-Reachability in Priority Duration Petrinets*, Workshop Concurrency, Specification & Programming, Czarna 25-27 September 2003, Tagungsband, Warsaw University 2003, pp. 561-574.

K. SCHMIDT: *Using Petri net Invariants for State Space Construction*. Proc. Int. Conf. Tools and Algorithms for the Construction and Analysis of Systems (TACAS), LNCS 2619, pp. 473-488, 2003.

K. SCHMIDT: *Distributed Verification with LoLA*. Fundamenta Informaticae 54 (2-3), 2003, pp. 325-335.

N. SCHWEIKARDT: *Die Ausdrucksstärke der Logik erster Stufe mit eingebauten Prädikaten*. In "Ausgezeichnete Informatik Dissertationen 2003" (D. Wagner et al., ed.), Vol. D-3 of Lecture Notes in Informatics - Dissertations, pp. 153-162, Gesellschaft für Informatik (GI), 2003.

Vorträge

M. GROHE: *Logic, Graph Theory, and Complexity*. Eingeladener Vortrag beim New Zealand Mathematics Meeting, New Plymouth, Januar 2003.

M. GROHE: *The Succinctness of Monadic Fixed-Point Logics on Trees*. Eingeladener Vortrag beim Workshop on Fixed-Point Logics in Computer Science, Warschau, April 2003.

M. GROHE: *The Succinctness of Monadic Logics over Words and Trees*. Invited Plenary Lecture beim Logic Colloquium 2003, Helsinki, August 2003.

M. GROHE: *Die Komplexität von Homomorphieproblemen*. AFM-Seminar über algorithmische Modelltheorie, Marburg, November 2003.

M. GROHE: *Die Komplexität von Homomorphieproblemen*. Berliner Logik Kolloquium, Berlin, Dezember 2003.

S. KREUTZER: *Once Upon A Time in the West*. Logic and Semantics Seminar, University of Cambridge, Mai 2003.

S. KREUTZER: *Pure and Applied Fixed-Point Logics*. GI-Dissertationspreis Kolloquium, Mainz, Juni 2003.

S. KREUTZER: *Once Upon A Time in the West*. Second Annual Meeting of the Research and Training Network „Games and Automata for Synthesis and Validation“, Wien, September 2003.

L. POPOVA-ZEUGMANN: *Non-Reachability in Priority Duration Petrinets*, Workshop Concurrency, Specification & Programming 2003, Czarna 25-27 September 2003.

N. SCHWEIKARDT: *The Crane Beach Conjecture*. Finite Model Theory Workshop FMT'03, Bedlewo, Polen, April 2003.

N. SCHWEIKARDT: *Das Query Containment Problem für Monadisches Datalog auf Bäumen*. Gastvortrag an der TU Wien, Juni 2003.

N. SCHWEIKARDT: *Die Ausdrucksstärke der Logik erster Stufe mit eingebauten Prädikaten*. GI-Dissertationspreis Kolloquium, Mainz, Juni 2003.

N. SCHWEIKARDT: *Das Query Containment Problem für Monadisches Datalog auf Bäumen*. Gastvortrag an der Philipps-Universität Marburg, Juli 2003.

N. SCHWEIKARDT: *Comparing the succinctness of monadic query languages over finite trees*. Conference of the European Association for Computer Science Logic (CSL'03), Wien, August 2003.

N. SCHWEIKARDT: *Die Succinctness der Logik erster Stufe auf Worten*. AFM-Seminar über algorithmische Modelltheorie, November 2003.

Sonstige Aktivitäten

Prof. Dr. Martin Grohe

- Mitherausgeber der Zeitschriften Journal of Symbolic Logic und Journal of Discrete Algorithms
- Mitglied des Programmkomitees des Kalmar Workshops on Logic and Computer Science und des 23rd ACM SIGMOD-SIGACT- SIGART Symposium on Principles of Database Systems
- Mitglied des Councils der Association for Symbolic Logic
- Diverse Gutachtertätigkeiten

Dr. Stephan Kreutzer

- Gutachtertätigkeiten für verschiedene Fachzeitschriften und Konferenzen
- Dissertationspreis 2002 der Gesellschaft für Informatik e.V.

Dr. Louchka Popova-Zeugmann

- Mitglied des Institutsrats
- Mitglied des Prüfungsausschusses
- Gutachten für CS&P 2003

Dipl.-Inf. Stephan Roch

- Gutachten, u.a. für Petri Net Newsletter, ICATPN 2004, Fundamenta Informaticae, International Conference on Software Engineering and Formal Methods SEFM 2003

Dr. Karsten Schmidt

- Editor für das Springer Journal of Universal Computer Science (J.UCS)
- Mitglied im örtlichen Wahlvorstand des Instituts für Informatik
- Mitglied der Leitung der FG "Petrietze und verwandte Systemmodelle" der GI
- Gutachten, u.a. für Fundamenta Informaticae, IJFCS Konferenzen CAV 2003

Dr. Nicole Schweikardt

- Mitglied der Habilitationskommission für Dr. Deryk Osthus
- Gutachtertätigkeiten für verschiedene Fachzeitschriften und Konferenzen
- Dissertationspreis 2002 der Gesellschaft für Informatik e.V.
- Forschungsförderpreis der Freunde der Universität Mainz

Lehr- und Forschungseinheit

Algorithmen und Komplexität

http://www.informatik.hu-berlin.de/Forschung_Lehre/algorithmen/

Leiter: PROF. DR. HANS JÜRGEN PRÖMEL

Tel.: (030) 2093 3188

E-Mail: proemel@informatik.hu-berlin.de

zusammen mit

Zufällige Strukturen und Algorithmen

http://www.informatik.hu-berlin.de/Forschung_Lehre/algorithmen/

Leiter: PROF. DR. ANUSCH TARAZ

Tel.: (030) 2093 3197

E-Mail: taraz@informatik.hu-berlin.de

Sekretariat

EVA SANDIG

Tel.: (030) 2093 3190

Fax: (030) 2093 3191

E-Mail: sandig@informatik.hu-berlin.de

Wissenschaftliche Mitarbeiter und Mitarbeiterinnen

DIPL.-INF. MICHAEL BEHRISCH

DR. AMIN COJA-OGHLAN

DR. CLEMENS GRÖPL

DR. STEFAN HOUGARDY

DR. MIHYUN KANG

DIPL.-INF. STEFAN KIRCHNER

DR. TILL NIERHOFF

DR. DERYK OSTHUS

DIPL.-MATH. MARTIN THIMM

Promotionsstipendiaten / Promotionsstudenten

DIPL.-INF. MANUEL BODIRSKY

DIPL.-INF. ANDRZEJ DUDEK

DIPL.-INF. DANIEL ROLF

DIPL.-INF. DIRK SCHLATTER

SHI LINGSHENG, M. S.

Techniker

DIPL.-MATH. RALF OELSCHLÄGEL

Tutoren und Tutorinnen

JULIA BÖTTCHER

DORATHA DRAKE

MATTHIAS FÜSSEL

SVEN HANKE

STEFAN KIRCHNER
IVO KÖTHING
LARS KUHTZ
PETER LISKE
MIKE LÖFFLER
NICOLAS ROCCA
CHRISTOPH WAGNER
MARIANO ZELKE
VALENTIN ZIEGLER

Zentrale Lehr- und Forschungsgegenstände der Arbeitsgruppe sind der Entwurf und die Analyse effizienter Algorithmen. In der Komplexitätstheorie werden Probleme hinsichtlich verschiedener Komplexitätsmaße, wie Laufzeit oder Speicherplatz, klassifiziert. Algorithmen zur Lösung von Problemen, die sich durch eine besonders kurze Laufzeit bzw. einen besonders geringen Speicherplatz auszeichnen, werden effizient genannt. Eine wichtige Grundlage für die Entwicklung von effizienten Algorithmen ist ein genaues Verständnis der den Problemen zugrunde liegenden Strukturen. Diese können in vielen Fällen als Graphen und Hypergraphen modelliert werden. Ein Schwerpunkt der Forschung liegt in der Untersuchung zufälliger Graphen und Hypergraphen und die Anwendung der dabei erzielten probabilistischen und asymptotischen Resultate beim Entwurf und der Analyse von Graphenalgorithmen. Zu den wichtigsten effizienten Verfahren gehören randomisierte Algorithmen, die vom Zufall Gebrauch machen, und approximative Algorithmen, die gute Näherungslösungen liefern.

Eine Umsetzung der theoretischen Erkenntnisse erfolgt im Rahmen von anwendungsorientierter Forschung, beispielsweise auf dem Gebiet der Bioinformatik (Algorithmen im Drug Design) in Zusammenarbeit mit der Charité und mit dem DFG-Forschungszentrum für Mathematik in Schlüsseltechnologien.

Lehre

Die Grundlagen der Gebiete Algorithmen und Komplexität werden im Grundstudium in den Vorlesungen *Theoretische Informatik 2* (Schwerpunkte: Algorithmen und Datenstrukturen) und *Theoretische Informatik 3* (Schwerpunkt: Komplexitätstheorie) vermittelt.

Der zentrale Forschungsgegenstand des Lehrstuhls, Graphentheorie und -algorithmen, wird in der Hauptstudiumsvorlesung *Graphen und Algorithmen* behandelt. Diese Lehrveranstaltung wird ergänzt durch Hauptstudiumsvorlesungen zu den Themen *Randomisierte Algorithmen und Probabilistische Analyse* sowie *Algorithmen in der Bioinformatik*. In den Seminaren werden spezielle Klassen von Algorithmen wie approximative, randomisierte und Online- Algorithmen untersucht. Abgerundet wird das Lehrangebot durch verschiedene Seminare und Workshops im Rahmen von Kooperationen, in denen aktuelle Themen aus der Forschung behandelt werden (siehe auch unter Tagungen, Workshops und wissenschaftliche Kooperationen). Das *Oberseminar Theoretische Informatik* dient dem Austausch mit den anderen Lehrstühlen des Schwerpunktes *Algorithmen und Logik* (Logik in der Informatik, Komplexität und Kryptografie).

Veranstaltungen im Grundstudium

- Theoretische Informatik 3 (T. NIERHOFF, D. OSTHUS, M. FÜSSEL, SS 03)
- Theoretische Informatik 2 (M. FÜSSEL, P. LISKE, WS 03/04)

Kernveranstaltungen (Halbkurse)

- Graphen und Algorithmen, Teil 2 (S. HOUGARDY, A. TARAZ, SS 03)
- Graphen und Algorithmen, Teil 1 (A. COJA-OGHLAN, WS 03/04)

Spezialveranstaltungen

- Algorithmen in der Bioinformatik (S. HOUGARDY, WS 03/04)
- Randomisierte Algorithmen und Probabilistische Analyse (A. TARAZ, WS 03/04)
- Kombinatorik und ihre Anwendungen (D. OSTHUS, WS 03/04)

Seminare

- Approximationsalgorithmen (T. NIERHOFF, SS 03)
- Markov-Ketten und Monte-Carlo Algorithmen (A. TARAZ, SS 03)
- Algorithmen und Komplexität (T. NIERHOFF, A. COJA-OGHLAN, SS 03, WS 03/04)
- Semidefinite Programmierung (A. COJA-OGHLAN, SS 03)
- The Strange Logic of Random Graphs (M. KANG, M. BODIRSKY, SS 03)
- Discrete Random Walks: Theory and Applications (M. KANG, WS 03/04)

Forschung

Projekt: „Analyse und Modellierung komplexer Netzwerke“ im DFG-Forschungszentrum für Mathematik in Schlüsseltechnologien

Ansprechpartner: DR. STEFAN HOUGARDY, PROF. DR. HANS JÜRGEN PRÖMEL, PROF. DR. ANUSCH TARAZ

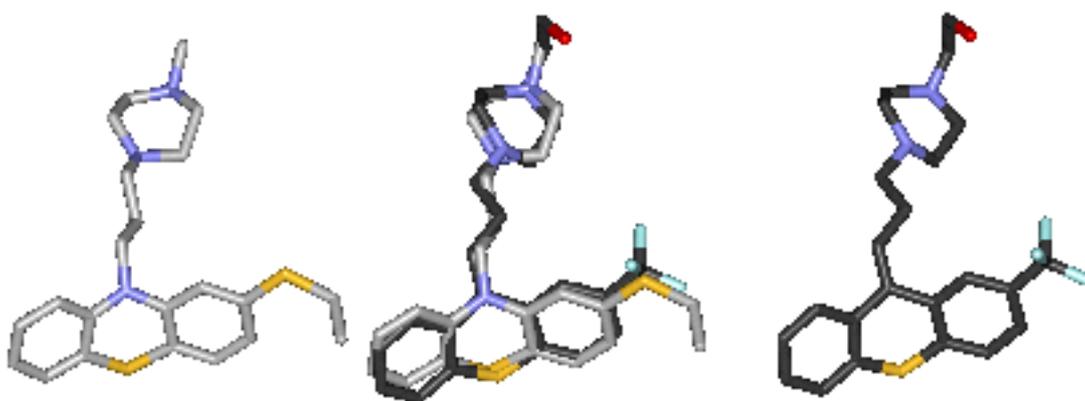
Beteiligte Mitarbeiter: DIPL.-INF. MICHAEL BEHRISCH, DIPL.-INF. STEFAN KIRCHNER, DR. TILL NIERHOFF, DIPL.-MATH. MARTIN THIMM

Zusammenarbeit: Freie Universität Berlin, Technische Universität Berlin, Konrad-Zuse-Zentrum für Informationstechnik Berlin, [Weierstraß-Institut für Angewandte Analysis und Stochastik](#)

Forschungsförderung: Deutsche Forschungsgemeinschaft

Ein Schwerpunkt unserer derzeitigen Forschung ist die möglichst exakte Berechnung der Ähnlichkeitsnetzwerke kleiner Moleküle. Dazu haben wir Algorithmen entwickelt, die sehr schnell die geometrische Ähnlichkeit dieser Moleküle nahezu optimal berechnen können. Diese Algorithmen kommen bei der Vorhersage von Wirkungen und Nebenwirkungen von Medikamenten zum Einsatz.

Viele Netzwerke aus den Lebenswissenschaften und der Informationstechnologie besitzen eine Reihe von sogenannten "kleine-Welt-Eigenschaften", d.h. sie haben lokale Cluster und gleichzeitig, trotz einer insgesamt nur geringen Anzahl von Verbindungen, einen kleinen Durchmesser. Ein weiteres Ziel dieses Projekts ist die Entwicklung realistischer stochastischer Modelle, die diese Eigenschaften gut widerspiegeln und dadurch Ansatzpunkte für schnelle randomisierte Datenbank-Verfahren bieten.



Die optimale "Überlagerung (mitte) des Antihistaminikums Thiethylperazin (links) mit dem Antipsychotikum Flupentixol (rechts)

Projekt: Forschergruppe „Algorithmen, Struktur, Zufall“

Ansprechpartner: PROF. DR. HANS JÜRGEN PRÖMEL

Beteiligte Mitarbeiter: DR. AMIN COJA-OGHLAN, DR. DERYK OSTHUS, DR. TILL NIERHOFF, PROF. DR. ANUSCH TARAZ, DR. STEFAN HOUGARDY, DR. CLEMENS GRÖPL

Zusammenarbeit: Technische Universität Berlin, Konrad-Zuse-Zentrum für Informationstechnik Berlin

Forschungsförderung: Deutsche Forschungsgemeinschaft

Der Entwurf und die Analyse von Algorithmen sind eng verknüpft mit Einsichten in die Struktur der Objekte, die die Algorithmen als Eingabe erhalten. Das zentrale Thema des Forschungsvorhabens ist es, diese Verknüpfung im Hinblick auf den Einfluss des Zufalls zu untersuchen - wie wirkt sich die Hinzunahme von Zufall auf algorithmische und strukturelle Fragestellungen in der Diskreten Mathematik aus?

Der Zufall ist hierbei gleichermaßen Forschungsobjekt wie Untersuchungsmethode. Im Hinblick auf Strukturkenntnisse wird einerseits nach Eigenschaften gesucht, die zufällige Objekte mit hoher Wahrscheinlichkeit besitzen, andererseits werden Objekte durch Benutzung des Zufalls charakterisiert. Und auch unter algorithmischen Aspekten setzt sich diese Dualität fort: Einerseits wird untersucht, wie sich Algorithmen auf zufälligen Eingaben verhalten, andererseits werden Verfahren analysiert, deren Entscheidungen zufällige Komponenten aufweisen.

Auf der algorithmischen Seite gilt das Interesse kombinatorischen Optimierungsproblemen. Dadurch treten ganz natürlich Polytope, Graphen und partielle Ordnungen als Forschungsobjekte von Strukturuntersuchungen in den Mittelpunkt. Die Forschergruppe bündelt dabei die individuellen Erfahrungen und Kompetenzen, die die beteiligten Arbeitsgruppen in den letzten Jahren mit unterschiedlichen Ansätzen in dem Spannungsfeld von Algorithmen, Struktur und Zufall gewonnen haben.

Projekt: Approximative, randomisierte und probabilistische Algorithmen für kombinatorische Optimierungsprobleme

Ansprechpartner: PROF. DR. HANS JÜRGEN PRÖMEL

Beteiligte Mitarbeiter: DR. CLEMENS GRÖPL, DR. STEFAN HOUGARDY, DR. TILL NIERHOFF

Forschungsförderung: Deutsche Forschungsgemeinschaft

Viele in der Praxis auftretende und theoretisch interessante Probleme sind NP-schwer, können also nach heutigem Kenntnisstand in polynomieller Zeit nicht exakt gelöst werden. Daher ist man an Algorithmen interessiert, die in polynomieller Zeit brauchbare Näherungslösungen finden.

In diesem Forschungsprojekt sollen zum einen möglichst gute Approximationsalgorithmen gefunden werden oder aber zum anderen soll bestimmt werden, welche Schranke die Approximationsgüte in polynomieller Zeit nicht überschreiten kann. Wir entwickeln und analysieren auch Approximationsalgorithmen, die auf den meisten Eingaben nachweisbar effizient und von hoher Güte sind. An einer solchen average-case Analyse ist man besonders dann interessiert, wenn ein Problem zwar NP-schwer, in der Regel aber gut lösbar ist, oder wenn ein exakter polynomieller Algorithmus zwar existiert, für die Praxis aber zu langsam ist. In diesem Zusammenhang widmen wir uns auch der effizienten Erzeugung von zufälligen Instanzen, die bestimmten Bedingungen genügen.

Projekt: Europäisches Graduiertenkolleg „Combinatorics, Geometry, and Computation“

Ansprechpartner: PROF. DR. HANS JÜRGEN PRÖMEL

Beteiligte Stipendiaten: DIPL.-INF. MANUEL BODIRSKY, DIPL.-INF. ANDRZEJ DUDEK (MARIE CURIE-STIPENDIAT), DIPL.-INF. DIRK SCHLATTER, SHI LINGSHENG, M. S., DR. MIHYUN KANG

Forschungsförderung: Deutsche Forschungsgemeinschaft

Das Graduiertenkolleg wird gemeinsam mit Partnern aus dem europäischen Ausland gestaltet. Es wird auf der Berliner Seite getragen von der Freien Universität Berlin, der Humboldt-Universität zu Berlin, der Technischen Universität Berlin und dem Konrad-Zuse-Zentrum für Informationstechnik Berlin. Partner sind die Arbeitsgruppen der Universitäten in Budapest, Eindhoven, Louvain-La-Neuve, Oxford, Prag, Poznań und Zürich. Die Forschungsschwerpunkte liegen auf den Gebieten Kombinatorik, Geometrie und Algorithmen. Das Graduiertenkolleg ist ein "Marie Curie Training Site": junge Doktoranden aus dem europäischen Ausland können ein Stipendium aus EU-Mitteln für einen Forschungsaufenthalt in Berlin für drei bis zwölf Monate erhalten.

Veröffentlichungen

Publikationen

M. BODIRSKY, J. NEŠETŘIL: *Constraint Satisfaction with Countable Homogeneous Templates*. Proceedings of Computer Science Logic (CSL03), 44-56, 2003.

M. BODIRSKY, C. GRÖPL, M. KANG: *Generating Labeled Planar Graphs Uniformly at Random*. Proc. 30th Colloquium on Automata, Languages, and Programming (ICALP 2003), Springer LNCS 2719, 1095-1107.

A. COJA-OGHLAN, M.-O. STEHR: *Revisiting the algebra of Petri net processes under the collective token philosophy*. Fundamenta Informaticae 54(2003) 151-164.

A. COJA-OGHLAN: *Finding large independent sets in polynomial expected time*. Proc. 20th International Symposium on Theoretical Aspects of Computer Science (STACS 2003), Springer LNCS 2607, 511-522.

- A. COJA-OGHLAN: *The Lovasz number of random graphs*. Proc. Approximation, randomization, and combinatorial optimization (APPROX+RANDOM 2003), Springer LNCS 2764, 228-239.
- A. COJA-OGHLAN, A. GOERDT, A. LANKA, F. SCHÄDLICH: *Certifying unsatisfiability of random $2k$ -SAT formulas using approximation techniques*. Proc. 14th International Symposium on Fundamentals of Computation Theory (FCT 2003), Springer LNCS 2751, 15-26.
- A. COJA-OGHLAN, S.O. KRUMKE, T. NIERHOFF: *A heuristic for the stacker crane problem on trees which is almost surely exact*. Proc. 14th Annual International Symposium on Algorithms and Computation (ISAAC 2003), Springer LNCS 2906, 605-614.
- A. COJA-OGHLAN, C. MOORE, V. SANWALANI: *MAX k-CUT and approximating the chromatic number of random graphs*. Proc. 30th Colloquium on Automata, Languages, and Programming (ICALP 2003), Springer LNCS 2719, 201-211.
- A. COJA-OGHLAN, A. TARAZ: *Coloring random graphs in expected polynomial time*. Proc. 20th International Symposium on Theoretical Aspects of Computer Science (STACS 2003), Springer LNCS 2607, 487-498.
- D.E. DRAKE, S. HOUGARDY: *Improved Linear Time Approximation Algorithms for Weighted Matchings*. Proc. Approximation, randomization, and combinatorial optimization (APPROX+RANDOM 2003), Springer LNCS 2764, 14-23.
- D.E. DRAKE, S. HOUGARDY: *Linear Time Local Improvements for Weighted Matchings in Graphs*. Proc. 2nd International Workshop on Experimental and Efficient Algorithms (WEA 2003), LNCS 2647, Springer 2003, 107-119.
- D.E. DRAKE, S. HOUGARDY: *A Simple Approximation Algorithm for the Weighted Matching Problem*. Information Processing Letters 85 (2003) 211-213.
- C. FRÖMME, C. GILLE, A. GOEDE, C. GRÖPL, S. HOUGARDY, T. NIERHOFF, R. PREISSNER, M. THIMM: *Accelerating screening of 3D protein data with a graph theoretical approach*. Bioinformatics 19 (2003), 2442-2447.
- M. KANG: *First hitting times of simple random walks on graphs with congestion points*. Int. J. Math. Math. Sci. 2003:30 (2003), 1911-1922.
- D. KÜHN, D. OSTHUS: *Partitions of graphs with high minimum degree or connectivity*. J. Combinatorial Theory Series B 88 (2003), 29-43.
- D. KÜHN, D. OSTHUS: *Minors in graphs of large girth*. Random Structures and Algorithms 22 (2003), 213-225.
- D. OSTHUS, H.J. PRÖMEL, A. TARAZ: *For which densities are random triangle-free graphs almost surely bipartite?* Combinatorica 23 (2003), 105-150.
- D. OSTHUS, H.J. PRÖMEL, A. TARAZ: *On random planar graphs, the number of planar graphs and their triangulations*. J. Combinatorial Theory Series B 88 (2003), 119-143.

Preprints

- M. BODIRSKY, D. DUCHIER, J. NIEHREN, S. MIELE: *A New Algorithm for Normal Dominance Constraints*. Erscheint in: ACM-SIAM Symposium on Discrete Algorithms (SODA04), New Orleans.

- M. BODIRSKY, C. GRÖPL, M. KANG: *Generating Labeled Planar Graphs Uniformly at Random*. Eingereicht, 2003.
- M. BODIRSKY, C. GRÖPL, M. KANG: *Generating unlabeled 2-connected planar graphs uniformly at random*. Eingereicht, 2003.
- M. BODIRSKY, M. KANG: *Generating Outerplanar Graphs Uniformly at Random*. Presented at the 1st Workshop on Algorithms for Listing, Counting, and Enumeration (ALICE03), Baltimore. Erscheint in: Journal of Combinatorics, Probability and Computing.
- M. BODIRSKY, T. GÄRTNER, T. VON OERTZEN, J. SCHWINGHAMMER: *Efficiently Computing the Density of Regular Languages*. Erscheint in: Proceedings of the Latin American Informatics (LATIN'04).
- B. BOLLOBÁS, Y. KOHAYAKAWA, V. RÖDL, A. TARAZ: *Essentially infinite colourings*. Humboldt-Universität zu Berlin, 2003.
- A. COJA-OGHLAN: *Solving NP-hard semirandom graph problems in polynomial expected time*. Humboldt-Universität zu Berlin, 2003.
- A. COJA-OGHLAN: *Improved heuristics for coloring semirandom graphs*. Humboldt-Universität zu Berlin, 2003.
- A. COJA-OGHLAN, S.O. KRUMKE, T. NIERHOFF: *A Hard Dial-a-Ride Problem that is Easy on Average*. Eingereicht, 2003.
- A. COJA-OGHLAN, A. TARAZ: *Exact and approximative algorithms for coloring $G(n,p)$* . Humboldt-Universität zu Berlin, 2003.
- A. COJA-OGHLAN, J. MICHALICEK: *Non-commutative random variables in connection with group actions on C^* -algebras*. Hamburger Beiträge zur Mathematik 167, 2003.
- A. COJA-OGHLAN, J. MICHALICEK: *Limit theorems for non-commutative random variables*. Hamburger Beiträge zur Mathematik 172, 2003.
- D.E. DRAKE, S. HOUGARDY: *On approximation algorithms for the terminal Steiner tree problem*. Erscheint in: Information Processing Letters.
- D.E. DRAKE, S. HOUGARDY: *A linear time approximation algorithm for weighted matchings*. Humboldt-Universität zu Berlin, 2003.
- A. GOEDE, S. HOUGARDY, R. PREISSNER, M. THIMM: *Comparison of 2D similarity and 3D superposition. Application to searching a conformational drug database*. Humboldt-Universität zu Berlin, 2003.
- J. GRAMM, T. NIERHOFF, R. SHARAN, T. TANTAU: *On the Complexity of Haplotyping via Perfect Phylogeny*. Erscheint in: Second RECOMB Satellite Workshop on Computational Methods for SNPs and Haplotypes.
- C.T. HOÀNG, S. HOUGARDY, F. MAFFRAY, N.V.R. MAHADEV: *On simplicial and co-simplicial vertices in graphs*. Erscheint in: Discrete Applied Mathematics.
- S. HOUGARDY: *Classes of perfect graphs*. Erscheint in: Berge memorial volume of Discrete Mathematics.
- S. HOUGARDY: *On a conjecture of Hoàng and Tu concerning perfectly orderable graphs*. Eingereicht, 2003
- S. HOUGARDY, S. KIRCHNER: *Lower Bounds for the Relative Greedy Algorithm for Approximating Steiner Trees*. Eingereicht, 2003.

S. HOUGARDY, I. KÖTHNIG: *Enumeration of best possible low degree expanders*. Eingereicht, 2003.

M. KANG: *Random walks on finite graphs with congestion points*. Erscheint in: Applied Mathematics and Computation.

M. KANG: *Efficiency test of pseudorandom number generators using random walks*. Eingereicht, 2003.

M. KANG, Y. KOH, T. ŁUCZAK, S. REE: *The connectivity threshold for the min-degree random graph process*. Eingereicht, 2003.

D. KÜHN, D. OSTHUS: *Forcing unbalanced complete bipartite minors*. Humboldt-Universität zu Berlin, 2003.

D. KÜHN, D. OSTHUS: *Extremal connectivity for topological cliques in bipartite graphs*. Humboldt-Universität zu Berlin, 2003.

D. KÜHN, D. OSTHUS: *Spanning triangulations in graphs*. Humboldt-Universität zu Berlin, 2003.

Poster

C. FRÖMME, A. GOEDE, C. GRÖPL, S. HOUGARDY, T. NIERHOFF, R. PREISSNER, M. THIMM: *Accelerating screening of 3D protein data with a graph theoretical approach*. Poster zur RECOMB Berlin, April 2003.

S. HOUGARDY, H.J. PRÖMEL, A. TARAZ, M. BEHRISCH, M. THIMM: *Komplexe Netzwerke*. Poster zum Tag der offenen Tür in der Urania, September 2003.

Vorträge

M. BODIRSKY: *Countable homogeneous structures, local clones, and constraint satisfaction*. Universität Potsdam, März 2003.

M. BODIRSKY: *Constraint Satisfaction with Countable Homogeneous Templates*. Technische Universität Dresden, April 2003.

M. BODIRSKY: *The Polymorphism Clones of Omega-Categorical Structures*. Universität Klagenfurt, Juni 2003.

M. BODIRSKY: *Generating Labeled Planar Graphs Uniformly at Random*. Technische Universität Eindhoven, Juli 2003.

M. BODIRSKY: *Homogeneous relational structures, clones from universal algebra, and combinatorial constraint satisfaction*. CGC, Humboldt-Universität zu Berlin, Juli 2003.

M. BODIRSKY: *Constraint Satisfaction with Countable Homogeneous Templates*. Technische Universität Wien, August 2003.

M. BODIRSKY: *Generating unlabeled cubic planar graphs uniformly at Random*. Karlsuniversität Prag, September 2003.

M. BODIRSKY: *Unrooting Planar Structures*. CGC Neustrelitz, Oktober 2003.

A. COJA-OGHLAN: *Approximieren der chromatischen Zahl zufälliger Graphen mittels semidefiniter Programmierung*. Technische Universität Chemnitz, Januar 2003.

A. COJA-OGHLAN: *Finding large independent sets in polynomial expected time*. STACS 2003, Freie Universität Berlin, März 2003.

- A. COJA-OGHLAN: *Spektraltechniken, semidefinite Relaxierungen und zufällige Graphen*. Universität Hamburg, April 2003.
- A. COJA-OGHLAN: *Certifying unsatisfiability of random $2k$ -SAT formulas*. FCT 2003, Universität Malmö, August 2003.
- A. COJA-OGHLAN: *The Lovasz number of random graphs*. RANDOM 2003, Princeton University, August 2003.
- A. COJA-OGHLAN: *Coloring random graphs*. Universität Göttingen, November 2003.
- S. HOUGARDY: *Enumeration of Best Possible Low Degree Expanders*. 1st Workshop on Algorithms for Listing, Counting, and Enumeration (ALICE03), Baltimore, Januar 2003.
- S. HOUGARDY: *Linear Time Local Improvements for Weighted Matchings in Graphs*. Second International Workshop on Efficient Algorithms (WEA 2003), Ascona, Mai 2003.
- S. HOUGARDY: *Approximationsalgorithmen für kombinatorische Optimierungsprobleme*. Institutskolloquium, Institut für Informatik, Humboldt-Universität zu Berlin, Juni 2003.
- S. HOUGARDY: *Linear Time Matching Algorithms*. 31. Berliner Algorithmentag, Berlin, Juli 2003.
- S. HOUGARDY: *Improved Linear Time Approximation Algorithms for Weighted Matchings*. APPROX/RANDOM 2003, Princeton, August 2003.
- S. HOUGARDY: *Three-Dimensional Similarity of Small Molecules*. Forschungsinstitut für Molekulare Pharmakologie, Berlin, November 2003.
- S. HOUGARDY: *Algorithmen im Drug-Design*. Campuswoche Adlershof, Berlin, Dezember 2003.
- S. HOUGARDY: *Biologische Netzwerke*. Netzwerktag des DFG-Forschungszentrums Mathematik für Schlüsseltechnologien, Berlin, Dezember 2003.
- S. HOUGARDY: *Efficient matching algorithms with applications in biology*. Institut für Pflanzengenetik und Kulturpflanzenforschung, Gatersleben, Dezember 2003.
- M. KANG: *Uniform generation of labeled planar graphs*. Konrad-Zuse-Zentrum für Informationstechnik Berlin, Februar 2003.
- M. KANG: *Evolution of random structures*. Workshop on random graphs, Pohang, August 2003.
- M. KANG: *Uniform generation of planar graphs*. 11th Conference on Random structures and algorithms, Poznań, August 2003.
- M. KANG: *Connectivity threshold of the min-degree random multi-graph process*. TU Poznań, September 2003.
- M. KANG: *Uniform generation of cubic planar graphs*. 10th workshop on graph theory: colourings, independence and domination, Karpacz, September 2003.
- M. KANG: *How big is the largest component in a random planar graph?* CGC, Neustrelitz, Oktober 2003.
- T. NIERHOFF: *Full components and approximation algorithms for the Steiner problem*. MPI für Informatik, Saarbrücken. März 2003.
- T. NIERHOFF: *Moving stacker cranes optimally over trees a.a.s.* Berlin-Poznań Workshop on Random Structures, Humboldt-Universität zu Berlin, Mai 2003.

T. NIERHOFF: *Approximationsalgorithmen für schwierige Optimierungsprobleme*. Ring-Vorlesung, Humboldt-Universität zu Berlin, Juli 2003.

D. OSTHUS: *Minors and subdivisions in locally sparse graphs*. Oberwolfach Conference on Graph Theory, Oberwolfach, Januar 2003.

D. OSTHUS: *Minors in graphs of large girth*. Conference on Extremal graph theory. Csopak, Juni 2003.

D. OSTHUS: *Large planar subgraphs and spanning triangulations*. 11th Conference on Random Structures and Algorithms, Poznań, August 2003.

D. OSTHUS: *Minors in graphs of large girth*. European Conference on Combinatorics, Prag, September 2003.

D. OSTHUS: *Large planar subgraphs and spanning triangulations*. Kolloquium über Kombinatorik, Magdeburg, November 2003.

D. OSTHUS: *Random graph models for complex networks*. Netzwerktag, TU Berlin, Dezember 2003.

H.J. PRÖMEL: *Random Planar Graphs*. Cambridge, August 2003.

H.J. PRÖMEL: *Random Planar Graphs*. Oxford, Dezember 2003.

A. TARAZ: *Large planar subgraphs in dense graphs*. Extremal Graph Theory, Budapest, Juni 2003.

A. TARAZ: *Essentially infinite colourings*. Random Structures and Algorithms, Poznań, August 2003.

A. TARAZ: *Colouring various random graphs in expected polynomial time*. Konrad-Zuse-Zentrum Berlin, Oktober 2003.

A. TARAZ: *Canonical colourings with many colours*. Kolloquium über Kombinatorik, Magdeburg, November 2003.

A. TARAZ: *Zufällige diskrete Strukturen: Eigenschaften und Algorithmen*. Institutskolloquium, Universität Freiburg, Dezember 2003.

M. THIMM: *3D-Similarity of Small Molecules*. BCB-Symposium, Berlin, Juni 2003.

Tagungen / Workshops

Berlin-Poznań Workshop on Random Structures

Veranstalter: M. Kang, M. Karoński (Poznań, Polen), A. Taraz

Dieses Blockseminar wird regelmäßig und gemeinsam mit der Arbeitsgruppe von Prof. Karoński an der Adam-Mickiewicz-Universität Poznań ausgerichtet. Es befasst sich mit Themen aus der diskreten Mathematik und der theoretischen Informatik und findet abwechselnd in Berlin und Poznań statt, in diesem Jahr am 30. Mai in Berlin mit etwa 30 Teilnehmern. Vortragende waren J. Grytczuk, T. Nierhoff, A. Ruciński und D. Schlatter.

31. Berliner Algorithmntag

Veranstalter: A. Coja-Oghlan, H.J. Prömel

Der Berliner Algorithmntag ist eine gemeinsame Veranstaltung der Freien Universität Berlin, der Humboldt-Universität zu Berlin, der Technischen Universität Berlin und des Konrad-Zuse-Zentrums für Informationstechnik Berlin, die jedes Semester stattfindet. Der 31. Berliner Algorithmntag wurde am 11. Juli 2003 vom Lehrstuhl Algorithmen und Komplexität ausgerichtet. Etwa 70 Wissenschaftler und Studierende nahmen daran teil. Den Hauptvortrag hielt Uriel Feige vom Weizmann-Institut in Rehovot (Israel).

Workshop „Randomness in the design and analysis of algorithms“

Veranstalter: A. Coja-Oghlan, S.O. Krumke

Dieser Workshop, der im Rahmen der DFG-Forschergruppe Algorithmen, Struktur, Zufall vom 18.-20.10.2003 stattfand, wendete sich an Wissenschaftler der Freien Universität Berlin, der Humboldt-Universität zu Berlin, der Technischen Universität Berlin und des Konrad-Zuse-Zentrums für Informationstechnik Berlin. Die Hauptvortragenden waren Dimitris Achlioptas (Microsoft Research, Redmond, USA) und Yossi Azar (Tel Aviv University, Israel).

Ausstellung

Das Steinerbaumspiel. Ein interaktives Exponat für die „Lange Nacht der Wissenschaften“, Berlin-Adlershof, Juni 2003.

Sonstige Aktivitäten

Michael Behrisch

- Mitglied der Haushaltskommission

Amin Coja-Oghlan

- Mitglied der Bibliothekskommission

Stefan Hougardy

- Mitglied der Berufungskommission für die Professur Systemarchitektur

Till Nierhoff

- Mitglied der Berufungskommission für die Professur Systemarchitektur

Ralf Oelschlägel

- Mitglied des Institutsrates

Deryk Osthus

- Mitglied der Bibliothekskommission
- European Prize in Combinatorics (gemeinsam mit D. Kühn), verliehen auf der Eurocomb 2003

Hans Jürgen Prömel

- Vizepräsident für Forschung der Humboldt-Universität zu Berlin
- Mitglied des Präsidiums der Deutschen Mathematiker-Vereinigung
- Sprecher der DFG-Forschergruppe „Algorithmen, Struktur, Zufall“
- Mitglied des Rates des DFG-Forschungszentrums "Mathematik für Schlüsseltechnologien"
- Mitglied der Bundesjury beim Wettbewerb „Jugend forscht“
- Mitglied des Kuratoriums des Max-Planck-Instituts für molekulare Genetik
- Vorsitzender des Verwaltungsrats des Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB)
- Vorsitzender des Aufsichtsrates der Multimedia Hochschulservice Berlin GmbH (MHSG)
- Mitglied des Aufsichtsrats der IPAL (Innovationen, Patente, Lizenzen) GmbH, Berlin
- Mitherausgeber der Zeitschriften *Random Structures and Algorithms*; *Electronic Journal of Combinatorial Number Theory*; *Combinatorics, Probability and Computing*; *SIAM Journal on Discrete Mathematics*

Anusch Taraz

- Mitglied der Kommission für Lehre und Studium
- Mitglied der Arbeitsgruppe zur Neuordnung der Lehramtsstudiengänge

Gäste am Lehrstuhl

DR. MICHAEL KRIVELEVICH, Tel-Aviv University, Dept. of Mathematics, Israel, Februar 2003.

DR. STEFANIE GERKE, Technische Universität München, Fakultät für Informatik, März 2003.

DR. JAROSLAW GRZYTCZUK, Adam-Mickiewicz-Universität Poznań, Institut für Mathematik, Polen, Mai 2003.

PROF. DR. MICHAŁ KAROŃSKI, Adam-Mickiewicz-Universität Poznań, Institut für Mathematik, Polen, Mai 2003.

PROF. DR. ANDRZEJ RUCIŃSKI, Adam-Mickiewicz-Universität Poznań, Institut für Mathematik, Polen, Mai und Oktober 2003.

PROF. DR. URIEL FEIGE, Weizmann Institute, Rehovot, Israel, Juli 2003.

PROF. DR. ANGELIKA STEGER, Fakultät für Informatik, Technische Universität München, Juli 2003.

PROF. DR. CRIS MOORE, University of New Mexico, Albuquerque, USA, Juli 2003.

VISHAL SANWALANI, University of New Mexico, USA, Albuquerque, Juli 2003.

PROF. DR. HANNO LEFMANN, TU Chemnitz, Fakultät für Informatik, Juli 2003.

DR. DIMITRIS ACHLIOPTAS, Microsoft Research, USA, Oktober 2003.

PROF. DR. YOSSI AZAR, Tel-Aviv University, Dept. of Computer Science, Israel, Oktober 2003.

PROF. DR. WILLIAM HESSE, Clarkson University, Division of Mathematics and Computer Science, Potsdam, USA, Dezember 2003.

DR. OLEG PIKHURKO, University of Cambridge, Großbritannien, Dezember 2003.

Dissertationen

SHI LINGSHENG: *Numbers and Topologies: Two Aspects of Ramsey Theory*. Juli 2003.

Diplomarbeiten

STEFAN KIRCHNER: *Ein Approximationsalgorithmus zur Berechnung der Ähnlichkeit dreidimensionaler Punktmengen*. September 2003.

DANIEL ROLF: *3-Sat is in Runtime 1.32971^n* . Februar 2003.

DIRK SCHLATTER: *On the evolution of random discrete structures: poset dimension and maximum planar subgraphs*. November 2003.

JAN SEIDEL: *3-Färbbarkeit regulärer Graphen und des 3-Core von zufälligen Graphen*. Juli 2003.

Lehr- und Forschungseinheit

Komplexität und Kryptografie

<http://www.informatik.hu-berlin.de/institut/struktur/algorithmenII>

Leiter

PROF. DR. JOHANNES KÖBLER

Tel.: (030) 2093 3189

E-Mail: kobler@informatik.hu-berlin.de

Sekretariat

EVA SANDIG

Tel.: (030) 2093 3190

Fax.: (030) 2093 3191

E-Mail: sandig@informatik.hu-berlin.de

Wissenschaftliche Mitarbeiter

DIPL.-MATH. OLAF BEYERSDORFF

DIPL.-INF. MATTHIAS SCHWAN

DIPL.-INF. LORENZ WEIZSÄCKER

Techniker

DIPL.-ING. NORBERT HEROLD

Tutoren

BENJAMIN BÄSSLER

CARSTEN SCHWARZ

Die Forschungsthemen des Lehrstuhls liegen vorwiegend in den Bereichen Komplexitätstheorie, Algorithmisches Lernen und Kryptografie. Die gegenwärtigen Interessen lassen sich entlang folgender Forschungslinien gliedern:

1. Eine Reihe von algorithmischen Problemstellungen weisen auf Grund ihrer algebraischen Struktur andere Komplexitätseigenschaften auf als die üblichen kombinatorischen Probleme. So lassen sich beispielsweise das Graphisomorphieproblem oder das Faktorisierungsproblem weder als effizient lösbar noch als NP-vollständig klassifizieren. Da diese Probleme sowohl aus theoretischer als auch aus praktischer Sicht eine bedeutende Rolle spielen, ist es wichtig, ihre strukturellen Eigenschaften (wie etwa Vollständigkeit oder Lowness für bestimmte Komplexitätsklassen) zu untersuchen.
2. Ein verwandtes Forschungsthema ist durch die Frage motiviert, ob die Verwendung von Zufallsentscheidungen und/oder Interaktion mit einem Prover (oder Orakel) eine Steigerung der Effizienz von Algorithmen für bestimmte Probleme ermöglicht. Einen vielversprechenden Ansatz bildet die Erforschung von Beziehungen zwischen Komplexitätsklassen, die auf der Basis unterschiedlicher Berechnungsmodelle wie etwa Turingmaschinen, kombinatorische Schaltkreise oder interaktive Beweis-

systeme definiert sind. Innerhalb dieser Forschungsrichtung sind wir beispielsweise an der Frage interessiert, ob NP-vollständige Probleme von Schaltkreisen polynomieller Größe berechnet werden können. Interessanterweise lassen sich hier enge Quer-bezüge zur Frage der Erlernbarkeit von spezifischen Konzeptklassen herstellen.

3. Weitere Forschungsinteressen der Gruppe liegen in den Bereichen Algorithmisches Lernen und Kryptografie. Diese Gebiete haben Fundamente in der Komplexitätstheorie und machen stark von komplexitätstheoretischen Begriffsbildungen und Methoden Gebrauch. Wie im letzten Absatz bereits angesprochen, gilt unser Hauptinteresse hierbei der Erforschung von Querbezügen zwischen den genannten Gebieten, wobei Angluins Modell des „Exakten Lernens durch Fragen“ und Valiants Modell des „PAC-learning“ (PAC = probably approximately correct) im Vordergrund stehen. So konnten wir beispielsweise eine enge Verbindung zwischen der Existenz von kryptografisch sicheren Pseudozufallsgeneratoren und der verteilungsspezifischen PAC-Erlernbarkeit aufzeigen. In einem eher praktisch orientierten Forschungsprojekt sollen Sicherheitsmodelle für IT-Sicherheitssysteme untersucht werden, die eine zuverlässige Evaluierung nach bestimmten Sicherheitskriterien ermöglichen.
4. Bekanntlich lassen sich komplexitätstheoretische Fragestellungen wie etwa $NP =? coNP$ mit der Beweislänge von Tautologien in aussagenlogischen Beweissystemen in Verbindung bringen. Damit verknüpft ist die Frage nach der Existenz von optimalen Beweissystemen für die Menge der aussagenlogischen Tautologien oder für andere Sprachen. Hier ist es gelungen, weitere Bezüge zu rein komplexitätstheoretischen Fragen wie die Existenz von vollständigen Problemen für Promise-Klassen herzustellen.
5. Schließlich gehen wir der Frage nach, ob sich das Paradigma der Parametrisierten Komplexität auf Gebiete wie Algorithmisches Lernen oder Kryptografie nutzbringend anwenden lässt.

Lehre

Veranstaltungen im Grundstudium

- Theoretische Informatik 2 (J. KÖBLER, O. BEYERSDORFF, C. SCHWARZ, WS 03/04)

Kernveranstaltungen (Halbkurse)

- Komplexitätstheorie (O. BEYERSDORFF, SS 03)
- Kryptologie 1 (J. KÖBLER, O. BEYERSDORFF, WS 03/04)
- Chipkartentechnologie 1 (E. G. GIESSMANN, WS 03/04)

Seminare

- IT-Sicherheit (J. KÖBLER, M. SCHWAN, WS 03/04)
- Internet-Standards (E. G. GIESSMANN, SS 03)

Forschung

Projekt: Untersuchung von Sicherheitsmodellen für IT-Sicherheitssysteme, die eine zuverlässige Evaluierung nach bestimmten Sicherheitskriterien erlauben

Ansprechpartner: PROF. DR. JOHANNES KÖBLER

Beteiligter Mitarbeiter: DIPL.-INF. MATTHIAS SCHWAN

Zusammenarbeit: ITC-Security der T-Systems International GmbH

Forschungsförderung: Deutsche Forschungsgemeinschaft

Eine grundlegende Forderung bei der Entwicklung von IT-Sicherheitssystemen ist die Zuverlässigkeit und Korrektheit einer durch das System realisierten Sicherheitspolitik, wie sie für elektronische Bezahlssysteme, Systeme für die biometrische Merkmalserkennung oder die Erstellung und Anwendung elektronischer Signaturen notwendig ist. Dabei ist die Formulierung vollständiger und konsistenter Regeln zur Abwehr der identifizierten Bedrohungen in Form eines Sicherheitsmodells ein wesentlicher Punkt für den Erfolg des Systems.

Das Projekt dient der Untersuchung der Beschreibungen solcher Modelle, die eine zuverlässige und verifizierbare Evaluierung nach bestimmten Sicherheitskriterien ermöglichen. Bisherige Sicherheitsmodelle verwenden überwiegend informelle und semiformale Beschreibungen, bei denen Implementationsfehler und Sicherheitslücken nur schwer erkannt werden. Im Projekt werden existierende generische formale Sicherheitsmodelle bezüglich ihrer Eigenschaften verglichen und bezüglich den Anforderungen der standardisierten internationalen Evaluationskriterien „Common Criteria“ (CC) untersucht. Darauf aufbauend wird für eine Anwendung aus dem Bereich der elektronischen Signatur und biometrischer Identifikationsverfahren ein vorhandenes formales Modell erweitert oder neu entwickelt. Das Projekt wird gemeinschaftlich von der ITC-Security der T-Systems International GmbH und der Humboldt-Universität zu Berlin bearbeitet und ist die Fortführung des Vorgängerprojektes „Semiformale Sicherheitsmodelle“.

Projekt: Polynomielle Fragekomplexität beim Algorithmischen Lernen

Ansprechpartner: PROF. DR. JOHANNES KÖBLER

Beteiligte Mitarbeiter: DIPL.-INF. LORENZ WEIZSÄCKER

Zusammenarbeit: Universität Ulm, Tokyo Institute of Technology, UPC Barcelona

Forschungsförderung: Deutsche Forschungsgemeinschaft

Ziel des Projekts ist es, deterministische und probabilistische Lernalgorithmen - vorwiegend im Kontext von Angluins Modell des „Exakten Lernens durch Fragen“ - zu konzipieren und ihre Komplexität zu analysieren. Eine wichtige offene Frage ist beispielsweise die effiziente (Nicht-)Erlernbarkeit boolescher Schaltkreise. Um einer Lösung näher zu kommen, soll systematisch untersucht werden, in welchen Fällen eine polynomielle Anzahl von Fragen ausreicht und unter welchen Voraussetzungen hieraus bereits auf die Erlernbarkeit in Polynomialzeit geschlossen werden kann. Dabei soll auch der Ressourcenverbrauch von Lernalgorithmen im „average-case“ untersucht werden, da eine „worst-case“-Analyse in manchen Fällen für die Praxis nur wenig aussagekräftig ist.

Weitere Ziele bestehen darin, algorithmische und stochastische Ansätze näher zusammenzubringen und neue Anwendungen von in der Erlernbarkeitstheorie entwickelten Lösungsstrategien in der Komplexitätstheorie zu finden.

Veröffentlichungen

Publikationen

B. JENNER, J. KÖBLER, P. MCKENZIE, J. TORÁN: *Completeness Results for Graph Isomorphism*. Journal of Computer and System Sciences (JCSS) 66, 549-566, 2003.

J. KÖBLER, J. MESSNER, J. TORÁN: *Optimal proof systems imply complete sets for promise classes*. Information and Computation 184, 71-92, 2003.

Preprints

J. KÖBLER, R. SCHULER: *Average-Case vs. Worst-Case Intractability*. Erscheint in: Information and Computation.

Vorträge

O. BEYERSDORFF: *A strong reduction between disjoint NP-pairs*. Fall School of the Logic Seminar Prague, Pec pod Snezkou, September 2003.

E. G. GIESSMANN: *Mehrfachsignaturen: Wie kann man Geschäftsprozesse mit mehreren Unterzeichnern elektronisch abbilden?* CAST-Forum Darmstadt, Januar 2003.

E. G. GIESSMANN: *Technische Fragen und die europäische Standardisierung*. Signatur-Workshop der Regulierungsbehörde für Telekommunikation und Post, September 2003.

M. SCHWAN: *Elektronisches Signieren mit Biometrie?* SymSec'03 Symposium IT-Security, Otto-von-Guericke-Universität Magdeburg, September 2003.

J. KÖBLER: *The Average-Case Complexity of NP and Other Problems*. Tokyo Institute of Technology, Tokio, September 2003.

L. WEIZSÄCKER: *Aggregate mit Komponentengröße 1 charakterisieren PSPACE*. 48. Workshop über Komplexitätstheorie, Datenstrukturen und Effiziente Algorithmen, Universität Hannover, Juni 2003.

Tagungen / Workshops

BITS - Berliner IT-Sicherheit

Veranstalter: E. G. Giessmann, J. Köbler, M. Schwan

Sichere IT-Systeme und sichere Kommunikation sind anspruchsvolle Aufgaben, die wissenschaftliche Zusammenarbeit und Kooperation erfordern. BITS ist dazu eine Berliner Veranstaltungsreihe, bei der man sich über Projekte, Probleme und Pläne austauschen kann. Es werden ausdrücklich alle Interessierten angesprochen, Forscher und Studenten, Entwickler und Anwender und auch Interessierte aus anderen Fachgebieten. Die BITS fand im Jahr 2003 im Februar und November statt.

<http://bits.informatik.hu-berlin.de>

Sonstige Aktivitäten

Prof. Dr. Johannes Köbler

- Mitglied der Haushaltskommission
- Mitglied im Prüfungsausschuss
- Mitglied der Wahlleitung des Instituts für Informatik

Matthias Schwan

- wissenschaftlicher Mitarbeiter in der ITS Security der T-Systems International GmbH
- Projektleiter im BMBF geförderten Projekt "Verisoft - Beweisen als Ingenieurwissenschaft"
- Mitglied im Normungsausschuss NI 27c des DIN (über die T-Systems)

Diplomarbeiten

B. BÄSSLER: *Entwicklung und Analyse eines biometrischen Authentifikations-Protokolls*.
Oktober 2003.

Lehr- und Forschungseinheit

Systemanalyse, Modellierung und Computersimulation » SAM «

<http://www.informatik.hu-berlin.de/sam>

Leiter

PROF. DR. SC. NAT. JOACHIM FISCHER
Tel.: (030) 2093 3109
e-mail: fischer@informatik.hu-berlin.de

Sekretariat

MARITA ALBRECHT
Tel.: (030) 2093 3111
e-mail: albrecht@informatik.hu-berlin.de

Wissenschaftliche Mitarbeiter

DR. RER. NAT. KLAUS AHRENS
DIPL.-INF. HARALD BÖHME
DIPL.-INF. ANDREAS KUNERT
DIPL.-INF. BERTRAM NEUBAUER
DIPL.-INF. TOBY NEUMANN
DIPL.-INF. MICHAEL PIEFEL
DIPL.-INF. FRANK STOINSKI

Technische Mitarbeiter

DIPL.-MATH. MANFRED HAGEN

Habilitanden

DR. RER. NAT. ECKHARDT HOLZ

Promovenden

DIPL.-INF. OLIVER BECKER
DIPL.-ING. DIEGO SANCHO LORAS
DIPL.-INF. TOM RITTER

Emeritus

PROF. DR. HABIL. GUNTER SCHWARZE

Tutoren

MARCUS ASSION (BIS MAI 2003)
GLENN SCHÜTZE (AB JUNI 2003)
KONRAD VOIGT (AB JUNI 2003)

Die Lehr- und Forschungseinheit SAM beschäftigt sich mit allgemeinen Grundlagen der computergestützten Modellierung und Simulation dynamischer Systeme und ihrer Anwendung auf praktische Fragestellungen bei der Softwareentwicklung verteilter und eingebetteter Systeme. Der Telekommunikationsbereich bildet seit vielen Jahren den Kern der praktischen Anwendungen und Werkzeugentwicklungen.



Im Berichtszeitraum wurden folgende thematische Zielstellungen in der Forschung verfolgt:

1. Stabilisierung der Konzepte einer seit vielen Jahren mit anderen internationalen Partnern entwickelten Spezifikationssprache *eODL* zur Entwicklung komponentenbasierter Softwaresysteme, die Übergänge zu existierenden Verhaltensbeschreibungssprachen wie SDL oder UML ermöglichen.
2. Fertigstellung wesentlicher Teile einer Laufzeitumgebung für CCM (*Corba Component Model*) als gemeinsame Forschungsleistung von SAM und dem Fraunhofer-Institut Fokus im Rahmen eines IST-Projektes.
3. Komplettierung der formalen Semantikdefinition von SDL-2000 durch die Ausgestaltung einer vorbereiteten Schnittstelle im bereits entwickelten ASM-Framework für die Modellierung und Interpretation der dynamischen Semantik von SDL-2000.

Zwei Ereignisse im Berichtszeitraum sollen besonders hervorgehoben werden. Ein Höhepunkt war der erfolgreiche Abschluss der Dissertation von Martin von Löwis. Mit dieser Arbeit wurde einerseits eine bestehende Lücke in der formalen Semantikdefinition von SDL-2000 geschlossen, die den Datentypenteil der Sprache betrifft. Zum anderen wurde damit ein entscheidender Beitrag zur semantischen Fundierung von komplexen Sprachen bei Einsatz geschlossener Werkzeugketten generell geleistet, der die Qualität formaler Sprachdefinitionen enorm verbessert. Am Beispiel der werkzeuggestützten Formalisierung der komplexen statischen Semantik von SDL-2000 auf der Basis von ASMs (*abstract state machines*) konnte dies eindrucksvoll nachgewiesen werden.

Ein weiterer Höhepunkt war die Annahme der Sprache *eODL* als ITU-Standard (*Recommendation Z.130*).

Im Rahmen der weiteren Konsolidierung von Forschungsschwerpunkten am Institut fühlte sich die LFE insbesondere federführend für die Vorbereitung des ersten Antrages eines Graduiertenkollegs verantwortlich, der die Bereiche *Modellbasierte Systementwicklung*

(Fischer, Holz, Malek, Reisig, Schlingloff) und *Große Datenmengen im Web* (Freytag, Leser, Reinefeld) im Themenkomplex *Modellbasierte Grid-Dienste* zusammenführte.

Durch die thematische Ausrichtung der Ringvorlesung am Institut, die im Berichtszeitraum von der LFE SAM unter der Thematik *Eingebettete Systeme* gestaltet worden ist, konnten potentielle Felder für verstärkte Kooperationen von LFEs des Instituts und Forschungsinstituten im Raum Berlin fokussiert werden.

Infolge der anhaltenden Vakanz des Lehrstuhls für Systemarchitektur (SAR) war die LFE temporär auch für die Führung der Rechnerbetriebsgruppe des Instituts verantwortlich. Da sämtliche SAR-Mitarbeiter in den organisatorischen Betrieb von SAM integriert worden sind, wurden deren Leistungen im Berichtszeitraum auch hier zusammengestellt. Dies betrifft folgende Mitarbeiter:

- DR. RER. NAT. JAN-PETER BELL
- DIPL.-ING. (FH) GABRIELE BÄRWOLFF
- DIPL.-INF. OLIVER BECKER (BIS OKTOBER 2003)
- DIPL.-ING. WOLFGANG GANDRE
- FRANK WOZOBULE

Lehre

Veranstaltungen im Grundstudium

- Vorlesung »Praktische Informatik III« (JOACHIM FISCHER, WS 2002/2003)
- Praktikum zur Vorlesung »Praktische Informatik III« (KLAUS AHRENS, MICHAEL PIEFEL, ANDREAS KUNERT, WS 2002/2003)
- Vorlesung »Einführung in UNIX« (ECKHARDT HOLZ, WS 2002/2003)
- Praktikum zur Vorlesung »Praktische Informatik II« (KLAUS AHRENS, SS 2003)
- Praktikum zur Vorlesung »Praktische Informatik I« (KLAUS AHRENS, WS 2003/2004)
- Übung zur Vorlesung »Praktische Informatik I« (OLIVER BECKER, WS 2003/2004)
- Praktikum zur Vorlesung »Praktische Informatik III« (ANDREAS KUNERT, WS 2003/2004)

Kernveranstaltungen im Hauptstudium

- Vorlesung »UNIX-Architektur« (JAN-PETER BELL, WS 2002/2003)
- Praktikum zur Vorlesung »UNIX-Architektur« (JAN-PETER BELL, WS 2002/2003)
- Vorlesung »UNIX-Werkzeuge« (JAN-PETER BELL, WS 2002/2003)
- Praktikum zur Vorlesung »UNIX-Werkzeuge« (JAN-PETER BELL, WS 2002/2003)
- Halbkurs »Verteilte Systeme – Architekturen und Konzepte« (ECKHARDT HOLZ, WS 2002/2003)
- Praktikum zum Halbkurs »Verteilte Systeme – Architekturen und Konzepte« (ECKHARDT HOLZ, WS 2002/2003)
- Halbkurs »Objektorientierte Modellierung in C++« (JOACHIM FISCHER, KLAUS AHRENS, SS 2003)
- Praktikum zum Halbkurs »Objektorientierte Modellierung in C++« (ANDREAS KUNERT, SS 2003)
- Kurs »Objektorientierte Modellierung, Spezifikation und Implementierung [OMSI]« (JOACHIM FISCHER, KLAUS AHRENS, WS 2003/2004)
- Praktikum zum Kurs »Objektorientierte Modellierung, Spezifikation und Implementierung [OMSI]« (MICHAEL PIEFEL, WS 2003/2004)
- Halbkurs »UML/SDL« (JOACHIM FISCHER, MARC BORN, SS 2003)
- Praktikum zum Kurs »UML/SDL« (MICHAEL PIEFEL, SS 2003)

- Vorlesung »UNIX-Systemadministration« (JAN-PETER BELL, SS 2003)
- Praktikum zur Vorlesung »UNIX-Systemadministration« (JAN-PETER BELL, SS 2003)
- Vorlesung »Heterogene und komplexe Systemarchitekturen« (ECKHARDT HOLZ, SS 2003)
- Praktikum zur Vorlesung »Heterogene und komplexe Systemarchitekturen« (ECKHARDT HOLZ, SS 2003)

Sonderveranstaltungen

- Ringvorlesung II - WS 2003/2004 (JOACHIM FISCHER)
Thema: »Schwerpunkte der Informatik: Eingebettete Systeme«

Seminare

- Seminar »Multimedia-Middleware – Plattformen und Systeme« (JOACHIM FISCHER, WS 2002/2003, SS 2003, WS 2003/2004)
- Seminar »Objektorientierter Entwurf und Analyse« (JOACHIM FISCHER, WS 2002/2003, SS 2003, WS 2003/2004)
- Seminar »XML für Fortgeschrittene« (OLIVER BECKER, SS 2003)
- Seminar »Compilergeneratoren« (ANDREAS KUNERT, SS 2003)
- Proseminar »Wissenschaftliches Publizieren mit L^AT_EX« (MICHAEL PIEFEL, SS 2003)

Forschung

Schwerpunkt: Plattform-Entwicklung für Komponentenbasierte Systeme der Telekommunikation

Ansprechpartner: PROF. DR. JOACHIM FISCHER

Beteiligte Mitarbeiter: HARALD BÖHME, JOACHIM FISCHER, MARTIN VON LÖWIS, FRANK STOINSKI, BERTRAM NEUBAUER

Zusammenarbeit: T-Systems Nova GmbH, Fraunhofer-Institut Fokus (Deutschland), Université Pierre et Marie Curie, Université des Sciences et Technologies de Lille, Centre National de la Recherche Scientifique, THALES Communications (Frankreich), ObjectSecurity Ltd (Großbritannien), Intracom (Griechenland), Lucent Technologies (Niederlande)

Forschungsförderung: EU IST

COACH

(Harald Böhme, Bertram Neubauer, Frank Stoinski)

Das bereits im Vorjahr begonnene europäische Forschungsprojekt COACH (*Component Based Open Source Architecture for Distributed Telecom Applications*) wurde unter Teilnahme von SAM erfolgreich fortgeführt. Dieses Projekt ist im Rahmen von IST (*Information Society Technologies - European Union's Fifth RTD Framework Programme*) gefördert und beschäftigt sich mit der Entwicklung einer Architektur für komponentenbasierte Systeme im Telekommunikationsbereich. Der Begriff Architektur umfasst hier alle wesentlichen Bereiche einer Komponententechnologie, beginnend vom Komponentenmodell für die Definition grundlegender Konzepte von Komponenten, über Methodiken und Werkzeuge zur Modellierung und Entwicklung von Komponenten und komponentenbasierten Anwendungen bis hin zu einer Laufzeitumgebung für Komponenten, einschließlich einer *Deployment*- und Testinfrastruktur. Die in COACH gewählten Basistechnologien sind CORBA und CCM. Die Implementierung der Laufzeitumgebung und entsprechender Werkzeuge erfolgt in Form einer *Open-Source*-Multiplattform-Implementierung, um eine möglichst uneingeschränkte Verwendung zu erlauben. Den Schwerpunkt

der Beteiligung von SAM bilden die Spezifikation und Implementierung der Laufzeitumgebung und damit verbundener Werkzeuge. Die dazu in Partnerschaft mit dem Fraunhofer-Institut Fokus entwickelte *Open-Source-Initiative Qedo (Quality enabled distributed objects)* konnte entscheidend weiterentwickelt und stabilisiert werden. Im Zentrum der Forschungsarbeit standen zum einen die Erweiterung des CCM um Datenstromkonzepte und zum anderen das *Deployment* von Komponenten, komponentenbasierten Applikationen und Diensten. SAM ist aktiv an der Standardisierung innerhalb der OMG im Rahmen der Revision des CCM Standards und bei der Spezifikation der Einbettung von Datenströmen in CCM beteiligt. Projektarbeit und gewonnene Projektergebnisse dienen der unmittelbaren Fortführung der bei SAM betriebenen Forschungsarbeiten auf dem Gebiet von Plattformen und bilden einen Rahmen für studentische Arbeiten und Graduierungen. Eine Diplomarbeit zum Thema der Transformation von *Deployment-Informationen* aus einer UML-Darstellung in eine XML-Darstellung gemäß CCM wurde erfolgreich abgeschlossen. Andere Arbeiten befassten sich mit der möglichen Visualisierung der in einem CCM-Container laufenden Komponenten und der Definition von Modelltransformationen auf der Metamodell-Ebene.

Integration kontinuierlicher Datenströme in das CORBA Component Model

(Joachim Fischer, Harald Böhme, Bertram Neubauer, Frank Stoinski)

Das CORBA *Component Model* (CCM) definiert ein Komponentenmodell auf der Basis von CORBA. CCM unterstützt hierbei die von CORBA bekannten operationelle Kommunikation und ereignisorientierte Kommunikation. Für eine Vielzahl von Anwendungen, vor allem im Bereich der Audio- und Videoverarbeitung, aber auch der Verarbeitung von Messwerten, wäre es wünschenswert auch eine dritte Kommunikationsform, kontinuierliche Datenströme, in dieses Modell zu integrieren. Im Rahmen einer Dissertationsarbeit wird dazu ein integrativer Ansatz entwickelt, um unter Nutzung der Vorteile von CCM verteilte komponenten-basierte Datenstrom-Applikationen zu bauen. Ergebnisse der Forschungen werden in den Standardisierungsprozess der OMG für die Integration kontinuierlicher Datenströme in CCM (*Streams for CCM RFP*) eingebracht.

Deployment von komponentenbasierten Anwendungen

(Joachim Fischer, Harald Böhme, Bertram Neubauer, Frank Stoinski)

Die Unterstützung für das *Deployment* von Komponenten und komponenten-basierten Anwendungen in CCM ist bereits im Standard enthalten, leidet jedoch unter Lücken und Fehlern in der Spezifikation. Für die tatsächliche Benutzbarkeit sind eine Reihe von Verbesserungen bzw. Erweiterungen erforderlich, die in der *Open-Source-Initiative Qedo (Quality enabled distributed objects)* in Partnerschaft mit dem Fraunhofer-Institut Fokus realisiert werden. Gegenwärtig existiert kein geeignetes Metamodell für das *Deployment*, um die Konzepte dieser Begriffswelt in ein umfassendes CCM-Metamodell zu integrieren. Die Repräsentation der *Deployment-Informationen* erfolgt lediglich mittels XML-Formaten, deren Bezug zu einer Modellbasis jedoch fehlt. Ein gemeinsames MOF-Metamodell würde die automatische Ableitung eines XML-Formates mit Hilfe von XMI erlauben und auf diese Weise Konsistenz garantieren können. Die spezifizierte *Deployment-Infrastruktur* ist in der jetzigen Fassung ausschließlich auf statische Aspekte ausgerichtet und bietet keine Unterstützung für dynamische Veränderungen von Komponentenanwendungen. Im Rahmen von COACH und zur Durchführung von Experimenten wurde die Infrastruktur für das *Deployment* in CCM erweitert und prototypisch implementiert. Dabei stehen auch die Einführung dynamischer Aspekte und die Untersuchung der Verwendbarkeit laufender Komponenten und Komponenten-*Assemblies* für die rekursive Konstruktion komplexer komponentenbasierter Systeme im Blickpunkt.

SDL-Technologien zur Komponentenentwicklung

(Joachim Fischer, Harald Böhme, Bertram Neubauer, Frank Stoinski)

Die Arbeiten an der Komponentenplattform **Qedo** im Rahmen des Projektes COACH konzentrieren sich auf die Realisierung einer CCM-Implementierung für C++. Entwicklung komponentenorientierter Software kann und soll aber auch im Softwareentwicklungsprozess ohne spezielle Festlegung auf die letztendliche Plattformtechnologie möglich sein. Die abstrakte Spezifikation von Komponenten in einer technologieneutralen Sprache ist dabei ein wichtiger Anfangsschritt. Die Entwicklung plattformunabhängiger Komponenten wird im Rahmen einer Dissertation behandelt. Dabei wird mit Hilfe von **eODL**, einer bei der ITU-T definierten Komponentensprache der strukturelle Entwurf des Softwaresystems vollzogen. Eine weitere Spezifikation des Verhaltens der einzelnen Komponenten kann dabei durch den Übergang in die Spezifikationsprache SDL (auch von der ITU-T entwickelt) erfolgen. In der Realisierungsphase des Softwaresystems muss nun eine Entscheidung bzgl. der Zielplattform getroffen werden. Diese ist aus gegebenen Gründen zugunsten von CCM getroffen worden. Nun können die Informationen aus dem strukturellen Entwurf (**eODL**) und der Verhaltensspezifikation (SDL) in die ausgewählte Plattformtechnologie übertragen werden. Das sind im vorliegenden Falle CCM und C++. Die Verhaltensaspekte können dabei analog der schon in der SITE-Werkzeugkette entwickelten Abbildung von SDL nach C++ übernommen werden. Somit ist ein fast vollständig automatischer Übergang von einer komponentenplattformunabhängigen Spezifikation eines Softwaresystems zur Realisierung in und mit **Qedo** möglich.

Nicht-funktionale Eigenschaften von Komponenten

(Joachim Fischer, Tom Ritter [Fraunhofer-Institut Fokus])

Das **CORBA Component Model (CCM)** definiert ein Komponentenmodell auf der Basis von CORBA. CCM unterstützt hierbei die von CORBA bekannten operationelle Kommunikation und ereignisorientierte Kommunikation. Für eine Vielzahl von Anwendungen wäre es wünschenswert, die Modellierung nicht-funktionaler Eigenschaften von Komponenten und deren Implementierung in das CORBA-Komponentenmodell zu integrieren. Im Rahmen einer Dissertationsarbeit wird dazu ein integrativer Ansatz entwickelt, um unter Nutzung der Vorteile von CCM verteilte komponentenbasierte Applikationen zu bauen, in denen nicht-funktionale Eigenschaften berücksichtigt werden. Ergebnisse der Forschungen werden in den Standardisierungsprozess der OMG für die Integration von *QoS* in CCM (*QoS for CCM RFP*) eingebracht.

Schwerpunkt: Semantische Fundierung von Modellbildungen

Ansprechpartner: PROF. DR. JOACHIM FISCHER

Zusammenarbeit: ITU-T, Microsoft Research (USA, Redmond), DResearch GmbH, Universität Kaiserslautern, (Deutschland) University of Posts & Telecommunication (China)

Forschungsförderung: Microsoft Research

Die langjährigen Aktivitäten der LFE SAM hinsichtlich einer neuen semantischen Fundierung des Datentypkonzeptes von SDL-2000 wurden erfolgreich abgeschlossen (vgl. <http://rn.informatik.uni-kl.de/projects/sdl/>). Die Ergebnisse fanden einerseits Eingang in die Neufassung des Standarddokumentes von SDL (Z.100, Annex F), liegen in Form eines normativen Referenzcompilers vor und wurden zudem in der Dissertation von Martin von Löwis wissenschaftlich aufgearbeitet.

Formale Semantik des Datentypmodells von SDL-2000

(Martin von Löwis of Menar)

Im Rahmen einer Dissertation wurde ein fundamentaler Beitrag zur semantischen Fundierung von komplexen Sprachen bei Einsatz geschlossener Werkzeugketten geleistet,

der die Qualität formaler Sprachdefinitionen enorm verbessert. An der Definition der formalen Semantik für die Gesamtsprache arbeitete seit 1998 eine vom SDL-Forum installierte Arbeitsgruppe unter Koordination der SG10 der ITU-T, bestehend aus universitären Mitgliedern und industriellen Einrichtungen. Martin von Löwis gehörte dieser Gruppe seit ihrer Gründung an, nachdem er sich bereits als international anerkannter SDL-Sprachexperte in der Konsolidierungsphase von SDL-96 in die Standardisierungsarbeit eingebracht hatte. Seine in der Dissertationsarbeit erbrachten Beiträge sind sowohl in diesem Forum als auch in der ITU-T auf hohe internationale Anerkennung gestoßen und haben nicht nur den Standard für SDL-2000, sondern auch die Entwicklung des Referenzcompilers maßgeblich beeinflusst. Für jeden der fünf Teile der informalen Semantikdefinition (Abstrakte Syntax 0, Korrektheitsbedingungen, Transformationsregeln zur Auflösung von *Short-Hand*-Notationen und abstrakten Modellkonstrukten, Abstrakte Syntax 1 und die Abbildung der Abstrakten Syntax 0 auf die Abstrakte Syntax 1) wurde eine ASM-Notation entwickelt, die von Werkzeugen verarbeitet werden kann. Die verwendeten Werkzeuge sind zum einen Standardwerkzeuge des Compilerbaus und zum anderen Werkzeuge, die innerhalb der LFE SAM entwickelt worden sind. Durch die werkzeuggestützte Formalisierung gelang es, zahlreiche Probleme in der Semantikdefinition deutlich zu machen und der ITU-T Lösungsvorschläge zu unterbreiten. Da Variablenänderungen in der SDL-Semantik allgemeine Zustandsänderungen (per Programmausführung der ASM-Agenten) bedeuten, ordnet sich die Betrachtung der Datentypsemantik zwangsläufig in die der dynamischen Gesamtsemantik von SDL ein. Die reine Datentypsemantik stellt sich selbst schon als eine äußerst komplexe Problematik aufgrund des Umfangs vordefinierter Datentypen und Operatoren in SDL-2000 dar. Hinzu kommt, dass die Abhängigkeiten zwischen unterschiedlichen Sichtbarkeitsbereichen in SDL-2000 enorm zugenommen haben und dass bei der dynamischen Semantik zudem von einer vorgegebenen funktionalen Schnittstelle auszugehen war. Trotz dieser erschwerenden Auflagen gelang es Herrn von Löwis nachzuweisen, dass eine solche Schnittstelle tatsächlich bei Implementierung der dynamischen Semantik in ASM im Prinzip bestätigt werden kann.

Schwerpunkt: Methoden und Werkzeuge für eine Modellgestützte Entwicklung komponentenbasierter Systeme

Ansprechpartner: PROF. DR. JOACHIM FISCHER

Beteiligte Mitarbeiter: JOACHIM FISCHER, ECKHARDT HOLZ

Zusammenarbeit: OMG, TINA-C, ITU-T, T-Nova (Darmstadt), Fraunhofer-Institut Fokus, DResearch GmbH, Microsoft Research (Redmond), Cinderellas ApS (Kopenhagen), DeTeWe Funkwerk Köpenick GmbH, Siemens AG (Berlin), IKV++ GmbH Berlin

Forschungsförderung: Fraunhofer-Institut Fokus, IKV++ Technologies AG, DeTeWe Funkwerk Köpenick GmbH

Modellierungstechniken und Modellbasierte Entwicklung

(Eckhardt Holz)

Der Einsatz von Modellen ist ein wesentliches Hilfsmittel beim Entwurf und der Entwicklung komplexer Softwaresysteme. Mit den Forschungsarbeiten zu Modellierungstechniken und zur Modellbasierten Entwicklung wurden drei verwandte Themenbereiche adressiert. Erstens wurden im Rahmen eines Habilitationsvorhabens für SAM modellbasierte Entwicklungsmethoden untersucht, die insbesondere die Kombination unterschiedlicher Modelle und Modellierungstechniken innerhalb des Softwareentwicklungsprozesses unterstützen. Zweitens entstand in Zusammenarbeit mit Olaf Kath (TU Berlin/IKV++) und Marc Born (Fraunhofer-Institut Fokus) eines der ersten deutschsprachigen Bücher zu UML 2 (<http://www.uml2-buch.de>). Dieses Buch bietet bis ins Detail eine vollständige Darstellung

der UML 2 und des dieser Sprache zugrunde liegenden Metamodells. Die verschiedenen Facetten der Modellierung mit UML werden an zahlreichen Beispielen leicht nachvollziehbar erläutert. Ein weiterer Schwerpunkt des Buches liegt auf der Anwendung des Prinzips Metamodellierung bei der Definition der Sprache UML 2, sowie bei deren Anpassung und Erweiterung im praktischen Einsatz. Ein separates Kapitel beschäftigt sich mit der Einbettung von UML 2 und weiteren in diesem Buch vorgestellten Techniken in die *Model Driven Architecture* (MDA).

Als dritte Aktivität wurde in Kooperation mit den LFE von Prof. Freytag, Prof. Leser, Prof. Malek, Prof. Reinefeld, Prof. Reisig sowie Prof. Schlingloff ein Antrag auf Einrichtung eines Graduiertenkollegs vorbereitet. Das Graduiertenkolleg soll die Grundlagen und Möglichkeiten einer dienstorientierten *Grid*-Architektur, daraus abgeleitete Anforderungen an die Entwicklung von Diensten, sowie konkrete Anwendungen der Architektur für konkrete Dienste erforschen. Dies stellt nicht nur ein technisches, sondern vor allem ein konzeptionelles Problem dar. *Grid*-Dienste sollten gegenüber den Nutzern alle technischen Details verbergen, was durch entsprechende Abstraktion erreicht werden kann. Ein Nutzer oder eine Anwendung kennt Modelle des *Grid*. Diese Modelle dienen zum einen der Realisierung von Diensten. Zum anderen kommen sie auch bei der Installation, der Konfiguration und der dynamischen Komposition komplexer Dienste zum Einsatz. Dabei werden bei der Dienstentwicklung und Dienstkomposition nicht nur funktionale, sondern auch nicht-funktionale Aspekte berücksichtigt. Auftraggeber, Entwickler, Betreiber, Nutzer, kurz: alle an einem *Grid*-System Beteiligten arbeiten mit solchen Modellen in Analogie zum wohlbekannten modellbasierten Software-Entwurf. Der Antrag wird im Frühjahr 2004 bei der Deutschen Forschungsgemeinschaft eingereicht.

SITE-Cinderella-Integration

(Toby Neumann)

Die SITE-Werkzeuge von SAM lassen sich über die graphische Oberfläche des kommerziellen SDL-Werkzeugs Cinderella-SDL verwenden, um aus einer Systemspezifikation lauffähigen Code zu erzeugen. Dieser Code basiert allgemein auf einer Bibliothek, im Besonderen auf der so genannten CORBA-Laufzeit-Bibliothek. Die damit zugängliche *Middleware*-Plattform CORBA erlaubt einer aus SDL generierten Anwendung die standardisierte Kommunikation mit anderen Anwendungen. Am Beispiel konkreter Ersetzung verschiedener Komponenten der Bibliothek wurden deren Austauschbarkeit und die Anpassungsfähigkeit der Bibliothek demonstriert. Die kommerzielle CORBA-Implementierung *Orbacus* wurde durch eine *OpenSource*-CORBA-Implementierung ersetzt und die kommerzielle Thread-Bibliothek JTC durch die BOOST *thread library*. Gleichzeitig war dies ein wichtiger Schritt zur Erhöhung der Flexibilität und Produktunabhängigkeit der SITE-Werkzeuge. Begleitet wurden diese Entwicklungsarbeiten durch eine erste vollständige Dokumentation des Bibliotheks-Quelltextes und eine umfassende Nutzer-Online-Dokumentation zur Verwendung der Werkzeuge.

SDL für eingebettete Systeme

(Diego Sancho Loras)

Im Laufe des Jahres wurden Möglichkeiten der Implementierung von hierarchischen Zustandsautomaten untersucht. Die Kriterien für diese Implementierung waren: die Effizienz bzgl. Laufzeit und Speicherplatzbedarf, die Verwendung von verbreiteten Programmiersprachen (C und C++), die Integration mit Echtzeitbetriebssystemen und die Lesbarkeit und Erweiterbarkeit für die zukünftige Wartung der Zustandsautomaten. Die Implementierung soll die Basis für die Code-Generierung von Systemen in SDL-2000 in andere Programmiersprachen sein, damit die Systeme in einer Hardware-Plattform ausführbar werden. Im Zusammenhang mit dem Projekt „Handfunkgerät nach Standard TETRA“ der Firma

DeTeWe Funkwerk Köpenick GmbH wurde eine Implementierung des *Man Machine Interface* (MMI) realisiert. Im Wintersemester wurde eine Untersuchung von *Middleware*-Infrastrukturen gestartet, die die Integration der Zustandsmaschinen in einem System ermöglichen. Die Kriterien für die Verwendung einer *Middleware*-Infrastruktur sind: Bereitstellung von Elementen für die (quasi-)parallele Abarbeitung der Zustandsmaschinen (Threads, Signal-Queues, Timers), Abstraktion des Betriebssystems und anderer hardware-spezifischer Komponenten (z.B. Treiber, Interrupts, Mutex, ...), die Möglichkeit der Integration „externer“ Software, die Effizienz bzgl. Laufzeit und Speicherverwaltung, der Kommunikation über Signale und der Zeitverwaltung. Auch diese *Middleware*-Infrastruktur soll eine Basis für die Code-Generierung von SDL-2000 in andere Programmiersprachen sein.

Schwerpunkt: Internationale Standardisierung

Ansprechpartner: PROF. DR. JOACHIM FISCHER

Zusammenarbeit: ITU-T, Fraunhofer-Institut Fokus (Berlin), IKV++ Technologies AG (Berlin)

eODL

(*Olaf Kath, Marc Born, Joachim Fischer, Harald Böhme, Bertram Neubauer*)

Heutige und künftige Telekommunikationssysteme bestehen aus Komponenten, die in einer Netzumgebung verteilt sind, und dabei Parallelitäts-, Autonomie-, Synchronisations- und Kommunikationsaspekte berücksichtigen. Die Entwicklung korrekter und performanter verteilter Telekommunikationsanwendungen ist eine komplexe und komplizierte Aufgabe, die den Einsatz von CASE-Werkzeugen zur Unterstützung von Entwurf und Modellierung dieser Anwendungen erfordert. Die Forschungen von SAM verfolgen die Vision der direkten Erzeugung von Anwendungen aus objektorientierten Entwurfsmodellen auf der Grundlage von Komponentenkonzepten, die insbesondere zu wieder verwendbaren und ausführbaren Komponenten führt. Ausführbare Komponenten müssen selbstverständlich Aspekte, die von der konkreten Laufzeitumgebung und der Technologie der *Middleware*-Plattform abhängen, mit Aspekten des applikationsspezifischen objektorientierten Entwurfsmodells vereinen. Die ausführbare Systemlösung kann durch Zusammenstellung und Konfiguration dieser Komponenten für eine Zielumgebung bereitgestellt werden. Basierend auf vorangegangenen Arbeiten, deren Ziel die Integration des objektorientierten Entwurfs mit der komponentenbasierten Fertigung von Telekommunikationsanwendungen in einem modellbasierten Ansatz war, wurde die Spezifikationsprache ***eODL*** der ITU-T entwickelt. Diese Sprache umfasst im Wesentlichen:

- wohldefinierte Konzepte für verteilte Telekommunikationssysteme,
- ein Metamodell, das die Semantik der Konzepte definiert,
- die Möglichkeit, beliebige Notationen zu definieren, die für verschiedene Spezifikationsmethoden und Werkzeuge geeignet sind.

Sowohl die Konzepte als auch Methoden sind weder an Netz- noch an Plattformtechnologien gebunden. Demzufolge ist die Modellierung von Systemen unabhängig von verwendeten Technologien. Für konkrete Technologien müssen Abbildungen der Konzepte auf deren Realisierung für konkrete Zielumgebungen definiert werden. Damit entspricht ***eODL*** im Grundsatz dem von der OMG verfolgten Ansatz zu *Model Driven Architectures* (MDA), einer allgemeinen Entwurfsmethodik, die im Wesentlichen durch zwei Dinge charakterisiert ist. Zum einen die Metamodellierung zur Widerspiegelung wichtiger Konzepte einer Domäne und zur Trennung dieser Konzepte von einer konkreten Notation. Zum anderen die Unterscheidung verschiedener Modellierungsebenen, namentlich zwischen plattformun-

abhängigen Modellen (PIM) und plattformspezifischen Modellen (PSM). Zunächst plattformunabhängige Modelle eines Systems werden je nach verwendeter Plattform in spezifische Modelle überführt.

In diesem Sinne bietet *eODL* eine plattformunabhängige Modellierung für Telekommunikationssysteme. Der Sprachstandard konnte im Rahmen der ITU-T unter dem Namen Z.130 verabschiedet werden.

CORBA Component Model

(Harald Böhme, Bertram Neubauer, Frank Stoinski)

Innerhalb der OMG waren Mitarbeiter von SAM im Rahmen der Standardisierung des *CORBA Component Model (CCM 1.1 Revision Task Force (CCM RTF))* aktiv. Frank Stoinski arbeitete weiterhin an der Standardisierung von Datenstromerweiterungen für das *CORBA Component Model (Streams for CCM RFP)* mit. Joachim Fischer setzte seine Funktion als Rapporteur innerhalb der ITU-T für Fragestellungen der Studiengruppe 17 fort, die sich mit der Standardisierung der von SAM mitentwickelten Entwurfssprache *eODL* auseinandersetzte.

Schwerpunkt: Werkzeuge und Methoden für Basistechnologien

Ansprechpartner: PROF. DR. JOACHIM FISCHER

Zusammenarbeit: Cinderella ApS, Dänemark

Kimwitu++

(Michael Piefel, Markus Scheidgen)

Das von SAM entwickelte Metawerkzeug *Kimwitu++* wurde weiterhin gepflegt. Um bereits existierende Projekte vollends von der Verwendung von *Kimwitu* auf das modernere Werkzeug *Kimwitu++* umzustellen, wurde bereits im vergangenen Jahr eine Studienarbeit (Markus Scheidgen) in Angriff genommen, die untersuchte, inwieweit die nötigen Anpassungen werkzeugunterstützt durchgeführt werden können. Als Ergebnis dieser Arbeit ist ein Werkzeug entstanden, welches es unter Ausnutzung allgemeiner Compilerbautechnologien und im Besonderen auch *Kimwitu++* selbst ermöglicht, auch größere Projekte wie SITE in kürzester Zeit nach *Kimwitu++* zu überführen.

ODEMx

(Klaus Ahrens, Joachim Fischer, Ralf Gerstenberger, Martin von Löwis, Michael Piefel)

Die von SAM entwickelte C++-Bibliothek zur Unterstützung einer ereignisorientierten Modellierung und Simulation hat in Lehre und Forschung breite Anwendung erfahren und wird ständig weiterentwickelt. Ergebnis der Aktivitäten aus dem Berichtszeitraum ist die Diplomarbeit von Ralf Gerstenberger. In dieser entwickelte der Autor *ODEMx* aus der ebenfalls an der LFE entstandenen Bibliothek *ODEM*. Bei der Implementierung von *ODEMx* kamen verstärkt moderne C++-Elemente wie Templates zum Einsatz, die den Einsatz in der Lehre in Kombination mit C++-Kursen nahe legen.

STX (Streaming Transformation for XML)*(Oliver Becker)*

Das im Vorjahr am Lehrstuhl für Systemarchitektur begonnene *Open-Source*-Projekt STX (*Streaming Transformations for XML*) wurde erfolgreich weitergeführt. Ziel dieses Projektes ist es, eine Sprache für die serielle Transformationen von XML-Daten zu entwickeln. Auf diese Weise lassen sich sowohl potenziell beliebig große XML-Dokumente als auch XML-Datenströme effizient transformieren. Der Sprachentwurf orientiert sich an der vom W3C spezifizierten Transformationssprache XSLT. Während die Weiterentwicklung der Sprache STX auf einer öffentlichen Mailingliste mit weiteren Autoren betrieben wird, wurde am Institut eine prototypische Implementierung in Java (*Joost*) von Oliver Becker entwickelt und auf *SourceForge.net* veröffentlicht. Teile dieser Implementierung entstammen einer im Vorjahr entstandenen Studienarbeit von Anatolij Zubow. *Joost* ist die einzige derzeit verfügbare STX-Implementierung in Java und eine von zwei Implementierungen überhaupt. Im Jahr 2003 konnte ein wachsendes Interesse von Seiten der Anwender beobachtet werden. Oliver Becker stellte STX und *Joost* auf zwei internationalen Konferenzen der Öffentlichkeit vor. Darüber hinaus wurde *Joost* in das *Publishing-Framework Cocoon* integriert. Die Sprache STX und Aspekte der Implementierung *Joost* werden Kernpunkte der Dissertation von Oliver Becker sein.

Minako*(Andreas Kunert)*

Bei *Minako* (*Multiple implementations of a nearly academic compiler*) handelt es sich um eine Sammlung von gegenwärtig 13 äquivalenten Compilern, die in verschiedenen Programmiersprachen unter Zuhilfenahme verschiedener Werkzeuge implementiert wurden. Begonnen hat das Projekt als Lehrcompiler für das Compilerbaupraktikum der Praktischen Informatik III. Dementsprechend wurde eine einfache Quellsprache (C0, eine selbst definierte Teilsprache von C) definiert und eine leicht vermittelbare Zielsprache (Jasmin, ein Assembler für die *Java Virtual Machine*) gewählt. Zwei Referenzimplementierungen wurden von Andreas Kunert in Java und C mit Hilfe der Parsergeneratoren *javacc* und *bison* vorgenommen, beide als Einpasscompiler. Im Sommersemester erfolgten dann im Rahmen des Seminars „Compilergeneratoren“ weitere Implementierungen (allerdings nur bis zur syntaktischen Analyse) durch die beteiligten Studenten. Anschließend wurden zwei Mehrpasscompiler implementiert: einer von den studentischen Mitarbeitern Glenn Schütze und Konrad Voigt unter Zuhilfenahme des am Lehrstuhl entstandenen Termprozessors *kimwitu++* und einer von Andreas Kunert mittels des Syntaxbaumgenerators *jjtree*. Ziel des Projekts ist es, eine möglichst umfangreiche Sammlung von äquivalenten Compilern zu erhalten, die man zum einen zur Bereicherung der Lehre des Compilerbaus und zum anderen als Nachschlagewerk für die Einarbeitung in die einzelnen Compilerwerkzeuge verwenden kann. Seit Beginn des Wintersemesters wird ein Compiler von *Minako* (ein mit *bison* in C implementierter Einpasscompiler) erfolgreich von der LFE Datenbanken und Informationssysteme für das Compilerbaupraktikum eingesetzt.

Projekt: Studentische Projekte**Ansprechpartner:** PROF. DR. JOACHIM FISCHER**Beteiligte Mitarbeiter:** JOACHIM FISCHER, MARTIN VON LÖWIS, MICHAEL PIEFEL, OLIVER BECKER, TOBY NEUMANN, HARALD BÖHME**SITE und ASN.1: Packed Encoding Rules***(Glenn Schütze, Konrad Voigt)*

Die SITE-Werkzeuge der LFE SAM unterstützen die Verwendung von ASN.1 in SDL-Spezifikationen. Die Sprache ASN.1 besteht aus zwei Schichten: zum einen aus der Syntax zur Definition abstrakter Datentypen und zum anderen aus verschiedenen standardisierten Sätzen von Codierregeln, die jeweils festlegen, wie ein abstrakter Wert eines Datentyps als Bitstrom repräsentiert ist, das heißt, wie solche Werte für die Datenübertragung codiert und decodiert werden. Bisher unterstützten die SITE-Werkzeuge den ältesten und am weitesten verbreiteten Satz an Codierregeln, genannt BER (*Basic Encoding Rules*). Nun wird aber mit zunehmender Verwendung von SDL und ASN.1 auch im Bereich der Spezifikation von Mobilfunk-Protokollen die Reduzierung der zu übertragenden Daten wichtiger. Eine solche Forderung unterstützt der Codierregel-Satz PER. Es wurde damit begonnen, die SITE-Werkzeuge so zu erweitern, dass sie auch anderer Codierregel-Sätze (insbesondere PER *Packed Encoding Rules*) unterstützen können. Basierend auf Arbeiten zur Codierung mit PER in Java wurden erste äquivalente C++-Routinen in die bestehende Laufzeit-Bibliothek integriert.

Mobiler Code

(Joachim Dorn)

Im Rahmen einer Studienarbeit wurde eine prototypische Implementierung von mobilem Code in Java realisiert. Dieser Ansatz soll zu einer Diplomarbeit ausgebaut werden. Im Rahmen der Koordination und Eingrenzung des Diplomthemas erfolgte eine Einarbeitung in verschiedene Themengebiete u.a. in verteilte Systeme nach RM-ODP. Ziel der Diplomarbeit, deren Abschluss 2004 geplant ist, ist die Beschäftigung mit dem Thema verteilter Systeme unter der Fragestellung, ob die Verteilung einer Anwendung immer komplett von der *Middleware* des verteilten Systems gesteuert werden sollte, oder wie und unter welchen Umständen Anwendungen selbst auf ihre Verteilung Einfluss nehmen können oder müssen. Dazu ist ebenfalls wieder eine, die theoretischen Überlegungen begleitende, prototypische Implementierung geplant.

SDL-2000 Metamodell

(Markus Scheidgen)

Nachdem in der ersten Jahreshälfte die Studienarbeit (Bau eines Konverters für Kimwitu nach Kimwitu++) fertiggestellt wurde, begann Markus Scheidgen Ende des Jahres im Rahmen seiner Diplomarbeit mit der Entwicklung eines Metamodells für SDL-2000. Dabei wurden Werkzeuge für eine programmatische Überführung der abstrakten SDL-2000 Syntax in ein MOF-Modell geschrieben. Bis Februar 2004 soll die Entwicklung des SDL-2000 Metamodells abgeschlossen sein.

Debugging Markup-basierter Transformationen am Beispiel STX*(Anatolij Zubow)*

Die Dissertation von Oliver Becker beschäftigt sich mit einer Transformationssprache für XML (genannt STX). Im Zuge dieser Arbeiten wurde u.a. ein Prototyp eines STX-Interpreters geschrieben. Anatolij Zubow hat im Kontext dieser Arbeiten einen Debugger für STX entwickelt, der auf dem Prototyp aufsetzt, und typische Debugger-Funktionalitäten wie Haltepunkte, Inspektion von Variablen, schrittweise Abarbeitung etc. unterstützt.

Metamodell-Transformation*(Andreas Wenskus)*

Im Rahmen der Diplomarbeit wurde ein Framework zur Transformation von Modellen realisiert. Dabei kommen Methoden der Metamodellierung zum Einsatz. Ausgangs- und Zielmodelle müssen entsprechend einem in MOF 1.3 definierten Metamodell vorliegen. Auch die Definition der Transformation erfolgt in Form eines Modells, das dem in der Arbeit entworfenen Metamodell für Transformationen genügt. Technologisch werden die Transformationsschritte dabei auf *Repository*-Ebene vollzogen. Zur Realisierung der Transformation wurde auf Java und das dort definierte MDR zurückgegriffen. Die eigentlichen Transformationsschritte im Transformationsmodell werden in Form von *F-Logic* Termen angegeben. *F-Logic* ist eine Erweiterung der standardisierten Programmiersprache Prolog. Für den Zugriff auf die Modellelemente in der Transformation wurde im Rahmen der Diplomarbeit eine Abbildung von MOF 1.3 auf *F-Logic* definiert.

Konzepte und Strategien für DBMS-Ressourcenmanager mit dem CORBA-Transaktionsdienst*(Hans-Peter Grimm)*

Transaktionen sind ein fundamentales Konzept der Informatik. Die XA-Schnittstelle standardisiert die Schnittstelle zwischen einem Datenbankmanagementsystem und einem Transaktionsmanager. Der CORBA-Transaktionsdienst realisiert Transaktionen für verteilte CORBA-Anwendungen. Das Zusammenspiel des CORBA-Transaktionsdienstes mit Datenbanken wird im zugehörigen Standard dieses Dienstes kaum geklärt. Die Diplomarbeit von Hans-Peter Grimm entwickelt Konzepte und Strategien zur Entwicklung und Verwendung von DBMS-Ressourcenmanagern im Zusammenspiel mit dem CORBA-Transaktionsdienst. Auf der Basis des im Jahr 2002 am Lehrstuhl entwickelten Prototyps eines CORBA-Transaktionsdienstes mit Namen „TRAX“ untersucht die Diplomarbeit Wege für eine schnelle und zweckmäßige Implementierung sowohl transaktionaler CORBA-Objekte als auch wiederherstellbarer CORBA-Objekte.

Modellierung von Geschäftsprozessen mit ODEMX*(Robert Sauer)*

Im Rahmen einer Diplomarbeit wurde eine Programmumgebung für die Modellierung und Simulation von Geschäftsprozessen konzipiert und bislang teilweise implementiert. Dazu war insbesondere die Portierung und Anpassung der Bibliothek ODEMX erforderlich. Zusätzlich wurde die Implementierung des CARNOT Metamodells komplettiert und mit *Java-Proxies* ausgestattet. Der gegenwärtig erreichte Stand ist nun, dass in einem Java-Programm zur Laufzeit Modelle erzeugt werden können, die durch ODEMX simuliert werden können. Die Testüberdeckung ist zwar noch nicht sehr groß, jedoch laufen *Split* und *Fork* von Aktivitätszweigen bzw. die Einbeziehung von *Performern* grundlegend. Im Rahmen der Implementierung wird als nächster Schritt angestrebt, reguläre Modelle aus dem CARNOT XML zu lesen und daraus die zu simulierenden Systeme automatisch zu erzeugen.

Veröffentlichungen

J. FISCHER (RAPPORTEUR Q.24/25): *eODL: Techniques for distributed software component development conceptual foundation, notations and technology mappings*, International Standard, ITU-T Recommendation Z.130.

O. BECKER: *Pfad-Finder – Navigation mit XPath*, Artikel in [JavaSPEKTRUM 06/2003](#), SIGS-DATACOM GmbH, Troisdorf, Nr. 48, S. 48 – 52.

O. BECKER: *An Introduction to Streaming Transformations for XML*, Artikel in [xml.com](#), O'Reilly, 26. Februar 2003.

E. HOLZ: *Strategies for the Support of Model Combination within the Software Engineering Process*, SFB-501 Kolloquium, SFB-Bericht 08/03, Kaiserslautern, April 2003.

E. HOLZ: *A Meta-Model based Approach for the Combination of Models in multiple Languages*, Proc. of. IASTED MS-2003, Palm Springs, Februar 2003.

O. KATH, ECKHARDT HOLZ, MARC BORN: *Softwareentwicklung mit UML 2.0*, Addison-Wesley, ISBN 3-8273-2086-0, November 2003.

F. STOINSKI: *Towards a Component Model for Universal Data Streams*, Proceedings of ISCC 2003, Kemer/Antalya, Turkey, July 2003.

T. RITTER, FRANK STOINSKI: *Initial Submission to the Streams for CCM RFP*, OMG document, mars/03-10-02.

Vorträge

E. HOLZ: *A Meta-Model based Approach for the Combination of Models in multiple Languages*, IASTED MS-2003, Palm Springs, Februar 2003.

O. BECKER: *Streaming Transformations for XML – STX*, Workshop XMIDX , FU Berlin, Februar 2003.

O. BECKER: *Transforming XML on the Fly*, „XML Europe“, London, UK, Mai 2003.

O. BECKER: *Extended SAX Filter Processing with STX*, „Extreme Markup Languages 2003“, Montreal Kanada, August 2003.

FRANK STOINSKI: *Presentation on Initial Submission to the Streams for CCM RFP*, OMG Technical Meeting, London, November 2003, OMG document, mars/03-11-08.

Sonstige Aktivitäten

Prof. Dr. Joachim Fischer:

- Rapporteur der Question 24 und 25 und der Studiengruppe 17 der ITU-T
- Programmkomitee SDL-Konferenz
- Diverse Gutachtertätigkeiten
- Instituts- und Fakultätsratsmitglied
- Vorsitzender der Haushaltskommission
- Vorsitzender der Berufungskommission „Systemarchitektur“
- Mitglied der Standardisierungsgruppe der Deutschen Telekom AG

Dr. Eckhardt Holz:

- Programmkomitee SDL Konferenz 2003

Dipl.-Inf. Oliver Becker:

- Teilnahme am *Namespaces*-Arbeitstreffen des W3C in St. Augustin, Köln-Bonn, Februar 2003
- Teilnahme am Berliner Kolloquium „Total vernetzt - Szenarien einer informatisierten Welt“, Mai 2003
- Teilnahme am erweiterten Mitgliedertreffen des W3C, Erfurt, September 2003

Jan-Peter Bell:

- Mitglied in der Haushaltskommission
- DV-Verantwortlicher des Instituts

Gastvorträge

ANDREAS METZGER (Universität Kaiserslautern): *Automatisierung der Entwicklung reaktiver Systeme durch Einsatz eines Produktmodells*, Institutskolloquium, Januar 2003.

OLAF KATH (IKV++ TECHNOLOGIES AG): *Repositories for MDA*, Vortrag im Forschungsseminar »Objektorientierter Entwurf und Analyse«, Januar 2003.

Graduierungsarbeiten

RALF SCHRÖDER: *SDL Datenkonzepte – Analyse und Verbesserungen*, Dissertation, März 2003.

MARTIN VON LÖWIS: *Formale Semantik des Datentypmodells von SDL-2000*, Dissertation, Dezember 2003.

Diplomarbeiten

ANDREAS MUSKE: *Vergleich verschiedener Techniken zur Generierung von Deployment-Deskriptoren aus UML-Modellen*, Januar 2003.

RALF GERSTENBERGER: *ODEMx – Neue Lösungen für die Realisierung von C++ Bibliotheken zur Prozesssimulation*, Februar 2003.

ANNA HOVHANNISYAN: *Synchronisation mobiler Daten*, September 2003.

ELIO LUIS TORRES VALDIVIA: *Interpretation of ASN.1 values and types in a TTCN-3 Runtime Environment*, Oktober 2003.

Lehr- und Forschungseinheit
Theorie der Programmierung

<http://www.informatik.hu-berlin.de/top>

Leiter

PROF. DR. WOLFGANG REISIG
Tel.: (030) 2093 3065
e-mail: reisig@informatik.hu-berlin.de

Sekretariat

BIRGIT HEENE
Tel.: (030) 2093 3066
e-mail: heene@informatik.hu-berlin.de

Wissenschaftliche Mitarbeiter

DIPL.-INF. ADRIANNA ALEXANDER
PROF. DR. BODO HOHBERG
DIPL.-INF. AXEL MARTENS
PRIV.- DOZ. DR. KARSTEN SCHMIDT (SEIT 1.8.2003)

Technikerin

DIPL.-CHEM. BIRGIT SCHIEFNER

Promotionsstipendiatin

BIXIA WU
DAVID GULUA
BAVER ACU

Tutoren

DIRK FAHLAND
PETER MASSUTHE
CHRISTIAN STAHL

In zunehmendem Umfang wird Software nicht mehr von Programmierern aus ihrem intuitiven Verständnis einer Aufgabenstellung heraus geschrieben, sondern aus Spezifikationen generiert. Die Forschungsarbeiten des Lehrstuhls konzentrieren sich deshalb auf den modellbasierten Software-Entwurf, insbesondere auf Konzepte und Methoden zur Modellierung und Verifikation verteilter Systeme und verteilter Algorithmen. Abstrakte Modelle sind eine entscheidende Hilfe zur Steigerung der Effizienz des Entwurfs und der Entwicklung komplexer verteilter Software. Modelle können verschiedene Aspekte eines Systems herausstellen und dadurch seine Verifikation und Validierung unterstützen. Modelle unterstützen auch Implementierung, Test, Konfiguration und Strukturierung. Insgesamt soll damit der Entwicklungsprozess von Software vereinfacht und die Qualität der erzeugten Software erhöht werden. Die Arbeiten decken den gesamten Entwurfsprozess ab, von grundlegenden theoretischen Fragestellungen bis hin zu anwendungsnahen Entwurfstechniken, Software-Werkzeugen und Fallstudien in Kooperation mit Partnern der Software-Industrie.

Lehre

In der Lehre konzentriert sich die LFE auf zentrale Aspekte des modellbasierten Software-Entwurfs mit den regelmäßig angebotenen Vorlesungen über Verteilte Algorithmen sowie Methoden und Modelle des Systementwurfs, insbesondere für verteilte Systeme und web-basierte Geschäftsprozesse. Daneben stehen vertiefende Seminare zu diesen Themen. Im Sommersemester 2003 und in den Wintersemestern 2002/2003 und 2003/2004 hat die LFE folgende Veranstaltungen angeboten:

Veranstaltungen im Grundstudium

- Praktikum zur Praktischen Informatik 1/2 (B. HOHBERG/K. AHRENS, WS 2002/2003, SS 2003)
- Praktikum zur Praktischen Informatik 1/2 (B. HOHBERG/M. RITZSCHKE, WS 2002/2003)
- Übung zur Praktischen Informatik 1 (K. SCHMIDT, WS 2003/04)
- Proseminar „Beauty is our Business“ (W. REISIG, WS 2002/2003, SS 2003, WS 2003/2004)

Kernveranstaltungen (Halbkurse)

- Geschäftsprozessmodellierung (A. MARTENS, WS 2002/2003, WS 2003/2004)
- Web Services (A. MARTENS, SS 2003)
- Verteilte Algorithmen (W. REISIG, SS 2003)
- Methoden und Modelle des Systementwurfs (W. REISIG, WS 2002/2003, WS 2003/2004)

Seminare

- Seminar „Spezifikation und Verifikation verteilter Systeme II“ (A. ALEXANDER, WS 2001/2002, WS 2003/2004)
- Seminar „Komposition“ (A. ALEXANDER, SS 2003)
- Seminar „E-commerce“ (A. MARTENS, SS 2003)
- Seminar „Systementwurf“ (W. REISIG, WS 2003/2004)
- Seminar/ Forschungsseminar „Angewandte Verteilte Algorithmen“ (W. REISIG, WS 2002/2003, SS 2003, WS 2003/2004)

Forschung

Projekt: Computergestützte Verifikation

Ansprechpartner: Dr. K. SCHMIDT, DIPL.-INF. AXEL MARTENS,

Zusammenarbeit: Université Paris VI: FABRICE KORDON, Universität Stuttgart: Javier Esparza, DLR Berlin: DR. MICHAEL WEBER.

Die Forschungsarbeiten des Lehrstuhls konzentrieren sich auf Methoden zur Modellierung und Verifikation verteilter Systeme und verteilter Algorithmen. Entwickelte Verifikationsmethoden sollen durch Werkzeuge unterstützt werden - insbesondere durch Petrinetz-Werkzeuge. Folgende Werkzeuge werden am Lehrstuhl (weiter-)entwickelt.

LoLA - Explizite Zustandsraumverifikation für Petrinetze

Das Werkzeug *LoLA* unterstützt die Analyse verteilter Systeme, die als Petrinetz modelliert sind. Basierend auf Suchalgorithmen auf dem Zustandsraum des Systems können neben einer Reihe von Standardeigenschaften (z.B. Deadlocks, Lebendigkeit, Reversibilität) auch die Erfüllbarkeit, Lebendigkeit oder Invarianz frei definierbarer Zustandsprädikate verifiziert werden. Außerdem bietet *LoLA* einen Modelchecker für die temporale Logik CTL. *LoLAs*

Algorithmen verwenden und kombinieren state-of-the-art-Techniken zur Dämpfung der für verteilte Systeme typischen Zustandsraumexplosion.

Im Jahr 2003 wurde die Sweep-Line-Methode in LoLA integriert, zusammen mit einer automatisierten Berechnung einer dazu notwendigen Progress Measure. LoLA wurde integriert in den an der Universität Stuttgart entstehenden Model Checking Kit, eine Plattform zur Integration verschiedenster Modellierungs- und Verifikationswerkzeuge.

<http://www.informatik.hu-berlin.de/~kschmidt/lola.html>

WOMBAT4WS - Workflow Modeling and Business Analysis Toolkit for Web Services

Im Rahmen der Forschungen zu verteilten Geschäftsprozessen ist eine prototypische Implementierung entstanden - das *Workflow Modeling and Business Analysis Toolkit for Web Services* (kurz *WOMBAT4WS*).

Das Ziel war zum einen die Validierung der entwickelten Konzepte und Methoden: Mit WOMBAT4WS kann die *Bedienbarkeit* eines Web Service effektiv nachgewiesen werden. Darüber hinaus ermöglicht WOMBAT4WS die Transformation eines gegebenen Web Service in ein *Public-View-Modell*.

Zum anderen stellt dieser Prototyp den Ausgangspunkt für eine integrierte Entwicklungsumgebung für verteilte Geschäftsprozesse auf Basis von Web Services dar: Neben der Implementierung weiterer Analyseverfahren wird eine Integration von WOMBAT4WS in die universelle Werkzeug-Plattform *eclipse* angestrebt.

Petrinetz-Kern

Der Petrinetz-Kern (PNK) ist eine Infrastruktur zum Bau von Petrinetz-Werkzeugen. Er wird durch einen Parameter auf einen gewünschten Petrinetz-Typ (S/T-Netze, High-level, DAWN-Netze) eingestellt. Damit stellt der PNK für den gewünschten Petrinetz-Typ Basisfunktionen wie das Verwalten, Abfragen und Modifizieren von Petrinetzen zur Verfügung. Dadurch können Algorithmen über Petrinetze mit geringem Aufwand implementiert werden.

PetriNetz-MarkupLanguage

Die PetriNetz-MarkupLanguage (PNML) ist ein Vorschlag für ein XML-basiertes Austauschformat für Petrinetze. PNML sollte ursprünglich als Datenformat für die Java Version des PetriNetzKerns dienen. Es stellte sich aber heraus, dass auch verschiedene andere Forschungsgruppen an einem XML-basierten Austauschformat für Petrinetze arbeiteten. Deshalb ist die PNML nur ein Beitrag für weiterführende Diskussionen sowie die Bemühungen ein XML-basiertes Format zu standardisieren.

Projekt: Geschäftsprozessmodellierung

Ansprechpartner: PROF. DR. WOLFGANG REISIG, DIPL.-INF. AXEL MARTENS

Das Projekt wird von einer Vorlesung im Hauptstudium begleitet

Die Optimierung einzelner Aktivitäten durch den Einsatz moderner Rechentechnik ist in vielen Bereichen der Wirtschaft an ihre Grenze gestoßen. Der Schlüssel zu höherer Effizienz und Qualität liegt in den Geschäftsprozessen, die die einzelnen Aktivitäten in Beziehung setzen. Schlagworte wie Business Process Reengineering und Workflow Management sind daher auch für kleine und mittlere Unternehmen von großem Interesse. Um die Geschäftsprozesse eines Unternehmens analysieren, optimieren bzw. automatisieren zu können, müssen diese angemessen beschrieben werden. Im Rahmen der DFG-Forschergruppe Petrinetz-Technologie wurde die prinzipielle Eignung von Petrinetzen als

Modellierungsmethode für betriebliche Abläufe anhand industrieller Fallstudien nachgewiesen. In diesem Projekt werden die Methoden zur angemessenen Beschreibung von Geschäftsprozessen weiterentwickelt und auf spezielle Anforderungen angepasst.

Szenarien

In diesem Projekt liegt der Ansatz in einer zweistufigen Vorgehensweise. Im ersten Schritt wird die Interaktion zwischen den einzelnen Partnern spezifiziert. Mit Hilfe von Szenarien werden mögliche Abfolgen der Kommunikation beschrieben. Damit kann das externe Verhalten einer jeden Organisation zugesichert werden. Im zweiten Schritt werden die Geschäftsprozesse innerhalb der lokalen Einheiten modelliert. Die lokalen Abläufe können gemäß der eigenen Firmenphilosophie gestaltet und mit gängigen Methoden der Geschäftsprozessmodellierung beschrieben werden. Dadurch erreicht man ein hohes Maß an Flexibilität. Um die Zuverlässigkeit zu belegen, wird die Übereinstimmung von internem und externem Verhalten nachgewiesen.

Komposition

Soll ein Geschäftsprozess durch die Komposition lokaler abgeschlossener Komponenten realisiert werden, bietet sich die Technologie der *Web Services* an. Dabei muss natürlich garantiert sein, dass die einzelnen Web Services zusammenarbeiten können, ohne zu verklemmen. Da die neue Technologie noch in den Kinderschuhen steckt, ist die semantische Kompatibilität von Web Services eine offene Frage. Das Ziel dieses Teilprojektes ist es daher, mit Hilfe der *Prozessalgebra CSS* ein Kriterium für die Kompatibilität von Web Services abzuleiten und eine Methode für dessen Nachweis zu entwickeln.

Transformation

Prozesse werden von verschiedenen Blickwinkeln aus unterschiedlich abgebildet. Während für die Ausführung eines Prozesses alle Schritte in der internen Sicht (genannt *Private-View-Modell*) detailliert dargestellt werden müssen, genügt für die Beschreibung des Prozesses nach außen (genannt *Public-View-Modell*) die Definition der Schnittstellen und der grundlegenden Strukturen. Es ist wünschenswert, die Details der internen Sicht bei der Kommunikation mit Partnern auszublenden. Das Ziel dieses Teilprojektes ist es daher, automatisch aus der internen Prozessbeschreibung eine externe Prozessbeschreibung abzuleiten, die nur die notwendigen Details enthält.

Lastanalyse

Dieses Projekt erweitert das Betrachtungsfeld: Komplexe kooperierende Systeme im Bereich der Informationstechnologie existieren in den unterschiedlichsten Ausprägungen. Ihnen allen gemein ist die Aufteilung in Komponenten, die über Nachrichten miteinander kommunizieren. Daher stellt sich die Aufgabe, die zu Grunde liegende Kommunikationsschicht in Dimension und Robustheit an die gegebenen Anforderungen anzupassen. In diesem Projekt wird der Austausch von Nachrichten mit Hilfe von Message-Queues realisiert. Ziel dieses Teilprojektes ist es, ein Verfahren zur geeigneten Dimensionierung der Message-Queues in einem komplexen System auf Basis der gegebenen Flussgraphen zu entwickeln.

Projekt: Verteilte Geschäftsprozesse mit Web Services**Ansprechpartner:** DIPL.-INF. AXEL MARTENS, PROF. DR. WOLFGANG REISIG**Beteiligte Mitarbeiter:** DIRK FAHLAND, CARSTEN FRENKLER, THOMAS HEIDINGER, CHRISTIAN STAHL, DANIELA WEINBERG, STEPHAN WEISSLEDER**Zusammenarbeit:** PROF. DR. FRANK LEYMANN, IBM Deutschland, Labor Böblingen**Berater:** PROF. DR. HOLGER SCHLINGLOFF, DR. KARSTEN SCHMIDT

Dieses Projekt ist eng verbunden mit der Entwicklung eines Werkzeuges zu Modellierung und Analyse verteilter Geschäftsprozesse auf Basis von Web Services.

Seit Beginn der neunziger Jahre sind die *Geschäftsprozesse* in den Mittelpunkt der informationstechnischen Betrachtung von Unternehmen gerückt. Modellierung, Analyse, Reorganisation und technische Realisierung von Geschäftsprozessen - Tätigkeiten, die unter dem Begriff *Workflow-Management* zusammengefasst werden - haben innerhalb der Unternehmen zu einer Steigerung sowohl der Produktivität als auch der Flexibilität geführt.

Durch den globalen Wettbewerb und die daraus resultierende verstärkte internationale Ausrichtung von Unternehmen entstand die Notwendigkeit, Geschäftsprozesse über organisatorische wie auch geographische Grenzen hinweg miteinander zu verzahnen. Unter diesen Rahmenbedingungen sind die erfolgreichen Methoden des klassischen Workflow-Managements praktisch nicht anwendbar. Erfolgversprechender scheint der Ansatz, den dezentralen Strukturen durch eine dezentrale Lösung entgegen zu kommen: Die einzelnen Teile des Geschäftsprozesses werden zu lokal abgegrenzten Teilprozessen, die wie Bausteine zu einem *verteiltern Geschäftsprozess* zusammengesetzt werden.

Führende Unternehmen der Softwareindustrie (darunter IBM, Microsoft, Sun Microsystems u.a.) definieren zur Zeit eine standardisierte Architektur für die Entwicklung, Bereitstellung und Komposition verteilter Systeme auf Basis von *Web Services*. Ein Web Service ist eine lokal abgegrenzte Komponente, auf die über eine definierte Schnittstelle zugegriffen werden kann und die mit weiteren Web Services zu einem verteilten System zusammengefügt wird. Ein verteilter Geschäftsprozess lässt sich somit durch die Komposition von Web Services implementieren.

Dieses Projekt beschäftigt sich mit der *Modellierung und Analyse* verteilter Geschäftsprozesse auf Basis von Web Services. Besonderes Augenmerk wird dabei auf die methodische Unterstützung der einzelnen Entwicklungsschritte gelegt. Dazu zählen insbesondere die systematische Ableitung eines Web Service und die Entscheidung, ob zwei gegebene Web Services zueinander kompatibel oder äquivalent sind. Die Arbeiten gruppieren sich in drei Teilprojekten:

Semantik

In diesem Teilprojekt werden die vorhandenen Standards analysiert und die Anforderungen an eine systematische Entwicklungsmethode abgeleitet. Mit der Definition der Sprache *BPEL4WS (Business Process Execution Language for Web Services)* wurde ein Rahmen vorgegeben, der jedoch noch etliche Unschärfen enthält. Für die Unterstützung der Analyse mit formalen Methoden ist es notwendig, jedem Konstrukt der Sprache eine genaue Bedeutung zu geben. In diesem Teilprojekt nähern wir uns dieser Sprache von verschiedenen Blickwinkeln aus und entwickeln so eine strukturierte, umfassende und dennoch verständliche Semantik für BPEL4WS.

Graphische Oberfläche

Es gibt derzeit keine einheitliche Darstellung eines in BPEL4WS notierten Prozesses. Ein kurzer Blick in das Internet unter dem Gesichtspunkt, mehr über Geschäftsprozessmodellierungssprache BPEL4WS zu erfahren, zeigt eine Flut verschiedenster graphischer Darstellungen. In diesem Teilprojekt wollen wir eine Möglichkeit schaffen, jeden in BPEL4WS notierten Prozess schnell und präzise graphisch darzustellen. Diese Graphik soll dabei einfach erfassbar sein, ohne dass der Leser die Sprache BPEL4WS selbst kennen muss. Neben der statischen Darstellung von Prozessen, sind die graphische Spezifikation von Eigenschaften und die Visualisierung von Zuständen und Abläufen Ziele in diesem Teilprojekt.

Analyse

Für die wesentlichen Ziele dieses Projekts, insbesondere für den effizienten Nachweis globaler Eigenschaften verteilter Geschäftsprozesse, müssen Analysemethoden auf einzelne Web Services herunter gebrochen werden. Das ist schwierig, weil ein Web Service ein offenes System ist, dessen Verhalten maßgeblich von seiner Umgebung bestimmt wird.

Eigene Vorarbeiten haben gezeigt, dass sich für abstrakte Modelle auch ohne Kenntnis der Umgebung lokale Modellierungsfehler finden, beseitigen oder sogar vermeiden lassen. Hierbei spielt der Begriff der *Bedienbarkeit* eines Web Services eine zentrale Rolle: Ein Web Service heißt bedienbar, wenn es mindestens einen anderen Web Service gibt, so dass das komponierte System ein vernünftiger verteilter Geschäftsprozess ist. In diesem Teilprojekt wird der Begriff der Bedienbarkeit substantiell ausgebaut, verallgemeinert und algorithmisch unterlegt. Darüber hinaus sollen weitere Methoden zur Analyse von Web Services entwickelt werden.

Task Force BPEL4WS

Für die Koordinierung der Teilprojekte wurde am Lehrstuhl eine *Task Force* gebildet, die konkrete Arbeitsschritte plant und mit Vertretern von IBM abstimmt. Erste Vorüberlegungen liegen vor. Studienarbeiten, Diplomarbeiten und weiterführende Forschungsarbeiten werden im Rahmen des Projektes auf die einzelnen Teilnehmer hin definiert.

Projekt: Ausdruckstärke von Abstract State Machines

Ansprechpartner: PROF. DR. WOLFGANG REISIG, DIRK FAHLAND

Um den Arbeits- und Kostenaufwand für den Entwurf und die Implementierung komplexer Softwaresysteme zu senken, wird üblicherweise die Korrektheit eines Systems zunächst mit theoretischen Verifikationsmethoden geprüft. Dafür muss das System erst auf einer geeigneten Ebene spezifiziert werden. Da insbesondere verteilte Systeme meist sehr komplex sind und gleichzeitig immer mehr an Bedeutung gewinnen, sind Spezifikationstechniken speziell für verteilte Systeme von besonderem Interesse. Zu den neueren Methoden zählen Lamports *Temporal Logic of Actions* und Gurevichs *Abstract State Machines*.

Im Projekt werden Entwurfs- und Analysetechniken entwickelt, die spezifische Vorteile vorhandener Techniken ausnutzen und so kombinieren, dass sie für spezielle Anwendungsgebiete, insbesondere spezielle verteilte Systeme, besonders nützlich sind. Dabei werden kausale Zusammenhänge zwischen Aktionen in Abläufen explizit modelliert und in der Analyse ausgenutzt. Das Projekt zielt letztendlich ab auf eine Spezifikationstechnik, die die Flexibilität von ASM, die kanonischen Ausdrucksmittel für Synchronisation und Kausalität von Petrinetzen und die Deduktionskraft von TLA kombiniert.

Projekt: Temporale Logik für kausal geordnetes Verhalten

Ansprechpartner: DIPL.-INF. ADRIANNA ALEXANDER, PROF. DR. WOLFGANG REISIG

Beteiligter Mitarbeiter: PETER MASSUTHE

Temporale Logik hat sich seit Beginn der achtziger Jahre zunehmend als ein geeignetes Werkzeug zur Spezifikation und Verifikation verteilter Systeme durchgesetzt. Die Idee, ein verteiltes System durch eine logische Formel zu repräsentieren, stammte ursprünglich von Amir Pnueli und wurde dann in verschiedenen Formalismen angewandt. Einer davon ist die Temporal Logic of Actions (kurz: TLA) von Leslie Lamport.

Im Rahmen des Projektes ist eine neue Temporale Logik entstanden, Temporal Logic of Distributed Actions (kurz: TLDA). Im Gegenteil zu TLA basiert sie auf kausal (halb-)geordneten Ereignissen. Dies erlaubt eine feinere Unterscheidung der Phänomene, die im zu spezifizierenden System auftreten, wie z.B. Nebenläufigkeit und Nichtdeterminismus. Die Komponierbarkeit der Spezifikationen bleibt dabei erhalten.

Veröffentlichungen

A. ALEXANDER, W. REISIG: *Logic of Involved Variables - System Specification with Temporal Logic of Distributed Actions*. Proc. of the 3rd International Conference on Application of Concurrency to System Design (ACSD'03), Guimaraes, Portugal, S. 167 - 176, 2003.

L. M. KRISTENSEN, K. SCHMIDT UND A. VALMARI: *Question Guided Stubborn Set Methods for State Properties*. Submitted to a journal.

A. MARTENS, CH. STAHL, D. WEINBERG, D. FAHLAND, TH. HEIDINGER: *Business Process Execution Language for Web services – Semantik, Analyse und Visualisierung*. Informatik-Berichte, Humboldt-Universität zu Berlin, erscheint demnächst.

A. MARTENS: *Analysis and re-engineering of Web Services*. Tagungsband: 6th International Conference on Enterprise Information Systems (ICEIS'04). Porto, Portugal, erscheint 2004.

A. MARTENS: *On Usability of Web Services*. Tagungsband: 1st Web Services Quality Workshop (WQW 2003). Rome, Italy, 2003.

A. MARTENS: *Compatibility of Web Services*. Tagungsband: 10th Workshop on Algorithms and Tools for Petri Nets (AWPN 2003). Eichstätt, Germany, 26.-27. September 2003.

A. MARTENS. *On Compatibility of Web Services*. Petri Net Newsletter, Nr. 65, S. 12-20, Oktober 2003.

A. MARTENS: *Verteilte Geschäftsprozesse - Modellierung und Verifikation mit Hilfe von Web Services*. Dissertation, Institut für Informatik, Humboldt-Universität zu Berlin, erscheint 2004.

W. REISIG: *On Gurevich's Theorem on Sequential Algorithms*. Acta Informatica 39, S. 273 - 305, Springer 2003.

W. REISIG: *The Computable Kernel of ASM*. In: Abstract State Machines 2003 (Börger, Gargantini, Riccobene Hrsg.), LNCS 2589, S. 421-422, Springer 2003.

W. REISIG: *The Expressive Power of Abstract-State Machines*. Computing and Informatics, Vol. 22, S. 209-219, 2003.

K. SCHMIDT: *Automated Generation of a Progress Measure for the Sweep-Line Method*. Accepted for TACAS 2004.

JOSÉ M. VIDAL, P. BUHLER, UND CH. STAHL. *Multiagent systems with workflows*. IEEE Internet Computing, 8(1), S. 76-82, Januar 2004.

H. WEBER, H. EHRIG, W. REISIG u.a.: The Petri Net Baukasten of the DFG Forschergruppe Petri Net Technology. In: H. EHRIG, W. REISIG, G. ROZENBERG, H. WEBER (Hrsg.): Petri Net Technology for Communication-Based Systems. LNCS 2472, Springer 2003.

Vorträge

A. ALEXANDER: *Logic of Involved Variables - System Specification with Temporal Logic of Distributed Actions*. 3rd International Conference on Application of Concurrency to System Design (ACSD 2003), Guimaraes, Portugal, Juni 2003.

A. MARTENS: *Bedienbare Web Services - Eine Dissertation zwischen Theorie & Praxis*. Klausurtagung Luhme IX, Groß Dölln, Mai 2003.

A. MARTENS: *Verteilte Geschäftsprozesse - Modellierung und Verifikation mit Web Services*. Promotionsverteidigung, Institut für Informatik, Humboldt-Universität zu Berlin, Juli 2003.

A. MARTENS: *Compatibility of Web Services*. 10th Workshop on Algorithms and Tools for Petri Nets (AWPN 2003). Eichstätt, Germany, September 2003.

A. MARTENS: *On Usability of Web Services - Modeling and Verification by help of Petri Nets*. Eingeladener Vortrag im IBM Labor Böblingen, November 2003.

A. MARTENS: *Verteilte Geschäftsprozesse - Modellierung und Verifikation mit Web Services*. Eingeladener Vortrag an der Universität Ulm, November 2003.

A. MARTENS: *On Usability of Web Services*. 1st Web Services Quality Workshop (WQW 2003). Rome, Italy, Dezember 2003.

A. MARTENS: *On Usability of Web Services - Modeling and Verification by help of Petri Nets*. Eingeladener Vortrag im IBM T.J. Watson Research Center, Hawthorne (NY), Dezember 2003.

W. REISIG: *Petrinetze und andere Vorschläge zur Fassung des "Implementierbaren"*. Kolloquium, Universität Frankfurt/ Main, Januar 2003.

W. REISIG: *The syntactical kernel of ASM*. 10th International Workshop, ASM 03 Taormina, Italien, März 03.

W. REISIG: *Modellbasierter Software-Entwurf für Web-basierte Geschäftsprozesse*. Labor Böblingen IBM Deutschland Entwicklung GmbH, Mai 2003.

W. REISIG: *Modellierung - Quo Vadis?* Modellierung 2003, Villach, Kärnten, Februar 2003.

W. REISIG: *Clocks, Mutex, and the Update of Variables*. Leslie Lamport Symposium, Universität Kiel, Juli 2003.

W. REISIG: *Tutorium: Informal Introduction to Petri Nets*. 4. Advanced Course of Petri Nets, Kath. Universität Eichstätt, September 2003.

W. REISIG: *Tutorium: Distributed algorithms*. 4. Advanced Course of Petri Nets, Kath. Universität Eichstätt, September 2003.

W. REISIG: *Verteilte Geschäftsprozesse und andere Modellierungsaufgaben*. Siemens AG, September 2003.

W. REISIG: *Busines Processes: Principles of Modelling and Implementation*. Deutsch-Chinesischer Workshop über Theoretische Informatik (CDWTI), Shanghai, China, Oktober 2003.

K. SCHMIDT: *Supporting Explicit State Space Verification by Transition Invariants*. Theorietag der FG Automaten und Formale Sprachen der GI in Herrsching bei München, September 2003.

K. SCHMIDT: *Supporting Explicit State Space Verification by Transition Invariants*. Kolloquium an der Universität Oldenburg, Oktober 2003.

K. SCHMIDT: *Automated Generation of a Progress Measure for the Sweep-Line Method*. Workshop AWPN, Eichstätt, September 2003.

K. SCHMIDT: *Computergestützte Verifikation einfach strukturierter Eigenschaften*. Kolloquium an der Universität Stuttgart, Dezember 2003.

Sonstige Aktivitäten

Prof. Dr. Wolfgang Reisig

- Mitglied im Steering-Committee „Conference on Application and Theory of Petri Nets“
- Course Director mit Jörg Desel, Grzegorz Rozenberg, 4. Advanced Course of Petri Nets , Eichstätt, September 2003
- Mitglied im Programmkomitee "Conference International Conference on ComputerSystems and Applications" ACS/IEEE ,2003, Tunisia
- Vice-chair Workshop "Distributed Algorithms" EuroPar'2003, Klagenfurth
- Mitglied der EG-Arbeitsgruppe „Formal Methods Europe“
- Gutachter für Graduiertenkollegs und Schwerpunktprogramme der DFG
- Gutachter für verschiedene Zeitschriften, Fachtagungen, DFG und NSF
- Editor für das "Journal on Software and Systems Modeling", (SoSym) Springer-Verlag
- Vertreter der HU Berlin im Fakultätentag Informatik
- Geschäftsführender Direktor seit Juli 2002
- Mitglied der Aufnahmekommission des Fakultätentages Informatik
- Mitglied des Beirates des BMFT-Projektes MMiSS (Multimedia-Instruktionen in Sicheren Systemen)

Dipl.-Inf. Adrianna Alexander

- Gutachterin für SEFM 2003

Prof. Dr. Bodo Hohberg

- Studienfachberatung Informatik
- Mitglied des Prüfungsausschuss des Instituts für Informatik

Dipl.-Inf. Axel Martens

- Organizing Chair des 1. Workshops: Business Process Execution Language for Web Services. Institut für Informatik, Humboldt-Universität zu Berlin, 22. Juli 2003
- Organizing Chair des Workshops "Technologies for e-services - TeS'03" im Rahmen der Tagung VLDB, Humboldt-Universität zu Berlin, 7./8. September 2003
- Organizing Chair des 2. Workshops: Business Process Execution Language for Web Services. Institut für Informatik, Humboldt-Universität zu Berlin, 10. September 2003
- Teilnahme 2. GI-Workshop "EPK 2003 - Geschäftsprozessmanagement mit Ereignis-gesteuerten Prozessketten", Bamberg, 08. Oktober 2003

- Teilnahme 11. Fachtagung "Modellierung betrieblicher Informationssysteme MobIs 2003" 9. - 10. Oktober 2003, Bamberg

Dr. Karsten Schmidt

- Gutachter im Promotionsverfahren von T. Junttila (Helsinki Univ. Technology)
- Mitglied in den Promotionskommissionen für T. Mailund (Univ. Aarhus), L.S. Shi (Humboldt-Universität)
- Editor für das Springer Journal of Universal Computer Science (J.UCS)
- Mitglied im örtlichen Wahlvorstand des Instituts für Informatik
- Mitglied der Leitung der FG "Petrietze und verwandte Systemmodelle" der GI
- Gutachten, u.a. für ICATPN 2004, FOSSACS 2004

Dissertationen

AXEL MARTENS: *Verteilte Geschäftsprozesse – Modellierung und Verifikation mit Hilfe von Web Services*. Promotion. Humboldt-Universität zu Berlin, August 2003.

Lehr- und Forschungseinheit

Softwaretechnik

<http://www.informatik.hu-berlin.de/institut/struktur/softwaretechnikII/>

Leiter

PROF. DR. KLAUS BOTHE

Tel.: (030) 2093 3007

E-mail: bothe@informatik.hu-berlin.de

Sekretariat

ULRIKE SCHOLZ

Tel.: (030) 2093 3008

E-mail: uscholz@informatik.hu-berlin.de

Wissenschaftlicher Mitarbeiter

DIPL.-INF. KAY SCHÜTZLER

Techniker

DIPL.-PHYS. ULRICH SACKLOWSKI

Tutoren

YVONNE GABRIEL

ANDREAS HIRTH

SAM JOACHIM

SKANDER MORGENTHALER

LUBOMIRA STANTCHEVA

TOBIAS THIEL

ANDREAS WENZEL

Die gegenwärtigen Arbeitsschwerpunkte der Gruppe liegen auf dem Gebiet der Softwaretechnik, insbesondere im Reverse Engineering, beim Software-Test sowie dem Projektmanagement. Dabei steht ein gemeinsames Projekt mit dem Institut für Physik der Humboldt-Universität zu Berlin auf dem Gebiet des Software Reverse Engineering im Mittelpunkt. Zur Ergänzung dieses Projekts wurde mit einer multimedialen Aufbereitung des Anwendungsbereichs begonnen.

Die Aktivitäten der Gruppe bei der Kooperation mit Universitäten in Südosteuropa wurden aufgrund des DAAD-Sonderprogramms "Akademischer Neuaufbau Südosteuropa" weiter ausgebaut.

Lehre

Veranstaltungen im Grundstudium

- Übungen „Praktische Informatik 1“ (K. SCHÜTZLER, WS 2003/2004)

Veranstaltungen im Hauptstudium

- Projektveranstaltungen „Software-Sanierung“ (K. BOTHE, WS 2002/2003, SS 2003)
- Projektveranstaltung „Visuelle Programmierung“ (K. SCHÜTZLER, WS 2002/2003)
- Halbkurs „Software Engineering“ (K. BOTHE, SS 2003, WS 2003/2004)
- Übungen zum HK „Software Engineering“ (K. SCHÜTZLER, SS 2003, WS 2003/2004)

Forschung

Projekt: Reverse Engineering für Steuerungssoftware in der Röntgenbeugung

Anrechnerpartner: PROF. KLAUS BOTHE

Zusammenarbeit: Institut für Physik der Humboldt-Universität zu Berlin

Am Institut für Physik der Humboldt-Universität ist ein Steuerprogramm entwickelt worden, das (trotz Nutzung von C++ als Implementationssprache) aufgrund einer ungünstigen SW-Architektur nicht mehr bzw. nur mit größerem Aufwand wartbar war. Das Ziel besteht in der Analyse sowie einer Restrukturierung der Software, um eine deutliche Erhöhung der Wartbarkeit zu erreichen.

Im Jahr 2003 wurden folgende Ergebnisse erzielt:

- Arbeiten an verschiedenen Subsystemen zur Abgrenzung und zur Analyse sowie Erweiterung ihrer Funktionalität
- Entwicklung eines Tools zur Ermittlung von Subsystemschnittstellen
- Erweiterung der Software um gewünschte Funktionalität (Automatische Protokollierung)
- Portierung der Software in eine moderne Entwicklungsumgebung und auf ein neueres Betriebssystem
- Erweiterung der umfangreichen Testfallsammlung für den Regressionstest
- Ausbau der Testumgebung ATOS (automatisches System für den Regressionstest)

Während des Berichtszeitraumes entstanden eine Studienarbeit und vier Diplomarbeiten (s.u.).

Projekt: Software Engineering – Education and Research Cooperation

Ansprechpartner: PROF. KLAUS BOTHE

Zusammenarbeit: Universität Novi Sad (Ser), Universität Skopje (Maz), Universität Plovdiv (Bul), Universität Kragujevac (Ser), Universität Nis (Ser), Universität Belgrad (Ser), Universität Zagreb (Kro), Universität Timisoara (Rum), Universität Tirana (Alb), Universität Sarajevo (Bos-Herz), Banja Luka (Bos-Herz)

Forschungsförderung: DAAD-Sonderprogramm „Akademischer Neuaufbau Südosteuropa“ im Rahmen des „Stabilitätspakt Südosteuropa“

Die Ziele unseres vom DAAD geförderten Projekts betreffen:

- Distance Education: Aufbau eines Internet-gestützten gemeinsamen Kurses „Software Engineering“
- Projektmanagement für die kooperative Entwicklung von Kursmaterialien
- Verteiltes Reverse Engineering eines Softwaresystems in der Experimentalphysik.

Zu den Projektaktivitäten zählen u. a.

- Gastaufenthalte südosteuropäischer Wissenschaftler in Berlin:
Prof. Katerina Zdravkova, Ph.D., Skopje
Prof. Zoran Budimac, Ph.D., Novi Sad
Dragan Bojic, Ph.D., Belgrad
Prof. Stanimir Stoyanov, Plovdiv
- Verbesserung der materiellen Bedingungen an den Partneruniversitäten (Hardware, Software, Literatur)
- Koordinationstreffen in Berlin: Vorbereitung des diesjährigen Workshops
- Workshop: Software Engineering Education and Reverse Engineering, Ohrid, 25. bis 30. August 2003

Auch in diesem Jahr stellte der Workshop unter Einbeziehung von Teilnehmern aus mittlerweile zwölf Universitäten aus sieben Ländern den Höhepunkt der Projektaktivitäten dar, wobei es schwerpunktmäßig um den Entwurf und die Organisation eines gemeinsamen Kurses im Gebiet Software Engineering und die Integration der in diesem Jahr neu dazugekommenen Teilnehmer ging. Dem Workshop ging diesmal aufgrund der wachsenden Projektgröße und Teilnehmerzahl ein Koordinationstreffen der Kernmitglieder des Projekts in Berlin voraus.

Projekt: XCTL-Projekt: Multimediale Präsentation

Ansprechpartner: DIPL.-PHYS. ULRICH SACKLOWSKI

Zusammenarbeit: Institut für Physik der Humboldt-Universität zu Berlin

Forschungsförderung: Multimedia-Förderprogramm 2003 "Neue Medien in Lehre und Forschung"

Grundlage ist das Projekt „Reverse Engineering für Steuerungssoftware in der Röntgenbeugung“. Der Anwendungsbereich dieses Projektes ist, bedingt durch seine Komplexität und Kompliziertheit, schwer zugänglich für Physik- und Informatikstudenten und für Partner aus Südost-Europa.

Dieses Defizit soll durch eine multimediale Aufbereitung von Projektteilen behoben werden. Gegenstand hierbei sind physikalische Grundlagen, Messabläufe und Messergebnisse. Sie werden mit multimedialen Mitteln wie Film, Animation, Fotos und Audio aufbereitet.

Im Jahr 2003 wurden folgende Ergebnisse erzielt:

- Multimediale Präsentation von Topographie-Messabläufen in englischer Sprache
- Erarbeitung einer System-Architektur

Projekt: Reengineering für Steuerungssoftware in der Photovoltaik

Ansprechpartner: DIPL.-INF. KAY SCHÜTZLER

Zusammenarbeit: Hahn-Meitner-Institut, Abt. Silizium-Photovoltaik, Berlin

In der Adlershofer Abteilung des Hahn-Meitner-Instituts wird eine selbstentwickelte Software zur Steuerung wichtiger Versuche in der Grundlagenforschung eingesetzt. Aufgrund einer anstehenden Erneuerung der Steuerungscomputer entstand der Wunsch, die vorliegende Software in Form einer Reimplementation in einer aktuellen visuellen Programmiersprache zu überarbeiten.

Im Jahr 2003 konnte auf der Grundlage der letztjährig ermittelten Anforderungen die Reimplementation erfolgreich durchgeführt werden. Die neuentwickelte Software wird mit Beginn des nächsten Jahres anstelle der alten Lösung eingesetzt werden.

Veröffentlichungen

S. LÜTZKENDORF, K. BOTHE: *Attributierte Klassifikationsbäume zur Testdatenbestimmung*. Treffen der GI-FG Test, Analyse und Verifikation von Software, Köln, Februar 2003, Softwaretechnik-Trends, Band 23, Heft 1.

J. HANISCH, J. LETZEL, K. BOTHE: *Ein Werkzeugsystem zur Automatisierung von GUI-Tests*. Treffen der GI-FG Test, Analyse und Verifikation von Software, Köln, Februar 2003, Softwaretechnik-Trends, Band 23, Heft 1.

K. SCHÜTZLER, T. THIEL: *Automatisierte Ermittlung von Subsystemschnittstellen*. 5. Workshop Software Reengineering, Bad Honnef, Mai 2003, Softwaretechnik-Trends, Band 23, Heft 2.

K. ZDRAVKOVA, K. BOTHE, Z. BUDIMAC: *SETT-Net: A Network for Software Engineering Training and Teaching*. ITI, Information Technology Interfaces, Cavtat, Kroatien, Juni 2003.

K. ZDRAVKOVA, K. BOTHE, Z. BUDIMAC: *The structure of SETT-Net*. Eurocon 2003, Ljubljana, Slowenien, September 2003.

K. BOTHE, K. SCHÜTZLER, Z. BUDIMAC, K. ZDRAVKOVA, D. BOJIC, S. STOYANOV: *Technical and Managerial Principles of a Distributed Cooperative Development of a Multi-Lingual Educational Course*. Proceedings of the 1st Balkan Conference in Informatics, Thessaloniki, Griechenland, November 2003.

Vorträge

K. BOTHE: *Three years of cooperation under the auspices of stability pact for SEE*. Ohrid, August 2003.

K. BOTHE: *The joint course on software engineering: past present, and future*. Ohrid, August 2003.

K. BOTHE: *On the structure of the course on software engineering*. Ohrid, August 2003.

K. BOTHE and Z. BUDIMAC: *Review Reports: A means to collect experience and feedback*. Ohrid, August 2003.

K. BOTHE: *Update policy in the joint course on software engineering*. Ohrid, August 2003.

K. SCHÜTZLER, U. SACKLOWSKI and K. BOTHE: *Progression of the XCTL system*. Ohrid, August 2003.

K. SCHÜTZLER: *Configuration management*. Ohrid, August 2003.

K. SCHÜTZLER and Z. BUDIMAC: *Assignments in a software engineering course*. Ohrid, August 2003.

U. SACKLOWSKI: *Visualization of the use of XCTL by multimedia means*. Ohrid, August 2003.

S. JOACHIM, K. BOTHE: *Current trends in e-learning and their impact on the JCSE*. Ohrid, August 2003.

Z. PUTNIK, Z. BUDIMAC, K. BOTHE: *International Educational Project – Some Experiences*. Conference on Computers in Education - RUN 03. Workshop on Teaching of Informatics "NaRa 03", Novi Sad, September, 2003.

Z. BUDIMAC, K. BOTHE: *Slides – Style and Usage Guidelines – Experiences from an International Project*. Conference on Computers in Education - RUN 03. Workshop on Teaching of Informatics "NaRa 03", Novi Sad, September, 2003.

K. BOTHE (presentation Z. BUDIMAC): *Invited tutorial: What is Software Engineering?*. Conference on Computers in Education - RUN 03. Workshop on Teaching of Informatics "NaRa 03", Novi Sad, September, 2003.

Wissenschaftliche Kooperationen

- DaimlerChrysler AG, Forschung und Technologie,
- Institut für Physik der Humboldt-Universität zu Berlin,
- Hahn-Meitner-Institut, Abteilung Silizium-Photovoltaik
- Universität Novi Sad (Serbien-Montenegro)
- Universität Skopje (Mazedonien)
- Universität Plovdiv (Bulgarien)
- Universität Kragujevac (Serbien-Montenegro)
- Universität Nis (Serbien-Montenegro)
- Universität Belgrad (Serbien-Montenegro)
- Universität Zagreb (Kroatien)
- Universität Timisoara (Rumänien)
- Universität Tirana (Albanien)
- Universität Sarajevo (Bosnien-Herzegowina)
- Universität Banja Luka (Bosnien-Herzegowina)

Gastwissenschaftler

KATERINA ZDRAVKOVA (Universität Skopje), Mai 2003

ZORAN BUDIMAC (Universität Novi Sad), Januar 2003, Mai 2003, September 2003

DRAGAN BOJIC (Universität Belgrad), Mai 2003

STANIMIR STOYANOV (Universität Plovdiv), Mai 2003

Sonstige Aktivitäten

Prof. Dr. Klaus Bothe

- Vorsitzender des Prüfungsausschusses des Instituts für Informatik
- Vorsitzender der Bibliothekskommission des Instituts für Informatik

Dipl.-Inf. Kay Schützler

- Mitglied der Kommission Lehre und Studium am Institut für Informatik

Dipl.-Phys. Ulrich Sacklowski

- Mitglied der Haushaltskommission des Instituts für Informatik

Diplomarbeiten

JENS KLIER: *Erweiterung eines Software-Altsystems um eine Komponente zur Automatisierung manueller Protokollierungsaufgaben.* Juni 2003.

RENÉ HARDER, ALEXANDER PASCHOLD: *Portierungsstrategie für ein Hardware-Steuerungsprogramm unter Anwendung von Reverse-Engineering-Techniken.* August 2003.

JAN PICARD: *Softwaresanierung durch Refactoring: Untersuchung von Methoden zur Verbesserung der Softwarequalität in einer Komponente eines Softwaresystems zur Kristallstrukturanalyse.* September 2003.

BERT SCHULZKI: *Techniken der Agilen Software-Entwicklung in Web-Projekten.* November 2003.

Lehr- und Forschungseinheit
Datenbanken und Informationssysteme

<http://www.dbis.informatik.hu-berlin.de>

Leiter

PROF. JOHANN-CHRISTOPH FREYTAG, PH.D.
Tel.: (030) 2093 3009
E-mail: freyttag@dbis.informatik.hu-berlin.de

Sekretariat

ULRIKE SCHOLZ
Tel.: (030) 2093 3008
E-mail: uscholz@informatik.hu-berlin.de

Wissenschaftliche Mitarbeiter

DIPL.-ING. DMITRI ASONOV
DIPL.-INF. CHOKRI BEN NECIB
DR. RAINER ECKSTEIN
DIPL.-INF. RALF HEESE
DIPL.-INF. SVEN HERSHEL
DR. STEPHAN HEYMAN
DIPL.-INF. HEIKO MÜLLER
PETER RIEGER
DIPL.-INF. DIETER SCHEFFNER
DIPL.-INF. (FH) KATJA THAM

Techniker

DIPL.-ING. HEINZ WERNER

Tutoren

JANA BAUCKMANN
ASHRAF EMAWI
TIMO MIKA GLÄSSER
MIKE HERZOG
MIRKO HERBIG
MARTIN KOST
FRANK LEGLER
MAREN LENK
CLAAS REIM
DIRK REMPE
HENNING ZAHN
GREGOR ZEITLINGER
MARKO BRIESEMAN

Für die Lehr- und Forschungseinheit Datenbanken und Informationssysteme (DBIS) war die Organisation der Very Large Databases (VLDB)-Konferenz zusammen mit der Organisation eines weiteren Workshops das absolute Highlight des Jahres. Zusammen mit der Organisation eines weiteren Workshops hat sich DBIS national und international stark engagiert und gezeigt, dass es gemeinsam mit anderen solch ein Großereignis (mehr als 600

Teilnehmer) erfolgreich gestalten und durchführen kann. Darüber hinaus hat DBIS auch in diesem Jahre in den verschiedenen Projekten den Forschungsfortschritt dokumentiert. Die Veröffentlichung des Buches „XML und Datenmodellierung – XML-Schema und RDF zur Modellierung von Daten und Metadaten einsetzen“ durch Herrn Dr. Eckstein als Mitautor ist als ein weiterer Höhepunkt hervorzuheben.

In der Lehre und Forschung hat DBIS weiterhin seine Arbeiten in den Bereichen XML, Kerndatenbanken und Bioinformatik/Life Science konsequent weiter vorangetrieben und trägt damit den aktuellen Herausforderungen in den Bereichen Datenbanken, Informationssysteme und angrenzenden Gebieten (wie dem Bereich E-Commerce und dem WWW) Rechnung. Diese Entwicklung spiegelt sich auch in der Einrichtung weiterer neuer und neu gestalteter Lehrveranstaltungen wider, die zeitnah Studenten mit den neuen Tendenzen in Forschung und Industrie vertraut machen sollen.

Die in diesem Bericht beschriebenen Arbeiten reflektieren auch die intensive Zusammenarbeit der Mitglieder der Forschungsgruppe mit lokalen, nationalen und internationalen Partnern. Neben dem Graduiertenkolleg „Verteilte Informationssysteme“ (GRK 316), gefördert durch die Deutsche Forschungsgemeinschaft, ist ein wesentlicher Anteil der Arbeiten in Zusammenarbeit mit industriellen Partnern entstanden. DBIS beteiligt sich außerdem an dem neuen, BMBF-geförderten Projekt „InterVal“ mit zwei Teilprojekten, das an allen drei Berliner Universitäten Partner hat. Des Weiteren konnte die Zusammenarbeit mit der Firma IBM Deutschland Entwicklung weiterentwickelt und ausgebaut werden. Für sein Engagement im Bereich der Genom-Forschung wurde dem Leiter der Lehr- und Forschungseinheit ein weiteres Mal (nun zum vierten Mal!) von der Firma IBM der Faculty-Award verliehen. Schließlich ist hervorzuheben, dass sich die Lehr- und Forschungseinheit mit der Ausrichtung eines Dagstuhl-Workshops im Bereich „Bioinformatik“ und einer Konferenz sowie zweier Workshops im Bereich „XML“ sowie eines DB2-Zertifizierungskurses für die Studenten (in Kooperation und mit Unterstützung der IBM Deutschland Entwicklung GmbH) - nun schon zum zweiten Mal - besonders engagiert hat.

Lehre

Veranstaltungen im Grundstudium

- Einführung in die Programmiersprache C (PROF. J.C. FREYTAG, WS 2003/04)
- Praktische Informatik III (Compilerbau) (PROF. J.C. FREYTAG, WS 2003/04)
- Praktikum zur Praktischen Informatik III (DR. R. ECKSTEIN, R. HEESE, WS 2003/04)

Kernveranstaltungen (Halbkurse) im Hauptstudium

- Grundlagen von Datenbanken (DBS I) (PROF. J.C. FREYTAG, WS 2002/03)
- Praktikum zu DBS I (C. BEN NECIB, K. THAM, S. HERSCHEL, WS 2002/03)
- Implementation von Datenbanksystemen (DBS II) (PROF. J.C. FREYTAG, SS 2003)
- Praktikum zu DBS II (C. BEN NECIB, D. SCHEFFNER, SS 2003)
- Grundlagen von Datenbanken (DBS I) (PROF. J.C. FREYTAG, WS 2003/04)
- Praktikum zu DBS I (C. BEN NECIB, S. HERSCHEL, K. THAM WS 2003/04)

Spezialveranstaltungen im Hauptstudium

- Ringvorlesung des Instituts: Schwerpunkte der Informatik (PROF. J.-C. FREYTAG, SS 2003)
- Bioinformatik I (PROF. J.-C. FREYTAG, WS 2002/03)
- Projekt zur Bioinformatik I (PROF. J.-C. FREYTAG, WS 2002/03)

- Bioinformatik II (PROF. J.-C. FREYTAG, DR. S. HEYMANN, SS 2003)
- Projekt zur Bioinformatik II (PROF. J.-C. FREYTAG, P. RIEGER, SS 2003)
- Methoden der Integration heterogener Datenbestände (DR. R. ECKSTEIN, WS 02/03)
- XML und Datenmodellierung (DR. R. ECKSTEIN, SS 2003)

Seminare im Hauptstudium

- XML und Datenbanken (D. SCHEFFNER, WS 2002/03)
- Grundlegende Aspekte des "Semantic Web" (DR. R. ECKSTEIN, WS 2002/03)
- Goya3-Entwicklung (S. HERSCHEL, WS 2002/03)

Forschung

Projekt: <xmlcity:berlin>

Teilprojekt: XML Clearinghouse für Berlin und Brandenburg

Ansprechpartner: DR. RAINER ECKSTEIN

Beteiligte Mitarbeiter: RALF HEESE, DIRK REMPE, MAREN LENK, FRANK LEGLER, MIRKO HERBIG, ASHRAF EMAWI, MIKE HERZOG

Zusammenarbeit: Freie Universität Berlin, Netzbasierte Informationssysteme

Forschungsförderung: Bundesministerium für Bildung und Forschung (BMBF)

Das XML Clearinghouse für Berlin und Brandenburg ist Teil des vom BMBF geförderten regionalen Wachstumskern <xmlcity:berlin>. Neben der Freien Universität Berlin, die an dem Teilprojekt mitwirkt, sind acht kleinere und mittlere Unternehmen (KMU) an dem gesamten Projekt beteiligt. Das XML Clearinghouse beobachtet als eine öffentlich zugängliche Dienstleistung die Entwicklung von XML-Technologien und bereitet sie auf, stellt die Anwendung von XML-Technologien dar und sorgt für einen Erfahrungsaustausch zwischen in der Region aktiven Firmen und Institutionen.

Im Rahmen des Teilprojekts XML Clearinghouse wurde der Workshop „XML-Technologien für Middleware – Middleware für XML Anwendungen“ (XMIDX 2003) mit großem Anklang in Berlin durchgeführt. In diesem Rahmen wurden auch halbtägig ein Doktorandenworkshop „Technologien und Anwendungen von XML“ und ein „Semantic Web Infotag“ veranstaltet. Kurz danach begann die Begutachtung der eingereichten Beiträge und das Erstellen des Programms für die ersten Berliner XML Tage. Die im Oktober mit großem Erfolg durchgeführte Konferenz bestand aus fünf wissenschaftlichen Workshops, einem Doktorandenworkshop und einem Wirtschaftsforum. Des Weiteren wurde die im letzten Jahr etablierte XML-Kolloquiumsreihe fortgesetzt, in der Referenten zu aktuellen Forschungsergebnissen und –projekten vortrugen. Das Informationsangebot des Webportals zu XML-bezogenen Standardisierungen, Veranstaltungen und Forschungsprojekten wurde erweitert und verbessert. Inzwischen wurde mit der Planung und Organisation der Berliner XML Tage 2004 begonnen, an denen sich wieder verschiedene wissenschaftliche Workshops mit XML-Bezug beteiligen; zudem ist auch wieder ein Wirtschaftsforum geplant.

Projekt: Internetökonomie InterVal**Teilprojekt: Wissensnetze**

Ansprechpartner: DR. RAINER ECKSTEIN, DIPL.-INF. RALF HEESE

Zusammenarbeit: Freie Universität Berlin, Lehrstuhl Netzbasierte Informationssysteme sowie Institut für Produktion, Wirtschaftsinformatik und OR

Forschungsförderung: Bundesministerium für Bildung und Forschung (BMBF)

Das Projekt untersucht die absehbare Entwicklung des zukünftigen Internets hin zu einem Netz des Wissens. Es analysiert anhand von zwei konkreten Szenarien aus dem Anwendungsbereich Automotive nach dem Potenzial und den wirtschaftlichen Auswirkungen eines globalen Semantic Web sowie wertschöpfungsketteninterner Wissensnetze. Durch diesen szenariobasierten Ansatz trifft das Teilprojekt nachhaltige Aussagen über die ökonomischen Auswirkungen.

In diesem Rahmen wurden zusammen mit der Freien Universität Berlin erste Szenarien entworfen und analysiert. Damit praxisbezogene Anwendungsszenarien entstehen, wurde bereits zu Kooperationspartnern aus der Wirtschaft Verbindung aufgenommen. Weiterhin etablierten sich Arbeitsgruppen innerhalb des Projekts, durch die der Wissenstransfer und die Zusammenarbeit zwischen den Teilprojekten gefördert wird. Im weiteren Verlauf werden die bisher entwickelten Szenarien detailliert fortgeführt bzw. mit Kooperationspartnern neue entwickelt.

Projekt: Datenbanktechnologien für das „Semantic Web“

Ansprechpartner: DR. RAINER ECKSTEIN

Das Web ist eine riesige vernetzte Sammlung von Seiten, deren Inhalte von menschlichen Nutzern interpretiert werden muss. Das Semantic Web reichert das bestehende World Wide Web um zusätzliche maschinenverarbeitbare Metadaten an, die den Web-Seiten aber auch beliebigen Ressourcen Bedeutung zuordnen. Mit Hilfe entsprechender Technologien kann so der Informationszugang von Rechnern unterstützt auf semantischer Ebene ablaufen und an die Stelle der Suche von Textmustern in Volltextsuchmaschinen tritt das Erschließen semantischer Konzepte und Wissen.

In den folgenden Teilprojekten wird der Einsatz von Datenbanktechnologien in Hinblick auf die Nutzbarkeit im Einsatz für das Semantic Web erforscht.

Teilprojekt: Performante RDF-Managementsysteme

Ansprechpartner: DIPL.-INF. RALF HEESE

Aufgrund der steigenden Bedeutung des Internets für das tägliche Leben können aussagekräftige Dokumente dazu beitragen, Aufwand und Kosten für deren Verarbeitung zu reduzieren. Durch semantische Annotation von Dokumenten können diese besser maschinell interpretiert und verarbeitet werden. Die Basistechnologien hierfür sind Spezifikationen wie RDF/S, DAML+OIL und OWL mit denen Konzepte, Taxonomien oder Ontologien dargestellt werden. Momentan werden bei RDF-Managementsystemen wie zum Beispiel Jena, Redland, Sesame oder RDFSuite vor allem objektrelationale Datenbanksysteme eingesetzt. Diese eignen sich jedoch nur bedingt zur Speicherung der Tripel-Datenmodelle von RDF, da im Allgemeinen zur Beantwortung von Anfragen viele kostenintensive Joins entstehen.

Das Ziel in diesem Projekt ist die Erforschung und Untersuchung von alternativen Speicher- und Anfragesystemen. Es sollen objektorientierte und deduktive Datenbanksysteme dahin-

gehend erweitert werden, dass eine effiziente Speicherung und performante Anfrage von RDF-Wissensbasen ermöglicht wird.

Teilprojekt: RDF-basierte semantische Suchmaschine

Ansprechpartner: DR. RAINER ECKSTEIN

Die Suche nach Informationen im Web gleicht bekanntermaßen der Suche einer Stecknadel im Heuhaufen. Die bekannten Suchmaschinen unterstützen dies derzeit im Wesentlichen auf einer syntaktischen Ebene durch Schlagwortsuche. Dies führt typischerweise zu einer Fülle von Ergebnissen, die häufig thematisch falsch sind. Die W3C-Empfehlung RDF (Resource Description Framework), die auf XML basiert, bietet hier Abhilfe, indem beliebige Ressourcen, also auch Webseiten, mit semantischen Informationen annotiert oder an separater Stelle beschrieben werden können. Unter Nutzung dieser semantischen Informationen wird eine agentenbasierte semantische Suchmaschine entwickelt. Hierfür werden die RDF-Daten in einer Datenbank abgelegt, die über das Jena-API angesprochen werden kann. Als Erweiterung wurde ein Inferenzmechanismus realisiert, mit dem die mit RDF-Schema und DAML+OIL aufgebauten Ontologien bei Anfragen durchsucht und semantische Zusammenhänge der Schlagworte aufgespürt werden können.

Teilprojekt: UML-basierte Modellierung von Ontologien

Ansprechpartner: DR. RAINER ECKSTEIN

Die Erstellung und Erweiterung von Ontologien ist eines der zentralen Probleme im Semantic Web, da dieses zur Zeit aufgrund überwiegend manueller Prozesse sehr kostenintensiv ist. Allerdings steckt viel Wissen über Ontologien in Modellen, die im Bereich der Software-Entwicklung oder der Geschäftsprozessmodellierung entstehen. Da häufig die Unified Modeling Language (UML) zur Modellierung eingesetzt wird, wurde ein Ansatz entwickelt, mit dem bestehende UML-Modelle (Klassendiagramme) in Ontologien überführt werden.

Projekt: Modell für einen Web-Service-Markt, basierend auf Abhängigkeiten

Ansprechpartner: RALF HEESE

Zusammenarbeit: Freie Universität Berlin, Arbeitsgruppe Netzbasierte Informationssysteme sowie Institut für Produktion, Wirtschaftsinformatik und OR

Die Komposition von Web Services aus vorhandenen Servicefragmenten erfordert eine semantische Beschreibung von den angebotenen und den gesuchten Web Services. Für deren Beschreibung werden Abhängigkeiten verwendet. Abhängigkeiten sind zwischen Web-Service-Fragmenten definiert, wobei technische, z.B. temporäre, und ökonomische (z.B. Preis/Leistungsverhältnis) unterschieden werden können. Auf Basis dieser semantischen Charakteristika können Web Services miteinander verglichen werden und eine Komposition von Web-Service-Fragmenten zu einem komplexen erfolgen. Ein Web-Service-Markt bildet die Plattform für die Komposition von komplexen Web Services aus einfachen. Die Komposition beruht hierbei auf Hierarchien von Prozessmodellen, in denen konkretere Prozesse durch Spezialisierung (Hinzufügen oder Ändern von Abhängigkeiten) entstehen. Eine Anfrage an das System besteht dabei aus einem abstrakten Prozessmodell, in dem eine Menge von gegebenen Abhängigkeiten spezifiziert sind. Basierend auf den Hierarchien der Prozessmodelle und unter Berücksichtigung aller gegebenen Abhängigkeiten wird ein konkretes Prozessmodell generiert, das in einem letzten Schritt zu einem ausführbaren komplexen Web Service führt. Abbildung 1 visualisiert die Gesamtarchitektur des beschriebenen Web-Service-Marktes.

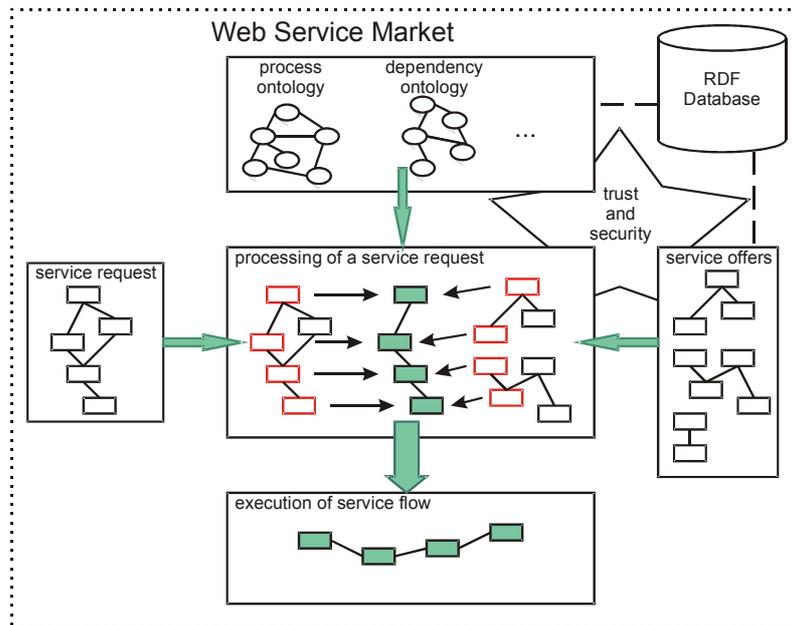


Abbildung 1: Architektur des Web-Service-Marktes

Ein weiterer zentraler Punkt in diesem Projekt ist die Modellierung von Vertrauens- und Sicherheitsaspekten. Nur unter Berücksichtigung dieser Aspekte kann ein solcher Web-Service-Markt erfolgreich realisiert werden. Die unterschiedlichen subjektiven Anforderungen an Vertrauen und Sicherheit sollen hierbei in den Web-Service-Markt integriert werden.

Projekt: Einsatz von Datenbanktechnologie für XML

Ansprechpartner: Dr. Rainer Eckstein

In den folgenden Teilprojekten werden verschiedene Aspekte der Datenbanktechnologie für deren Einsatz im Bereich von XML erforscht.

Teilprojekt: Datenmodellierung und XML

Ansprechpartner: Dr. Rainer Eckstein

Im Sinne der Datenbankmodellierung wurde der entwickelte Ansatz verfeinert, mit dem die Datenschemata, als die man die Dokument-Typ-Definitionen (DTD) und die XML-Schema-Spezifikationen ansehen kann, konzeptionell modelliert werden können. Hierfür haben wir die Unified-Modeling-Language (UML) verwendet, deren Erweiterungsmöglichkeiten an manchen Stellen genutzt werden konnte. Die bisherigen Ansätze wurden einer kritischen Überprüfung unterzogen, korrigiert und ergänzt sowie auf XML-Schema übertragen. Der alternative Ansatz geht auf das Meta-Modell der UML zurück, in dem ein eigenes DTD-Profil und ein XML-Schema-Profil entwickelt wurde.

Zu einem Ansatz wurde ein webfähiges Werkzeug für die automatische Transformation entwickelt. Der XMI-Export (XML Metadata Interchange) der gängigen UML-Modellierungswerkzeuge wird mit Hilfe von XSLT in eine DTD überführt. Zur Zeit wird dieses Werkzeug um alternative Ansätze erweitert. Die Erzeugung von XML-Schema ist in Planung.

Darüber hinaus wurde die Modellierung von Metadaten stärker berücksichtigt. Hierzu wird auf das Resource Description Framework (RDF) bzw. die Web Ontology Language (OWL) zurückgegriffen, das XML-Strukturen verwendet, um Metadaten zu Dokumenten standardisiert zu verwalten.

Teilprojekt: XML-basierte Integration molekularbiologischer Datenbanken

Ansprechpartner: DR. RAINER ECKSTEIN, DR. STEPHAN HEYMANN

Die Zusammenführung molekularbiologischer Datenbanken, die unabhängig voneinander für unterschiedliche Zwecke entwickelt wurden (s. a. Projekt Gene-EYE GDW), erfordert die Aufbereitung fremdkuratierter XML-basierter Datenquellen aus öffentlich zugänglichen Quellen. Deren Informationsgehalte sind in zahlreichen uneinheitlichen, teils widersprüchlichen Beschreibungssprachen (Markup Languages) abgefasst. Die große Herausforderung, die semantische Integration im Gesamtkontext voranzutreiben, wird in diesem Teilprojekt durch Remodellierung bestehender XML-Dokumenttypen umgesetzt. Die entstehenden UML-Klassendiagramme repräsentieren die existierenden Strukturen auf einer abstrakteren Ebene und schaffen dadurch die Voraussetzungen für eine Reale-Welt-nahe Abbildung komplexer Zusammenhänge aus den Lebenswissenschaften. Innewohnende Graphstrukturen sollen mittels GraphML expliziert und mit dem GeneViator (s. Teilprojekt Visualisierung von Graphen) navigierbar gemacht werden. Auf diese Weise wird es (auch) Nutzern aus den Lebenswissenschaften ermöglicht, Recherchen über den gesamten Datenbestand flexibel auszuführen, ohne an die starren Dokumenten- und Datenstrukturen primärer Daten- und Wissensquellen gebunden zu sein.

Teilprojekt: XML Query Execution Engine (XEE)

Ansprechpartner: DIPL.-INF. DIETER SCHEFFNER, PROF. J.C. FREYTAG, PH.D.

Beteiligte Mitarbeiter: GREGOR ZEITLINGER

XML hat große Bedeutung für den Austausch und die Repräsentation von Daten und Dokumenten im WWW erlangt. Hieraus ergibt sich die Notwendigkeit, große Datenbestände, die in Form von XML-Dokumenten vorliegen, verwalten zu können. Die Verwaltung derartiger Daten, die im allgemeinen nicht vollständig strukturiert sind, bereitet "herkömmlichen" DBMSen, wie objektorientierten, objektrelationalen bzw. relationalen DBMSen, häufig Probleme. Es sind deshalb neue Verfahren zur Speicherung, Veränderung und Anfrage von XML-Dokumenten in Datenbanken notwendig. Das Projekt XML Query Execution Engine (XEE) bildet als prototypisches Anfragesystems einen Rahmen für die Forschung auf diesem Gebiet. Das wesentliche Ziel im Projekt ist die Entwicklung von Speicherstrukturen, die ein effizientes Anfragen und Verändern von großen XML-Dokumenten zulassen, ohne dass hierbei die Dokumente bestimmten Strukturvorgaben genügen müssen. Darüber hinaus werden sowohl Text-Retrieval- als auch strukturelle Anfragen gleichermaßen unterstützt.

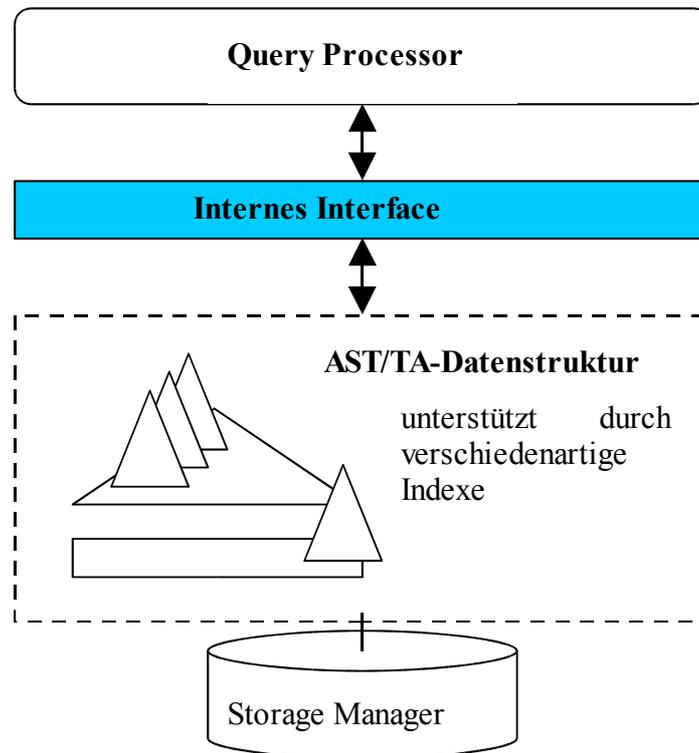


Abbildung 2: Architektur der XML Query Execution Engine (XEE)

Um die oben genannten Ziele zu erreichen, wurde für XEE bereits in der Vergangenheit die Datenstruktur "Access Support Tree & TextArray" (AST/TA) entwickelt. Die Datenstruktur basiert auf der Idee, den gesamten Inhaltstext - der Teil eines XML-Dokuments, der für den Leser eines Dokuments "sichtbar" und auch primär von Wichtigkeit ist - als zusammenhängende Zeichenkette zu speichern. Die logische Struktur, gegeben durch die Tag-Hierarchie des Dokuments, wird hierbei auf eine korrespondierende Baumstruktur abgebildet, die lediglich Verweise auf die entsprechenden Textteile in der Zeichenkette enthält. Die AST/TA-Datenstruktur wird derzeit für den Sekundärspeicher implementiert. Im Rahmen der bisherigen Implementierungsarbeiten wurde begonnen, eine für die Datenstruktur spezifische Schnittstelle als Erweiterung der DOM-Schnittstelle zu entwickeln und zu implementieren. Erste Untersuchungen des Verhaltens der Datenstruktur zeigen, dass der gewählte Ansatz die geforderten Eigenschaften unterstützt. Nächste Schritte im XEE-Projekt sehen die Untersuchung der Indexierungsmöglichkeiten und die Erweiterung der Schnittstelle der AST/TA-Datenstruktur vor.

Projekt: Privater Datenzugriff

Ansprechpartner: DIPL.-ING. DMITRI ASONOV, PROF. JOHANN-CHRISTOPH FREYTAG, PH.D.

Zusammenarbeit: IBM Research T.J. Watson Lab. New York, IBM Deutschland Entwicklung GmbH, TU Berlin, FU Berlin, TU Cottbus

Forschungsförderung: Deutsche Forschungsgemeinschaft (GRK 316) – Graduiertenkolleg „Verteilte Informationssysteme“

Mit zunehmender Nutzung des Web wird es immer notwendiger, Einzelpersonen in ihrer Privatsphäre beim Zugriff auf Daten im Web zu schützen. Der Bereich „Private Information Retrieval“ (PIR) befasst sich mit dem Problem des Datenzugriffs „ohne Beobachtung“, d.h.

Individuen sollen auf Datenquellen zugreifen können, ohne dass ein anderer feststellen kann, welche Daten „von Interesse“ waren. Die bisherigen Ergebnisse sind von hoher Komplexität und in vielen Fällen theoretischer Natur ohne Berücksichtigung von Anforderungen aus Datenbanksicht.

In diesem Projekt werden ausgehend von existierenden Verfahren neue Ansätze des „Private Information Retrieval“ entwickelt, die sowohl Anforderungen aus dem Datenbankbereich berücksichtigen als auch die bisherigen Algorithmen vereinfachen. Bei den in diesem Projekt entwickelten Algorithmen werden u. a. Vorverarbeitungsphasen vorgeschlagen („offline work“), um beim tatsächlichen Zugriff auf die Daten („online work“) den Aufwand zu reduzieren. Darüber hinaus werden Fragen des privaten Zugriffs im Kontext von E-Commerce untersucht.

Um unsere Hypothesen mit experimentellen Daten untermauern zu können, wurde ein Prototyp entwickelt und die während ca. 30 Tagen gewonnenen Testergebnisse analysiert. Der Prototyp wurde auf Basis einer von IBM entwickelten Spezialhardware und Spezialsoftware implementiert. Ein weiterer Fokus unserer Forschung liegt auf der Generalisierung des Problems „Privater Datenzugriff“. Damit erlauben wir, jeden beliebigen Kompromiss zwischen dem Umfang des privaten Zugriffs und der Komplexität des Algorithmus zu bestimmen und auszuwählen.

Projekt: Unvollständige Workflows – Konzepte und Methoden zur Weiterführung

Ansprechpartner: DIPL.-INF. (FH) KATJA THAM

Forschungsförderung: Promotionsstipendium der Fachhochschule für Technik und Wirtschaft (FHTW) Berlin

Traditionelle Ansätze zum Workflow-Management beschränken sich weitgehend auf vorab vollständig modellierbare Prozesse. Sie verfolgen deshalb eine strikte Trennung zwischen Modellierung und Ausführung eines Workflows. Der Einsatz im Umfeld wissenschaftlicher Forschung wird hierdurch stark eingeschränkt, da das explorative Vorgehen eine anfänglich unvollständige Beschreibung des Gesamtprozesse bedingt.

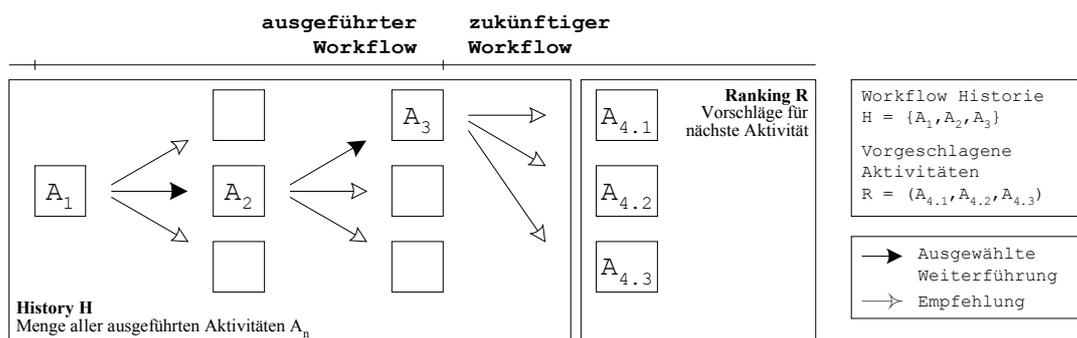


Abbildung 3: Vereinfachte Darstellung eines unvollständigen Workflows

Das vorgestellte Projekt verfolgt eine Erweiterung traditioneller Ansätze zur Anpassung von Workflow-Management-Systemen an die Anforderungen an empirische Prozesse im wissenschaftlichen Umfeld. Trifft ein unvollständiger Prozess seine vordefinierten Grenzen, kennzeichnet ihn neben der Abfolge der bis dato ausgeführten Aktivitäten auch bisherige Ergebnisse/Daten und global definierte Beschränkungen.

Dieser Arbeit liegt die Hypothese zugrunde, dass der aktuelle Zustand eines unvollständigen Prozesses Hinweise auf die eigene Weiterführung beinhaltet. Idealerweise folgt der Beschreibung eines Prozesszustandes (vergleichbar einer Augenblicksaufnahme) eine Gegenüberstellung, Einschränkung und Auswahl von potenziell nächsten Aktivitäten.

Ziel ist es, aus einem einmal durchgeführten Forschungsprozess das Wissen über die Abfolge der Einzelschritte zu extrahieren und für nachfolgende, spätere Prozesse zu nutzen. Zur Verifikation und Falsifikation von Hypothesen sind dabei erfolgreiche wie fehlgeschlagene Prozessketten relevant, da sie genutzt werden können, um eine Wiederholung zu empfehlen oder von ihr abzuraten.

Neben einer Verschmelzung von Workflow-Modellierung und –Ausführung soll eine maschinell unterstützte, wissensbasierte Entscheidungsfindung zur Fortsetzung unvollständiger Prozesse realisiert werden. Die entwickelte Systematik ermöglicht es, Workflowkonzepte und deren Vorteile in völlig neuen Anwendungsbereichen, wie der Genomforschung einzusetzen, in denen bisherige Workflowlösungen aufgrund ihrer mangelnden Flexibilität nicht bestehen konnten.

Projekt: Genome Data Warehouse

Beteiligte Mitarbeiter: DR. STEPHAN HEYMANN, PETER RIEGER, PROF. JOHANN-CHRISTOPH FREYTAG, PH.D., DIPL.-INF. CHOKRI BEN NECIB, DIPL.-INF. HEIKO MÜLLER, DIPL.-INF. (FH) KATJA THAM, TIMO MIKA GLÄSSER

Zusammenarbeit: Berlin Center for Genome Based Bioinformatics (BCB), IBM Deutschland Entwicklung GmbH, Bioinformatik-Zentrum Jena, Napier University Edinburg, UK

Forschungsförderung: IBM Deutschland Entwicklung GmbH, Deutsche Forschungsgemeinschaft (DFG), Bundesministerium für Bildung und Forschung

Der Aufbau eines Genome Data Warehouse macht die Integration heterogener verteilter Datenquellen unterschiedlicher Größe und Qualität erforderlich. Die vorgestellte Architektur (Gene-EYE) dient zur Beschreibung der Transformationen dieser Datensammlungen in biologisch sinnvolle Daten hoher Qualität. Die einzelnen Ebenen bilden den formalen Rahmen, in dem durch Einsatz geeigneter Werkzeuge das Ziel einer integrierten Genomdatenbank erreicht werden kann. Die Architektur besteht aus drei Ebenen. Auf der unteren Ebene des *Data Store Layer* werden die vorhandenen Datenquellen (z.B. EMBL, SWISSPROT/TrEMBL, ENSEMBL) transformiert und in einer relationalen Datenbank integriert. Das resultierende Schema ist stark beeinflusst durch das logische Schema der jeweiligen Datenquelle. Dies ermöglicht es, eine große Menge der ursprünglichen Informationen beizubehalten. Auf dieser Ebene erfolgt auch das syntaktische Data Cleansing. Der *Genome Database Layer* ermöglicht die Abbildung der in der relationalen Datenbank verwalteten Daten auf biologische Entitäten (z.B. GEN). Die Definitionen und Abbildungen werden in Zusammenarbeit mit Domänenexperten vorgenommen. Hierbei erfolgt zusätzlich das semantische Data Cleansing. Unterstützt wird dieser Schritt durch die Verwendung des von IBM entwickelten Werkzeugs CLIO zum semi-automatischen Erstellen von Schemaabbildungen.

Auf der höchsten Ebene, dem *Data Warehouse Layer*, können die zuvor definierten Entitäten in biologische Abläufe (z.B. PATHWAY) zum Zwecke der Prozesssimulation eingebunden werden. Außerdem können mit Hilfe von Data-Mining-Methoden weitere Erkenntnisse über biologische Gesetzmäßigkeiten und Abläufe gewonnen werden

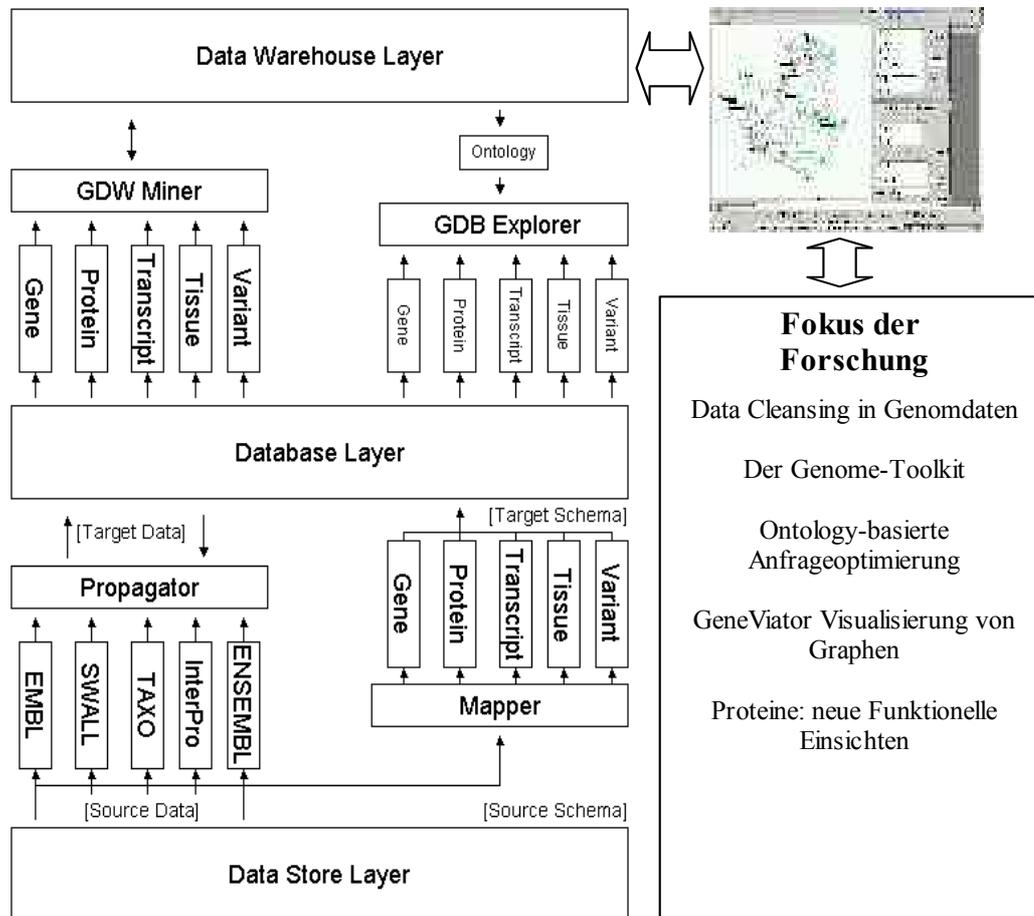


Abbildung 4: Gene-Eye-Architektur, Nutzung, Ergebnisaufbereitung

Teilprojekt: GEM – Das Gene-EYE Metadatenmodell

Ansprechpartner: PETER RIEGER

Neben dem eigentlichen Ziel von Gene-EYE, eine Vielzahl heterogener Datenbestände in einen einheitlichen Verwaltungs- und Abframerahmen zu integrieren, gilt ein weiterer Forschungsschwerpunkt der Automatisierung der erforderlichen Verarbeitungsschritte, die erforderlich sind, den oben genannten Zielzustand zu erreichen. Dazu ist es erforderlich, weit mehr als lediglich die formale Struktur der zu integrierenden Datenquellen zu beschreiben. In dem Gene-EYE Metadatenmodell werden daher für jede Datenquelle der Mechanismus der Datenbereitstellung, die Zuordnung lokaler Ressourcen für Speicherung und Verarbeitung sowie die Struktur und Syntax jeder Datenquelle erfasst.

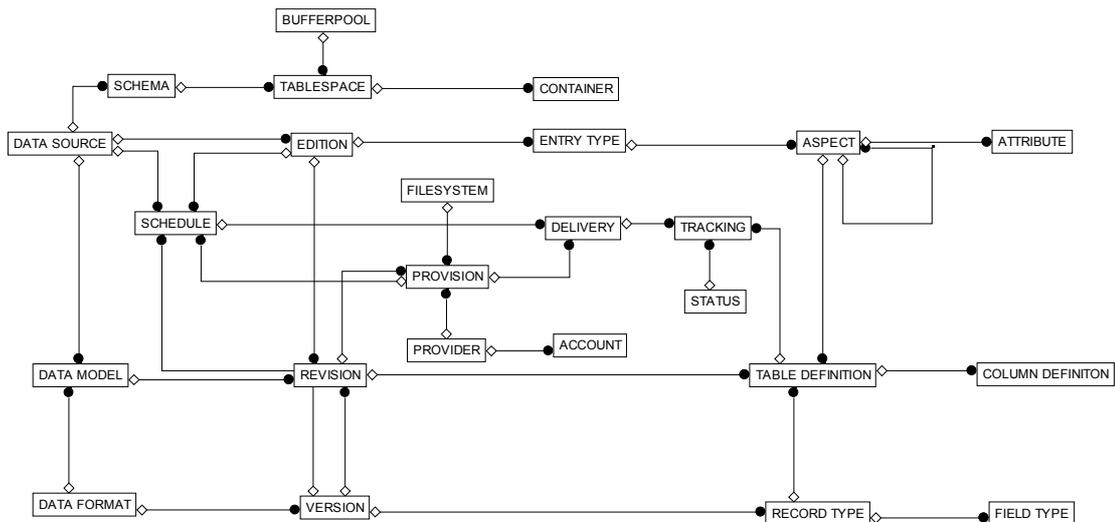


Abbildung 5: Gene-Eye Metadatenmodell

Die Daten werden verwendet, um die gesamte Integrationsprozesskette zu unterstützen und den Fortschritt der einzelnen Verarbeitungsschritte zu dokumentieren. Um den Umgang mit den Metadaten systematisch abzubilden und zu dokumentieren, wurde das Eclipse Modeling Framework (EMF) eingesetzt. Damit ist es möglich, die Objekte und Zusammenhänge im Bereich der Metadaten in UML-Modellen zu erfassen und diese automatisch in Editoren für die Erfassung der Metadaten zu überführen.

Teilprojekt: GeneViator – Visualisierung von Graphen

Ansprechpartner: DR. STEPHAN HEYMANN, PETER RIEGER

Zusammenarbeit: Berlin Center for Genome Based Bioinformatics (BCB)

Eine der wesentlichen Eigenschaften von Genomdaten ist ihre Verknüpfung durch Hyperlinks, die Beziehungen zwischen zwei oder mehreren Objekten darstellen. Diese Verknüpfungen können in Form von Graphen modelliert werden. Die große Anzahl an Datensammlungen und die Vielzahl ihrer Einträge führen zu sehr großen Strukturen bezogen auf die Anzahl an Knoten und Kanten. Das Visualisierungs- und Analysewerkzeug GeneViator eignet sich zur Darstellung von und Navigation in solchen Graphen, die aus datenhärenten semantischen Beziehungen zwischen biologischen Entitäten abgeleitet werden. Es erlaubt über seine diversen Filterfunktionen (siehe Abbildung 6) die Definition von Teilmengen dieser Graphen sowie die Suche nach Korrelationen innerhalb und zwischen diesen Teilmengen.



Abbildung 6: Beispielhafte Darstellung von verknüpften Datenquellen im Web (SRS) mit Hilfe des GeneViators

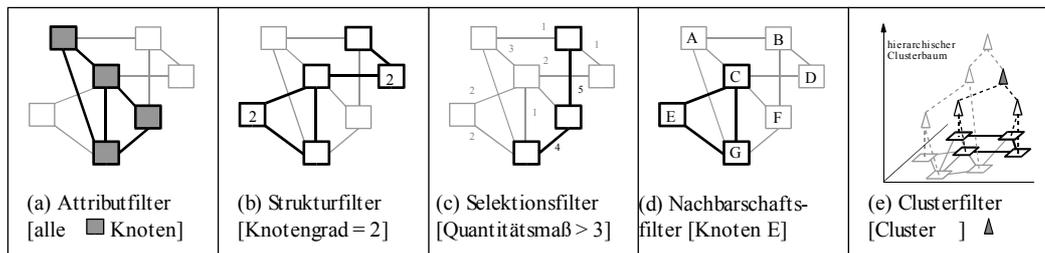


Abbildung 7: Einige der Filterfunktionen des GeneViators

Die Anwendung des GeneViators in der akademischen Forschung (u.a. Group Lacroix, University of Arizona Tucson, USA; Group Butler, Concordia University Montreal, Kanada) erstreckte sich zunächst auf Graphen, deren Knoten gewöhnlich Objekte einer Klasse repräsentieren. In erweiterten Pilotanwendungen kamen k-partite Graphen hinzu. Gegenwärtig ist der Schwerpunkt der Funktionserweiterungen auf die Navigation in Graphen mit *Kanten unterschiedlicher Kennung* (siehe Abbildung 8) gerichtet. Das schließt Darstellungsmöglichkeiten von *Multigraphen* mit ein.

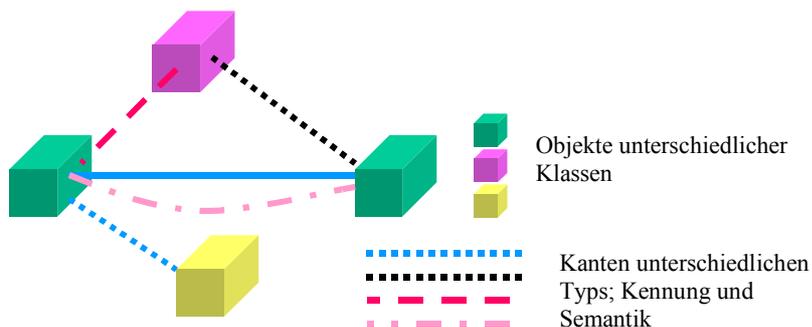


Abbildung 8: Der Gen-Viator als Darstellungshilfe komplexer Sachverhalte

Damit ist gleichzeitig die Basis für eine Breitenanwendung der Software in Vorhaben des Berliner Centrums für genom-basierte Bioinformatik (BCB) gegeben, etwa für die Navigation in Literaturrechercheergebnissen zu Protein-Protein-Wechselwirkungen und für die vertiefte semantische Analyse verknüpfter Informationsgehalte in web-gestützten Datenbankeinträgen.

Teilprojekt: Proteine - Neue funktionelle Einsichten

Ansprechpartner: DR. STEPHAN HEYMANN, PETER RIEGER, TIMO MIKA GLÄSSER

Zusammenarbeit: Bioinformatik-Zentrum Jena

Forschungsförderung: Zugang zu Grid-Computing Kapazitäten der IBM, Deutsche Forschungsgemeinschaft (DFG)

Für die überwiegende Mehrheit der aus mehreren Exons bestehenden menschlichen Gene werden alternative Spleißformen vermutet. Hierbei handelt es sich um verschiedene lineare Kombinationen der Exons. Die entstehenden Spleißformen bilden Vorlagen für die Synthese von Proteinen, welche dann in aller Regel unterschiedliche bzw. abweichende Funktionalität besitzen. Ziel des Projekts ist das Auffinden bisher unbekannter Spleißformen mit Hilfe

computergestützter Methoden. Wir konzentrieren uns dabei auf Spleißformen, bei denen durch Auslassen von Exons neue funktionelle Motive entstehen, d.h. Abschnitte des Proteins, denen eine biochemische Funktion zugewiesen werden kann. Ursache hierfür können die entstehenden Sequenzabschnitte an den Spleißstellen sein oder Verschiebungen im Leseraster, welche durch die Länge des entfallenden Exons bedingt sind.

Sämtliche lineare Kombinationen distanter Erbinformation werden mit Hilfe einer Modifikation des Viterbi-Algorithmus (dynamische Programmierung) berechnet [M. HILLER, R. BACKOFEN, S. HEYMANN, A. BUSCH, T. M. GLÄSSER, J.-C. FREYTAG: *Efficient Prediction of Alternative Spliceforms Using Protein Domain Homology* Submitted for publication in: *In silico Biology*]. Die entstehenden künstlichen Spleißformen werden anhand des genetischen Codes in die Aminosäuresequenz des resultierenden Proteins übersetzt. Jede Sequenz wird dann auf das Vorkommen neuer funktioneller Motive untersucht. Hierzu bedienen wir uns der Datensammlung PFam und ihrer angebotenen Werkzeuge. Die einzelnen Motive sind dort in Form von Hidden-Markov-Modellen (HMM) repräsentiert. Mit Hilfe des Werkzeugs HMMSearch kann jede Aminosäuresequenz gegen die komplette Sammlung verglichen und auf funktionelle Motive hin untersucht werden.

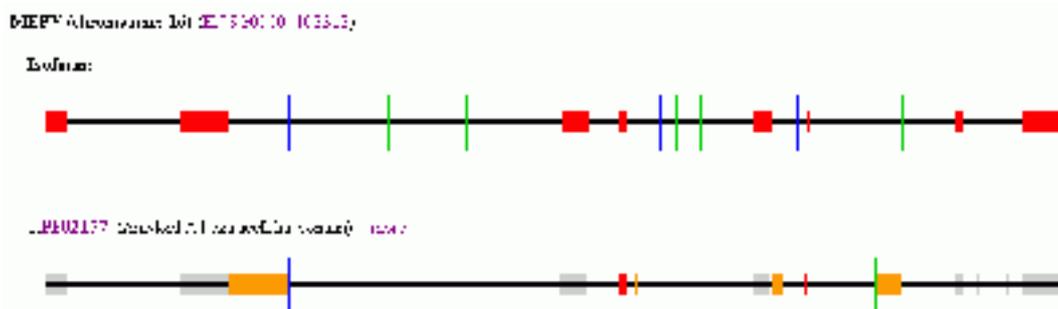


Abbildung 9: Eine Alternative Spleißform des Marenostrins (unteres Piktogramm) ist auf Grund der Domänenverwandtschaft mit bekannten plaque-bildenden Proteinen Ansatzpunkt für Untersuchungen von Proteinaggregationsvorgängen

Die verschiedenen Komponenten des Workflows, die für die Vorverarbeitung (Auffinden von zusätzlichen Transkriptionsstartsignalen, Donor- und Akzeptorpaaren etc.) und die PFam-Suche zuständig sind, greifen über eine SOAP-Schnittstelle auf die Datenbank zu, wie aus Abbildung 9 ersichtlich ist.

Die aufgrund der neu auftretenden Motive als geeignete Kandidaten angesehenen Spleißformen werden dann in Zusammenarbeit mit Biologen auf Existenz und molekular-ätiologische Relevanz hin untersucht. Abbildung 9 zeigt am Beispiel des Marenostrins – Mutanten dieses Gens sind kausal für Familiäres Mittelmeerfieber (FMF) –, wie über lineare Konkatenierung kodierender Genabschnitte Alternative Spleißformen generiert und auf Funktionsindikatoren hin untersucht werden. Im betreffenden Fall erkannten die Experten für FMF eine plausible arbeitshypothetische Erklärung für Dysfunktion dieses Proteins in der Niere der Patienten.

Teilprojekt: Semantisches Data Cleansing von Genomdaten

Ansprechpartner: DIPL.-INF. HEIKO MÜLLER

Unter Genomdaten verstehen wir Informationen über die sequentielle Abfolge der Bausteine in Biomolekülen, kurz Sequenz, sowie Beschreibungen über strukturelle und funktionale Eigenschaften der Sequenzen (Annotationen). Bedingt durch unsicheres Wissen und ein hohes Maß an Datenabhängigkeiten im Analyseprozess ist im Bereich der Genomdaten die Fehlerhäufigkeit besonders hoch, angefangen von einfachen Syntaxfehlern bis hin zu konfliktreichen und somit zweifelhaften Annotationen von Genstruktur und Proteinfunktionalität. Ziel des Semantischen Data Cleansing ist es, die Korrektheit existierender Genomdaten zu validieren bzw. zu gewährleisten. Dies erfordert ein Erkennen und Korrigieren fehlerhafter Daten. In diesem Zusammenhang untersuchen wir momentan zwei unterschiedliche Vorgehensweisen (Abbildung 10).

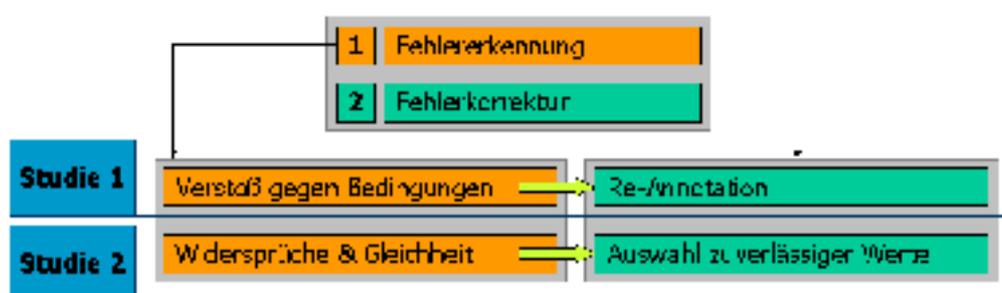


Abbildung 10: Vorgehensmodell beim Semantic Data Cleansing

In unserer ersten Studie erfolgt ausgehend von einem speziellen Datenmodell für Genomdaten eine formale Beschreibung des Annotationsprozesses. Teil dieser Beschreibung sind Bedingungen an die Daten, die in einzelnen Schritten erfüllt sein müssen. Diese Bedingungen werden dann zum Auffinden von Fehlern in existierenden Daten verwendet. Die Korrektur erfolgt selektiv anhand der formalen Prozessbeschreibung. Somit verhindern wir eine komplette Re-Annotation, wie sie heutzutage oftmals vollzogen wird, die zeitaufwendig ist und aufgrund der Datenabhängigkeiten bei Änderungen in den verwendeten Quellen jeweils wiederholt werden müsste.

Im Rahmen einer zweiten Studie sollen existierende Redundanzen in den verfügbaren Datensammlungen ausgenutzt werden, um eine zuverlässige integrierte Instanz aus diesen abzuleiten. Durch den Vergleich inhaltlich überlappender Datenquellen werden die übereinstimmenden und damit als zuverlässig angesehenen Teile identifiziert. Des Weiteren sollen Konflikte und Widersprüche in den Daten aufgezeigt und vom Experten bewertet werden, um so eine Bewertung der widersprechenden Werte zu erhalten. Im Ergebnis entsteht eine integrierte Datensammlung, welche die jeweils zuverlässigen Teile der einzelnen Quellen vereint.

Teilprojekt: The Genome-Toolkit – Erweiterung relationaler Datenbankfunktionalität um Algorithmen der Bioinformatik

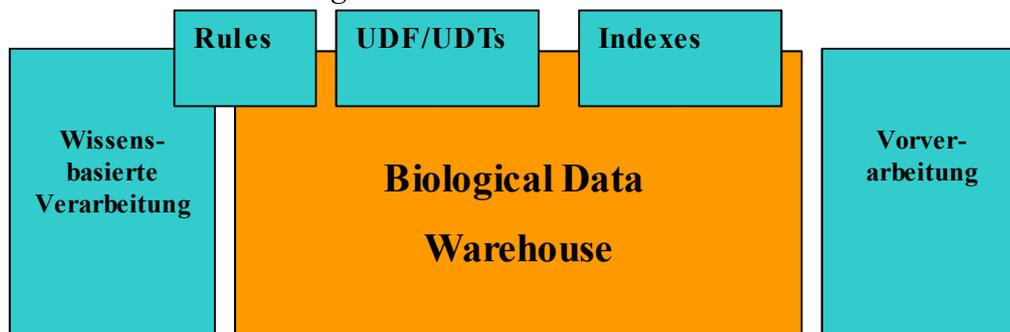
Ansprechpartner: DIPL.-INF. CHOKRI BEN NECIB, PETER RIEGER

Zusammenarbeit: IBM Deutschland Entwicklung GmbH

Sequenzanalyse spielt eine zentrale Rolle in der molekularen Biologie. Dabei werden Eigenschaften von Protein- und Nukleotidsequenzen (DNA) identifiziert und charakterisiert. Die verwendeten Algorithmen basieren heutzutage weitestgehend auf Textdateien zur Speicherung der benötigten Daten. Das Genome-Toolkit stellt eine Erweiterung relationaler Datenbankfunktionalität um eine Menge von Funktionen dar, die eine effiziente Datenanalyse ermöglichen sollen. Zu diesen Funktionen gehören die gängigen Algorithmen zum Sequenzvergleich, die u.a. eine Ähnlichkeitssuche auf Datenbanken ermöglichen.

In einem ersten Schritt wurden die beiden populären Programmpakete *BLAST* und *FASTA* implementiert und an eine objektrelationale Datenbank angepasst. Zwei Anpassungsvarianten wurden verwendet: Die erste Variante basiert auf der Programmierschnittstelle *CLI (Call Level Interface)*, während die zweite Variante die Algorithmen in der Datenbank als *UDFs (User Defined Functions)* realisiert. In einem zweiten Schritt wurden zwei Algorithmen zur paarweisen Alignierung von Sequenzen als UDFs implementiert: *NeedelmanWunsch*- und *SmithWaterman*-Algorithmus. Der erste Algorithmus liefert eine globale Alignierung, der zweite Algorithmus eine lokale Alignierung. Zusätzlich wurde die Programmpakete *TRANSLATE* für die Übersetzung einer Nuklein- in einer Proteinsequenz und umgekehrt in das Toolkit integriert.

Abbildung 11: Genomic-Toolkit-Rahmenwerk



Für das Projekt wird das DBMS IBM DB2/UDB eingesetzt, das durch Parallelisierung von Anfragen deren Ausführungszeiten erheblich verbessert. Des Weiteren sollen neue Indizierungsverfahren entworfen und realisiert werden, um die Ausführung zusätzlich zu beschleunigen. Zu diesem Zweck wurde eine neuartige Indexstruktur, der *String B-Tree*, implementiert und untersucht. Im nächsten Schritt soll die Leistung der UDFs in Bezug auf die Laufzeit und die lokalen Ressourcen verbessert werden. Weiterhin soll der Nutzen von *String B-Tree's* für die Suche in biologischen Sequenzen untersucht werden.

Teilprojekt: Ontologiebasierte Anfragenverarbeitung

Ansprechpartner: Dipl.-Inf. Chokri Ben Necib

Die leistungsfähige Anfragenbearbeitung ist eine der vorrangigsten Aufgaben des DBMS. Besonders für große Datenbanken, wie wir sie im Bereich der Genomdaten antreffen, ist dies eine Herausforderung.

In den letzten Jahren haben *Ontologien* gezeigt, dass sie gute semantische Unterstützungen für Datenverarbeitung anbieten können. Ontologien stellen abstrakte und semantisch reichere Beschreibungen der unterliegenden Daten in der Datenbank dar. Dies wird durch eine hierarchische Struktur von relevanten Konzepten und deren Beziehungen realisiert. Unser Ziel ist es, einen ontologie-basierten Ansatz für die semantische Anfrageverarbeitung zu entwickeln. Dieser Ansatz soll das Hinzufügen neuer semantischer Regeln ermöglichen, die aus einer gegebenen Ontologie abgeleitet werden. Diesen Regeln werden für die Umformulierung der Benutzeranfrage verwendet. Dies führt zu einer neuen Anfrage, die für den Benutzer sinnvolle Ergebnisse von der Datenbank liefern kann.

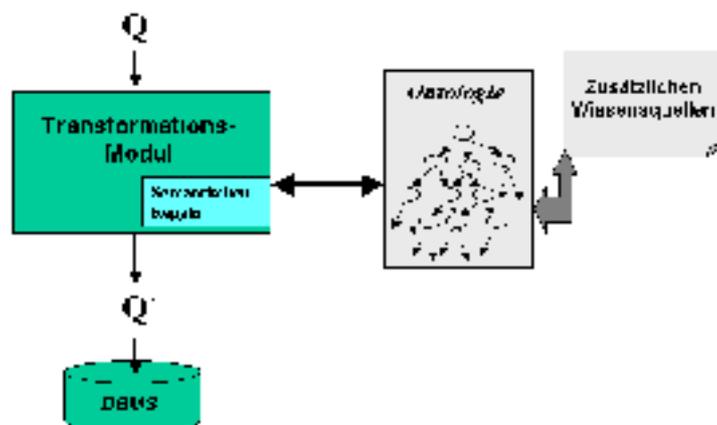


Abbildung 12: Nutzung von Ontologien für die Anfragebearbeitung

Projekt: Goya

Ansprechpartner: DIPL.-INF. SVEN HERSCHEL

Goya3 ist ein Verwaltungssystem zur Unterstützung des universitären Lehr- und Übungsbetriebes. Es umfasst eine Kernfunktionalität zur Erfassung von Leistungen im Rahmen beliebiger Studiengänge. Zusätzliche Module des Systems sind das Veranstaltungsmanagement zur Verwaltung des Übungsbetriebes, das Aufgabenmanagement zur Unterstützung von regelmäßigen Übungsaufgaben im Rahmen von Veranstaltungen sowie das Prüfungsmanagement zur Verwaltung von Prüfungen.

Den Kern von Goya3 bildet das Creditmanagement. Dabei werden alle Leistungen der Studierenden einschließlich ihrer Bewertung erfasst. Diese Leistungen werden im Rahmen von modularisierten Studiengängen erbracht und können zu größeren Modulen zusammen-

gefasst werden. Leistungen entstehen entweder direkt durch Eingabe durch das Prüfungsamt oder durch Ablegen der Leistung im Rahmen einer Veranstaltung oder Prüfung.

Veranstaltungsmanagement umfasst die Anlage von Lehrveranstaltungen und Veranstaltungsterminen, die Generierung eines Vorlesungsverzeichnisses, die Einschreibung in Veranstaltungstermine sowie die Verwaltung des Studienfortschrittes im Rahmen der Veranstaltung. Am Ende einer Veranstaltung kann ein Credit generiert werden, welcher automatisch in die Leistungsverwaltung des Systems übernommen wird.

Analog zu Veranstaltungen kann ein Credit auch aufgrund einer Prüfung entstehen. Das Prüfungsmanagement kümmert sich um das Anlegen und Verwalten von Prüfungsterminen, die Anmeldung von Studierenden zu Prüfungen sowie die Bewertung der Prüfung durch den Prüfer. Weitere Möglichkeiten sind Abmeldung von Prüfungen innerhalb bestimmter Termine sowie Stornierung von Ergebnissen durch das Prüfungsamt.

Das Aufgabenmanagement setzt das Veranstaltungsmanagement voraus. Es können für die Studierenden einer Veranstaltung Übungsaufgaben online ins Netz gestellt werden, diese können die Lösungen ebenfalls online einreichen. Die Lösungen werden dann in einem Pool gesammelt und den Korrektoren zur Verfügung gestellt, die dann die Punkte für die Lösungen vergeben können. Möglichkeiten für die Abgabe von Lösungen sind per File-Upload direkt in Goya, per Papier außerhalb von Goya oder per HTTP-Post über eine externe Webseite, z.B. ein am Lehrstuhl lokal erstelltes Formular.

Das Goya3 - System wird am Institut für Informatik der Humboldt-Universität stetig weiterentwickelt, um den steigenden Anforderungen an ein solches System gerecht zu werden.

Veröffentlichungen

S. HEYMAN, K. THAM, P. RIEGER J.-C. FREYTAG: *Rechnergestützte Suche nach Korrelationen in komplexen Datensätzen der Biowissenschaften*. Tagungsband der 10. BTW Konferenz, Leipzig, Februar 2003.

R. TOLKSDORF, C. BIZER, R. HEESE: *A Web Service Market Model based on Dependencies*. In Proceedings of The Twelfth International World Wide Web Conference, Poster Sessions. Budapest, Ungarn, Mai 2003.

U. LESER, P. RIEGER: *Integration molekularbiologischer Daten*. Datenbankspektrum, Ausgabe 6, Juni 2003.

J.-C. FREYTAG, P. C. LOCKEMANN, S. ABITEBOUL, M.L J. CAREY, P. G. SELINGER, A. HEUER: Editors of the *Proceedings of 29th International Conference on Very Large Data Bases*, September 9-12, 2003, Berlin, Germany. Morgan Kaufmann, September 2003.

I. CRUZ, V. KASHYAP, S. DECKER, R. ECKSTEIN: *Semantic Web and Data Bases SWDB 2003*. CD-Proceedings zum Workshop, an der Humboldt-Universität zu Berlin, September 2003.

R. ECKSTEIN, R. TOLKSDORF (Hrsg.): *XMIDX – XML-Technologien für Middleware, Middleware für XML-Anwendungen*. GI-Edition Lecture Notes in Informatics, Nr. 24, Köllen Verlag. September 2003.

H. MÜLLER, J.-C. FREYTAG: *Problems, Methods, and Challenges in Comprehensive Data Cleansing*. Informatik-Berichte 164, September 2003.

H. MÜLLER: *Semantic Data Cleansing in Genome Databases*. Proceedings of the VLDB 2003 PhD Workshop, Berlin, September 2003.

H. MÜLLER, F. NAUMANN, J.-C. FREYTAG: *Data Quality in Genome Databases*. Proceedings of the Conference on Information Quality (IQ 03), Boston (USA), Oktober 2003.

R. TOLKSDORF, R. ECKSTEIN (Hrsg.): *Berliner XML-Tage 2003*. Humboldt-Universität und Freie Universität Berlin, Oktober 2003.

M. HILLER, R. BACKOFEN, S. HEYMANN, A. BUSCH, T. M. GLÄSSER, J.-C. FREYTAG: *Efficient Prediction of Alternative Spliceforms Using Protein Domain Homology*. Submitted for publication: In Silico Biology, Oktober 2003.

C. BEN NECIB, J.-C. FREYTAG: *Ontology based Query Processing in Database Management Systems*. Proceedings of the International Conference on Ontologies, Databases and Applications of Semantics (ODBASE), Catania, Italy, November 2003.

R. TOLKSDORF, C. BIZER, R. ECKSTEIN, R. HEESE: *Business to Consumer Markets on the Semantic Web*. Proceedings of OTM 2003 Workshops, Lecture Notes in Computer Science, Volume 2889, S. 816-828, Catania, Italy, November 2003.

R. ECKSTEIN, S. ECKSTEIN: *XML und Datenmodellierung – XML-Schema und RDF zur Modellierung von Daten und Metadaten einsetzen*. dpunkt.verlag, Heidelberg, November 2003.

S. HERSCHEL: *Goya - Vom Programmierwettbewerb zum Verwaltungssystem* in Klaus R. Dittrich, Wolfgang König, Andreas Oberweis, Kai Rannenber, Wolfgang Wahlster (Eds.): INFORMATIK 2003 - Innovative Informatikanwendungen, Band 2, Beiträge der 33. Jahrestagung der Gesellschaft für Informatik e.V. (GI), Frankfurt am Main, Oktober 2003.

Vorträge

J.-C. FREYTAG: *Processing Genome Data using Scalable Database Technology*. IBM Almaden Research Center, USA, März 2003.

J.-C. FREYTAG: *Processing Genome Data using Scalable Database Technology*. Yorktown Heights, NY, USA, März 2003.

J.-C. FREYTAG: *Processing Genome Data using Scalable Database Technology*. HP-Labs, CA, USA, August 2003.

J.-C. FREYTAG: *Privacy in Databases*. University of Bergen, Norwegen, August 2003.

R. ECKSTEIN: *Semantic Web - Die Zukunft des Dokumentenmanagements im Internet?* Eingeladener Vortrag XML-Konferenz – Dokumentenmanagement, ZGDV, Darmstadt, Juli 2003.

S. HEYMANN: *Correlation Mining in Large Heterogeneous Datasets*. Dagstuhl-Seminar (03051) „Information and Process integration: A Life Science Perspective“, Schloss Dagstuhl, Wadern, Januar 2003.

S. HEYMANN: *Alternative Spleißformen auf dem Prüfstand*. Ehrenkolloquium für J.C. Freytag aus Anlass der Verleihung des IBM Faculty Awards 2003, Berlin, Oktober 2003.

S. HEYMANN: *Spleißformen eines Neurotensin-Rezeptors*. Einladungsvortrag am Forschungszentrum Rossendorf, November 2003.

J.-C. FREYTAG, S. HEYMANN: *Data & Process Integration and Alternative Splicing*. BCB-Tagung, Berlin, Juni 2003.

C. BEN NECIB, H. MÜLLER, P. RIEGER, K. THAM, S. HEYMANN, T. M. GLÄSSER, J.-C. FREYTAG: *Genome Data Warehouse: Ready for Use*. Interne Evaluation des BCB, Berlin, November 2003.

S. HEYMANN, T. M. GLÄSSER, M. HILLER, P. RIEGER, R. BACKOFEN, J.-C. FREYTAG: *Plurivalent Information Chopping & Concatenation in silico: Search for More Protein Phenotypes Encoded by One Gene*. Interne Evaluation des BCB, Berlin, November 2003.

C. BEN NECIB: *Ontology based Query Processing in Database Management Systems*. The International Conference on Ontologies, Databases and Applications of SEMantics (ODBASE) Catania, Italy, November 2003.

H. MÜLLER: *Semantic Data Cleansing in Genome Databases – Exemplified on Gene Prediction Data*. GK-Workshop, Hubertusstock, April 2003.

H. MÜLLER: *Semantic Data Cleansing in Genome Databases*. VLDB 2003 PhD Workshop, Berlin, September 2003.

H. MÜLLER: *Data Quality in Genome Databases*. Conference on Information Quality, Boston (USA), Oktober 2003.

H. MÜLLER: *Comparing Data Sources for Reliability-based Merging*. GK-Workshop, Berlin, November 2003.

Gastvorträge

STEFAN DECKER, USC Information Sciences Institute: *TRIPLE - eine regelbasierte Sprache für die Verarbeitung von Metadaten mit heterogener Semantik*. XML-Kolloquium, Januar 2003.

FRANK LEYMAN, IBM Deutschland Entwicklung GmbH: *Web Services: Distributed Applications Without Limits*. XML-Kolloquium, Mai 2003.

RONALD MELSTER, Fraunhofer First: *Wissensvisualisierung auf der Basis von XML Topic Maps*. XML-Kolloquium, Juli 2003.

OLIVER BECKER, Humboldt-Universität zu Berlin: *Streaming Transformations for XML – STX*. XML-Kolloquium, Oktober 2003.

ELKE KASIMIR, EsPresto AG: *μ - eine Sprache für den medienübergreifenden Schema-Entwurf*. XML-Kolloquium, November 2003.

Besuch an der Lehr- und Forschungseinheit

ELA HUNT, University of Glasgow, März 2003.

B. KUCKLICK, IBM Entwicklung Deutschland GmbH, Vorstellung des IBM Discovery Links, Februar 2003.

PROF. DR. FRANK LEYMAN, IBM Entwicklung Deutschland GmbH, Mai 2003.

Sonstige Aktivitäten

Prof. Johann-Christoph Freytag, Ph.D.

- Mitglied des VLDB-Endowments
- Mitglied des Beirates des Berliner Zentrums für Genomforschung (BCB), Berlin
- Mitglied des Beirates des Zentrums für Biophysik und Bioinformatik, Humboldt-Universität zu Berlin

- Organisator des Dagstuhl-Workshops Information and Process Integration: A Life Science Perspective (zusammen mit Thure Etzold, Lion, England, Carole Goble, University of Manchester, England, Peter Schwarz, IBM Almaden Research Center, CA, USA und Rolf Apweiler, EBI, England), Januar 2003
- Teilnahme an der Konferenz Datenbanksysteme für Business, Technologie und Web (BTW), Leipzig, 2003
- Organisator der der „29th International Conference on Very Large Data Bases“, Berlin, September 2003 (zusammen mit H. Schewpe und P. Lockemann)
- Arbeitstreffen an der Universität Jena, Prof. Backofen, März 2003
- Teilnahme am IBM Workshop „Life Science“, IBM Böblingen, März 2003
- NSF-Gutachter, Washington DC, USA, Mai 2003
- Teilnahme GK-Workshop „Verteilte Informationssysteme“, Jagdschloss Hubertusstock, April 2003
- Einladung der IBM zum 50. Geburtstag der IBM Deutschland, September 2003
- Erhalt des IBM-Faculty-Awards, Oktober 2003
- Arbeitstreffen an der Universität Jena, Prof. Backofen, Oktober 2003
- Teilnahme am “Workshop on Data Provenance and Annotation”, Edinburgh, Dezember 2003

Dr. Rainer Eckstein, Dipl.-Inf. Ralf Heese

- Organisation des Workshops „XML-Technologien für Middleware – Middleware für XML Anwendungen XMIDX 2003“, Berlin, im Rahmen des XML Clearinghouse, Februar 2003
- Organisation der Konferenz „Berliner XML Tage 2003“, Berlin, im Rahmen des XML Clearinghouse, Oktober 2003
- Organisation des „Semantic Web Infotages“, Berlin, im Rahmen des XML Clearinghouse, Februar 2003
- Organisation des Infotages „XML-Technologien im E-Business“, Berlin, im Rahmen des XML Clearinghouse, Dezember 2003

Dr. Rainer Eckstein

- Organisation des Workshops Semantic Web and Databases, (Colocated zur VLDB) Berlin, (mit Isabel Cruz, Vipul Kashyap, Stefan Decker), September 2003
- Teilnahme an Konferenz VLDB, September 2003 Berlin
- PC-Mitglied Workshop Practical and Scalable Semantic Systems PSSS, Oktober 2003
- Gutachter Datenbankspektrum (Organ der Fachgruppe Datenbanken der Gesellschaft für Informatik e.V.)

Dipl.-Inf. Ralf Heese

- Teilnahme am Workshop „Web Services meet Semantic Web“, Darmstadt, Juni 2003
- Teilnahme an der „International Conference on Very Large Databases“, Berlin, September 2003
- Teilnahme an der Konferenz „International Semantic Web Conference 2003“, Sanibel Island, USA, Oktober 2003
- Teilnahme am IBM Scholars DB2 Tech Update Event, Böblingen, November 2003.

Dipl.-Inf. Heiko Müller

- Teilnahme am Workshop des GKs „Verteilte Informationssysteme“, Hubertusstock, April 2003

- Teilnahme an der „29th International Conference on Very Large Databases“, Berlin, September 2003
- Teilnahme an dem „VLDB 2003 PhD Workshop“, Berlin, September 2003
- Teilnahme an der „Conference on Information Quality“, Boston (USA), Oktober 2003
- Teilnahme am Workshop des GKs „Verteilte Informationssysteme“, Berlin, November 2003

Dipl.-Inf. Chokri Ben Necib

- Teilnahme an der Konferenz „Very Large Databases“, Berlin, September 2003
- Teilnahme an der Konferenz „Ontologies, Databases and Applications of Semantics“, Catania, November 2003

Dr. Stephan Heymann

- Teilnahme am „International BCB-Workshop on Statistics and Cancer Genomics“, Berlin, Magnus-Haus, August 2003
- Vorlesung Bioinformatik, Technische Fachhochschule Wildau, Dualer Studiengang Bioprozesstechnik/Bioinformatik, Wildau, SS 2003
- Ko-Betreuung einer externen Belegarbeit (Uni Potsdam) „Modellierung der Gene Expression Markup Language (GEML) mit den Mitteln der Unified Modelling Language (UML)“, September 2003
- Teilnahme an der 29. Konferenz „Very Large Databases“, Berlin, September 2003

Peter Rieger

- Teilnahme an der Konferenz „Very Large Databases“, Berlin, September 2003
- Organisation der IBM DB2 Zertifizierung am Lehrstuhl DBIS, Oktober 2003
- Arbeitstreffen an der Universität Jena, Prof. Backofen, Oktober 2003
- Teilnahme am IBM DB2 Scholars Program Technical Update, Böblingen, November 2003

Dipl.-Inf. (FH) Katja Tham

- Einwerbung eines Promotionsstipendium: Graduiertenprogramm der Fachhochschule für Technik und Wirtschaft (FHTW) Berlin
- Teilnahme an der Konferenz Datenbanksysteme für Business, Technologie und Web (BTW), Leipzig, Februar 2003
- Teilnahme an der Konferenz Very Large Databases (VLDB), Berlin, September 2003

Dipl.-Inf. Sven Herschel

- Teilnahme an der Konferenz Very Large Databases (VLDB), Berlin, September 2003
- Teilnahme an der Jahrestagung der GI, Frankfurt a.M., Oktober 2003

Diplomarbeiten

KARSTEN LÜCKE: *Indexierung von XML-Dokumenten: Entwurf und Implementation für die AST/TA-Datenstruktur*. März 2003.

MICHAEL BELL: *Entwurf und Implementation des persistenten Access Support Tree*. März 2003.

ANNETT RUMPELT: *Eine RDF-basierte semantische Suchmaschine – Konzeptuelle Entwicklung, Inferenzmechanismen*. April 2003.

Lehr- und Forschungseinheit

Wissensmanagement in der Bioinformatik

<http://www.informatik.hu-berlin.de/wbi>

Leiter

PROF. DR. ULF LESER

Tel.: (030) 2093 3902

E-mail: leser@informatik.hu-berlin.de

Sekretariat

CORNELIA MISPELHORN

Tel.: (030) 2093 3901

E-mail: mispel@informatik.hu-berlin.de

Wissenschaftliche Mitarbeiter

DIPLOM.-INF. JÖRG HAKENBERG

SILKE TRISSL, MSc

DR. LUKAS FAULSTICH

Studentische Hilfskräfte

CONRAD PLAKE

RAPHAEL BAUER

MUHAMMAD GHIYAS

Techniker

ROBERT MIELKE

Der Lehrstuhl beschäftigt sich mit Wissens- und Datenmanagement in der molekularbiologischen und biomedizinischen Forschung. Die Forschungsarbeiten konzentrieren sich auf die Bereiche Integration biomedizinischer Daten, Modellierung und Performanz von molekularbiologischen Datenbanken, und Methoden zur automatischen Analyse von biomedizinischen Veröffentlichungen (TextMining). In der Lehre werden Themen aus dem Bereich Datenbanken, Datenanalyse und der Bioinformatik behandelt.

Die Professur wurde am 1.10.2002 eingerichtet. Das Jahr 2003 war deshalb stark vom Aufbau der neuen Gruppe geprägt. Der erste wissenschaftliche Mitarbeiter nahm im Februar seine Tätigkeit auf, und bis Ende April waren alle Stellen besetzt. Zusätzlich zum Abschluss des personellen Aufbau des Lehrstuhls konnten wir im Dezember, nach mehrmaligen Umzügen, bedingt durch den Umbau der unteren Stockwerke von Haus IV, auch unsere endgültigen Räumlichkeiten beziehen.

Inhaltlich hat die Gruppe im Laufe des Jahres eine Reihe interessanter Forschungsprojekte begonnen. Alle unsere Projekte sind hochgradig interdisziplinär; es bestehen Kooperationen mit Molekularbiologen, Biochemikern, Biophysikern und Linguisten. Zusammen mit Forschern des Max-Planck-Instituts für molekulare Genetik werden Fragen des TextMinings in der Systembiologie untersucht, in einer Kooperation mit der Charité und verschiedenen anderen Instituten in Berlin arbeiten wir an der Erstellung einer integrierten Proteindatenbank und mit dem Institut für deutsche Sprache der Humboldt Universität wurde ein

interdisziplinärer Forschungsverbund zur Untersuchung von Abstammung und Vererbung am Beispiel von molekularbiologischen Sequenzen und geschriebener Sprache gegründet. Darüber hinaus wurden im Rahmen des Institutsschwerpunkts „Große Datenmengen in webbasierten Umgebungen“ verschiedene Aktivitäten begonnen bzw. intensiviert. So wurde mit der Gruppe von Prof. Naumann ein gemeinsames Forschungsseminar initiiert und mit der LFE von Prof. Freytag wurden die existierenden Kooperationen im Bereich der Integration molekularbiologischer Daten fortgeführt.

Im Jahr 2003 wurde auch die technische Infrastruktur der LFE beschafft und gegen Ende des Jahres erfolgreich installiert. Es stehen nun zwei Dual-Prozessorserver mit insgesamt 3 Terabyte Plattenplatz unter Linux, zwei Fileserver und entsprechende Workstations zur Verfügung. Die Server arbeiten im Wesentlichen als Datenbankserver, sowohl für die Forschung als auch für die Lehre.

Eine große Kraftanstrengung für die gesamte Gruppe war die Mitorganisation der 25. Conference on Very Large Databases (VLDB) im September in Berlin. Sämtliche Mitglieder des Lehrstuhls waren, neben vielen weiteren Kollegen, über mehrere Wochen voll eingebunden bei der Bereitstellung der technischen Infrastruktur, dem Betrieb des Konferenzbüros und der Sicherstellung eines reibungslosen Ablaufs aller Vorträge. Nicht zuletzt Dank dieser Hilfe wurde die Konferenz zu einem großen Erfolg.

Lehre im Hauptstudium

Halbkurse

- Vorlesung „Bioinformatik“ (PROF. LESER, WS 2003/2004)

Spezialveranstaltungen

- Vorlesung „Molekularbiologische Datenbanken“ (PROF. LESER, SS 2003)
- Übung zur Vorlesung „Molekularbiologische Datenbanken“ (PROF. LESER, SS 2003)
- Vorlesung „Data Warehousing“ (PROF. LESER, SS 2003)
- Übung zur Vorlesung „Data Warehousing“ (PROF. LESER, SS 2003)

Seminare

- Seminar „Integration molekularbiologischer Daten“ (PROF. LESER, WS 2002/2003)
- Seminar „Advanced Data Warehousing“ (PROF. LESER, WS 2003/2004)

Lehre im Zentrum für Biophysik und Bioinformatik

- Vorlesung „Bioinformatik für Biophysiker“, mit Softwarepraktikum, PROF. LESER, WS 2003/2004 (interdisziplinäre Vorlesung zusammen mit Prof. Frömmel, Dr. Preissner, Dr. Meisel (alle Charité) und Prof. Herzel (Institut für Biologie).

Forschung

Projekt: Berlin Center for Genome-Based Bioinformatics

Ansprechpartner: PROF. DR. ULF LESER

Zusammenarbeit: Max-Planck-Institut für Molekulare Genetik, Universitätsklinikum Charité, Konrad-Zuse-Zentrum für Informationstechnik, Max-Delbrück Centrum für molekulare Medizin, Freie Universität Berlin

Forschungsförderung: Bundesministerium für Bildung und Forschung

Das „Berlin Center for Genome-Based Bioinformatics“ (BCB) hat das Ziel, die Lücke zwischen der genomorientierten Grundlagenforschung und der medizinischen Anwendung molekularbiologischer Erkenntnisse zu schließen. Das BCB besteht aus 3 Professuren und 7 Nachwuchsforschergruppen, die sich in interdisziplinären Verbänden um die folgenden Themen gruppieren: (a) Annotation und Wissensmanagement von biomedizinischen Daten, (b) Untersuchung von Struktur und Funktion von Genprodukten und (c) der Modellierung von Zellen, Krankheiten und Stoffwechselwegen.

Projekt: Columba - A Database of Protein Structure Annotation

Ansprechpartner: PROF. DR. ULF LESER, SILKE TRISSL

Beteiligte Mitarbeiter: SILKE TRISSL, RAPHAEL BAUER

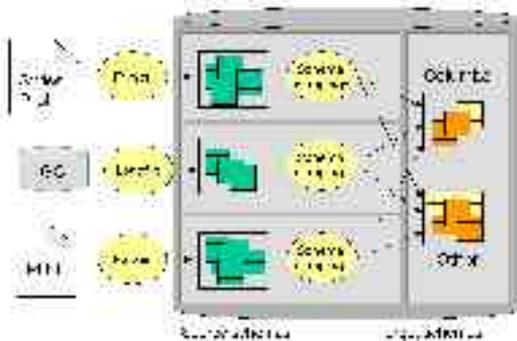
Zusammenarbeit: Universitätsklinikum Charité: PROF. FRÖMMEL, DR. PREISSNER, K. ROTHER; Konrad Zuse Zentrum für Informationstechnik: DR. STEINKE; Technische Fachhochschule Berlin: PROF. KOCH; Humboldt-Universität Berlin: H. MÜLLER

Forschungsförderung: Im Rahmen des BCB

In biologischen Fragestellungen sind Forscher an mehr als nur Proteinsequenzen bzw. -strukturen interessiert. Auch Informationen über deren Funktion, ihre biologische Bedeutung, verwandte Proteine, oder über Aufgaben der Proteine im Stoffwechsel sind von Bedeutung. Die entsprechenden Informationen sind im Web auch frei verfügbar, sind aber verteilt über eine große Anzahl heterogener Datenbanken.

Ziel des Projektes Columba ist es, diese verteilten Daten in einer zentralen relationalen Datenbank zu sammeln und semantisch zu integrieren. Diese Integration verfolgt mehrere Ziele: Zum einen ist die Abfrage der Daten für Benutzer wesentlich einfacher als eine aufwändige Suche im Web, und zum anderen können durch das Ausnutzen von Redundanz, d.h. doppelt auftretenden Daten, Rückschlüsse über die Qualität der Ausgangsdatenbanken gezogen werden.

Kernstück von Columba sind Daten über Proteinstrukturen aus der Protein Data Bank (PDB). Diese werden ergänzt mit Informationen über Sekundärstrukturelemente (DDSP), Faltungstyp des Proteins (CATH und SCOP), Beschreibung der Proteinfunktion (SwissProt und GeneOntology), Beteiligung des Proteins an Stoffwechselfvorgängen oder Signalketten in Zellen (Enzyme und KEGG), aktive Flächen auf der Proteinoberfläche (DIP) und geometrischen Eigenschaften der Proteinstruktur (PTGL). In Columba sind zur Zeit 12 verschiedene Datenquellen integriert, von denen auch 8 über das Webinterface abfragbar sind.



Architektur von Columba. Die Daten werden auf verschiedene Weisen in eine relationale Datenbank eingebracht. Die initialen Schema werden in einem zweiten Schritt, der die eigentliche Integration vornimmt, in das Columba Schema transformiert. Benutzer richten ihre Anfragen nur gegen dieses integrierte Schema.

Webinterface von Columba. Das Interface basiert auf Query Refinement, d.h., ausgehend von einer ersten Anfrage wird die Ergebnismenge iterativ durch Hinzufügen weiterer Bedingungen verfeinert. Der Benutzer erhält laufend Feedback über die aktuellen Ergebnisse und kann dadurch unsinnig große oder leere Ergebnismengen frühzeitig erkennen und verhindern.

Columba dient auch zur Untersuchung der Qualität von molekularbiologischen Datenbanken. Beispielsweise werden Daten oft durch fehlerhafte Parser, falsche Transformationen oder nichtzutreffende Beschreibungen verunreinigt.

	P(A)	Columba	MMS
NAME	0,999	2.299	19.736
YEAR_PDB_DEPOSITION	0,004	33	34
DEPOSITION_DATE	0,006	3.755	3.949
RELEASE_DATE	0,572	1.320	1.184
STRUCTURE_METHOD	0,183	135	96
RESOLUTION	0,451	289	356
R_VALUE	0,999	1	851
R_FREE	0,999	1	1.492
REFINEMENT PROGRAM	1	1	485

Wechselseitige Konsistenz von Proteinstrukturdaten. Columba/PDB und MMS sind zwei Datenbanken die auf dem selben Datensatz beruhen, der aber in den beiden Systemen unterschiedliche Bearbeitungsschritte durchläuft. P(A) gibt an, in wie vielen Prozent aller Fällen sich die Werte in den beiden Datenbanken unterscheiden.

Das Projekt wird in einer Kooperation mit Partner an der Charité, dem Zuse-Zentrum und der Technischen Fachhochschule durchgeführt.

Projekt: Text Mining für die Systembiologie

Ansprechpartner: PROF. DR. ULF LESER

Beteiligte Mitarbeiter: JÖRG HAKENBERG, CONRAD PLAKE

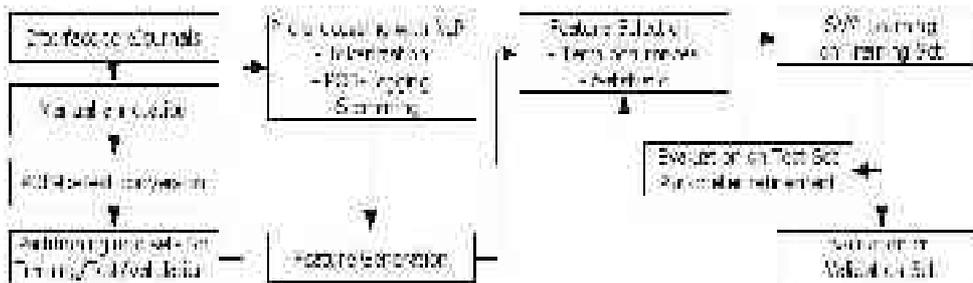
Zusammenarbeit: Max-Planck-Institut für molekulare Genetik: DR. KLIPP, DR. KOWALD

Forschungsförderung: Im Rahmen des BCB

Das Projekt beschäftigt sich mit dem Erkennen von Fachpublikationen, die für die kinetische Modellierung biochemischer Systeme relevant sind. In diesen Publikationen finden sich wichtige Daten wie Enzymkonzentrationen und Reaktionskonstanten und -gleichungen, die experimentell im Labor beobachtete Stoffwechselvorgänge beschreiben. Leider werden Art und Ausrichtung vormals einschlägiger Fachzeitschriften in den letzten Jahren zunehmend

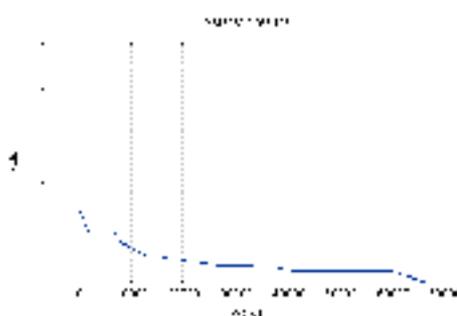
diffuser und auch die reine Anzahl von neu erscheinenden Artikeln und Journals macht eine manuelle Suche nach interessanten Artikeln unmöglich.

In dem Projekt wird ein Verfahren entwickelt, das automatisch eine thematische Einsortierung von Fachartikeln vornimmt und damit dem Forscher wertvolle Zeit erspart. Wir benutzen dazu maschinelle Lernverfahren, die Vorhersagen bezüglich der Relevanz von noch ungelesenen Fachartikeln treffen. Die Artikel werden durch eine ganze Reihe von Verarbeitungsschritten aufbereitet, analysiert und schließlich automatisch bewertet. Die entsprechende Software wird am Lehrstuhl in einer JAVA Bibliothek zusammengefasst.

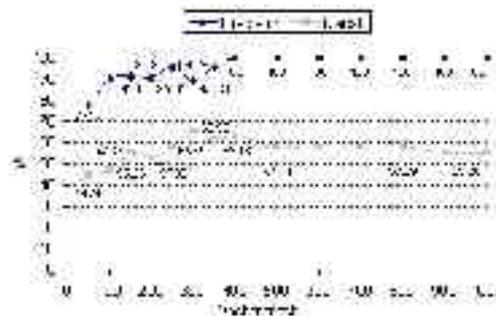


Schritte der Dokumentenverarbeitung, beginnend vom Zugreifen auf Artikel über das Internet, die Transformation von PDF Dokumenten in Textdateien, die Auswahl relevanter Wörter und das Trainieren eines maschinellen Lernverfahrens auf den gesammelten Dokumenten.

Zum Einsatz kommen Support Vector Machines, die aus einer hand-annotierten Menge von Trainingsbeispielen ein Modell zur Klassifizierung neuer Texte generieren. Sämtliche Artikel überführen wir dazu in eine Vector Space Model-Repräsentation, die eine räumliche Anordnung und damit die Berechnung einer Trennebene ermöglicht. Die bisher erzielten Resultate übertreffen die manuelle und mühsame Suche nach relevanten Publikationen bei weitem.



Bewertung der Diskriminierungsfähigkeit von Wörtern. Es wird klar deutlich, dass der Großteil aller Wörter in allen Dokumenten, egal welcher Klasse sie zugeordnet sind, gleich verteilt sind. Nur ein kleiner Prozentsatz ist tatsächlich diskriminierend.



Qualität der Klassifizierungsergebnisse in Abhängigkeit von der Anzahl zur Diskriminierung benutzter Wörter. Precision gibt an, in wie vielen Fällen das Verfahren eine Klasse richtig vorhersagt; Recall ist ein Maß für die Wahrscheinlichkeit, mit der ein Dokument richtig klassifiziert wird.

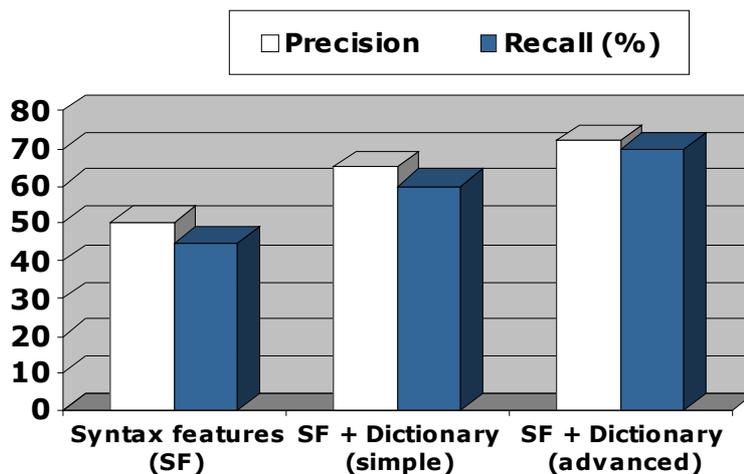
Projekt: Critical Assessment of Information Extraction Systems in Biology

Ansprechpartner: Prof. Dr. Ulf Leser

Beteiligte Mitarbeiter: Jörg Hakenberg, Conrad Plake, Lukas Faulstich

Zusammenarbeit: Humboldt-Universität Berlin: Prof. Scheffer

Der BioCrAtiVe-Cup ist ein internationaler Wettbewerb zu verschiedenen offenen Problemen und Fragestellungen aus dem Text Mining in den Lebenswissenschaften (Medizin, Biologie, Biochemie, Bioinformatik). Anhand fest vorgegebener Aufgabenstellungen, klaren Zielsetzungen und gemeinsam verwendeter, identischer Datensätze sollen unterschiedliche Lösungsansätze ausprobiert und einander gegenübergestellt werden. Die Teilaufgaben in diesem Jahr beschäftigten sich mit der Named Entity Recognition (Erkennung von Protein- und Gennamen), Entity Normalization (Zuordnung von Namen zu eindeutigen Bezeichnern) und der Automatic Functional Annotation (Zuordnung von Genen zu ihrer Gene-Ontology-Annotation). In jeweils zwei verschiedenen Kategorien konnten dabei Systeme getestet werden, die zum einen nur die aus den vorliegenden Texten extrahierte Daten verwenden oder sich zum anderen beliebiger Datenquellen bedienen dürfen. Unsere Gruppe beschäftigte sich mit der ersten Aufgabe, Namen von Genen und Proteinen in Freitexten zu identifizieren.



Schrittweise Verbesserung der Vorhersagequalität durch Definition neuer Merkmale von Proteinennamen und unter Einbeziehung eingeschränkter oder offener Wörterbücher mit bekannten Namen.

Anhand der zur Verfügung gestellten hand-annotierten Beispieldaten konnten wir eine Support Vector Machine trainieren, die Wörter in den neuen, ungesehenen Daten aus der Aufgabenstellung klassifizieren kann. Die dazu benötigten Merkmalsvektoren setzen sich zusammen aus verschiedenen Klassen von Merkmalen, deren Einfluss sukzessive ausprobiert wurde. So spielen beispielsweise die in Gennamen enthaltenen und für diese typischen Buchstabenkombinationen (q-Gramme) eine sehr wichtige Rolle. Innerhalb kurzer Zeit konnte ein System erstellt werden, dass eine Vorhersagegenauigkeit von ca. 72 % erreicht und damit viele in der Literatur beschriebene Verfahren bei weitem übertrifft. Das Projekt wurde zusammen mit der Gruppe von Prof. Scheffer durchgeführt.

Projekt: Interdisziplinärer Forschungsverbund Linguistik - Bioinformatik

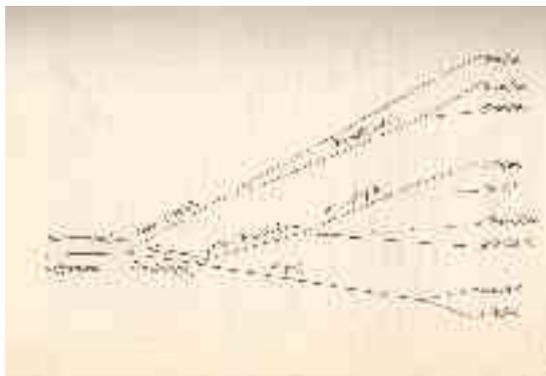
Ansprechpartner: PROF. DR. ULF LESER

Beteiligte Mitarbeiter: DR. LUKAS FAULSTICH

Zusammenarbeit: Institut für Deutsche Sprache, Humboldt-Universität: PROF. LÜDELING, PROF. DONHAUSER

Forschungsförderung: Senatverwaltung der Stadt Berlin

Linguistik und Bioinformatik haben überschneidende Fragestellungen vor allem in zwei Bereichen: der Erarbeitung, Modellierung und Speicherung komplexer Annotationen auf Zeichenketten sowie der vergleichenden Analyse von Zeichenketten zur Ermittlung von Änderungsprozessen in historischen Zeiträumen. Ziel des Forschungsverbundes ist es, Methoden und Erkenntnisse aus dem Bereich Biologie/Bioinformatik auf den Bereich Linguistik/Korpuslinguistik zu übertragen und umgekehrt.



Abstammungsbäume von Sprachfamilien (Links, aus einem Manuskript aus dem 19. Jahrhundert) und in der Biologie (rechts, Teil des Stammbaums von Primaten). Stammbäume werden auf verschiedene Weisen hergestellt; in der Sprachwissenschaft durch den Vergleich struktureller Eigenschaften, in der Biologie durch morphologische Ähnlichkeit und durch den computerbasierten Vergleich der DNA Sequenzen verschiedener Spezies.

Verwandtschaft und Abstammungsbeziehungen werden in der modernen Biologie zunehmend durch den computerbasierten Vergleich von DANN-Sequenzen ermittelt. Diese Methoden sollen innerhalb des Forschungsverbundes auf Texte unterschiedlicher Sprachfamilien des Deutschen übertragen werden. Dazu wird als empirische Basis eine repräsentative Sammlung digitalisierter und annotierter Texte benötigt, auf denen die Vergleichsalgorithmen operieren. Eine entsprechende Sammlung existiert zur Zeit nicht.

Der Forschungsverbund widmet sich deswegen in seiner ersten Phase insbesondere der Modellierung und Implementierung von Datenbanken, die entsprechende Textsammlungen speichern und recherchierbar machen. Dazu beteiligt er sich intensiv an der Vorbereitung des Projekts DDD (siehe nächsten Abschnitt).

Am Forschungsverbund sind neben unserer LFE noch die Gruppe von Prof. Lüdeling und Prof. Donhauser, beide vom Institut für deutsche Sprache der Humboldt-Universität, beteiligt.

Projekt: Deutsch. Digital. Diachron.**Ansprechpartner:** PROF. DR. ULF LESER**Beteiligte Mitarbeiter:** DR. LUKAS FAULSTICH**Zusammenarbeit:** Institut für Deutsche Sprache, Humboldt-Universität Berlin: PROF. LÜDELING, PROF. DONHAUSER

Bundesweit weiter 15 Gruppen im Bereich Linguistik, Germanistik und Philologie

Texte in deutscher Sprache sind seit dem 9. Jahrhundert in Handschriften und Drucken überliefert. Sie spiegeln einen wesentlichen der europäischen (Kultur-)Geschichte und sind von großem historischen und kulturwissenschaftlichen Interesse. Bislang sind diese Zeugnisse überwiegend in gedruckten Editionen verfügbar, digitalisierte Texte sind auf Einzelkorpora verteilt, die nur zum Teil öffentlich sind und unterschiedlich gut aufbereitet wurden.

Das Projekt DDD ist eine bundesweite Initiative einer interdisziplinären Forschergruppe, die sich zusammengeschlossen hat, um ein Digitales Referenzkorpus des Deutschen zu entwickeln. Zielsetzung ist die Erstellung einer Digitalen Bibliothek, die Faksimiles von Handschriften und Drucken sowie deren Inhalte in Originalsprache und Übersetzung bereitstellt. Arbeitsgruppen an verschiedenen Universitäten Deutschlands werden je eine Sprachepoche bearbeiten und die philologisch geprüften und mit linguistischen Annotationen versehenen elektronischen Texte an eine technische Arbeitsstelle (in Berlin) weiterleiten. Diese gliedert die Texte in einer für alle Sprachstufen übereinstimmenden Form in die Digitale Bibliothek ein und verwaltet sie.

Der Forschungsverbund fungiert innerhalb des DDD als Vorprojekt. Es wurden verschiedene Vorstudien zur Konvertierung unterschiedlicher XML-Repräsentationen von Korpora und deren Annotationen in ein einheitliches Modell durchgeführt, ein Modell zur Speicherung und Retrieval beliebiger, verschränkter und miteinander in Beziehung stehender Annotationsebenen für digitalisierte Texte entworfen sowie eine Architektur für die Implementierung der zentralen Korpusdatenbank erstellt. Die Ergebnisse wurden im Dezember auf einer nationalen Konferenz zur Vorbereitung des DDD-Projektes vorgestellt und werden als technische Grundlage der Antragstellung dienen.

Veröffentlichungen

F. NAUMANN, J.-C. FREYTAG, U. LESER: *Completeness of Integrated Information Sources*. Journal on Information Systems (to appear).

U. LESER, P. RIEGER: *Integration molekularbiologischer Daten*. Datenbankspektrum 6, pp. 56-66.

J. HAKENBERG, A. NUSSBAUM, ST. STEFANOVIC, H.-G. RAMMENSEE, CH. KUTTLER, H.-G. HOLZHÜTTER, P.-M. KLOETZEL, ST. H.E. KAUFMANN, H.-J. MOLLENKOPF: *MAPPP - MHC-I Antigenic Peptide Processing Prediction*. Applied Bioinformatics, 2(3).

F. NAUMANN, J.-C. FREYTAG, U. LESER: *Completeness of Information Sources*. Workshop on Data Quality in Cooperative Information Systems, Siena, Italy.

S. SCHMEIER, J. HAKENBERG, A. KOWALD, E. KLIPP, U. LESER: *Text Mining for Systems Biology Using Statistical Learning Methods*. 3. Workshop des Arbeitskreises Knowledge Discovery, Karlsruhe, Germany.

L. C. FAULSTICH, P. F. STADLER, C. THURNER, C. WITWER: *litsift: Automated text categorization in bibliographic search*. In Data Mining and Text Mining for Bioinformatics, Workshop at the ECML / PKDD 2003.

T. SCHEFFER, U. LESER (EDITORS): *Data Mining and Text Mining for Bioinformatics*. Proceedings of the European Workshop in conjunction with ECML/PDCK, Dubrovnik, Croatia.

K. ROTHER, S. TRISSL, H. MÜLLER, P. MAY, R. HEEK, R. PREISSNER, T. STEINKE, I. KOCH, U. LESER, C. FRÖMMEL: *Columba - A Database of Annotations of Protein Structure*. 2nd European Conference on Computational Biology, Paris (Poster).

K. ROTHER, H. MÜLLER, S. TRISSL, I. KOCH, T. STEINKE, R. PREISSNER, C. FRÖMMEL, U. LESER: „*COLUMBA: Multidimensional Data Integration of Protein Annotations*“. International Workshop on Data Integration in the Life Sciences, Leipzig, Germany (to appear).

Vorträge

U. LESER: *Completeness of Information Sources*. Workshop on Data Quality in Cooperative Information Systems, Siena, Italien.

U. LESER: *Data Quality in Bioinformatics*. Panel on "Data Quality in Cooperative Information Systems", at DQCIS Workshop, Siena, Italien.

U. LESER: *Data Quality in Life Science Data Integration*. Dagstuhl Seminar "Data and Process Integration: A Life Science Perspective", Dagstuhl.

U. LESER: *Ontologien in der Bioinformatik*. 31.1.2003, Institutskolloquium des Fraunhofer Instituts für System- und Softwaretechnik, Berlin.

U. LESER: *Der Berliner Centrum für Genombasierte Bioinformatik*. 1.4.2003, BioTOP Statusseminar Bioinformatik, Investitionsbank Berlin.

U. LESER: *Ontologien in der Bioinformatik*. Workshop "Systembiologie mikrobieller Organismen", Bielefeld.

U. LESER: *Wissensmanagement in der Bioinformatik*. 1. Symposium des Berlin Center for Genombased Bioinformatics, Berlin.

U. LESER: *Integrating Protein Data*. Tag der Wissenschaft des Zentrums für Biophysik und Bioinformatik, Berlin.

U. LESER: *Data Integration in Life Sciences*. EMBO Course on „Sequence Analysis and Data Mining“, Peru, Lima.

U. LESER: *Mining Biomedical Literature*. EMBO Course on „Sequence Analysis and Data Mining“, Peru, Lima.

J. HAKENBERG, U. LESER: *Was ist Bioinformatik?* Schülertag des Instituts für Informatik der Humboldt-Universität, Berlin.

U. LESER: *Verwandtschaft und Abstammung – Linguistics meets Bioinformatics* Interdisziplinäres Symposium zum Verhältnis Geistes- zu Naturwissenschaften, Berlin.

J. HAKENBERG: *Text Mining for Systems Biology Using Statistical Learning Methods*. 3. Workshop des Arbeitskreises Knowledge Discovery, Universität Karlsruhe.

J. HAKENBERG: *Statistical Approaches to Analyze Positional Dependencies in Protein Domains*. BCB-Seminar, Humboldt-Universität Berlin.

J. HAKENBERG: *Ein statistischer Ansatz zur Beschreibung von Strukturen in Proteindomänen*. DBIS Forschungsseminar, Humboldt-Universität Berlin.

C. FAULSTICH: *litsift: Automated text categorization in bibliographic search*. Workshop on Data Mining and Text Mining for Bioinformatics, Cavtat, Kroatien.

Tutorials

U. Leser: *An Introduction to Relational Databases*, EMBO Course on „Sequence Analysis and Data Mining“, Peru, Lima

U. Leser: *Molecular Biology Databases*, EMBO Course on „Sequence Analysis and Data Mining“, Peru, Lima

Gastvorträge und Gastwissenschaftler

- Fouzia Moussouni, Hospital de Pontchaillou, Rennes (Gastwissenschaftlerin)
- Dr. Kai-Uwe Sattler, Universität Illmenau (Vortrag)
- Prof. Backofen, Universität Jena (Vortrag)
- Dr. Dieterich, GBF Braunschweig (Vortrag)
- Christian Piepenbrock, Epigenomics AG Berlin (Vortrag)
- Kim Henrik, Melford John, European Bioinformatics Institute (EBI), UK (Vortrag)
- Lutz Müller (Praktikant)

Sonstige Aktivitäten

Prof. Dr. Ulf Leser:

- Aufnahme in das Zentrum für Biophysik und Bioinformatik (BPI) der Humboldt-Universität
- Mitglied des Programmkomitees für IQ 2003 und DILS 2004
- Gutachter für die DFG
- Koorganisation des Workshops "Data- und Text-Mining in Bioinformatics" (mit Prof. Scheffer, Humboldt-Universität)
- Technical Chair der VLDB 2003, Berlin
- Koorganisation des Projektsymposiums „Deutsch. Digital. Diachron.“, Berlin (mit Prof. Lüdeling und Prof. Donhauser, Humboldt-Universität)

Diplomarbeiten

SEBASTIAN SCHMEIER (Bachelorarbeit Bioinformatik): *Klassifizierung von biochemischen Texten mittels statistischer Lernverfahren*, Gemeinsame Betreuung mit dem Max-Planck-Institut für Molekulare Genetik, Berlin.

Lehr- und Forschungseinheit

Künstliche Intelligenz

<http://www.ki.informatik.hu-berlin.de/>

Leiter

PROF. DR. HANS-DIETER BURKHARD

Tel.: (030) 2093 3167

E-Mail: hdb@informatik.hu-berlin.de

Sekretariat

RENATE ZIRKELBACH

Tel.: (030) 2093 3167, Fax.: (030) 2093 3168

E-Mail: zirkel@informatik.hu-berlin.de

Wissenschaftliche Mitarbeiter

DIPL.-INF. JOSCHA BACH

DIPL.-MATH. MANFRED HILD

DIPL.-PHYS. JAN HOFFMANN

DR. GABRIELA LINDEMANN-VON TRZEBIATOWSKI

DIPL.-INF. MIRJAM MINOR

M. SC. DAGMAR MONETT DÍAZ

DIPL.-INF. (FH) ALEXANDER OSHERENKO

DIPL.-MATH. OLGA SCHIEMANGK

DIPL.-INF. KAY SCHRÖTER

DIPL.-KFM. DIPL.-INF. DIEMO URBIG

Promotionsstudenten

M. SC. BASMAH EL-HADDAD

DIPL.-INF. MOHAMMED MAYYAS

DIPL.-INF. HELMUT MYRITZ

M. SC. JESÚS EMETERIO NAVARRO BARRIENTOS

DIPL.-INF. JAN WENDLER

Gäste

B. SC. VLADIMIR KURBALJA, Universität Novi Sad, Serbien-Montenegro

PROF. DR. MIRJANA IVANOVIC, Universität Novi Sad, Serbien-Montenegro

M. SC. RAFAL LATKOWSKI, Universität Warschau, Polen

M. SC. FARLY NOBRE, Universität Birmingham, Großbritannien

M. SC. VICTOR EMANUEL UC CETINA, Universität Merida, Yukatan, Mexiko

Tutoren

BENJAMIN ALTMAYER

CHRISTINA BELL

RALF BERGER

BIRGER BRUNSWIEK

UWE DÜFFERT

DANIEL GÖHRING

MICHAEL GOLLIN

PETER HAMMELS

CLAUDIA HÄRTEL
MATTHIAS JÜNGEL
GORDON KRAMER
MARTIN LÖTZSCH
ROBIN MALITZ
CHRISTIAN WIECH



Die Forschungsgruppe Künstliche Intelligenz

Verständnis wächst mit aktiver Auseinandersetzung: Etwas zu „machen“, zu beherrschen, bedeutet zugleich besseres Verstehen. Angewandt auf die Erforschung geistiger Prozesse führt das auf die Nachbildung intelligenten Verhaltens mit Maschinen. So ist „Künstliche Intelligenz“ unter drei Aspekten zu sehen: Modellierung von Intelligenz mit dem Ziel, sie besser zu verstehen, Ausnutzung maschineller Leistungsfähigkeit zur Erledigung intelligenter Aufgaben, sowie Kooperation von Mensch und Maschine.

Wichtiges Anliegen unserer Arbeit ist die Erprobung und Verwertung der Erkenntnisse in praktischen Anwendungen. Das erfordert die Einbettung intelligenter Verfahren in umfassende Lösungen und betrifft insbesondere Probleme der Skalierbarkeit, Robustheit, Dauerhaftigkeit und Anpassungsfähigkeit. Natürliche Systeme haben hier vielfältige Formen entwickelt, die auch für technische Umsetzungen zunehmend interessant werden.

Schwerpunkte der Gruppe sind Fallbasiertes Schließen, Agenten-Orientierte Techniken, Verteilte Künstliche Intelligenz, Sozionik und Kognition mit deren Anwendungen für Wissensmanagement, eCommerce, Medizin und Robotik.

Im Rahmen des internationalen RoboCup-Wettbewerbs konnte das GermanTeam, an dessen Entwicklung der Lehrstuhl beteiligt ist, den diesjährigen „RoboCup Challenge“ der „Sony four-legged Robot League“ gewinnen. Der für das Fraunhofer Institut für Rechnerarchitektur und Softwaretechnik (FIRST) tätige Sören Sonnenburg erhielt für seine im Jahr 2002 fertig

gestellte und vom Lehrstuhl mitbetreute Diplomarbeit zum Thema „New Methods for Splice Site Recognition“ den Hugo-Geiger-Preis.

Im März 2003 führte der Lehrstuhl seine erste dreitägige Frühjahrsschule in Caputh bei Potsdam durch. Ziel der Veranstaltung war ein breiter Erfahrungsaustausch, sowie die Vorstellung der aktuellen Forschungsprojekte der wissenschaftlichen und studentischen Mitarbeiterinnen und Mitarbeiter, der Doktorandinnen und Doktoranden, und der Diplomandinnen und Diplomanden. Durch den auswärtigen Veranstaltungsort wurde sichergestellt, dass für ausführliche Diskussionen von Ideen und Ergebnissen genügend Zeit zur Verfügung stand. Ein besonderer Höhepunkt war ein Kurztraining in Teamarbeit und Vortragsgestaltung, welches von einer Mitarbeiterin der Weiterbildungsabteilung der Humboldt-Universität durchgeführt wurde.

Die Lehr- und Forschungseinheit beteiligte sich an den beiden DFG-Schwerpunktprogrammen „Sozionik“ und „Verteilte Systeme in dynamischen Umgebungen (RoboCup)“. Langjährige Kooperationsvereinbarungen existieren mit den Firmen empolis und RoboWatch.

Lehre

Veranstaltungen im Grundstudium

- Vorlesung „Praktische Informatik II“ (PROF. H.-D. BURKHARD, SS 2003)
- Übungen „Praktische Informatik II“ (DR. G. LINDEMANN- V. TRZEBIATOWSKI, DIPL.-INF. M. MINOR, DIPL.-INF. K. SCHRÖTER, DIPL.-INF. (FH) A. OSHERENKO SS 2003)
- Übungen „Praktische Informatik I“ (DR. G. LINDEMANN- V. TRZEBIATOWSKI, WS 2003/2004)

Kernveranstaltungen (Halbkurse)

- Halbkurs „Moderne Methoden der Künstliche Intelligenz“ (PROF. H.-D. BURKHARD, DR. H.-M. VOIGT, SS 2003)
- Praktikum „Moderne Methoden der Künstlichen Intelligenz“ (DIPL.-INF. J. BACH, DIPL.-PHYS. J. HOFFMANN, SS 2003)
- Projektkurs „Kognition und Robotik“ (PROF. H.-D. BURKHARD, DIPL.-INF. J. BACH, DIPL.-MATH. M. HILD, DIPL.-PHYS. J. HOFFMANN, SS 2003)
- Übung/Praktikum „Kognition und Robotik“ (DIPL.-INF. J. BACH, DIPL.-MATH. M. HILD, DIPL.-PHYS. J. HOFFMANN, SS 2003)
- Vorlesung „Einführung in die Künstliche Intelligenz“ (PROF. H.-D. BURKHARD, WS 2003/2004)
- Übungen „Einführung in die Künstliche Intelligenz“ (DIPL.-INF. M. MINOR, WS 2003/2004)
- Projektkurs „Kognition und Robotik II“ (PROF. H.-D. BURKHARD, DIPL.-INF. J. BACH, DIPL.-MATH. M. HILD, DIPL.-PHYS. J. HOFFMANN, WS 2003/2004)

Seminare

- „Evolution und Robotik II“ (DIPL.-MATH. M. HILD, SS 2003)
- „Evolution und Robotik“ (DIPL.-MATH. M. HILD, WS 2003/2004)
- „Simulation sozialer Systeme“ (DR. G. LINDEMANN- V. TRZEBIATOWSKI, DIPL.-KFM. DIPL.-INF. D. URBIG, WS 2003/2004)
- Proseminar „Verarbeitung natürlicher Sprache“ (DIPL.-INF. K. SCHRÖTER, WS 2003/2004)

Forschung

Fallbasiertes Schließen, Informationssysteme und Wissensmanagement

Fallbasiertes Schließen modelliert und implementiert das Handeln aus Erfahrung. Bekannte Lösungen vergangener Probleme werden genutzt für die Bewältigung aktueller Aufgaben. Ein zentrales Problem ist dabei das Erinnern an relevante Sachverhalte. Die dafür am Lehrstuhl entwickelte Technik der Case Retrieval Netze erlaubt vielfältige Anwendungen für Suchmaschinen im Wissensmanagement und im eCommerce. Neben entscheidungsunterstützenden Systemen haben sich in den letzten Jahren auch Knowledge-Management-Systeme etabliert. In jedem Einzelfall muss geklärt werden, wie Erfahrungswissen in Form von Fällen repräsentiert, kommuniziert und aktuell gehalten wird. In verschiedenen Projekten werden hierzu am Lehrstuhl sowohl theoretische Arbeiten geleistet als auch reale Anwendungen erstellt. Dabei arbeitet der Lehrstuhl seit Jahren erfolgreich mit der Firma empolis (ehemals tec:inno) zusammen.

Agenten-Orientierte Techniken und Verteilte KI

Die Zusammenarbeit intelligenter Systeme erfordert kooperatives und autonomes Verhalten, wobei der Begriff des „intelligenten Agenten“ eine zentrale Rolle spielt. Wesentliche Merkmale von Agenten sind aus unserer Sicht die dauerhafte autonome Arbeit in Wechselwirkung mit der Umgebung, wobei rationales Handeln durch Techniken der KI für Architektur, Programmierung und Interface erreicht werden soll. Die autonome Arbeit erfordert „Wissen“ über die Situation, „Fähigkeiten“ zur Lösung von Aufgaben und „Entscheidungen“ über beabsichtigte Handlungen. Agenten-Architekturen modellieren die dafür erforderlichen Strukturen, „Multi-Agenten-Systeme“ die Mechanismen der Koordination. Agenten-Orientierte Techniken sind Bestandteil in zahlreichen weiteren Projekten des Lehrstuhls.

Sozionik

Sozionik ist eine interdisziplinäre Forschungsrichtung von Soziologie und Informatik, insbesondere der Verteilten KI. Ziel ist die Erforschung und Modellierung künstlicher Sozialität. Soziale Konzepte sollen für Multi-Agenten-Systeme im Hinblick auf Robustheit, Skalierbarkeit und Adaptivität nutzbar gemacht werden. Umgekehrt ermöglichen informatische Plattformen das experimentelle Studium des dynamischen Wechselspiels zwischen Individuum und Gesellschaft und der Emergenz sozialen Verhaltens in verschiedenen Kontexten. In zukünftigen Anwendungsszenarien werden sich „hybride“ Gemeinschaften aus künstlichen Agenten und menschlichen Nutzern mit verteilter Handlungsträgerschaft herausbilden. Der Lehrstuhl kooperiert mit der TU Berlin und der Siemens-Forschung in München.

Medizin

Seit mehreren Jahren arbeiten wir an Projekten, die sich mit dem Einsatz intelligenter Systeme in der Medizin befassen. Krankenhäuser bieten mit umfangreichen Datenbeständen, heterogener Technik und komplexen organisatorischen Abläufen ein äußerst anspruchsvolles Einsatzgebiet. Umfangreiche offene Systeme wie Klinikinformationssysteme sind aus unserer Sicht nur inkrementell auf der Basis kooperierender Agenten zu entwickeln. Wenn solche Systeme Organisations- und Planungsaufgaben übernehmen, ist es wichtig für die Akzeptanz und Funktion, dass die vorhandenen Entscheidungsstrukturen und Kompetenzen auf adäquate Weise abgebildet werden (sozionische Problematik hybrider Systeme). Der Lehrstuhl kooperiert mit der Charité – Universitätsmedizin Berlin.

Kognitive Robotik

Die Kognitionswissenschaft untersucht Grundlagen und Funktionsweisen menschlicher Geistestätigkeiten wie z.B. Wahrnehmung, Handeln, Denken, Lernen. Sie ist ein interdisziplinäres Projekt insbesondere unter Beteiligung von Anthropologie, Informatik, Psychologie, Philosophie, Linguistik und Neurobiologie. Die Ergebnisse der Kognitionswissenschaft sollen für die Gestaltung künstlicher intelligenter Systeme (Roboter, Computer) genutzt werden. Gegenwärtiger Schwerpunkt der Untersuchungen sind kognitive Architekturen (Wahrnehmung, rationales Handeln, Emotionsmodelle) und ihre Anwendungen.

RoboCup

Im Rahmen des Testfeldes Computer-Fußball können Orientierung und Handeln von Agenten in komplexen dynamischen Umgebungen untersucht werden, wobei die internationale RoboCup-Initiative einen Rahmen für Kooperationen, Wettbewerbe und wissenschaftlichen Austausch schafft, und so die Forschung und Lehre in den Gebieten Künstliche Intelligenz und Robotik fördert. Durch die einheitliche Aufgabenstellung „Fußballspiel“ können verschiedenste Techniken ausprobiert, integriert und in den seit 1997 jährlich stattfindenden RoboCup-Weltmeisterschaften verglichen werden. Ergebnisse aus der RoboCup-Forschung lassen sich vielfach übertragen, insbesondere für die Steuerung von Multi-Agenten-Systemen, den Umgang mit dynamischen Umwelten, Sensorik und Aktorik, Servicerobotik.

Projekte

Projekt: Integration kooperationsfähiger Agenten in komplexen Organisationen – INKA

Ansprechpartner: PROF. DR. HANS-DIETER BURKHARD, DR. GABRIELA LINDEMANN-V. TRZEBIATOWSKI

Beteiligte Mitarbeiter: DIPL.-MATH. MANFRED HILD, M. SC. DAGMAR MONETT DÍAZ, DIPL.-INF. (FH) ALEXANDER OSHERENKO, DIPL.-INF. KAY SCHRÖTER, DIPL.-KFM. DIPL.-INF. DIEMO URBIG, NORA HANS, CLAUDIA HÄRTEL, ROBIN MALITZ, CHRISTIAN WIECH

Zusammenarbeit: PROF. DR. WERNER RAMMERT, Technische Universität Berlin, Institut für Soziologie, Fachgebiet Techniksoziologie; PROF. GERT BAUMANN, Klinik für Kardiologie, Angiologie und Pulmologie, Charité – Universitätsmedizin Berlin

Forschungsförderung: DFG, im Rahmen des Schwerpunktprogramms „Sozionik“

Informationen: <http://www.ki.informatik.hu-berlin.de/inka>

Ziel des Projektes ist die Untersuchung der konzeptionellen und technischen Grundlagen für offene agenten-basierte Systeme, die mit den Problemen der Inkohärenz und Heterogenität in komplexen Organisationen erfolgreich umgehen können. Solche Systeme sind als sozionische Systeme unter dem Gesichtspunkt der hybriden Organisation der Aktivitäten menschlicher Akteure und künstlicher Agenten zu betrachten. Als eine adäquate Problemstellung wurden Verhandlungen über den Tausch von Arbeitsschichten im Krankenhaus gewählt.

Die Zuschreibung von Sozialtypen, in Form von Selbst- und Fremdbildern, ermöglicht typisierte Erwartungen hinsichtlich des Verhaltens anderer Handlungsträger in Verhandlungen. Diese Konzepte wurden in einem Multi-Agenten-System umgesetzt. Zur Strukturierung und Modularisierung der durch die Agenten zu treffenden Entscheidungen während

einer Verhandlung wurde das abstrakte C-IPS Framework entwickelt und für die Architektur und Implementation der INKA-Agenten angewendet

Die Umsetzung erfolgte unter Nutzung der Java-basierten Multi-Agenten-Plattform JADE. Zurzeit wird das System genutzt, um Simulation von Verhandlungen zum Schichttausch zu analysieren. Die Agenten, deren Architektur bereits eine „Graduelle Agency“ vorsieht, werden um Schnittstellen für menschliche Handlungsträger erweitert. Diese können dann einzelne oder alle Entscheidungen bezüglich der Verhandlung übernehmen. In Interaktivitätsexperimenten kann so die Funktionalität des Gesamtsystems in einem hybriden Szenario untersucht werden.



Das INKA-Team auf dem Jahrestreffen des DFG Schwerpunktprogrammes Sozionik in Kloster Seeon.

Projekt: Simulation von Einstellungsveränderungen in Gesellschaften (Attitude and Opinion Dynamics)

Ansprechpartner: DIPL.-KFM. DIPL.-INF. DIEMO URBIG

Zusammenarbeit: PROF. DR. CHRISTIAN SCHADE, Humboldt Universität zu Berlin, Institut für Entrepreneurship/Innovationsmanagement

Einstellungen bilden im Marketing eine Grundlage zur Prognose von Verhalten. Nach einer Analyse der Muster von Einstellungsentwicklungen bei beschränkten Verbalisierungsfähigkeiten wurde im Rahmen einer Diplomarbeit erforscht, wie interdependente Einstellungen im Rahmen eines Simulationsansatzes modelliert werden können. Dabei wird auf eine explizite Modellierung der Kommunikation und Repräsentation von Wissen um den Zusammenhang zwischen verschiedenen Objekten fokussiert. Dabei wurden Konzepte der Theorie des Konsumentenverhaltens in auf Multi-Agenten-Systemen basierenden Simulationen umgesetzt. Dieses Kooperationsprojekt wird im Rahmen eines Promotionsvorhabens fortgesetzt.

Projekt: Architekturprinzipien und Lernen auf der Basis mentaler Modelle

Ansprechpartner: PROF. DR. HANS-DIETER BURKHARD, DIPL.-PHYS. JAN HOFFMANN

Forschungsförderung: DFG, im Rahmen des Schwerpunktprogramms „Verteilte Systeme in dynamischen Umgebungen (RoboCup)“

Ziel des Projektes ist die Untersuchung und Benutzung Agenten-Orientierter und Fallbasierter Techniken zur Programmierung kooperativen Verhaltens und zum Online-

Lernen in dynamischen Umgebungen. Eine hierarchische Agentenarchitektur wurde entwickelt, die auf allen Ebenen realzeitfähig ist. Mit ihrer Hilfe werden im weiteren Verlauf des Projektes langfristige kooperative Pläne implementiert und Standardsituationen/-lösungen als Fälle gespeichert werden. Anwendungsszenario ist der Roboterfußball (RoboCup).

1 German Team / AIBO Team Humboldt (RoboCup „SONY-Legged-Robots“)

Ansprechpartner: PROF. DR. HANS-DIETER BURKHARD, DIPL.-PHYS. JAN HOFFMANN

Beteiligte Mitarbeiter: UWE DÜFFERT, MATTHIAS JÜNGEL, MARTIN LÖTZSCH,
DANIEL GÖHRING

Zusammenarbeit: SONY; GI-AK RoboCup; Arbeitsgruppen des „German Team“ an den Universitäten Bremen, Darmstadt, Dortmund; TU München; DFG-Schwerpunktprogramm „RoboCup“; PROF. DR. TIM LÜTH, Klinik für Mund-, Kiefer- und Gesichtschirurgie - Klinische Navigation und Robotik, Charité – Universitätsmedizin Berlin

Informationen: <http://www.robocup.de/germanteam/>
<http://www.aiboteamhumboldt.com/>
<http://www.ki.informatik.hu-berlin.de/XABSL/>

In der „Sony Four-Legged Robot League“ des RoboCup treten Teams völlig autonom agierender, vierbeiniger Roboter gegeneinander an. Dabei spielen pro Team vier Roboter. Als Hardware kommt ein handelsüblicher Sony Aibo zum Einsatz. Um die Arbeit insgesamt zu stärken und weiteren deutschen Universitäten die Beteiligung zu ermöglichen, haben wir uns 2001 mit den oben genannten Universitäten für den internationalen RoboCup-Wettbewerb zur deutschen Nationalmannschaft „German Team“ zusammengeschlossen.

Die Roboter nehmen ihre Umgebung mit Hilfe verschiedener Sensoren (Kamera usw.) wahr. Ein Forschungsschwerpunkt ist daher der Wahrnehmungsprozess des Roboters, also die Datenaufnahme, ihre Analyse und Auswertung und eine darauf aufbauende Steuerung der Aufmerksamkeit. Darüber hinaus wurde eine Architektur entwickelt, die die verteilte Softwareentwicklung im Rahmen des „German Team“ ermöglicht. Grundprinzipien dieser Architektur sollen auch für unser Team in der Simulationsliga angewendet werden. Die von uns entwickelte Sprache zur Beschreibung des Verhaltens autonomer Agenten XABSL („Extensible Agent Behavior Specification Language“) kam beim RoboCup erfolgreich zur Anwendung. In dieser Sprache wird sowohl reaktives als auch deliberatives Verhalten des Agenten über Zustandsmaschinen spezifiziert und in XML notiert. Weitere Schwerpunkte der Projektarbeit sind verteiltes Handeln, Kooperation, die Verfolgung langfristiger Pläne und die Anwendung fallbasierter Methoden.

Bei der diesjährigen GermanOpen in Paderborn wurde unsere Mannschaft deutscher Vize-Meister. Bei den Weltmeisterschaften in Padua, Italien, erreichte das GermanTeam das Viertelfinale. Neben dem Fußball-Wettbewerb gibt es beim RoboCup die „Technical Challenge“, in der jedes Jahr für die Weiterentwicklung der Roboter wichtige Aspekte in Form von klar umrissenen Aufgaben getestet werden. Diesen Wettbewerb hat das GermanTeam souverän gewonnen.



*German Open 2003 in Paderborn:
Unser Team der Roboter-Liga (l); ein Spiel unserer Roboter (r)*

Wie auch in den vergangenen Jahren wurden die Roboter bei diversen Wissenschaftsveranstaltungen, wie zum Beispiel der „Langen Nacht der Wissenschaft“, präsentiert. Unsere Software-Entwicklungsumgebung für den Roboter findet sowohl in einem Projektkurs („Kognitive Robotik“) als auch an einem Gymnasium Anwendung.

2 AT Humboldt (RoboCup Simulationsliga)

Ansprechpartner: PROF. DR. HANS-DIETER BURKHARD, RALF BERGER

Beteiligte Mitarbeiter: RALF BERGER, BIRGER BRUNSWIEK, MICHAEL GOLLIN, CHRISTINA BELL, DIPL.-INF. JOSCHA BACH

Zusammenarbeit: GI-AK RoboCup, DFG-Schwerpunktprogramm „RoboCup“

Informationen: <http://www.ki.informatik.hu-berlin.de/RoboCup/index.shtml>

In der Simulationsliga treten je 11 autonome Programme (Agenten), sowie ein Coach-Programm, aus zwei Teams auf einem virtuellen Spielfeld gegeneinander an. Die halbphysikalische Raum-/Zeitsimulation des Spielgeschehens erfolgt durch den „Soccerserver“. Er nimmt die Kommandos der einzelnen Agenten entgegen und sendet ihnen Informationen über ihre lokale Sicht. Anfang dieses Jahr wurde die Entwicklung unserer neuen Agentenplattform abgeschlossen. Diese beinhaltet vor allem die in den letzten Jahren konzeptuell entwickelte „Doppelpassarchitektur“, erste Strukturen zur Nutzung von Methoden des Fallbasierten Schließens für langfristige Entscheidungsfindung, sowie einige softwaretechnisch besonders interessante Ansätze.

Ziel der Entwicklung dieser neuen Agentenarchitektur ist die Möglichkeit zur Planung längerfristiger kooperativer Absichten und die parallele, unabhängige Ausführung und Generierung von dazu konformen Aktionen in Echtzeit. Dies wird erreicht durch die Kombination einer zentralen hierarchischen Optionsstruktur und von zwei unabhängig voneinander darauf arbeitenden Planungs- und Ausführungskomponenten. Die parallel dazu entwickelte Beschreibungssprache hat inzwischen ebenfalls eine stabile Formulierung erreicht.

Die neuen Konzepte wurden während der German Open in Paderborn und während der Weltmeisterschaft in Padua eingesetzt. Sie sind derzeit Gegenstand weiterer Untersuchungen.



Unser Roboter- und Simulationsteam während der RoboCup Weltmeisterschaft 2003 in Padua, Italien.

Projekt: Theorie und Techniken des fallbasierten Schließens

Ansprechpartner: PROF. DR. HANS-DIETER BURKHARD

Beteiligte Mitarbeiter: DIPL.-INF. MIRJAM MINOR, DIPL.-INF. KAY SCHRÖTER, CARSTEN HILBIG, ALEXANDER MOCZKO, SANDRO KÖPPEN

Eine Kernaufgabe des fallbasierten Schließens ist das Wiederfinden von gespeichertem Wissen („Erinnern“), das sogenannte Retrieval. Hierbei spielen sowohl Fragen des effizienten als auch des flexiblen Zugriffs auf die Datenbestände eine gravierende Rolle. Am Lehrstuhl für KI werden sogenannte „*Case Retrieval Nets*“ entwickelt, die eine spezielle Art von Assoziativspeichern darstellen. Mit Hilfe dieser Techniken kann ein in hohem Maße effizientes Retrieval realisiert werden. Techniken zur Akquisition, Kommunikation und Wartung von Erfahrungswissen bilden einen weiteren Schwerpunkt der aktuellen Forschungsarbeiten. Die Ergebnisse werden in den anwendungsorientierten Projekten intensiv genutzt.

1 Fallbasiertes Schließen zur geologischen 3D-Strukturanalyse

Geologische Strukturmodelle werden derzeit meist von ausgebildeten Geowissenschaftlern mittels Auswertung und Interpretation einer oft umfangreichen Datengrundlage entwickelt. In diesem Teilprojekt soll ein fallbasiertes System entstehen, das aus Daten von geologischen Aufschlüssen (z.B. Bohrprofilen, geologischen Karten) ein 3D-Modell über den Aufbau eines Untergrundes erstellt. Dabei soll nicht nur zwischen den einzelnen Aufschlüssen interpoliert werden, sondern bereits bekannte Strukturmodelle mit ähnlichen Gegebenheiten aus der Fallbasis hinzugezogen werden. Bestehende Verfahren und Systeme zur automatischen Strukturanalyse sollen damit verglichen und bewertet werden.

2 Vages Matchen in Datenbanken

Methoden des Textuellen Fallbasierten Schließens wurden so in ein Datenbankmanagementsystem integriert, dass damit ein vages Matching auf domänenspezifischen Daten implementiert wird. Die Daten können sowohl Texte enthalten als auch in Attribut-Werte-Form vorliegen. Als Ergebnis dieses Teilprojekts wurde mit dem Datenbankmanagementsystem Firebird ein lauffähiger Prototyp namens CaseMaker implementiert, der Case Retrieval Netze für beliebige kompositorische Ähnlichkeitsfunktionen ausschließlich mit Standardtechniken relationaler Datenbanken wie *stored procedures* und Funktions-

bibliotheken realisiert. Alternativ dazu wurde untersucht, wie dieselbe Funktionalität rein mit SQL-Befehlen emuliert werden kann.

3 Fallbasiertes Schließen zur geologischen 3D-Strukturanalyse

Geologische Strukturmodelle werden derzeit meist von ausgebildeten Geowissenschaftlern mittels Auswertung und Interpretation einer oft umfangreichen Datengrundlage entwickelt. In diesem Teilprojekt soll ein fallbasiertes System entstehen, das aus Daten von geologischen Aufschlüssen (z.B. Bohrprofilen, geologischen Karten) ein 3D-Modell über den Aufbau eines Untergrundes erstellt. Dabei soll nicht nur zwischen den einzelnen Aufschlüssen interpoliert werden, sondern bereits bekannte Strukturmodelle mit ähnlichen Gegebenheiten aus der Fallbasis hinzugezogen werden. Bestehende Verfahren und Systeme zur automatischen Strukturanalyse sollen damit verglichen und bewertet werden.

4 Vages Matchen in Datenbanken

Methoden des Textuellen Fallbasierten Schließens wurden so in ein Datenbankmanagementsystem integriert, dass damit ein vages Matching auf domänenspezifischen Daten implementiert wird. Die Daten können sowohl Texte enthalten als auch in Attribut-Werte-Form vorliegen. Als Ergebnis dieses Teilprojekts wurde mit dem Datenbankmanagementsystem Firebird ein lauffähiger Prototyp namens CaseMaker implementiert, der Case Retrieval Netze für beliebige kompositorische Ähnlichkeitsfunktionen ausschließlich mit Standardtechniken relationaler Datenbanken wie *stored procedures* und Funktionsbibliotheken realisiert. Alternativ dazu wurde untersucht, wie dieselbe Funktionalität rein mit SQL-Befehlen emuliert werden kann.

Projekt: Fallbasiertes Schließen für intelligente Dokumenten-Recherche

Ansprechpartner: PROF. DR. HANS-DIETER BURKHARD

Beteiligte Mitarbeiter: DIPL.-INF. MIRJAM MINOR, GORDON KRAMER, PETER HAMMELS, JOSE CARLOS DEL PRADO

Zusammenarbeit: empolis: DR. MARIO LENZ, DIPL.-INF. WOLF-DIETRICH MATERNA.

In vielen Bereichen werden Wissen und Erfahrung in textueller Form dokumentiert (z.B. Arztbrief, Dokumentation von Geschäftsvorgängen, FAQ-Listen in Newsgroups etc.). Bisherige Methoden zur Recherche in diesen Dokumenten-Sammlungen beschränkten sich meist auf traditionelle Verfahren des Information Retrieval, bei denen jedoch in erster Linie statistische Verfahren (term frequency) zum Einsatz kommen. Demgegenüber zielt das Fallbasierte Schließen zwar auf die direkte Nutzung des Wissens in Fallbasen ab, behandelt traditionell jedoch strukturierte Dokumente und weniger textuelle Beschreibungen. Im Rahmen der Forschungsarbeiten an der LFE wurden Konzepte entwickelt, um fallbasierte Methoden auch auf Textdokumente anwenden zu können. Im Unterschied zu Methoden des Information Retrieval können damit Systeme erstellt werden, die das Wissen einer Anwendungsdomäne ausnutzen, um sinnvolle Beziehungen zwischen den Objekten der Domäne herzustellen. Das von empolis entwickelte Werkzeug CBR-Answers wird in zahlreichen Anwendungsprojekten benutzt.

1 Multilinguales Fallbasiertes Schließen

In diesem Teilprojekt wurde untersucht, wie sich fallbasierte Methoden für multilinguale Fallbasen anwenden lassen. Für eine Anfrage kann eine beliebige Sprache gewählt werden, die dem System bekannt ist. Tests mit einer Trainingsfallbasis (Veröffentlichungen der EU in mehreren Übersetzungen) haben gezeigt: In technischen Anwendungsgebieten reicht eine Eins-zu-Eins-Übersetzung von Begriffen aus, während in allgemeinsprachlichen Domänen

differenziertere multilinguale Modellierungen mit Begriffsspektren nötig sind. Um diese Begriffsspektren zu modellieren, werden halbautomatische Techniken untersucht, die externe Wissensressourcen wie elektronische Lexika einbeziehen. Am Beispiel verschieden-sprachiger Fußball-Texte aus dem Web wurde ein Prototyp entwickelt.

2 *OntoDigger*

OntoDigger ist ein System zur Textanalyse, um automatische oder halbautomatische Unterstützung für das Erstellen von Domänenontologien zu bekommen, wie sie auch im Fallbasierten Schließen eine Rolle spielen. Die untersuchten Algorithmen extrahieren wichtige Konzepte und Beziehungen zwischen den Konzepten aus einem deutschen Textkorpus. Am Beispiel einer Internet-Textsammlung aus dem Kulturjournalismus werden die Ergebnisse ausgewertet.

3 *Lernen für ThemeSearch*

In diesem Teilprojekt werden Konzeptknoten von Wendy Lehnert für die Klassifikation von Newsgroup-Beiträgen eingesetzt. Die aktuellen Arbeiten beschäftigen sich mit dem Verstehen natürlichsprachlicher Texte mit Hilfe einer flachen Grammatikanalyse, um Lernverfahren für das automatische Erzeugen von Konzeptknoten aus Textkorpora zu untersuchen.

Projekt: Intelligente Techniken im Knowledge Management

Ansprechpartner: PROF. DR. HANS-DIETER BURKHARD, DIPL.-INF. MIRJAM MINOR

Beteiligte Mitarbeiter: M. SC. DAGMAR MONETT DÍAZ, BENJAMIN ALTMAYER, CHRISTINA BELL, FABIAN MÜLLER, MIKE WERNICKE

Zusammenarbeit: Zentrum für Biomaterialien der Universität Havanna.

Forschungsförderung: Rosa-Luxemburg-Stiftung

In diesem Projekt werden Fallbasierte und Agenten-Orientierte Methoden für das Wissensmanagement eingesetzt. Knowledge Management ist kein einheitliches Konzept, sondern umfasst eine Vielzahl von Techniken, um mit der Ressource Wissen umzugehen. Neben der Akquisition von Wissen spielen Organisation, Wartung, Wiederverwendung und Transfer von Wissen eine wesentliche Rolle. Das Web eröffnet neue Kommunikationswege und bietet riesige Informationsbestände, die mit Hilfe intelligenter Techniken als Wissen nutzbar gemacht werden können. Der Schwerpunkt unserer Arbeiten im Wissensmanagement liegt darauf, Wissen zu finden, geeignet zu strukturieren und es in der Form von Fällen oder Diensten anwendbar zu machen.

1 *Agenten-Orientierte Techniken auf der Basis von Java*

In diesem Teilprojekt wird die Programmiersprache Java mit Konzepten und Methoden ergänzt, um eine effiziente agenten-orientierte Programmierung zu ermöglichen. Als Beispielanwendung wird ein Assistenzprogramm für die Benutzung und Installation von Software entwickelt. Der Assistenz-Agent verwaltet eine Menge derartiger Dienste, zu denen jeweils auch eine detaillierte Beschreibung gehört. Bei Anforderung wählt der Agent die passenden Dienste aus. Findet er zu einer Anfrage keinen passenden Dienst, so wendet er sich an andere Agenten im Netz, die ihm dann gegebenenfalls ihre Dienste anbieten. Der Agent kann auch gefundene Dienste in seinen Funktionsumfang aufnehmen. Die zwei wesentlichen Aspekte der Anwendung sind somit das Wiederfinden und Wiederbenutzen von Informationen (Diensten) und die Verteilung der Agenten zum Austausch von Diensten. Die aktuelle Arbeit beschäftigt sich mit einer Verbesserung der Benutzerfreundlichkeit durch eine natürlichsprachliche Interaktion mit den Benutzern auf der Grundlage fallbasierter

Methoden, für die Dienstbeschreibungen als Fälle betrachtet werden und verteilte Wissensbestände getreu dem Multiagentenprinzip integriert werden können.

2 *ExperienceBook*

Sowohl bei der Systemadministration als auch im alltäglichen Umgang mit Rechnernetzen wird Erfahrungswissen für die Behandlung von Störungen und Problemen eingesetzt. Mit Hilfe von Daten konkreter Störfälle wurde das fallbasierte System ExperienceBook entwickelt. Systemadministratoren und andere Benutzer können via Internet auf bereits gelöste Problemfälle zugreifen oder die Fallbasis durch neue Fälle erweitern. Die aktuelle Arbeit untersucht die Möglichkeiten zur besseren Akquisition, Wartung und Pflege von Wissensinhalten sowie das Langzeitverhalten eines solchen Systems im praktischen Einsatz.

3 *ExperienceBook für Erstsemestler*

In diesem Teilprojekt wurde das fallbasierte System ExperienceBook erweitert und Studierenden der Informatik im ersten Semester zur Verfügung gestellt, um einen Erfahrungsaustausch über technische Fragen, Fragen zur Vorlesung Praktische Informatik sowie andere studentische Belange zu ermöglichen. Ein klassisches Diskussionsforum für offene Fragen ergänzt die fallbasierte Technik, die eine intelligente Suche über bereits beantwortete Fragen der Studierenden ermöglicht. Aktuelle Arbeiten beschäftigen sich mit dem Wissenstransfer vom Forum zum ExperienceBook sowie Patenschaftsmodellen für die Pflege der einzelnen Wissensbereiche. Eine Kooperation mit dem Goya-Team zum Einsatz des ExperienceBook als kontextsensitives Hilfesystem für Goya ist geplant.

4 *Agenten-basierte Konfiguration von metaheuristischen Algorithmen*

Als Fortsetzung des Themas „Werkzeuge für die mathematische Modellierung chemischer Prozesse mittels intelligenter Agenten und evolutionärer Techniken“ ist die *Agenten-basierte Konfiguration von metaheuristischen Algorithmen* von großer Bedeutung. Dieses Teilprojekt betrachtet die aus der Literatur bekannten existierenden metaheuristischen Ansätzen unter spezieller Beachtung evolutionärer Algorithmen. Dabei spielt die Beziehungen zwischen Evolutionärer Computation und Künstliche Intelligenz eine wichtige Rolle, wenn verschiedene Mechanismen beider Bereiche enthalten sind und miteinander interagieren. Der Schwerpunkt der Arbeit liegt sowohl auf der Theorie zur Konfiguration von Metaheuristiken, als auch auf ihrer Agenten-basierten Automatisierung. Bestandteile sind die Konzeption, Struktur, Implementierung und Realisation des Multi-Agenten-System +CARPS (Multiagent System for Configuring Algorithms in Real Problem Solving). Verschiedene Experimente sind geplant worden, um die praktische Umsetzung des Agenten-basierten Ansatz zu ermöglichen und die Konfiguration von metaheuristischen Algorithmen zu unterstützen und zu automatisieren.

5 *Grundlagen und Anwendungen des Experience Management*

Experience Management ist eine Spezialform von Wissensmanagement, die sich mit Erfahrungswissen beschäftigt. Das Anwenden vorhandener Erfahrung trägt dazu bei, Probleme erfolgreich zu lösen. Ziel von Experience Management ist es, personengebundenes Erfahrungswissen in eine personenunabhängige Form zu bringen und überall dort zur Verfügung zu stellen, wo es für die Lösung neuer Probleme nutzbringend eingesetzt werden kann. Dazu werden moderne Methoden der Künstlichen Intelligenz benötigt, wie zum Beispiel Agenten, Fallbasiertes Schließen oder Ontologien. Sie unterstützen hauptsächlich die folgenden Prozesse: Extrahieren, Editieren, Speichern, Wiederfinden, Teilen, Wiederverwenden, Warten und Evaluieren von Erfahrungswissen.

6 *Weblexikon*

Im Rahmen der Forschungsarbeiten an der LFE wurden klassische Modelle aus dem betriebswirtschaftlichen Wissensmanagement in der Literatur studiert (z.B. Nonaka und Takeuchi, Probst, Lehner, Morey) und auf ihre Übertragbarkeit auf Experience-Management-Systeme hin untersucht. Die Ergebnisse dieser theoretischen Untersuchungen werden am Beispiel eines intelligenten Weblexikons umgesetzt, das verschiedene KI-Methoden für das Experience Management in kooperierenden Forschungsgruppen miteinander kombiniert. Ungesichertes, aktuelles Wissen aus e-mails werden dafür mit gesichertem Wissen aus Manuals und wissenschaftlichen Arbeiten automatisch verlinkt und in einem fallbasierten System zur Verfügung gestellt.

Projekt: Verteilte Sensor-/Aktor-Infrastrukturen für sozionische Konzepte

Ansprechpartner: PROF. DR. HANS-DIETER BURKHARD, DR. GABRIELA LINDEMANN-V. TRZEBIATOWSKI

Beteiligte Mitarbeiter: CLAUDIA HÄRTEL, ROBIN MALITZ

Förderung: Siemens AG, München

Im Rahmen einer Studie wurden aktuelle Trends im Bereich von Sensor-/Aktor-Infrastrukturen vorgestellt. Es zeigte sich, dass bisher vor allem die Vernetzung von Sensoren und informationsliefernden Medien im Mittelpunkt stand. Dazu gehörten insbesondere Fragen der Übertragungsmedien mit unterschiedlicher Leistungsfähigkeit und der Kommunikation zwischen Daten- und Verarbeitungsquellen. Zentrale Instanzen wurden genutzt zur Kommunikation, zur Datenauswertung und zur Übermittlung der Resultate. Vielfältig vernetzte Strukturen können die Nutzung zentraler Instanzen oft ersetzen. Möglich sind auch hybride Systeme, in denen menschliche Akteure und künstliche Agenten interagieren. Insgesamt sind die Szenarien als hintereinander geschaltete Prozesse zu verstehen, die in großer Zahl parallel ablaufen. Vereinfacht lässt sich dies durch sechs Teilprozesse beschreiben: Datenerfassung und lokale Aufbereitung durch Sensoren, Kommunikation zu lokalen oder zentralen Instanzen, Datenfusion, Auswertung, Verteilung, Auslösen von Aktionen.

Projekt: Verhaltensmodellierung

Ansprechpartner: DIPL.-INF. JAN WENDLER

Kooperation: Graduiertenkolleg „*Verteilte Informationssysteme*“

Forschungsförderung: DFG

Informationen: <http://www.wiwi.hu-berlin.de/gkvi/>

In Multi-Agenten-Systemen (MAS) kooperieren und konkurrieren Agenten um ihre jeweiligen Ziele zu erreichen. Für optimierte Agenten-Interaktionen sind Kenntnisse über die aktuellen und zukünftigen Handlungen anderer Agenten (Interaktionspartner, IP) hilfreich. Bei der Ermittlung und Nutzung solcher Kenntnisse kommt dem automatischen Erkennen und Verstehen sowie der Vorhersage von Verhalten der IP auf Basis von Beobachtungen eine besondere Bedeutung zu.

Dieses Projekt beschäftigt sich mit der automatischen Bestimmung und Vorhersage von Verhalten der IP durch einen Modellierenden Agenten (MA). Der MA generiert fallbasierte, adaptive Verhaltens-Modelle seiner IP und verwendet diese zur Vorhersage ihrer Verhalten. Als Anwendungsszenario wird mit dem virtuellen Fußballspiel des RoboCup ein komplexes und populäres MAS betrachtet.

Projekt: Analyse medizinischer Daten mit Methoden des Fallbasierten Schließens**Ansprechpartner:** DIPL.-INF. KAY SCHRÖTER**Beteiligte Mitarbeiter:** HELMUT MYRITZ, DANILO SCHMIDT**Zusammenarbeit:** Medizinische Klinik und Poliklinik mit Schwerpunkt Nephrologie, Charité – Universitätsmedizin Berlin: PROF. DR. H.-H. NEUMAYER, DR. LUTZ FRITSCHKE

Der durch den Einsatz einer Elektronischen Patientenakte gewonnene Datenbestand soll für die Entscheidungsunterstützung im Bereich der Nephrologie genutzt werden. Dabei werden Techniken des Fallbasierten Schließens angewendet. Basierend auf möglichst vielen relevanten Daten sollen allgemeine Aussagen zur Ähnlichkeit zweier Fälle gemacht werden. Dabei ist zunächst zu klären, wie sich ein Fall in der nephrologischen Domäne darstellt. Die Qualität des Systems soll sowohl anhand verschiedener medizinischer Fragestellungen automatisch analysiert werden, als auch direkt von den Medizinern beurteilt werden. Vorbereitend wurden Methoden der Entscheidungsfindung in der Medizin untersucht.

Projekt: Agententechnologie in der Medizin**Ansprechpartner:** DIPL.-INF. HELMUT MYRITZ**Kooperation:** SIEMENS Medical Solutions, Erlangen

Dieses Projekt beschäftigt sich mit der Anwendung der Agententechnologie im medizinischen Kontext. An der TU München (Frauenklinikum) soll ein Krankenhaus-Informationssystem entstehen, in dem ein Multi-Agenten-System (MAS) sinnvoll Anwendung findet und die Ärzte bei Pflege und Behandlung der Patienten sowie in der medizinischen Forschung unterstützt. Diagnoseüberprüfung anhand vorher festgelegter Leitlinien und stete Behandlungskontrolle sind hierbei zentrale Aufgabenfelder der Agenten.

Projekt: Klinisches Multi-Agenten-System**Ansprechpartner:** M. SC. BASMAH EL-HADDAD**Kooperation:** Institut für Nationale Planung, Kairo, Ägypten**Forschungsförderung:** DAAD

Ziel des Projektes ist es, ein klinisches Multi-Agenten-System mit unterschiedlichen Agenten zu implementieren, das bestimmten Sicherheitsrichtlinien und Prozessflüssen der Patientenversorgung folgt. Dies könnte in den Rettungsstellen-, Intensiv-Pflege- oder Operations-Räumen genutzt werden. Um die Prozessflüsse entsprechend den sich ändernden Anforderungen in einer optimalen Art anzupassen, sind zur Implementierung dieses Multi-Agenten-Systems sorgfältig auf das Krankenhaus abgestimmte Methoden notwendig.

Gastwissenschaftler**Prof. Dr. Mirjana Ivanovic, Universität Novi Sad, Serbien-Montenegro,
(10-11/2003)****Förderung:** DAAD

Im Rahmen der langjährigen Zusammenarbeit der Lehrstühle für „Künstliche Intelligenz“ und „Softwaretechnik“ mit der Universität Novi Sad besuchte uns Frau Prof. Mirjana Ivanovic. Sie forscht und lehrt unter anderem auf dem Gebiet der Multi-Agenten-Systeme.

B. Sc. Vladimir Kurbalija, Universität Novi-Sad, Serbien-Montenegro (bis 1/2003)**Förderung: DAAD**

Herr Kurbalija hat während seines Aufenthaltes an unserem Lehrstuhl ein Teilthema des Fallbasierten Schließens bearbeitet. In vielen entscheidungsunterstützenden Systemen spielen Zeitverläufe von Werten eine Rolle. Für Voraussagen möglicher Zeitverläufe in der Zukunft wurden bekannte Zeitdiagramme genutzt. Es wurde ein Verfahren implementiert, welches es erlaubt mittels Spline-Interpolation Aussagen zur Ähnlichkeit zweier Zeitkurven zu treffen. Das darauf beruhende Ähnlichkeitsmaß ist in vielen Anwendungsdomänen praktikabel.

M. Sc. Rafal Latkowski, Universität Warschau, Polen (bis 9/2003)

Herr Latkowski beschäftigt sich mit Data Mining und Knowledge Discovery in Databases. In seiner Forschung konzentriert er sich auf das Problem des Schließens aufgrund von Daten mit fehlenden Attributwerten. Im Jahr 2003 hat er einen Komplexitätsexplosion für Entscheidungsbaumealgorithmen auf Daten mit vielen fehlenden Attributwerten bewiesen. Dies kommt z.B. in der Domäne des Roboterfußballspiel vor. Herr Latkowski forscht über die Verbindungen zwischen Fallbasiertem Schließen, Fuzzy Sets und Rough Sets im Bereich der Begriffe: Ähnlichkeit, Distanz, Akzeptanz und fehlende Information. Er interessiert sich auch für Multi-Agenten-Systeme und Anwendungen der Verteilten Künstlichen Intelligenz in Data Mining und Knowledge Discovery.

Dipl.-Inf. Mohammed Mayyas, externer Promotionsstudent

Herr Mayyas beschäftigt sich mit „Experience Management mit CBR (Case-Based Reasoning) Methoden“. Er arbeitet auf dem Gebiet Fallbasiertes Schließen, Informationssysteme und Wissensmanagement. Ziel seiner Arbeit ist es, personengebundenes Erfahrungswissen in eine personenunabhängige Form zu bringen und dort zur Verfügung zu stellen, wo es für die Lösung neuer Probleme nutzbringend eingesetzt werden kann. Zusätzlich erforscht er den möglichen Einsatz von CBR Methoden für multilinguales Experience Management.

M. Sc. Jesús Emeterio Navarro Barrientos, Centro de Investigacion en Matematicas, Guanajuato, Guanajuato, Mexiko**Förderung: DAAD**

Herr Navarro forscht im Bereich der Multi-Agenten-Systeme und der Agent-Based Computational Economics (ACE). Speziell entwickelte er Simulations-Programme um intelligente Investitionen und Koalition-Netzwerke zwischen Agenten zu simulieren. Ziel seiner Forschung ist die Aufstellung und Entwicklung von Interaktion-Netzwerke zu verstehen und die globale Regelmäßigkeiten, die vom Bottom-up entstehen könnten, bei den mehrmaligen Interaktionen zwischen autonomen Agenten zudemonstrieren.

M. Sc. Farley Nobre, Universität Birmingham, Großbritannien (bis 8/2003)**Förderung:** CNPq Stiftung Brasilien

Im Rahmen seines Ph.D. Programms beschäftigt sich Herr Nobre mit der einer vereinheitlichten Theorie von organisationalen Systemen und mit deren Anwendung in Betrieben (A Unified Theory on Organisational Systems and its Application to Manufacturing). Hier bestehen insbesondere Anknüpfungspunkte zum Sozionik-Projekt. Herr Nobre interessiert sich für die Forschungsgebiete Soziale und Organisationale Systeme, Soft Computing, Allgemeine System Theorie und Kognitionswissenschaften.

M. Sc. Victor Emanuel Uc Cetina, Universität Merida, Yukatan, Mexico (8-9/2003)

Herr Uc Cetina forscht auf dem Gebiet der Multi-Agenten-Systeme. Speziell interessiert er sich für verteilte Lernalgorithmen in heterogenen Umgebungen, die Kombination von Reinforcement-Lernmethoden mit Klassifizierungsverfahren und spieltheoretischen Ansätzen. Anwendungen dafür finden sich in hoch-dynamischen Umgebungen, wie z.B. beim RoboCup. Mit Hilfe der zu entwickelnden Verfahren sollen die Koordinationsfähigkeiten einzelner Agenten bzw. Roboter und das globale Verhalten des Gesamtsystems verbessert werden.

Veröffentlichungen

J. BACH: *Emotionale Virtuelle Agenten auf der Basis der Dörnerschen Psi-Theorie*. In H.-D. Burkhard, T. Uthmann, G. Lindemann (Eds.): Beiträge des ASIM-Workshops „Modellierung und Simulation menschlichen Verhaltens“. Informatik-Berichte, ISSN 0863-095X, Nr. 163, S. 1-10, Humboldt-Universität zu Berlin, 2003.

J. BACH: *The MicroPsi Agent Architecture*. In Proceedings of Fifth International Conference on Cognitive Modeling (ICCM 2003). pp. 15-20 University Bamberg, 2003.

J. BACH: *Die Grenzen des Verstehens - Voraussetzungen der Kommunikation zwischen Intelligenten Systemen*. In M. Schetsche (Ed.): Der maximal Fremde. Freiburg, 2004 (in Vorbereitung).

J. BACH AND THE AEP GROUP: *Connecting MicroPsi Agents to Virtual and Physical Environments – Applications of the AEP toolkit*. In P. Dittrich, J. Kim (Eds.): Proceedings of the European Conference for Artificial Life (ECAL 2003), Part: Workshops and Tutorials. pp. 128-132, University of Dortmund, 2003.

J. BACH, R. VUINE: *Designing Agents with MicroPsi Node Nets*. In A. Günter, R. Kruse, B. Neumann (Eds.): KI 2003: Advances in Artificial Intelligence – Proceedings of the 26th Annual German Conference on Artificial Intelligence. LNAI 2821, pp. 164-178, Springer-Verlag Berlin Heidelberg, 2003.

J. BACH, R. VUINE: *The AEP Toolkit for Agent Design and Simulation*. In M. Schillo et al. (Eds.): Multiagent System Technologies, Proceedings of the First German Conference (MATES 2003). LNAI 2831, pp. 38-49, Springer-Verlag Berlin Heidelberg, 2003.

J. BACH, R. VUINE: *The MicroPsi Architecture for Cognitive Agents*. Poster, In Proceedings of EuroCogSci 2003. p. 370, University Osnabrück, 2003.

R. BERGER, M. GOLLIN: *Auf dem Weg zu menschlichem Verhalten – Simulation einer virtuellen Fußballumgebung*. In H.-D. Burkhard, T. Uthmann, G. Lindemann (Eds.):

- Beiträge des ASIM-Workshops „Modellierung und Simulation menschlichen Verhaltens“. Informatik-Berichte, ISSN 0863-095X , Nr. 163, S. 20-25, Humboldt-Universität zu Berlin, 2003.
- H.-D. BURKHARD: *Autonome Agenten und Roboter*. In T. Christaller, J. Wehner (Hrsg.): *Autonome Maschinen*. Westdeutscher Verlag, 2003.
- H.-D. BURKHARD, M. ASADA, A. BONARINI, A. JACOFF, D. NARDI, M. RIEDMILLER, C. SAMMUT, E. SKLAR, M. VELOSO: *RoboCup: Yesterday, Today, and Tomorrow*. In: 7th International Workshop on RoboCup 2003 (Robot World Cup Soccer Games and Conferences). LNAI, Springer-Verlag, 2004 (to appear).
- H.-D. BURKHARD, G. LINDEMANN, K. SCHRÖTER, D. URBIG: *Die Evolution von Regelpraxis in offenen Netzwerken*. Abstract für die Tagung „Agentengesellschaften außer Kontrolle? Sozialtheoretische Mechanismen für den Bau von künstlichen Sozialsystemen“, TU Hamburg-Harburg, 2003.
- H.-D. BURKHARD, H.-A. MARSISKE: *Endspiel 2050*. Telepolis (Heise), 2003.
- B. EL-HADDAD: *Towards building Multi-Agent Systems as Supporters in the Health Care Domain*. In L. Czaja (Ed.): *Proceedings of the Workshop on Concurrency, Specification and Programming (CS&P 2003)*. ISBN 83-88374-71-0, Vol. 1, pp. 213-224, Warsaw University, Poland, 2003.
- R. GERSTL, A. HANFT, S. MÜLLER, M. HAHNE, M. MEISTER, D. MONETT DIAZ: *Modellierung der praktischen Rolle in Verhandlungen mit einem erweiterten Verfahren des Fallbasierten Schließens*. Technical University - Technology Studies Working Papers, TUTS-WP-1-2003, Technische Universität Berlin, Institut für Soziologie, 2003. <http://www.tu-berlin.de/~soziologie/Tuts/Wp/TUTS_WP_1_2003.pdf>
- J. HOFFMANN, U. DÜFFERT: *Frequency Space Representation and Transitions of Quadruped Robot Gaits*. In V. Estivill-Castro (Ed.): *Proceedings of the 27th Australasian Computer Science Conference (ACSC 2004)*. *Conferences in Research and Practice in Information Technology*, Vol. 26, University of Otago, New Zealand, 2004 (to appear)
- M. JÜNGEL, J. HOFFMANN, M. LÖTZSCH: *A real-time auto-adjusting vision system for robotic soccer*. In: 7th International Workshop on RoboCup 2003 (Robot World Cup Soccer Games and Conferences). LNAI, Springer-Verlag, 2004 (to appear).
- F. KLÜGL, G. LINDEMANN, D. MOLDT: *Social Life (Preface)*. In P. Dittrich, J. Kim (Eds.): *Proceedings of the European Conference for Artificial Life (ECAL 2003)*, Part: Workshops and Tutorials. pp. 62-63, University of Dortmund, 2003.
- R. LATKOWSKI: *High Computational Complexity of the Decision Tree Induction with many Missing Attribute Values*. In L. Czaja (Ed.): *Proceedings of the Workshop on Concurrency, Specification and Programming (CS&P 2003)*. ISBN 83-88374-71-0, Vol. 2, pp. 318-325, Warsaw University, Poland, 2003.
- R. LATKOWSKI: *On Decomposition for Incomplete Data*. *Fundamenta Informaticae* 54(1), pp. 1-16, IOS Press, 2003.
- G. LINDEMANN, M. MEISTER: *Sozial agierende Agenten in hybriden Multi-Agentensystemen – Das INKA-Projekt*. In H.-D. Burkhard, T. Uthmann, G. Lindemann (Eds.): *Beiträge des ASIM-Workshops „Modellierung und Simulation menschlichen Verhaltens“*. Informatik-Berichte, ISSN 0863-095X , Nr. 163, S. 92-99, Humboldt-Universität zu Berlin, 2003.

- M. LÖTZSCH, J. BACH, H.-D. BURKHARD, M. JÜNGEL: *Designing agent behavior with the extensible agent behavior specification language XABSL*. In: 7th International Workshop on RoboCup 2003 (Robot World Cup Soccer Games and Conferences). LNAI, Springer-Verlag, 2004 (to appear).
- M. MEISTER, D. URBIG, K. SCHRÖTER, R. GERSTL: *Agents Enacting Social Roles. Balancing Formal Structure and Practical Rationality in MAS Design*. Technical University - Technology Studies Working Papers, TUTS-WP-6-2003, Technische Universität Berlin, Institut für Soziologie, 2003.
<http://www.tu-berlin.de/~soziologie/Tuts/Wp/TUTS_WP_6_2003.pdf>
- M. MINOR: *Making Use of a System's Treasury of Experiences*. In Beiträge der GI-Workshopwoche „Lehren-Lernen-Wissen-Adaptivität“. S. 210–215, Universität Karlsruhe, 2003.
- M. MINOR: Buchrezension zu *Experience Management*. Zeitschrift KI, 3/03, S. 78, 2003.
- M. MINOR, M. WERNICKE: *The Exchange of Retrieval Knowledge about Services between Agents*. In U. Reimer et al. (Hrsg.): Beiträge der 2. Konferenz Professionelles Wissensmanagement – Erfahrungen und Visionen (WM2003). Lecture Notes in Informatics, Vol. P-28, S. 295 – 303, Gesellschaft für Informatik (GI), Bonn, 2003.
- M. MINOR, M. WERNICKE: *Experience Management with knowledge sharing agents*. In F. McGrath, D. Remenyi (Eds.): Proceedings of the Fourth European Conference on Knowledge Management. pp. 669-676, MCIL, Reading, United Kingdom, 2003.
- D. MOLDT, G. LINDEMANN, M. PAOLUCCI (Eds.): *Regulated Agent-Based Social Systems: Theories and Applications*. LNCS 2934, Springer-Verlag, 2004 (to appear).
- D. MONETT: *Configuration of Metaheuristics: Overview and Theoretical Approach*. In L. Czaja (Ed.): Proceedings of the Workshop on Concurrency, Specification and Programming (CS&P 2003). ISBN 83-88374-71-0, Vol. 2, pp. 353-364, Warsaw University, Poland, 2003. (Eingereicht für das Journal Fundamenta Informaticae)
- J. E. NAVARRO, F. SCHWEITZER: *The Investors Game: A Model for Coalition Formation*. In L. Czaja (Ed.): Proceedings of the Workshop on Concurrency, Specification and Programming (CS&P 2003). ISBN 83-88374-71-0, Vol. 2, pp. 369-381, Warsaw University, Poland, 2003.
- J. E. NAVARRO, F. SCHWEITZER: *Formation of Coalitions in an InvestorGame*. Poster auf der Frühjahrstagung des Arbeitskreis Festkörperphysik bei der Deutschen Physikalischen Gesellschaft e. V., Dresden, 24. März 2003.
- A. OSHERENKO: *Modeling the Negotiation Object in Social Systems*. In G. Lindemann, C. Jonker (Eds.): Proceedings of the KI-2003 Workshop on Modelling Artificial Societies and Hybrid Organizations (MASHO'2003). pp. 29-39, Universität Hamburg, 2003.
- A. OSHERENKO: *Modeling the Negotiation Object in Social Systems*. In R. von Lüde, D. Moldt, R. Valk (Hrsg.): Sozionik Aktuell. Ein internationales Journal zur Sozionik. Ausgabe 2003, ISSN 1617-2477, S. 24-34, Universität Hamburg, 2003.
- A. OSHERENKO: *Modeling the Negotiation Object in Social Systems*. In Proceedings of the Second International Workshop on Regulated Agent-Based Social Systems: Theories and Applications (RASTA'03), University of Edinburgh, United Kingdom, 2003.
- T. RÖFER, I. DAHM, U. DÜFFERT, J. HOFFMANN, M. JÜNGEL, M. KALLNIK, M. LÖTZSCH, M. RISLER, M. STELZER, J. ZIEGLER: *GermanTeam 2003*. In 7th International Workshop on RoboCup 2003

(Robot World Cup Soccer Games and Conferences). LNAI, Springer-Verlag, 2004 (to appear).

T. RÖFER, M. JÜNGEL: *Fast and robust edge-based localization in the Sony four-legged robot league*. In 7th International Workshop on RoboCup 2003 (Robot World Cup Soccer Games and Conferences) . LNAI, Springer-Verlag, 2004 (to appear).

T. RÖFER, M. JÜNGEL: *Vision-Based Fast and Reactive Monte-Carlo Localization*. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA-2003). pp. 856-861, Taipei, Taiwan, 2003.

K. SCHRÖTER, D. URBIG, D. MONETT DIAZ: *Introducing the IPS framework to model altruistic relation-building agents in negotiations*. In G. Lindemann, C. Jonker (Eds.): Proceedings of the KI-2003 Workshop on Modelling Artificial Societies and Hybrid Organizations (MASHO'2003). pp. 40-50, Universität Hamburg, 2003.

K. SCHRÖTER, D. URBIG, D. MONETT DIAZ: *Introducing the IPS framework to model altruistic relation-building agents in negotiations*. In R. von Lüde, D. Moldt, R. Valk (Hrsg.): Sozionik Aktuell. Ein internationales Journal zur Sozionik. Ausgabe 2003, ISSN 1617-2477, S. 35-45, Universität Hamburg, 2003.

K. SCHRÖTER, C. WIECH: *Die Entwicklung sozialer Agenten für hybride Systeme am Beispiel der Schichtplanverhandlung*. Poster auf dem Jahrestreffen des DFG-Schwerpunktprogramms „Sozionik“, Kloster Seeon, 27.-30. Juni 2003.

D. URBIG: *Attitude Dynamics With Limited Verbalisation Capabilities*. Journal of Artificial Societies and Social Simulation (JASSS). Vol. 6, No. 1, 2003.
<<http://jasss.soc.surrey.ac.uk/6/1/2.html>>.

D. URBIG: *The value chain's values: Interpretations and Implications for industry and firm analysis*. In Proceedings of Perspectives in Business Informatics Research. Shaker Verlag, 2003.

D. URBIG, D. MONETT DÍAZ: *A general negotiation scenario and a multilateral negotiation mechanism with directed concessions*. In L. Czaja (Ed.): Proceedings of the Workshop on Concurrency, Specification and Programming (CS&P 2003). ISBN 83-88374-71-0, Vol. 2, pp. 524-538, Warsaw University, Poland, 2003.

D. URBIG, D. MONETT DIAZ, K. SCHRÖTER: *The C-IPS Agent Architecture for Modeling Negotiating Social Agents*. In M. Schillo et al. (Eds.): Multiagent System Technologies, Proceedings of the First German Conference (MATES 2003). LNAI 2831, pp. 217-228, Springer-Verlag Berlin Heidelberg, 2003.

D. URBIG, K. SCHRÖTER, D. MONETT DIAZ: *Modellierung strategisch verhandelnder Agenten in hybrider Sozialität*. In H.-D. Burkhard, T. Uthmann, G. Lindemann (Eds.): Beiträge des ASIM-Workshops „Modellierung und Simulation menschlichen Verhaltens“. Informatik-Berichte, ISSN 0863-095X , Nr. 163, S. 122-133, Humboldt-Universität zu Berlin, 2003.

J. WENDLER, J. BACH: *Case based behaviour assessment in simulated soccer*. In H.-D. Burkhard, T. Uthmann, G. Lindemann (Eds.): Beiträge des ASIM-Workshops „Modellierung und Simulation menschlichen Verhaltens“. Informatik-Berichte, ISSN 0863-095X , Nr. 163, S. 134-147, Humboldt-Universität zu Berlin, 2003.

C. WIECH, G. LINDEMANN: *Agents in the INKA-System based on the JADE-Platform*. In P. Dittrich, J. Kim (Eds.): Proceedings of the European Conference for Artificial Life (ECAL 2003), Part: Workshops and Tutorials. pp. 121-127, University of Dortmund, 2003.

C. WIECH, G. LINDEMANN: *The Description of INKA-Agents based on the JADE-Platform*. In L. Czaja (Ed.): *Proceedings of the Workshop on Concurrency, Specification and Programming (CS&P 2003)*. ISBN 83-88374-71-0, Vol. 2, pp. 575-581, Warsaw University, Poland, 2003.

Vorträge

J. BACH: *Emotionale Virtuelle Agenten auf der Basis der Dörnerschen Psi-Theorie*. ASIM-Workshop „Modellierung und Simulation menschlichen Verhaltens“, Humboldt-Universität zu Berlin, 24.-25. März 2003.

J. BACH: *Designing Agents with MicroPsi Node Nets*. 26th Annual German Conference on Artificial Intelligence (KI 2003), Universität Hamburg, 15.-18. September 2003.

J. BACH: *Connecting MicroPsi Agents to Virtual and Physical Environments – Applications of the AEP toolkit*. First International Workshop on Social Life (SOLI'03), held in Conjunction with the European Conference for Artificial Life (ECAL 2003), Dortmund, 14. September 2003.

J. BACH: *The MicroPsi Agent Architecture*. Fifth International Conference on Cognitive Modeling (ICCM 2003), Bamberg, 10.-12. April 2003.

J. BACH: *MicroPsi*. Kolloquium des Max-Planck Instituts für Adaptive Behavior and Cognition, Berlin, 16. April 2003.

J. BACH: *Action Control and Planning in MicroPsi Agents*. Workshop on Plan-Based Control of Robotic Agents, Schloss Dagstuhl, 22.-27. Juni 2003.

J. BACH: *Virtuelle Robotik mit MicroPsi-Agenten*. AIS, Schloß Birlinghoven, St. Augustin, 7. Juli 2003.

J. BACH: *Gespräch mit einer Künstlichen Intelligenz*. Fachtagung „Der maximal Fremde“ des IGPP und des Instituts für Soziologie der Universität Freiburg, 4.-6. Dezember 2003.

J. BACH, R. VUINE: *The AEP Toolkit for Agent Design and Simulatio*. First German Conference on Multiagent System Technologies (MATES 2003), Erfurt, 22.-24. September 2003.

R. BERGER: *Auf dem Weg zu menschlichem Verhalten – Simulation einer virtuellen Fußballumgebung*. ASIM-Workshop „Modellierung und Simulation menschlichen Verhaltens“, Humboldt-Universität zu Berlin, 24.-25. März 2003.

H.-D.BURKHARD: *Autonomous Robots - Programmed Rationality*. Workshop on Bounded Intelligent Extended Rationality, Berlin, 18.-19. Januar 2003.

H.-D.BURKHARD: *Fußball, Roboter und Intelligenz*. Faszination Wissenschaft, Konrad-Adenauer-Stiftung, Berlin, 4. Februar 2003.

H.-D.BURKHARD: *Wissensmanagement: Fußball, Agenten und Intelligenz*. IT & Versicherungstag, Empolis, Kaiserslautern, 16.-17. Juli 2003.

H.-D.BURKHARD: *Roboter spielen Fußball*. Veranstaltung „E-Fußball“, IFA Berlin, 1. September 2003.

H.-D.BURKHARD: *Software-Architekturen für Agenten und Autonome Mobile Roboter*. Fraunhofer Institute for Experimental Software Engineering (IESE), Kaiserslautern, 22. Oktober 2003.

- H.-D.BURKHARD: *Wenn Fußball-Roboter das Logische Schließen approximieren*. Symposium „Logik und Approximation“, Dagstuhl, 23.-24. Oktober 2003.
- H.-D.BURKHARD: *Architekturkonzepte für Autonome Systeme in komplexen dynamischen Umgebungen*. Workshop „Selbstoptimierung und Adaption“ (SFB 614), Paderborn, 24.-25.November 2003.
- B. EL-HADDAD: *Towards building Multi-Agent Systems as Supporters in the Health Care Domain*. Workshop on Concurrency, Specification and Programming (CS&P 2003), Czarna, Polen, 25.-27. September 2003.
- J. HOFFMANN: *A real-time auto-adjusting vision system for robotic soccer*. 7th International Workshop on RoboCup 2003, Padua, Italien, 10.-11. Juli 2003.
- R. LATKOWSKI: *High Computational Complexity of the Decision Tree Induction with many Missing Attribute Values*. Workshop on Concurrency, Specification and Programming (CS&P 2003), Czarna, Polen, 25.-27. September 2003.
- G. LINDEMANN: *Frauen in der Informatik – Von der Antike bis zur Gegenwart*. Schülerinformationstag an der Humboldt-Universität zu Berlin, 22. Mai 2003.
- G. LINDEMANN, M. MEISTER: *Sozial agierende Agenten in hybriden Multi-Agenten systemen – Das INKA-Projekt*. ASIM-Workshop „Modellierung und Simulation menschlichen Verhaltens“, Humboldt-Universität zu Berlin, 24.-25. März 2003.
- M. LÖTZSCH: *Designing agent behavior with the extensible agent behavior specification language XABSL*. 7th International Workshop on RoboCup, Padua, Italien, 10.-11. Juli 2003.
- M. MINOR: *The Exchange of Retrieval Knowledge about Services between Agents*. 2. Konferenz „Professionelles Wissensmanagement – Erfahrungen und Visionen“ (WM 2003), Luzern, Schweiz, 2.-4.April 2003.
- M. MINOR: *Experience Management with knowledge sharing agents*. 4th European Conference on Knowledge Management (ECKM-2003), Oxford, Großbritannien, 18. -19. September 2003.
- M. MINOR: *Making Use of a System's Treasury of Experiences*. FGWM-Workshop „Wissens- und Erfahrungsmanagement“ (FGWM 2003), Karlsruhe, 6.-8. Oktober 2003.
- D. MONETT: *Agent-Based Configuration of Metaheuristics -Ongoing PhD thesis-*. Technische Universität Berlin, Forschungsgebiet Bionik und Evolutionstechnik, 4. Juni 2003.
- D. MONETT: *Configuration of Metaheuristics: Overview and Theoretical Approach*. Workshop on Concurrency, Specification and Programming (CS&P 2003), Czarna, Polen, 25.-27. September 2003.
- D. MONETT: *A general negotiation scenario and a multilateral negotiation mechanism with directed concessions*. Workshop on Concurrency, Specification and Programming (CS&P 2003), Czarna, Polen, 25.-27. September 2003.
- J. E. NAVARRO: *The Investors Game: A Model for Coalition Formation*. Workshop on Concurrency, Specification and Programming (CS&P 2003), Czarna, Polen, 25.-27. September 2003.
- A. OSHERENKO: *Modeling the Negotiation Object in Social Systems*. Second International Workshop on Regulated Agent-Based Social Systems: Theories and Applications (RASTA'03), co-located to the 9th International Conference on Artificial Intelligence and Law (ICAAIL 2003), Edinburgh, Großbritannien, 23. Juni 2003.

A. OSHERENKO: *Modeling the Negotiation Object in Social Systems*. KI-2003 Workshop on Modelling Artificial Societies and Hybrid Organizations (MASHO 2003), Universität Hamburg, 16. September 2003.

K. SCHRÖTER: *Introducing the IPS framework to model altruistic relation-building agents in negotiations*. KI-2003 Workshop on Modelling Artificial Societies and Hybrid Organizations (MASHO 2003), Universität Hamburg, 16. September 2003.

K. SCHRÖTER, M. MEISTER: *Das Design rollenbasierter Agenten für hybride Organisationen*. Jahrestreffen des SPP Sozionik, Kloster Seeon, 27.-30. Juni 2003.

K. SCHRÖTER, D. URBIG: *Modellierung strategisch verhandelnder Agenten in hybrider Sozialität*. ASIM-Workshop „Modellierung und Simulation menschlichen Verhaltens“, Humboldt-Universität zu Berlin, 24.-25. März 2003.

D. URBIG: *Simulation von Einstellungsdynamiken im Kontext der Innovationsdiffusion*. Technische Universität Berlin, Institut für Betriebswirtschaftslehre, Lehrstuhl für Marketing I, 30. April 2003.

D. URBIG: *The C-IPS Agent Architecture for Modeling Negotiating Social Agents*. First German Conference on Multiagent System Technologies (MATES 2003), Erfurt, 22.-24. September 2003.

D. URBIG: *Die Evolution von Regelpraxis in offenen Netzwerken*. Tagung „Agentengesellschaften außer Kontrolle? Sozialtheoretische Mechanismen für den Bau von künstlichen Sozialsystemen“, TU Hamburg-Harburg, 7.-8. November 2003.

D. URBIG, E. LETTKEMENN, R. GERSTL: *Implementierung, Simulation und Hybridisierung des INKA-MAS*. Jahrestreffen des SPP Sozionik, Kloster Seeon, 27.-30. Juni 2003.

D. URBIG: *The value chain's values: Interpretations and Implications for industry and firm analysis*. Conference on Perspectives in Business Informatics Research (BIR 2003), Berlin, 18.-20. September 2003.

J. WENDLER, J. BACH: *Case based behaviour assessment in simulated soccer*. ASIM-Workshop „Modellierung und Simulation menschlichen Verhaltens“, Humboldt-Universität zu Berlin, 24.-25. März 2003.

C. WIECH, G. LINDEMANN: *Agents in the INKA-System based on the JADE-Platform*. First International Workshop on Social Life (SOLI'03), held in Conjunction with the European Conference for Artificial Life (ECAL 2003), Dortmund, 14. September 2003.

C. WIECH, G. LINDEMANN: *The Description of INKA-Agents based on the JADE-Platform*. Workshop on Concurrency, Specification and Programming (CS&P 2003), Czarna, Polen, 25.-27. September 2003.

Präsentationen zum RoboCup

Demonstrationsspiele mit den Robotern bei folgenden Veranstaltungen:

- *Techfest (Technologie Festival)* in Bombay, Indien (30.1.-4.2.2003)
- *Faszination Wissenschaft: Computer - Roboter - künstliche Intelligenz*. Berlin, Akademie der Konrad-Adenauer-Stiftung (3.2. und 4.2.2003)
- *Lange Nacht der Wissenschaft* in Berlin, 15.5.2003
- *Lange Nacht der Wissenschaften* in Erlangen, 25.10.2003

- *Campuswoche Adlershof, 2.12.2003*

Kooperationen:

Innerhalb der Humboldt-Universität

- Charité – Universitätsmedizin Berlin, Klinik für Kardiologie, Angiologie und Pulmologie, PROF. DR. GERT BAUMANN
- Charité – Universitätsmedizin Berlin, Klinik für Mund-, Kiefer- und Gesichtschirurgie - Klinische Navigation und Robotik, PROF. DR. TIM LÜTH
- Charité – Universitätsmedizin Berlin, Medizinische Klinik und Poliklinik mit Schwerpunkt Nephrologie, PROF. DR. H.-H. NEUMAYER, DR. LUTZ FRITSCHKE
- Institut für Entrepreneurship/Innovationsmanagement, PROF. DR. CHRISTIAN SCHADE

Mit anderen Universitäten

- Technische Universität Berlin, Institut für Soziologie, Fachgebiet Techniksoziologie, PROF. DR. WERNER RAMMERT
- Technische Universität München
- Universität Bremen, Fachbereich Mathematik und Informatik, Arbeitsbereich „Kognitive Robotik“, Arbeitsgruppe des „German Team“
- Universität Darmstadt, [Fachbereich Informatik](#), Fachgebiet Simulation und Systemoptimierung, Arbeitsgruppe des „German Team“
- Universität Dortmund, Computer Engineering Institute, Robotics Section, Arbeitsgruppe des „German Team“
- University of Technology and Information Management, Rzeszow, Polen
- Universität Warschau, Institute für Informatik und Mathematik, Polen
- Zentrum für Biomaterialien der Universität Havanna, Kuba

Mit außeruniversitären Forschungseinrichtungen und Partnern in der Industrie

- Empolis nGmbH Kaiserslautern
- Siemens AG, München
- SONY, Japan

Weitere Aktivitäten:

Prof. Dr. Hans-Dieter Burkhard

- Mitglied im Board of Trustees der RoboCup Federation
- ECCAI Fellow
- Sprecher des AK „RoboCup“ der GI
- Mitglied der Koordinationsgruppe des DFG-SPP 1125 „Kooperierende Teams mobiler Roboter in dynamischen Umgebungen“
- Mitgliedschaft in Programmkomitees: AOIS2003, ASIM03, BIS2003, EUMAS03, IAT2003, ICAAI03, ICCBR03, IDA2003, IEEE-WIC2003, KI2003, MATES2003, RSFDGrC2003
- Mitgliedschaft in Organisations-Komitees: German Open 2003, CS&P2003
- Mitglied im Fakultätsrat
- Mitglied im Institutsrat

Dipl.-Inf. Joscha Bach

- Mitglied im Programmkomitee der 5. International Conference on Cognitive Modelling (ICCM 03), Bamberg
- Mitglied im Programmkomitee des German Workshop on Artificial Life (GWAL04)

Dipl.-Phys. Jan Hoffmann

- Mitglied im Programmkomitee der 27th Australasian Computer Science Conference (ACSC2004), Dunedin, New Zealand

Dr. Gabriela Lindemann-v. Trzebiatowski

- Frauenbeauftragte der Mathematisch Naturwissenschaftlichen Fakultät II
- Mitglied der Fachgruppenleitung „VKI“ der Gesellschaft für Informatik
- Stellvertreterin „Akademischer Mittelbau“ in der Kommission für Frauenförderung der Humboldt Universität zu Berlin (seit 12/2003)
- Co-Chair des Workshops Regulated Agent-Based Social Systems: Theories and Applications (RASTA'03), co-located to the 9th International Conference on Artificial Intelligence and Law (ICAIL 2003), Edinburgh, Scotland, United Kingdom
- Co-Chair des Workshops Modelling Artificial Societies and Hybrid Organizations (MASHO'03), 26th German Conference on Artificial Intelligence (KI 2003), Hamburg
- Co-Chair des Workshops Concurrency, Specification & Programming (CS&P 2003), Czarna, Poland
- Co-Editorin “Fundamenta Informaticae”, Vol. 55, Nr. 2, ISSN 0169-2968, FUINE 55(2) 101-242 (2003), IOS Press, The Netherlands
- Mitgliedschaft in Programmkomitees: 5th International Conference on Enterprise Information Systems (ICEIS 2003), RASTA'03, MASHO'03, CS&P 2003, ASIM 2003, First German Conference on Multi-Agent System Technologies (MATES'03), 6th International Conference on Business Information Systems (BIS2003)
- Mitglied in den Berufungskommissionen C4-Professur für Systemarchitektur und S-C3-Professur “Computer Vision”

Dipl.-Inf. Mirjam Minor

- Stellvertretende Sprecherin der GI-Fachgruppe „Wissensmanagement“.
- Co-Chair des ICCBR-Workshops “Long-Lived CBR Systems”, Trondheim, Norwegen, 2003.
- Programmkomitees: GWEM-2003, FGWM-2003.
- Mitglied der Kommission Lehre und Studium des Instituts für Informatik
- Mitglied der Wahlkommission für die zentrale Frauenbeauftragte der HU Berlin

Dipl.-Kfm. Dipl.-Inf. Diemo Urbig

- Mitglied der Academy of Management (AoM)

Diplomarbeiten

MIKE WERNICKE: *Textuelles Fallbasiertes Schließen für die Benutzerschnittstelle von Agenten*. Februar 2003.

JOSÉ CARLOS DEL PRADO: *Multilingual – ein mehrsprachiges Fallbasiertes Informationssystem zur Fußball-Domäne im Web*. Februar 2003.

PETER HAMMELS: *OntoDigger – Erfassen von Ontologien aus deutschen Texten mit Hilfe von Wortanalyse und regulären Ausdrücken*. Juli 2003.

GORDON KRAMER: *Lernen für ThemeSearch – Erweiterung einer themenspezifischen Suchmaschine und semiautomatische Konstruktion der Wissensbasis*. September 2003.

HELMUT MYRITZ: *Fallbasierte Suche ähnlicher Fälle in einer nephrologischen Domäne*. November 2003.

ALEXANDER MOCZKO: *Anwendbarkeit fallbasierter Ähnlichkeitsfunktionen für vages Matching in Datenbanken*. Dezember 2003.

INGMAR DITTMANN: *Webdienste und Weboberflächen Internetzugriff auf „ChariTime“*. Dezember 2003.

DIPL.-KFM. DIEMO URBIG: *Agentenarchitektur zur Simulation von auf Informationsdiffusion beruhenden Dynamiken mehrerer interdependenter Einstellungen*. Dezember 2003.

CARSTEN HILBIG: *Retrievalmechanismen für BCRNs in Datenbanken*. (eingereicht)

Dissertationen

DIPL.-INF. JAN WENDLER: *Automatisches Modellieren von Agenten-Verhalten – Erkennen, Verstehen und Vorhersagen von Verhalten in komplexen Multi-Agenten-Systemen*. August 2003.

DIPL.-ING. STEFAN WEBER: *Ein navigierter Bildbetrachter für medizinische Bilddaten*. (eingereicht)

Lehr- und Forschungseinheit

Wissensmanagement

http://informatik.hu-berlin.de/Forschung_Lehre/wm/

Leiter

PROF. DR. TOBIAS SCHEFFER (SEIT 06.2003)

Tel.: (030) 2093 3107

E-Mail: scheffer@informatik.hu-berlin.de

Sekretariat

SABINE DZIWSZ

Tel.: (030) 2093 3044

Fax: (030) 2093 3045

E-Mail: dziwisz@informatik.hu-berlin.de

Wissenschaftliche Mitarbeiter

DIPL. WIRTSCH.-INFORM. STEFFEN BICKEL (SEIT 10. 2003)

DIPL.-INFORM. ULF BREFELD (SEIT 10. 2003)

DIPL.-INF. (FH) ISABEL DROST (SEIT 12. 2003)

Wir entwickeln Verfahren, die Wissen in großen Datenbanken - oder großen Textsammlungen - entdecken und nutzbar machen. Dazu gehören statistische, maschinelle Lernverfahren, Sampling- und aktive Lernalgorithmen, Assoziationsregel- und Subgruppenverfahren. Wir untersuchen Anwendungen vor allem im Dokumentenmanagement und in der Bioinformatik.

Zu den Herausforderungen des maschinellen Lernens gehört die Analyse sehr großer Datenbanken. Sampling-Verfahren können Datenbanken verarbeiten, die so groß sind, dass es unmöglich wäre, über alle Einträge zu iterieren, und liefern dennoch Optimalitätsgarantien. Eine Herausforderung beim Klassifizierungslernen liegt in der effektiven Ausnutzung unklassifizierter Beispieldaten. Wir untersuchen Verfahren, die Klassifikatoren aus wenigen klassifizierten und vielen unklassifizierten Beispielen lernen können.

Bei vielen Wissensentdeckungsproblemen liegen Hintergrundinformationen in Form von Texten vor - etwa im Web. Wir entwickeln Text-Mining-Verfahren, die Wissen in Dokumentensammlungen entdecken und nutzen; beispielsweise um Dokumentenverarbeitungsprozesse zu unterstützen. Dokumentenmanagement (spezieller: Email-Management) ist eines der interessantesten Anwendungsgebiete des maschinellen Lernens. Viele Unternehmen - aber auch Personen - investieren viel Zeit in das Beantworten von Emails und verwenden dabei häufig wiederkehrende Formulierungen, Sätze oder Emails. Wir untersuchen, wie sich das Verfassen von Dokumenten in bestimmten, thematisch eingeschränkten Gebieten durch Software unterstützen lässt.

Eine weitere Herausforderung für Text-Mining-Ansätze liegt in der Bioinformatik: um biologische Modelle zu gewinnen, die etwa die Funktion bestimmter Gene vorhersagen, müssen Informationen berücksichtigt werden, die über zahllose wissenschaftliche Veröffentlichungen verteilt sind.

Lehre

Veranstaltungen im Grundstudium

- Proseminar: Maschinelles Lernen (T. SCHEFFER, S. BICKEL, WS 2003/2004).

Veranstaltungen im Hauptstudium

- Vorlesung: Maschinelle Sprachverarbeitung, Halbkurs (T. SCHEFFER, WS 03/04).
- Übung: Maschinelle Sprachverarbeitung (U. BREFELD, WS 2003/2004).
- Forschungsseminar: Machine Learning Journal Club (T. SCHEFFER, WS 2004/2004).

Forschung

Projekt: "Text Mining: Wissensentdeckung in Textsammlungen und Effizienz von Dokumentenverarbeitungsprozessen"

Ansprechpartner: PROF. DR. TOBIAS SCHEFFER

Beteiligte Mitarbeiter: STEFFEN BICKEL, ULF BREFELD, ISABEL DROST

Mitarbeiter mit Werkvertrag: ANDREAS LÜNEBURG

Forschungsförderung: Deutsche Forschungsgemeinschaft

Das Volumen an Dokumenten, die in Archiven und im Web verfügbar sind, wächst explosionsartig. Daraus entsteht ein Bedarf an Verfahren, die große Textmengen automatisch analysieren, klassifizieren und filtern können, und die das in ihnen enthaltene Wissen extrahieren, darstellen und nutzbar machen können.

Viele Arbeitsprozesse bestehen operational aus der Verarbeitung – dem Lesen und Verfassen – von Texten. Es gibt bereits viele Werkzeuge für die *Verwaltung* von Texten, etwa Dateisysteme, Datenbanken, Dokumentenmanagementsysteme und Suchmaschinen. Die eigentlichen *Verarbeitungsprozesse* – dazu zählt etwa das Verfassen von Texten – sind jedoch wissensintensiv; eine Unterstützung dieser Prozesse durch Software ist ungleich schwieriger.

Gegenstand des Forschungsvorhabens sind die Entwicklung und Untersuchung von Text-Mining-Verfahren, die Wissen in Dokumentensammlungen entdecken, und dieses gewonnene Wissen nutzen, um künftige Dokumentenverarbeitungsprozesse zu unterstützen und effizienter zu gestalten.

- Ein Ziel besteht in der Weiterentwicklung und Untersuchung der Eigenschaften effizienter Algorithmen zum aktiven Lernen von Sequenzmodellen aus großen Textmengen. Solche statistischen, lernfähigen Verfahren ermöglichen es, Textanalysemodelle automatisch aus Sammlungen von Beispieltextrn zu generieren. Sie sind in der Lage, Texte zu klassifizieren, zu segmentieren und Informationen daraus zu extrahieren und in strukturierter Form abzulegen. Auf der Grundlage statistischer Textmodelle eröffnen sich zahlreiche Möglichkeiten, die Verarbeitung von Dokumenten zu unterstützen und dadurch Effizienz zu gewinnen.
- Während Text-Mining-Verfahren wie Markov-Modelle es ermöglichen, Informationen aus Texten zu extrahieren und zu strukturieren, können Data-Mining-Verfahren in strukturierten Datensammlungen Wissen in Form von Regeln, Mustern oder Zusammenhängen gewinnen. Ein weiteres Ziel ist die Entwicklung und Untersuchung von Verfahren, die eine Verzahnung dieser beiden Schritte leisten und es damit ermöglichen, Wissen in unstrukturierten Textsammlungen zu entdecken.

- Umfassendes Ziel des Projektes ist die Kombination von Wissensentdeckung in Textsammlungen und der Unterstützung von Dokumentenverarbeitungsprozessen. Erst das durch die Analyse archivierter Texte gewonnene Wissen ermöglicht die Unterstützung zukünftiger Prozesse durch neu zu entwickelnde Verfahren. Exemplarisch soll dafür eine Komplettierungsfunktion für natürlichsprachliche Texte entwickelt und untersucht werden. Basierend auf gespeicherten, in der Vergangenheit geschriebenen Texten soll das Text-Mining-System ein statistisches Textmodell generieren, das häufig verwendete Formulierungen und deren semantischen Kontext beschreibt. In der Anwendungsphase soll das System den Inhalt eines Textfragmentes analysieren und feststellen, ob aufgrund des semantischen Kontextes und in der Vergangenheit geschriebener Texte die Fortsetzung einer Textpassage mit hoher Konfidenz vorhergesagt und dem Benutzer vorgeschlagen werden kann. Eine zu untersuchende Anwendung ist der Einsatz in einem Email-Client. Hier wird ein statistisches Textmodell aus gespeicherten Emails gelernt. Nach Analyse einer eingehenden Email (und eventuell eines Antwort-Fragmentes) soll das System einen Antworttext vorschlagen, der in ähnlicher Form häufig auf Emails ähnlichen Inhalts erwidert wurde.

Projekt: "European Network of Excellence in Knowledge Discovery"

Netzwerk-Koordination: DR. MICHAEL MAY, Fraunhofer-Institut für autonome, intelligente Systeme, Bonn

Ansprechpartner für den Netzwerk-Knoten Humboldt-Universität zu Berlin: PROF. DR. BETTINA BERENDT, PROF DR. TOBIAS SCHEFFER

Forschungsförderung: Europäische Union

Das „KD-Net“ ist ein Europäisches Exzellenz-Netzwerk, an dem sich Partner aus Wissenschaft, Industrie und öffentlichem Sektor beteiligen. Ziele dieses internationalen Projektes sind es, neue Anwendungen des maschinellen Lernens zu identifizieren und in die Forschung einfließen zu lassen, sowie in Zusammenarbeit die zukünftige inhaltliche Struktur des Forschungsgebietes „Wissensentdeckung und Data Mining“ zu gestalten.

Wir haben im Jahr 2003 einen vom KD-Net finanziell unterstützten Workshop „Data Mining und Text Mining for Bioinformatics“ organisiert.

Veröffentlichungen

M.-A. KROGEL, T. SCHEFFER: *Effectiveness of information extraction, multi-relational and semi-supervised learning for predicting functional properties of genes*. In Proceedings of the IEEE International Conference on Data Mining. 2003.

T. SCHEFFER, U. LESER, editors: Proceedings of the European Workshop on Data Mining and Text Mining for Bioinformatics. 2003.

M.-A. KROGEL, T. SCHEFFER: *Effectiveness of information extraction, multi-relational and multi-view learning for predicting gene deletion experiments*. In Proceedings of the Third ACM SIGKDD International Workshop on Data Mining for Bioinformatics. 2003.

M.-A. KROGEL, M. DENECKE, M. LANDWEHR T. SCHEFFER: *Using Data and Text Mining Techniques for Yeast Gene Regulation Prediction: A Case Study*. In SIGKDD Explorations 4 (2), 2003.

M. KOCKELKORN, A. LÜNEBURG, T. SCHEFFER: *Using transduction and multi-view learning to answer emails*. In Proceedings of the European Conference on Principles and Practice of Knowledge Discovery in Databases. 2003.

M. KOCKELKORN, A. LÜNEBURG, T. SCHEFFER: *Learning to answer emails*. In Proceedings of the International Symposium on Intelligent Data Analysis. 2003.

U. BREFELD, P. GEIBEL, F. WYSOTZKI: *Support Vector Machines with Example Dependent Costs*. In: N. LAVRAC, D. GAMBERGER, L. TODOROVSKI, H. BLOCKEEL (Eds.), Proceedings of the 14th European Conference on Machine Learning, ECML 2003, LNAI 2837, Springer, 2003.

P. GEIBEL, U. BREFELD, F. WYSOTZKI: *Learning Linear Classifiers Sensitive to Example Dependent and Noisy Costs*. In: F. PFENNING, M. R. BERTHOLD, H.-J. LENZ, E. BRADLEY, R. KRUSE, C. BORGELT, (Eds.), *Advances in Intelligent Data Analysis - 5th International Symposium on Intelligent Data Analysis, IDA 2003*, LNCS 280, Springer, 2003.

Vorträge

S. BICKEL: *Optimierung von Sicherheitsbeständen in Supply Chains mit Simulation*. Operations Research 2003.

U. BREFELD: *Support Vector Machines with Example Dependent Costs*. European Conference on Machine Learning.

T. SCHEFFER: *Aktive Lernalgorithmen für die Wissensentdeckung in Textsammlungen*. Jahrestreffen des GI-Beirats der Universitätsprofessoren.

T. SCHEFFER: *Effectiveness of information extraction, multi-relational and semi-supervised learning for predicting functional properties of genes*. IEEE International Conference on Data Mining.

T. SCHEFFER: *Effectiveness of information extraction, multi-relational and multi-view learning for predicting gene deletion experiments*. Third ACM SIGKDD International Workshop on Data Mining for Bioinformatics.

T. SCHEFFER: *Using transduction and multi-view learning to answer emails*. European Conference on Principles and Practice of Knowledge Discovery in Databases.

T. SCHEFFER: *Learning to answer emails*. International Symposium on Intelligent Data Analysis.

T. SCHEFFER: *Aktive Wissensentdeckung für die Wissensentdeckung in Datenbanken*. Ludwig-Maximilians-Universität München.

T. SCHEFFER: *Wissensentdeckung in Textarchiven: Überblick und Anwendungen des Text Mining*. Eingeladener Vortrag bei den Data-Mining-Cup-Anwendertagen der Prudsys AG, Chemnitz.

Workshop-Organisation

“European Workshop on Data Mining and Text Mining for Bioinformatics” im Rahmen der European Conference on Machine Learning, Dubrovnik. Gemeinsame Organisation mit PROF. ULF LESER. 22.09.2003.

Kooperationen

- Mit der Westfalia AG: Pilotprojekt zum Management von Kundenanfrage-E-mails.

- Mit der GSD – Gesellschaft für Systemforschung und Dienstleistungen im Gesundheitswesen mbH: Pilotprojekt zur automatischen Beantwortung technischer Anfragen.
- STEFAN WROBEL, MICHAEL MAY, Fraunhofer-Institut für autonome, intelligente Systeme und Universität Bonn: Entwicklung eines Email-Management-Werkzeuges.
- MARK KROGEL, Otto-von-Guericke-Universität Magdeburg: Multirelationales Data Mining.
- MARCO LANDWEHR, Leibniz-Institut für Neurobiologie Magdeburg: Text Mining in der Bioinformatik.
- Mit der Arbeitsgruppe Wissensmanagement in der Bioinformatik (PROF. ULF LESER), Named-Entity-Recognition in biomedizinischen Veröffentlichungen.

Sonstige Aktivitäten

Dipl.-Wirtsch.-Inform. Steffen Bickel

- Gutachten für Intelligent Data Analysis Journal

Dipl.-Inf. Ulf Brefeld

- Gutachten für Intelligent Data Analysis Journal

Dipl.-Inf. Isabel Drost

- Gutachten für Intelligent Data Analysis Journal

Prof. Dr. Tobias Scheffer

- Vorsitzender des Programmkomitees, European Workshop on Data Mining and Text Mining for Bioinformatics
- Mitglied des Programmkomitees, European Conference on Machine Learning
- Mitglied des Programmkomitees, International Symposium on Intelligent Data Analysis
- Mitglied des Programmkomitees, International Conference on Discovery Science
- Mitglied des Programmkomitees, SIGKDD Workshop on Data Mining in Bioinformatics
- Mitglied des Programmkomitees, International Workshop on Biological Data Management
- Gutachter, Machine Learning Journal
- Gutachter, Journal of Machine Learning Research
- Gutachter, Information Processing Letters
- Gutachter, Intelligent Data Analysis Journal
- Gutachter, International Fuzzy Systems Association World Congress
- Gutachter, GI-Fachgruppentreffen Maschinelles Lernen
- Mitglied der Berufungskommission für C3-Professur “Praktische Informatik” an der Otto-von-Guericke-Universität Magdeburg

Lehr- und Forschungseinheit

Datenanalyse

<http://www.informatik.hu-berlin.de/lehrstuehle/datenanalyse/>

Leiter

PROF. DR. EGMAR RÖDEL

Tel.: (030) 2093 3079

E-Mail: roedel@informatik.hu-berlin.de

Sekretariat

BIRGIT EISENMANN

Tel.: (030) 2093 3080

Fax: (030) 2093 3081

E-Mail: eisenman@informatik.hu-berlin.de

Mitarbeiter

PRIV. DOZ. DR. WOLFGANG KÖSSLER

Externe Mitarbeiter

DR. KLAUS-PETER NEUENDORF

Tutoren

FRANK HAMBACH

ANDRE ROSIN

Die Lehr- und Forschungseinheit vertritt die Gebiete „Stochastische Aspekte der Informatik“ und „Computergestützte Statistik“.

Lehre

Vorlesungen

- Stochastik für Informatiker (E. RÖDEL, WS 2002/03)
- Mathematik für Informatiker (L) (E. RÖDEL, WS 2002/03)
- Einführung in die Informations- und Kodierungstheorie (E. RÖDEL, SS 2003)
- Forschungsfreiemester (E. RÖDEL, WS 2003/04)
- Werkzeuge der empirischen Forschung (W. Kössler, SS 2003)
- Zirkel der Mathematischen Schülergesellschaft, 12./13. Klasse (W. Kössler, WS 2002/03, SS 2003 und WS 2003/04)
- Theorie und Praxis des E-Learnings (K.-P. NEUENDORF, SS 2003)

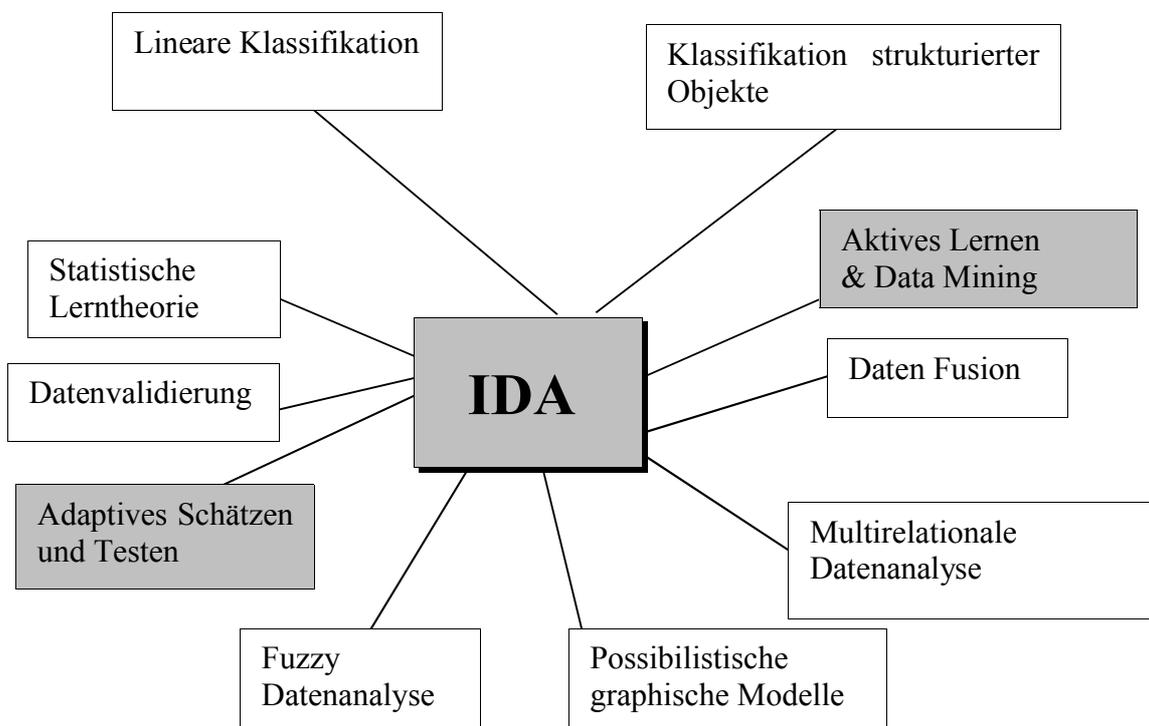
Forschung

Projekt: Intelligente Datenanalyse (IDA)

Ansprechpartner: Prof. Dr. Egmar Rödel

Intelligente Datenanalyse ist ein gutes Beispiel, wie Computer und Netzwerke Praxis und Wissenschaft verändert haben und weiter verändern werden. Heutzutage sind Probleme der Datenanalyse lösbar, die vor Dekaden außerhalb des „Machbaren“ lagen, insbesondere im Bereich personeller Computer. Ermöglicht wurde dies durch extreme Fortschritte in der Computertechnologie, Rechengeschwindigkeit und Speicherkapazität. Die modellorientierte Sicht der Statistik ergänzt die datengetriebene Sicht der Informatik: Skalierbarkeit von Problemen, Lernverfahren und die Effizienz von Algorithmen und Datenstrukturen treten gleichberechtigt neben Fragen nach Unverzerrtheit, Konsistenz, Robustheit und Adaptivität von Schätz- und Testverfahren.

Vertreter aus der TU Berlin, der FU Berlin, der TU Magdeburg und unser Bereich bilden seit drei Jahren eine Arbeitsgruppe, die sich dieser Thematik stellt. „Intelligente Datenanalyse“ wird dabei nicht vornehmlich als Datenanalyse durch „intelligente Maschinen“ gesehen. Das Ziel ist es vielmehr, Forschung und Lehre auf die „intelligente“ Verbindung von Informatik, Statistik und Mathematik abzustellen. Auf ausgewählten Problemfeldern (s. Diagramm) sollen diese drei Sichten verschmolzen werden zu einer einheitlichen („intelligenten“) Vorgehensweise moderner Datenanalyse.



Ausgewählte Problemfelder der IDA, Beteiligung der HUB (schattiert)

Dazu gehören Datenvalidierung, Schätzen und Lernen sowie das Aufdecken relevanter unbekannter (latenter) Strukturen sowie Klassifikationsverfahren. Fragen der Skalierbarkeit,

der Laufzeit-Komplexität der Algorithmen und der (finiten bzw. asymptotischen) Eigenschaften neuerer statistischer Verfahren sind neben Interpretierbarkeit der Ergebnisse wesentliche Teilziele.

Unser Schwerpunkt in diesem Projekt liegt in der Entwicklung und Analyse adaptiver Test- und Schätzverfahren, einem typischen Ansatz der intelligenten Datenanalyse. In vielen Anwendungen ist die Zugrundelegung eines bestimmten, auf sehr restriktiven Modellannahmen basierenden Schätz- oder Testverfahrens nicht zu rechtfertigen, der praktizierende Statistiker hat nun einmal in der Regel keine gesicherte Erkenntnis über den seine Daten erzeugenden Prozess. Was liegt dann also näher, als vorab „einen Blick auf die Daten zu werfen“ und nach Ausnutzung spezifischer Informationen aus diesem Datensatz dann ein geeignetes Verfahren auszuwählen.

Der Initiative der Arbeitsgruppe ist es mit zu verdanken, dass die internationale Konferenz „IDA“ im Jahr 2003 in Berlin stattfand.

Projekt: Optimale Datentransformationen und funktionelle Abhängigkeiten

Ansprechpartner: PROF. DR. EGMAR RÖDEL

Multivariaten Roh-Daten mangelt es meistens an Transparenz. Durch geeignete Datentransformationen kann dieser Zustand verbessert werden. Die erforderlichen Transformationen hängen insbesondere von den Informationen ab, die hinterfragt werden sollen. Zum Erkennen funktioneller Abhängigkeiten werden optimale Datentransformationen ermittelt. Diese Transformationen basieren auf multivariaten Dichteschätzungen und daraus resultierenden Eigenproblemen. Für positiv abhängige Merkmale wurde die Struktur dieser Eigenprobleme erarbeitet und dementsprechende statistische Methoden wurden entwickelt und implementiert. Hierbei bestätigten sich erneut die guten numerischen Eigenschaften der in den letzten Jahren entwickelten Algorithmen für die Projektion auf konvexe Kegel und der rangbasierten ACE-Verfahren (Alternating Conditional Expectation). Diese Verfahren besitzen hervorragende Konvergenzeigenschaften, eine hohe Robustheit gegenüber Modellabweichungen und sind nicht so rechenintensiv wie herkömmliche Optimierungsverfahren.

Die Anwendung der Methoden in der Finanzmathematik und im Data Mining werden vorbereitet und bereits in Zusammenarbeit mit dem Institut für Wirtschaftsinformatik der FU Berlin erprobt.

Projekt: Statistische Prüfung von Daten

Ansprechpartner: PROF. DR. EGMAR RÖDEL

Zwischen den Merkmalen großer Datenmengen bestehen im Allgemeinen logische, arithmetische, probabilistische oder, noch allgemeiner, strukturelle Beziehungen. Wenn die Daten mit Fehlern behaftet sind, so werden diese inhärenten strukturellen Beziehungen verletzt. Das Hauptziel der Datenprüfung besteht nun darin, Widersprüche zwischen Daten und a priori gegebenen strukturellen Beziehungen aufzudecken. Eine besondere Rolle spielen hierbei Kovarianzstrukturen und Testverfahren für Hypothesen über diese Strukturen. Die Güte dieser Testverfahren hängt stark von der Verteilung der Daten ab, die im Allgemeinen unbekannt ist. So kann die Anwendung eines bestimmten Testverfahrens zu ungenauen oder sogar falschen Entscheidungen führen. Um die Güte verschiedener Tests unter verschiedenen Verteilungsannahmen einschätzen zu können, eignen sich Simulationsstudien. Eine solche Studie wurde für Unabhängigkeitstests durchgeführt. Neben üblichen bekannten Tests wurden verschiedene verteilungsunabhängige neue adaptive Verfahren miteinander verglichen.

Projekt: Konsultationszentrum „Statistische Methoden und ihre Anwendungen in der empirischen Forschung“**Ansprechpartner:** DR. WOLFGANG KÖSSLER

Im Laufe der letzten Jahre wurde die Lehr- und Forschungseinheit Datenanalyse zu einem Anlaufpunkt für Ratsuchende in der Anwendung statistischer Methoden bei der Auswertung empirischer Datenerhebungen. Viele dieser Ratsuchenden wurden vom Universitätsrechenzentrum an uns verwiesen. Insbesondere wenden sich Promovenden und Habilitanden naturwissenschaftlicher und medizinischer Einrichtungen unserer Universität an uns. Wir beraten sowohl bei der Methodenauswahl als auch bei der Nutzung von Soft- und Hardware.

Im vergangenen Jahr wurden u.a. folgende Projekte unterstützt:

- Chemometrische Verfahren in der pharmazeutischen Analytik (Institut für Pharmazie)
- Habitatverhalten von Schleiereulen (Institut für Biologie)
- Habitatverhalten von Hyänen (Institut für Biologie)
- Studien zur stimmlichen Erkennung von afrikanischen Wildhunden (Institut für Biologie)
- Untersuchungen zur Zelldichte (Institut für Biologie)
- Charakterisierungen immun-modulierender Wirkungen (Institut für Biologie)
- Genotyp und Leistungsfähigkeit von Schweinen (Institut für Nutztierwissenschaften)
- Analyse von Kraftdefiziten nach Sportverletzungen (Institut für Sportwissenschaften)
- Anatomische Veränderungen bei Zahnspangebehandlungen (Institut für Zahnmedizin).

E-Learning Projekt: TIM-Tutor –Theoretische Informatik mit *Mathematica***Ansprechpartner:** DR. KLAUS-PETER NEUENDORF**Studentische Mitarbeiter:** FRANK HAMBACH, TINO NAPHTALI, GREGOR LÄMMEL

Im Rahmen des Multimedia-Projekts TIM-Tutor, gefördert unter **MM-F03-38**, wurde ein Lehrauftrag zur Theorie und Praxis des E-Learning realisiert, sowie eine Sammlung von *Mathematica* – Moduln und Web-Moduln erstellt, die der multimedialen Darstellung und dem interaktiven Training von Grundbegriffen der Mathematik und der Theoretischen Informatik dienen. Nachfolgend zwei Beispiele aus dem TIM-Tutor Projekt:

Logik mit Minesweeper

Das Minenfeld wird durch genau eine Elementarkonjunktion repräsentiert und diese gilt es zu finden. Die Aussagenvariable p_{ij} repräsentiert das Feld mit den Koordinaten (i, j) und hat die Semantik, das Feld (i, j) ist frei.

Welches sind die logischen Regeln mit denen hier Schlüsse gezogen werden können?

Im TIM-Tutor System werden die Aussagen der Studenten überprüft und steuern den Spielverlauf.

	2	*		1
	3	*		
	2	*	4	
	1	1	2	*
			1	1

	2	*	3	1
	3	*	*	1
	2	*	4	2
	1	1	2	*
			1	1

folgt zwingend

Zum Markieren von Minen den Button aktivieren und danach wieder deaktivieren!

Anzahlaufgaben mit Venn-Diagrammen

Gegeben sind die Anzahlen der Elemente für folgende Mengen:

$$|\bar{C}| = 8, |B| = 5, |\bar{A} \cap B| = 5, |\bar{A} \cap \bar{C}| = 4, |\bar{A}| = 5, |A \cap C| = 5, |A \cap \bar{C}| = 4, |\bar{A} \cap B| = 0$$

Gesucht sind die möglichen Anzahlen für die restlichen Mengen bzw. Teilmengen:

$$|A \cap B \cap C|, |A \cap B \cap \bar{C}|, |A \cap B|, |A \cap \bar{B} \cap C|, |A \cap \bar{B} \cap \bar{C}|, |A \cap \bar{B}|, |A|, \\ |\bar{A} \cap B \cap C|, |\bar{A} \cap B \cap \bar{C}|, |\bar{A} \cap \bar{B} \cap C|, |\bar{A} \cap \bar{B} \cap \bar{C}|, |\bar{A} \cap C|, |B \cap C|, \\ |B \cap \bar{C}|, |\bar{B} \cap C|, |\bar{B} \cap \bar{C}|, |\bar{B}|, |C|, |\text{allSet}|.$$

Veröffentlichungen

E. RÖDEL: *Einführung in die Informations- und Kodierungstheorie*. Lehrbuch, eingereicht.

E. RÖDEL: *R-Estimation of Optimal Transformations via Copulas and ACE*. Statistics, im Druck.

E. RÖDEL, W. KÖSSLER: *Linear rank tests for independence in bivariate distributions – power comparisons by simulation*. Computational Statistics & Data Analysis, im Druck.

W. KÖSSLER: *Nichtparametrische Lokationstests bei eingeschränkten Alternativen*. Habilitationsschrift, Humboldt-Universität zu Berlin.

Die Lösung wird mit *Mathematica* berechnet und ist im nebenstehenden Venn-Diagramm durch die Exponenten der Bezeichner repräsentiert, wobei die Parameter p und q als natürliche Zahlen so gewählt werden müssen, dass alle Anzahlen nichtnegativ sind.

Mit dem TIM-Tutor System können Aufgabenstellungen diesen Typs randomisiert erzeugt werden und die Antworten der Studenten überprüft werden.

W. KÖSSLER: *Some c-Sample Rank Tests for Homogeneity Against Ordered Alternatives Based on U-Statistics*, eingereicht.

W. KÖSSLER: *Some c-Sample Rank Tests for Homogeneity Against Umbrella Alternatives With Unknown Peak*, eingereicht.

Vorträge

E. RÖDEL: *Optimal Transformation, Alternating Conditional Expectation and Tests of Independence*. International Conference of the Cuban Mathematical Society, Sancti Spiritus, November 2003.

E. RÖDEL: *Singular Cases for Optimal Transformations*. Universität Havanna, November 2003.

W. KÖSSLER: *Annahmeprüfung für stetige Verteilungen des Qualitätsmerkmals*. Hochschule Mittweida (FH) März 2003.

W. KÖSSLER: *Explizite robuste Skalenschätzungen*. Habilitationsvortrag, Humboldt-Universität zu Berlin, Mai 2003.

Sonstige Aktivitäten

Prof. Dr. Egmar Rödel, Dr. Wolfgang Kössler

- Gutachtertätigkeit für Zeitschriften: Statistics, Biometrical Journal, Statistical Papers, Allgemeines Statistisches Archiv
- Beratung von Anwendern statistischer Methoden

Prof. Dr. Egmar Rödel

- Direktor für Studienangelegenheiten (bis April 2003)
- Studienfachberater für die Lehramtsstudiengänge Informatik
- Mitglied im Institutsrat
- Mitglied der Internationalen Association for Statistical Computing (IASC)
- Mitglied der Fachgruppe Stochastik der DMV
- Mitglied des Programmkomitees IDA 2003

Dr. Wolfgang Kössler

- Statistische Erfassung und Auswertung der Studentenfragebögen
- Mitglied der Deutschen Statistischen Gesellschaft
- Mitglied der Kommission Lehre und Studium

Wissenschaftliche Kooperationen

- Institut für Mathematik und Informatik der Universität Havanna/Kuba
- Institut für Wirtschaftsinformatik der Freien Universität Berlin
- Institut für Statistik und Ökonometrie der Freien Universität Berlin
- Institut für Informatik der TU Berlin
- Universität Magdeburg

Diplomarbeiten

MAIHUB DAHDAL: *Numerische Verfahren zur Berechnung von Kanalkapazitäten*.

DIRK-JOHANN DETERS: *Ein SAS-Bediensystem zum Vergleich von Lokationsparametern*.

Lehr- und Forschungseinheit
Informatik in Bildung & Gesellschaft

<http://waste.informatik.hu-berlin.de/>

Leiter

PROF. DR. WOLFGANG COY

Tel. : (030) 2093 3166

E-Mail: coy@hu-berlin.de

Sekretariat

RENATE ZIRKELBACH

Tel.: (030) 2093 3167 Fax: (030) 2093 3168

E-Mail: zirkel@informatik.hu-berlin.de

Wissenschaftliche Mitarbeiter

DIPL. INF. PETER BITTNER

DR. HABIL. CHRISTIAN DAHME

DR. VOLKER GRASSMUCK (DFG)

DR. JOCHEN KOUBEK

DIPL. INF. ROLAND KUBICA

Technische Unterstützung

FRANK WOZOBULE

TUTOREN

HEINZ-GÜNTER KUPER

CONSTANZE KURZ

JENS-MARTIN LOEBEL

Die Arbeitsgruppe „Informatik in Bildung und Gesellschaft“ untersucht die Rolle der Informatik auf dem Weg zur Informationsgesellschaft, insbesondere die sozialen und kulturellen Wechselwirkungen der Informatik. Zu einer zentralen Herausforderung für den Bereich Informatik und Gesellschaft wird die entstehende globale „Informationsgesellschaft“, in der die Informatik als technische Grundlagenwissenschaft eine definierende Rolle spielen kann. Dies bildet die Ausgangslage für die Forschungen der Arbeitsgruppe.

Ein besonderer Schwerpunkt liegt in Analyse und Entwicklung Digitaler Medien. Zum einen betrifft dies Techniken der Digitalisierung, die tradierte Medien verändern, zum anderen die Entwicklung und Einsatz Neuer Medien mit einem Akzent auf dem Einsatz in Bildung und Lehre.

Zu den aktuellen Forschungsthemen der Arbeitsgruppe zählen einerseits technikspezifische Grundlagen der Informationsgesellschaft, dabei vor allem Nutzung und Dienste in globalen Rechnernetzen und die in diesem Kontext entstehenden Digitalen Medien, aber auch Fragen der Sicherheit und der Beherrschbarkeit der Informationstechnik sowie Fragen des politischen und gesellschaftlichen Umgangs mit dieser neuen Technik, wie z.B. Fragen des Urheberrechts oder der Internet-Ökonomie. Um diese Prozesse kritisch zu begleiten und zu

reflektieren sind grundlegende theoretische und methodische Fragen der Wissenschaft Informatik zu identifizieren und zu klären.

Besondere Aufmerksamkeit in dem durch die Informatik verursachten Prozess kulturellen Wandels beansprucht die Stellung der Informatik in Erziehung, Schule, Hochschule oder in der Fort- und Weiterbildung. Der Arbeitsbereich *Informatik in Bildung und Gesellschaft* fühlt sich damit nicht nur der Diplombildung, sondern auch der Informatikausbildung von Magister- und Lehramtstudierenden verpflichtet.

Die Gesellschaft für Informatik (GI) hat zu ihrer Jahrestagung in Frankfurt am Main satzungsgemäß die „Ethischen Leitlinien“ von 1994 überarbeitet. Drei Mitglieder der Arbeitsgruppe „Informatik und Verantwortung“ (W. Coy, P. Bittner und C. Kurz) waren an dieser etwa zweijährigen Überarbeitung aktiv beteiligt. Als ein weiteres Ergebnis dieser Arbeit wurde eine GI-Fachgruppe „Ethik“ gegründet.

Neue Akzente in der Forschung

Die Forschungsschwerpunkte der Arbeitsgruppe haben sich in die Richtungen „Informatik und Informationsgesellschaft“ und „Digitale Medien“ vertieft. Dies sind transdisziplinäre Themen. Die Arbeitsgruppe hat sich deshalb am „HERMANN VON HELMHOLTZ-Zentrum für Kulturtechnik“ beteiligt, wo sie im Rahmen der DFG-Forschergruppe „Bild-Schrift-Zahl“ ein Projekt zur „Visuellen Argumentation“ im Kontext der Digitalen, rechnergestützten Medien verfolgt und ein neues Projekt zu Urheberschaft und geistigen Eigentum für die nächsten drei Jahre beantragt hat.

Die Arbeitsgruppe vertritt die Humboldt-Universität im Berliner Stiftungsverbundkolleg der Alcatel-Stiftung. Mit Unterstützung der Stiftung wurde im Januar ein Festkolloquium aus Anlaß des 80. Geburtstages von Joseph Weizenbaum gefeiert und im Dezember eine internationale Tagung „100 Jahre Johann von Neumann 1903-2003“ veranstaltet. Im Rahmen dieser „Johann von Neumann“-Tagung stellte unsere *Hörspielgruppe* unter Leitung von Dr. Jochen Koubek erstmals eine eigene Produktion öffentlich vor.

<http://waste.informatik.hu-berlin.de/Tagungen>



100 Jahre Johann von Neumann – Einleitungsvortrag von William Aspray.

Innovationen in der Lehre

Die Arbeitsgruppe hat in zwei Richtungen neue Wege in der Lehre verfolgt. Seit SS 1999 werden regelmäßige Telelehrveranstaltungen durchgeführt, die inzwischen mit Hilfe der Multimediaabteilung des zentralen Rechenzentrums verstetigt sind. In diesem Rahmen werden Vorlesung und Übungen zeitgleich an zwei Standorten, nämlich BERLIN-MITTE (Seminarerbäude) und BERLIN-ADLERSHOF (im Informatikgebäude) durchgeführt. Die schon lange multimedial angelegte Lehrveranstaltung „Informatik & Informationsgesellschaft I: Digitale Medien“ sowie die Vorlesung „Informatik & Informationsgesellschaft II: Technik, Kontext und Geschichte“ mit ihren Übungen wurden entsprechend angepasst. Damit wird es Studierenden im Lehramt- und Magisterstudium, die ihren Studienschwerpunkt in Berlin-Mitte haben, erleichtert, an Vorlesungen, Seminaren und Übungen, die in Adlershof stattfinden, teilzunehmen.

Durch den studentischen Streik gegen die Hochschulpolitik des Landes Berlin bedingt, hat die Lehre im Semester gelitten, da das Seminarerbäude mehrere Wochen besetzt und unzugänglich war. Auch die Vorlesung *Zwischen „Information Sharing“ und „Information Control“* von Peter Bittner wurde vom Streik betroffen – sie fand zweimal in Form einer Begehung der Berliner Innenstadt statt, wo sich mit einer Fülle installierter Videokameras reichhaltiges Anschauungsmaterial für die aktuellen Trends der Überwachungstechnik im öffentlichen Raum finden lässt. Da solche Veranstaltungen nach §14 Versammlungsgesetz anmelde-, aber nicht genehmigungspflichtig sind, war dies wohl die erste Vorlesung des Instituts, die unter Polizeiaufsicht stattfand. Nachdem die Beamten über ihre Rechte aufgeklärt wurden, hat sich dabei eine vertrauensvolle Kooperation entwickelt.

Die Kooperation der Arbeitsgruppe mit den Kultur- und den Kunstwissenschaften wurde im Rahmen der Arbeiten im HERMANN V. HELMHOLTZ-ZENTRUM vertieft. Im WS 2002/2003 ein Seminar „Geist & Körper – Programm & Maschine“ mit Prof. HORST WENZEL (HU Berlin Germanistik) und im SS 2003 zusammen mit Prof. JOCHEN BRÜNNING (HU Mathematik) ein Seminar „Diagrammatik“ veranstaltet, in beiden Fällen mit Unterstützung der HEINZ-NIXDORF-STIFTUNG und der HANNS-MARTIN-SCHLEYER-Stiftung veranstaltet. Diese disziplinübergreifenden Kooperationen unterschiedlicher Studiengänge wurden von den Studierenden gut aufgenommen.



Beim Kolloquium aus Anlaß seines 80. Geburtstages: Joseph Weizenbaum mit Wolfgang Coy und dem Vizepräsidenten der Humboldt-Universität Prof.Dr. Hans-Jürgen Prömel

Disziplinübergreifende Vernetzung

Die disziplinübergreifende Forschungs Kooperation wurde wie in den Vorjahren fortgesetzt. Mit der wachsenden Bedeutung der Informatik als Teil einer in Umrissen erkennbaren Informationsgesellschaft wird die Notwendigkeit einer Vernetzung mit anderen Sichten und Herangehensweisen in Forschung und Wissenschaft offensichtlich. Die langjährig in gemeinsamen Forschungsprojekten gewachsenen Bindungen zu den Kulturwissenschaften haben 1999 zur Gründung eines zentralen „HERMANN VON HELMHOLTZ-Zentrums für Kulturtechnik“ geführt, an dem die Arbeitsgruppe als Gründungsmitglied, neben Kulturwissenschaftlern, Kunstwissenschaftlern, Mathematikern, Altorientalisten und Philosophen beteiligt ist. Weitere interdisziplinären Zentren sollen künftig im Rahmen der Strukturreform der Universität verstärkt eingerichtet werden.

<http://www2.rz.hu-berlin.de/kulturtechnik/bsz/bsz.htm>

Dies hat zum Aufbau einer DFG-Forschergruppe „Bild-Schrift-Zahl“ geführt (Sprecher: WOLFGANG COY), die seit März 2001 gefördert wird.

<http://waste.informatik.hu-berlin.de/Forschung/>

Eine solche Kooperation besteht auch mit dem BMBF-Forschungsverbund „New Economy“ (Förderprogramm „Multimedia in der Lehre“), einem Forschungsverbund von ... Universitäten und Forschungsinstituten, dessen Geschäftsstelle an der Freien Universität (CEDIS, DR. APOSTOLOPOULOS).

Für die MULTIMEDIA-HOCHSCHULSERVICE GMBH (MHSB) der Berliner Hochschulen wurde eine Produktion zur Multimediatechnik erstellt.



Teleteaching zwischen Adlerhof und Berlin-Mitte

Lehre

URL: <http://waste.informatik.hu-berlin.de/Lehre/>

Veranstaltungen im Grundstudium

- Praktische Informatik I (W. COY WS 2002/2003)
- Übungen zu Praktische Informatik I (R. KUBICA, WS 2002/2003)
- Übungen zu Praktische Informatik I (P. BITTNER, WS 2002/2003)
- Informatik und Gesellschaft (W. COY, SS 2003)
- Unterrichtspraktikum (CH. DAHME, WS 2002/2003)
- Unterrichtspraktikum (CH. DAHME, SS 2003)
- Übungen zu Praktische Informatik I (R. KUBICA, WS 2003/2004)
- Unterrichtspraktikum (CH. DAHME, WS 2003/2004)

Kernveranstaltungen im Hauptstudium

- Informatik und Informationsgesellschaft II: Technik, Geschichte, Kontext (W. COY mit J.. KOUBEK, WS 2003/2004)
- Informatik und Informationsgesellschaft I: Digitale Medien (W. COY mit J.. KOUBEK, WS 2003/2004)

Spezialveranstaltungen

- Geschichte der Informatik (W. COY, WS 2002/2003)
- Kooperatives Prototyping (CH. DAHME, WS 2002/2003)
- Konzepte der frühen Phasen der Softwareentwicklung (CH. DAHME, SS 2003)
- Kooperatives Prototyping (CH. DAHME, WS 2003/2004)
- Zwischen „Information Sharing“ und „Information Control“ (P. BITTNER, WS 2003/2004)

Seminare

- 3D-Grafik (R. KUBICA WS 2002/2003)
- Internet Governance – Wer regiert das Internet? (P. BITTNER, Proseminar SS 2003)
- Bewußte evolutionäre Softwareentwicklung (CH. DAHME, SS 2003)
- 3-D Grafik (R. KUBICA, Proseminar SS 2003)
- Geschichte der frühen Informatik (W. COY, WS 2003/2004)
- 3D-Grafik (R. KUBICA WS 2003/2004)

Auswärtige Veranstaltungen

- Profilkurs Informatik an der OSZH, Berlin, Oktober bis Dezember 2003.
- PHP. Blockseminar, Universität Siegen (bei Prof. Dr. Peter Matussek), am 18. und 19.7.2003.



Humboldts und Helmholtz im Streik

Forschung

URL: <http://waste.informatik.hu-berlin.de/Forschung/>

Projekt: Projekt „Visuelle Argumentationen“

Ansprechpartner: PROF. DR. WOLFGANG COY

Beteiligte Mitarbeiter: DR. VOLKER GRASSMUCK

Zusammenarbeit: Hermann v. Helmholtz Zentrum für Kulturtechnik, DFG-Forschergruppe „Bild-Schrift-Zahl“

Forschungsförderung: DFG (seit 2001, zunächst bis 2003)

URL: <http://waste.informatik.hu-berlin.de/Forschung/bsz/>

Ein Kern wissenschaftlicher Tätigkeit besteht in der Überzeugungsarbeit: Eine Erkenntnis so aufzubereiten, daß andere ihr zustimmen - sie mit Argumenten zu vermitteln, die möglicher Kritik begegnen. In der *Gutenberg-Galaxis* der modernen Wissenschaft nimmt dies vor allem die Form wissenschaftlicher Texte an. Wissenschaftliche Texte sind in den Einzelwissenschaften weniger die Verkettung umgangssprachlicher Worte: sie werden wesentlich getragen von Fachsprachen, ergänzt durch formelhafte Notationen und durch Bilder. Idealtypisch kommt die korrekt angewandte Formelsprache der formalen Logik ohne weitere sprachliche oder bildliche Erläuterung aus, um aus Definitionen, Axiomen und Regelanwendungen neue Erkenntnisse in Form von Theoremen zu beweisen. Bilder werden dagegen in vielen Wissenschaften als eher nebensächliche, illustrative Elemente angesehen, deren logischer Status eher unsicher scheint, so dass sie in reinen Textwissenschaften wie etwa der Jurisprudenz, den Sprachwissenschaften oder der Theologie fast völlig fehlen. In der Mathematik und den Natur- und Technikwissenschaften haben Bilder jedoch eine essentielle Bedeutung behalten. Diese wird mit den modernen bildgebenden Verfahren, aber auch mit den digitalen Modellierungs- und Simulationsverfahren sogar weiter ausgebaut, so dass gelegentlich schon von einer *Science of Imaging* gesprochen wird. Untersucht wird die

argumentative Kraft wissenschaftlich-technischer Bilder sowie die Techniken ihres Einsatzes.

Veröffentlichungen

Viele der Veröffentlichungen können als PDF- oder HTML-Datereien unter URL: <http://Waste.Informatik.HU-Berlin.de/> kopiert werden.

L. HEMPEL, P. BITTNER: *Zur Evaluation von Videoüberwachung*. Deutsches Polizeiblatt für die Aus- und Fortbildung der Polizeibeamten (DPolBl), 1/2003, S. 27-32.

P. KLAUS BITTNER: *Informatik (anders) denken ... Über „gute“ Disziplinarität, Kritische Theorie und Informatik*. In: BÖHME, GERNOT; MANZEL, ALEXANDRA (HRSG.): *Kritische Theorie der Technik und der Natur*. München: Wilhelm Fink Verlag, 2003, S. 155-172.

P. BITTNER: *Unser aller Profession gib uns heute ... oder die Frage nach einer mäeutischen Informatik*. EMISA Forum, Jg. 23, Heft 2 (August 2003), S. 62-66. [Überarbeitete Fassung des Beitrages aus den Proceedings zum Interdisziplinären Symposium: Objekt-, subjekt oder handlungsorientiert? Perspektiven der Informatik, 10.-11.07.2003, Marburg/Lahn, S. 86-94].

AK VIDEOÜBERWACHUNG UND BÜRGERRECHTE: *Consultation of the Art. 29 Data Protection Working Party on Video Surveillance*. 30.05.2003; abgedruckt in: FIF-Kommunikation, 3/2003, S. 50-51. [Ko-Autor und Mitglied des Arbeitskreises].

P. BITTNER: *Videoüberwachung und Bürgerrechte*. Ein Jahr überregionaler Arbeitskreis. FIF-Kommunikation, 3/2003, S. 49-50.

P. BITTNER, E. HORNECKER, U. TWISSELMANN, K. WEBER: *Die berufliche Situation informatisch Handelnder in der ethischen Reflexion*. In: Dittrich, Klaus; König, Wolfgang; Oberweis, Andreas; Rannenber, Kai; Wahlster, Wolfgang (Hrsg.): *Informatik 2003 - Innovative Informatikanwendungen*. Beiträge der 33. Jahrestagung der Gesellschaft für Informatik e.V. (GI) 29.09.-02.10.2003 in Frankfurt am Main, Bd. 2, Bonn: Köllen Verlag [GI-Edition - Lecture Notes in Informatics (LNI), P-35], 2003, S. 176-181.

E. HORNECKER, P. BITTNER: *Informatik als Profession? Neue Wege für die Professionalisierungsdebatte*. In: DITTRICH, KLAUS; KÖNIG, WOLFGANG; OBERWEIS, ANDREAS; RANNENBERG, KAI; WAHLSTER, WOLFGANG (Hrsg.): *Informatik 2003 - Innovative Informatikanwendungen*. Beiträge der 33. Jahrestagung der Gesellschaft für Informatik e.V. (GI) 29.09.-02.10.2003 in Frankfurt am Main, Bd. 2, Bonn: Köllen Verlag [GI-Edition - Lecture Notes in Informatics (LNI), P-35], 2003, S. 182-187.

P. BITTNER: *Auf der Suche nach einer Patentlösung? Über Software und Patente*. FIF-Kommunikation, 4/2003, S. 3-5. [Editorial und Redaktion (mit SABRINA GEISSLER) des Themenschwerpunktes Softwarepatente].

P. BITTNER: *Rezension zu Müller-Heidelberg, Till u.a. : Grundrechte-Report 2003. Zur Lage der Bürger- und Menschenrechte in Deutschland*. FIF-Kommunikation, 4/2003, S. 25.

AK INFORMATIK UND VERANTWORTUNG: *Ethische Leitlinien der GI*. Entwurf des Arbeitskreises „Verantwortung“. *Informatik-Spektrum*, Band 26, Heft 6 (Dezember 2003), S. 418-422. [Ko-Autor und Mitglied des Arbeitskreises].

W. COY: *Eure Rede aber sei: 0, 0; 1, 1 – Vom Detail bei der Digitalisierung*, in W. SCHÄFFNER, S. WEIGEL, TH. MACHO, *Der liebe Gott steckt im Detail*, München: Fink 2003.

- W. COY: *Candidate Event – Wahrheit und Lüge im wissenschaftlichen Sinn*. In PETER MATUSSEK, NATASCHA ADAMOWSKY (Hg.), *Auslassungen – Festschrift für Hartmut Böhme*, Würzburg: Königshausen & Neumann 2004, (im Satz).
- W. COY: *Digital/Analog - Zur Klärung eines Sachverhaltes*. In M. WARNKE, W. COY, C. THOLEN, *HyperKult 2 - Virtuelles & Digitales*, (in Vorbereitung für 2004).
- W. COY: *Hardware, Software, Content*. In: *FAktisch – Festschrift zum 60. Geburtstag von Friedrich Kittler*, München: Fink-Verlag 2003.
- W. COY: *Was ist Informatik? Zur Entstehung des Faches an deutschen Universitäten*, in HANS DIETER HELDIGE, *Geschichten der Informatik*, Berlin-Heidelberg-New York: Springer 2004.
- W. COY, *Argumentatives Nachdenken*. Zeitschrift für Germanistik, Neue Folge 3, 2003, S.551-559
- W. COY: Von Gutenberg zu www.gutenberg.net, in: ULRICH SCHMITZ UND HORST WENZEL, *Wissen und neue Medien - Bilder und Zeichen von 800 bis 2000*, Berlin: Erich Schmidt Verlag, 2003.
- W. COY: *Eine Einheit der Wissenschaften - Die Wahrheit ist zwischen den Disziplinen versteckt*. DER TAGESSPIEGEL 14.4.2003, Nr.18092, S.B1).
- W. COY: Die Konstruktion technischer Bilder – eine Einheit von Bild, Schrift und Zahl, in SYBILLE KRÄMER, HORST BREDEKAMP, *Bild-Schrift-Zahl*, München: Fink, 2003.
- W. COY: *Informationsgesellschaft - Betaversion 0.9*. In D.KLUMPP, H.KUBICEK, H.ROSSNAGEL U.A.: *Jahrbuch Telekommunikation*, Stuttgart:Thalheimer 2003.
- W. COY: *Analog/Digital*. In DETLEV SCHÖTTKER, *Mediengebrauch und Erfahrungswandel*, Göttingen: Vandenhoeck & Ruprecht, 2003.
- W. COY: *Die protestantische Ethik und der „Geist“ der Informatik*. In Karl-Heinz Rödiger, *Algorithmik–Kunst–Semiotik*, Heidelberg: Synchron, 2003.
- AK INFORMATIK UND VERANTWORTUNG: Ethische Leitlinien der GI. Entwurf des Arbeitskreises „Verantwortung“. Informatik-Spektrum, Band 26, Heft 6 (Dezember 2003), S. 418-422. [Ko-Autor und Mitglied des Arbeitskreises].
- W. COY: *Eine Intellektueller als Wissenschaftler verkleidet*. In "...Joseph Weizenbaums zu ehren...", Stuttgart, 2003 (im Druck)
- W. COY: *Brauchen wir eine Disziplin Medienwissenschaft?* Beitrag zu einer Podiumsdiskussion auf der Jahrestagung Medienwissenschaften *Konzeptionen der Medienwissenschaft*, Siegen 21.-23.11.02 (im Satz).
- CH. DAHME: Activity Theory as a Basis for the Design of Useful Software. In: B. VIEHWEGER (ed.): *Perspectives in Business Informatics Research, Proceedings of the BIR-2003-Conference*, September 18-20, 2003, Berlin , Germany, SHAKER Verlag, pp. 113-119.
- CH. DAHME: Historische, wissenschaftstheoretische und kulturelle Wurzeln der angewandten Informatik - Versuch einer Bestimmung in Form von Thesen. In: NAKE, FRIEDER; ROLF, ARNO; SIEFKES, DIRK: *Wozu Informatik? Theorie zwischen Ideologie, Utopie und Phantasie*. Materialien zu einer Arbeitstagung in Bad Hersfeld März 2002, Berlin: TU Berlin [Forschungsberichte der Fakultät IV - Elektrotechnik und Informatik, Bericht 2002-25], S. 50-53.

VOLKER GRASSMUCK, *Der zweite Korb dient der Allgemeinheit!*, (Panelbeitrag auf "Urheberrecht in der Informationsgesellschaft Auftakt zum zweiten Korb", München 16.10.2003), *ZUM - Zeitschrift für Urheber- und Medienrecht*, 12/03.

JOCHEN KOUBEK, *Vernetzung als kulturelles Paradigma*. Dissertation 2003.



*Der GI-Arbeitskreis „Informatik und Verantwortung“
bei der Erarbeitung der Neufassung der „Ethischen Leitlinien“ im Herbst 2003 in der
Humboldt-Universität*

Vorträge

P. BITTNER: *Informationelle Selbstbestimmung*. Vortrag im Rahmen der Projektwoche „Neue Medien“ der Anna-Seghers-Oberschule, Berlin-Adlershof, 28.01.2003

P. BITTNER: *Des Menschen technisches Bild: Biometrie und Gesichtserkennung*. Oberseminar „Diagramme und technische Bilder“, Reichenow, 17.05.2003

P. BITTNER: *Anonymität im Netz?!* Vortrag im Rahmen des Profilkurses „Kryptologie“ der Anna-Seghers-Oberschule, Berlin-Adlershof, 18.06.2003

P. BITTNER: *Unser aller Profession gib uns heute ... oder die Frage nach einer mäeutischen Informatik*. Interdisziplinäres Symposium „Objekt-, subjekt oder handlungsorientiert? Perspektiven der Informatik“, Marburg/Lahn, 11.07.2003

P. BITTNER: *Informatik als Profession? Neue Wege für die Professionalisierungsdebatte*. 33. Jahrestagung der GI, Workshop „Ethik in der Informatik“, Frankfurt/Main, 02.10.2003. [gemeinsamer Beitrag mit EVA HORNECKER]

P. BITTNER: *e-identity: Wenn der Mensch vermessen zur Information wird!*. 19. FIF-Jahrestagung „Die e-lektrisierte Gesellschaft“, Beitrag im gleichnamigen Workshop, Bad Hersfeld, 22.11.2003

W. COY: *Ein Intellektueller als Informatiker verkleidet*, Workushop „...Joseph Weizenbaum zu Ehren...“. Humboldt-Universität zu Berlin, 10.1.03

W. COY: *Multimediatechnik als universitäre Dienstleistung*. DINI-Jahrestagung, Berlin, 27.3.03

- W. COY: „*Ob aber jemand was verstanden hat, kann man nicht sehen*“. Universität Linz, 29.3.03
- W. COY: *On Memory, Storage, and Computers*. Workshop „The Renaissance of the Theatre of Memory“ im SFB „Medien des Performativen“, FU Berlin 11.4.03
- W. COY: *Lernen mit Multimedia*. Multimediahochschultage der HU Berlin; Veranstaltung der Multimediahochschulservice GmbH MHS, Berlin, 27.6.03
- W. COY: *Über Karl v. Frischs „Die Sprache der Bienen*“. Lange Nacht des wissenschaftlichen Films, Humboldt-Universität zu Berlin, 25.6.03
- W. COY: *Analog/Digital – Zur Klärung eines Sachverhaltes*, Hyperkult XII, Universität Lüneburg, 25.7.03
- W. COY: *Geistiges Eigentum – einige Nebenbemerkungen*. Tagung der Grünen Akademie der Heinrich-Böll-Stiftung, Kloster Neustift, Brixen/Südtirol, 28.7.03
- W. COY: *Communicating Insights Visually*. International University of Bremen, 9.9.03
- W. COY: Diskussionsleitung auf dem Workshop WS 6 „*Die Zukünfte des Internet: eine Retrospektive*“. GI-Jahrestagung, Frankfurt am Main, 2.10.2003
- W. COY: *Bleibt die Informationsgesellschaft bloß Vaporware?* Festvortrag aus Anlaß des 25jährigen Bestehens des Studiengangs Informatik der Universität Bremen, 10.10.2003
- W. COY: „*Was wir wissen, wissen wir aus den Massenmedien...*“. Internationale INST-Tagung „Das Verbindende der Kulturen“, Wien, 7.11.03
- W. COY: *Die von-Neumann-Architektur*. Tagung 100 Jahre Johann von Neumann 1903-2003, Humboldt-Universität zu Berlin, 3.12.03
- W. COY: *Zum Streit der Fakultäten*. 50 Jahre Informationstechnische Gesellschaft - ITG, Museum für Kommunikation, Berlin, 19.12.03
- V. GRASSMUCK: *DRM und neues Urheberrecht für Forschung und Lehre*. Multimedia Forum, Demoraum des Computer- und Medienservice, HU Berlin, 5.2.03
- V. GRASSMUCK: Panel „*Eigentum*“, mit Mercedes Bunz und Sebastian Luetgert, mod. Pit Schultz und Stefan Heidenreich, 4 Gespräche veranstaltet vom Lehrstuhl für Ästhetik, Kulturwissenschaften, Humboldt-Universität und Bootlab e.V., Raum3, Berlin, 5.2.03
- V. GRASSMUCK: *DRM*. Kollegiatentag der SEL Alcatel Stiftung im Hotel Schloss Reichenow, 7.-8.2.03
- V. GRASSMUCK: *Urheberrecht: Geschichte, Stand und Ausblick*. Auf Bundeskongress Grüne Jugend, Paulinum Gymnasium, Münster, 8.3.03
- V. GRASSMUCK: *Für einen Wissensumweltschutz*. Auf: Symposium „Material Re Material (remix & copyright)“ auf Maerz-Musik, Berliner Festspiele, 22.3.03
- V. GRASSMUCK: *Technological Copyright Enforcement and Open Access* Auf: Open Archives Forum Workshop, HU Berlin, 29.3.03
- V. GRASSMUCK: *Japans Westen*. Lesung aus „Geschlossene Gesellschaft“, Start der Reihe „New West“, German Theater Abroad (GTA), Schröderstr.1, Berlin, 3.4.03
- V. GRASSMUCK: Panel-Statement auf „*Trusted Computing*“, org. Zentrum für Europäische Integrationsforschung, Rheinische FriedrichWilhelmsUniversität Bonn, 9.5.03

- V. GRASSMUCK: *Scanning the Body*. CTs als Diagramme, auf Bild-Schrift-Zahl Blockseminar Diagramme, Schloß Reichenow, 17.5.03
- V. GRASSMUCK: *Der tote Autor und die konnektive Intelligenz*. Auf Originalkopie. Praktiken des Sekundären, SFB Medien und kulturelle Kommunikation, Uni Köln, 22.5.03
- V. GRASSMUCK: *Rechtekontrolle für mobile Medienanwendungen*. In „Entwicklungsperspektiven mobiler Medienkommunikation“, Vortrags- und Case Study-Seminar der Alcatel SEL Stiftung, org. Valerie Feldmann, FU Berlin, 16.6.03
- V. GRASSMUCK: *Wider den digitalen Stacheldraht. Digital Restrictions Management und die Alternativen*. Im Rahmen der Kritischen Uni, org. Stefan Merten, TU Kaiserslautern, 26.6.03
- V. GRASSMUCK: *Datenherren contra freie Information*. Im Rahmen von tell.net, org. Johannes Auer, Stadtbücherei Stuttgart, 27.6.03
- V. GRASSMUCK: *Reputation in Freier Software und Wissenschaft*. Auf „Kathedrale und Bazar: Das Phänomen freie Software“, org. Alexander von Gernler, TheodorHeussAkademie, Gummersbach, 28.6.03
- V. GRASSMUCK: *PC => TC*. Vortrag im Workshop „Medienwirtschaft, DRM und TCG“ auf dem Symposium „Trusted Computing Group“ (TCG), Bundesministerium für Wirtschaft und Arbeit, Berlin, 3.7.03
- V. GRASSMUCK: *Reputation in Wissenschaft und freier Software*. Im Rahmen des Forum Transmedien an der Hochschule für bildende Künste Hamburg, 11.7.03
- V. GRASSMUCK: *Der zweite Korb dient der Allgemeinheit!* Panelbeitrag auf "Urheberrecht in der Informationsgesellschaft Auftakt zum zweiten Korb", BMJ und Institut für Urheberrecht, Literaturhaus München, 16.9.03
- V. GRASSMUCK: *Urheberrechtliche Fragen des digitalen Publizierens - Gegenwart und Ausblick*. FIGARO-Workshop, Gästehaus der Universität Hamburg, 23.-24.9.03
- V. GRASSMUCK: *Entwicklung des Internet von einer offenen Wissensgesellschaft zur geschlossenen Copyright-Gesellschaft*. Trusted Computing und Digital Restrictions Management, auf WS 6 „Die Zukünfte des Internet: eine Retrospektive“, GI Jahrestagung, Frankfurt am Main, 2.10.2003
- J. KOUBEK: *Organisationskultur*. Disputationsvortrag , 10.2.2003
- J. KOUBEK: *Visuelle Strategien der Betriebswirtschaftslehre*. Schloß Reichenow, 18.05.2003
- J. KOUBEK: *Die Technik der Analog/Digital-Wandlung*, Hyperkult XIII, 24.7.03
- J. KOUBEK, C. KURZ, J.-M. LOEBEL, H.-G. KUPER, J. KUNZE, H. KESPER, C. KURZ, D. APELT, J. BÖTTCHER: *John von Neumann. Ein Audio-Feature* im Rahmen der Tagung „100 Jahre Johann von Neumann 1903-2003,“ 1.12.2003.
- R. KUBICA: *Die Skalierbarkeit virtueller Welten*. Seminar „Informatik und Gesellschaft“, Lanke 19.12.03



*Öffentliche Vorlesung von Peter Bittner zur Videoüberwachung
während des Streiks im Wintersemester 2003/04*

Sonstige Aktivitäten

Dipl.-Inf. Peter Bittner

- Mitglied der Bibliothekskommission des Instituts
- Stellvertretender Vorsitzender des Forum InformatikerInnen für Frieden und gesellschaftliche Verantwortung (FIF) e.V.
- Fachexperte im Leitungsgremium des Fachbereichs „Informatik und Gesellschaft“ (FB IuG) der Gesellschaft für Informatik (GI)
- Mitglied im Arbeitskreis „Informatik und Verantwortung“ des FB 8 der GI
- Mitbegründer des bundesweiten Arbeitskreises „Videoüberwachung und Bürgerrechte“
- Mitglied der Arbeitsgruppe Informations- und Kommunikations-Technologie zum Wiederaufbau Afghanistans (AITA)
- Mitarbeit in den Netz-Initiativen *privatkopie.net* und *informationsfreiheiten.de*
- Mitbegründer und Reihenherausgeber (gemeinsam mit Jens Woinowski, München) der Buchreihe „Kritische Informatik“ beim LIT-Verlag
- Assoziiertes Mitglied des Graduiertenkollegs „Technisierung und Gesellschaft“ des FB Gesellschafts- und Geschichtswissenschaften der Technischen Universität Darmstadt
- Ehrenmitglied der Fachschaft Informatik der Universität Kaiserslautern

Prof. Dr. Wolfgang Coy

- Mitglied des *Konzils der Humboldt-Universität*
- Vorsitzender der zentralen *Medienkommission des Akademischen Senats*
- *Informations- und Kommunikationsbeauftragter* der Humboldt-Universität
- Mitglied im *Institutsrat des Instituts für Informatik*
- Sprecher der *DFG-Forschergruppe „Bild-Schrift-Zahl“* im Hermann v. Helmholtz-Zentrum der Humboldt-Universität zu Berlin
- Deutscher Vertreter im Technical Committee 9 ›Computers and Society‹ der *International Federation of Information Processing Societies (IFIP)*
- Fellow der *Gesellschaft für Informatik*
- Mitglied der *Grünen Akademie* der Heinrich-Böll-Stiftung, Sektion „Zukunftstechnologien“

- Mitglied im wissenschaftlichen Beirat des *Forum InformatikerInnen für Frieden und gesellschaftliche Verantwortung* (FIF)
- Mitglied des Vorstands des *Berliner Studienverbundkollegs* der Alcatel-Stiftung
- Mitglied im Herausgebergremium des *Informatik-Spektrum* (Springer Verlag)
- Projektteilnehmer am „Projekt Zukunft – Der Berliner Weg in die Informationsgesellschaft“ des Berliner Senats
- Organisator des *Multimedia-Arbeitskreises* der Humboldt-Universität
- Fachexperte des Fachbereichs FB8 „Informatik und Gesellschaft“ der *Gesellschaft für Informatik*
- Mitglied im Arbeitskreis „Informatik und Verantwortung“ des FB 8 der GI
- Mitglied im Programmkomitee des GI-Workshops *HyperKult XII*, Universität Lüneburg
- Gutachter des Hochschulwettbewerbs „Digital Sparks“ der FhG-GMD St. Augustin
- *Gutachtertätigkeiten* u. a. für die Deutsche Forschungsgemeinschaft (DFG), den Fond zur Förderung der wissenschaftlichen Forschung (FWF), Wien und das Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (BMBF), Bonn.

Dr. Christian Dahme

- Mitglied des Akademischen Senats
- Mitglied des Konzils
- Mitglied des Konzilvorstandes
- Prodekan für Lehre und Studium – Studiendekan der Mathematisch-Naturwissenschaftlichen Fakultät II
- Mitglied des Fakultätsrates der Mathematisch-Naturwissenschaftlichen Fakultät II
- Mitglied der Kommission Lehre und Studium des Akademischen Senats
- Stellvertretender Vorsitzender der Kommission Lehre und Studium des Akademischen Senats
- Mitglied der gemeinsamen Kommission aus EPK, LSK und FNK im Rahmen der Strukturplanung
- Mitglied der Arbeitsgruppe zur Reform der Lehramtsausbildung in Informatik

Dr. Volker Grassmuck

- Mitglied des Expertenbeirats des Bundesministeriums für Wirtschaft und Technologie für Open Source Software

Dipl.-Inf. Roland Kubica

- *Referenzimplementierung* einer navigierbaren Foto-realistischen VR-Umgebung; informelle Forschungsk Kooperation mit der UK der Bauhaus Universität Weimar

Dr. Jochen Koubek

- Mitglied der Arbeitsgruppe zur Reform der Lehramtsausbildung in Informatik

Constanze Kurz

- Mitglied im Arbeitskreis „Informatik und Verantwortung“ des FB 8 der GI

Diplomarbeiten, Dissertationen

JOCHEN KOUBEK: *Vernetzung als kulturelles Paradigma*. Dissertation im Studiengang Kulturwissenschaften, 2003

HENDRIK PANTLE: *Kulturtechnische Aspekte interaktiven Fernsehens und MHP-Authoringtools*. Magisterarbeit im Studiengang Kulturwissenschaften, 2003

TORSTEN OETTEL: *Informationstechnologie in der Entwicklungszusammenarbeit. Konzeption eines nachhaltigen Entwicklungsprojektes zur Etablierung eines informationstechnischen Ausbildungszentrums in „Ban Sane Vang Gnay“, Volksrepublik Laos.* Diplomarbeit im Studiengang Informatik, 2003

DANIEL OHST: *Einsatz elektronischer Signaturen und Zeitstempel für die Sicherung digitaler Dokumente.* Diplomarbeit im Studiengang Informatik, 2003

Lehr- und Forschungseinheit
Parallele und Verteilte Systeme

<http://www.zib.de/csr/>

Leiter

PROF. DR. ALEXANDER REINEFELD

Tel.: (030) 84185-130

E-Mail: ar@zib.de

Sekretariat

PETRA FEHLHAUER

Tel.: (030) 84185-131

E-Mail: fehlhauer@zib.de

Wissenschaftliche Mitarbeiter

DR. ARTUR ANDRZEJAK

SEBASTIAN HEIDL

FELIX HUPFELD

PATRICK MAY

THOMAS RÖBLITZ

FLORIAN SCHINTKE

THORSTEN SCHÜTT

DR. THOMAS STEINKE

DR. JAN WENDLER

Studentische Mitarbeiter

GUILLERMO ARIAS DEL RIO

JANIS BARDINS

MEHMET CEYRAN

MINOR GORDON

RENÉ HEEK

THILO KOCH

JÖRG MELTZER

MONIKA MOSER

KARL PAULS

SIMON RIECHE

JOCHEN WITTE

Die wissenschaftliche Arbeit des Lehrstuhls konzentriert sich im wesentlichen auf zwei Schwerpunkte: die Entwicklung effizienter Systemsoftware für innovative Cluster-Systeme und die Konzeption und Implementation von Methoden zur koordinierten Nutzung verteilter Computerressourcen. Der Lehrstuhlinhaber leitet zugleich den Bereich Computer Science (mit Hochleistungsrechenzentrum) am *Zuse-Institut Berlin (ZIB)*. Dadurch können theoretische Erkenntnisse schnell in die Praxis umgesetzt und Erfahrungen des wissenschaftlichen Hochleistungsrechnens auf direktem Wege in den Lehr- und Forschungsbetrieb der Humboldt-Universität eingebracht werden. Die Lehrveranstaltungen und Forschungsprojekte finden institutsübergreifend an den Standorten HU/Adlershof und ZIB/Dahlem statt.

Das Zuse-Institut Berlin (www.zib.de) ist eine außeruniversitäre Forschungseinrichtung des Landes Berlin auf dem Gebiet der angewandten Mathematik und Informatik. Es betreibt parallele Hochleistungsrechner (Cray T3E und IBM p690) und bietet den Nutzern Fachberatung zur effektiven Nutzung der Rechnersysteme an. Ein unabhängiger wissenschaftlicher Ausschuss entscheidet über die Vergabe der Rechenzeit, die ausschließlich Spitzenforschern im Berliner und norddeutschen Raum zugute kommt. Derzeit betreibt das ZIB im Rahmen des *Norddeutschen Verbunds für Hoch- und Höchstleistungsrechnen (HLRN)* einen parallelen Hochleistungsrechner p690 der Firma IBM, der mit einer Spitzenleistung von über 4 TeraFlop/s zu den leistungsfähigsten in Deutschland und Europa gehört. Mit seinen zwei Standorten in Berlin und Hannover ist dieser Rechner zudem der erste verteilte Supercomputer in Deutschland. Technische Einzelheiten des HLRN-Systems stehen im Web unter www.hlrn.de.

Der Lehrstuhl Parallele und Verteilte Systeme deckt ein breites Forschungsspektrum ab, das von systemnahen Arbeiten in Clustern bis zur Konzeption von Grid Systemen für die kooperative Nutzung geographisch verteilter Computer reicht. Das verbindende Element zwischen diesen beiden Themenbereichen liegt in der effizienten Nutzung der zugrunde liegenden Hardware- und Softwarekomponenten für verteilte Berechnungen. Auf den untersten Schichten aufbauend entwickeln wir höherwertige Dienste, die den effizienten und benutzerfreundlichen Einsatz paralleler und verteilter Systeme ermöglichen.

Im Forschungsschwerpunkt Grid Computing liegt unser Fokus auf der Entwicklung von Methoden zum Management geographisch verteilter Ressourcen. Dabei umfassen Ressourcen nicht nur Rechner, sondern auch verteilte Datei- und Archivsysteme, Netzwerke, Visualisierungskomponenten, Softwarepakete und beliebige Datenquellen. Unser Ziel ist es, diese Ressourcen in einem überregionalen Lastverbund verfügbar zu machen, so dass die Nutzer möglichst einfach darauf zugreifen können. Das Gesamtsystem soll sich so verhalten, als wären alle Teilkomponenten direkt am Arbeitsplatz verfügbar. Auf diese Weise werden den Nutzern über die gewohnten Zugriffsmechanismen qualitativ höherwertige Dienste zur Verfügung gestellt und Systembetreiber erzielen eine höhere Auslastung ihrer kostenintensiven Hard- und Software.

Ein wichtiges Teilprojekt unserer Arbeitsgruppe ist die zuverlässige, sichere und effiziente Speicherung großer Datenmengen in verteilten Grid-Systemen. Da Speicherknoten im Grid jederzeit (transient oder permanent) ausfallen können, ist eine geeignete Datenreplikation mit effizienter Datenplatzierung notwendig. Wir entwickeln Methoden für Caching, Staging, Routing, Scheduling und Synchronisation.

Zum praktischen Test unserer Grid-Methoden und Algorithmen für das Datenmanagement dienen komplexe Anwendungen der Bioinformatik, die gleichfalls in unserer Gruppe entwickelt werden. Die im Lehrstuhl angesiedelte Forschergruppe „Alignment und Threading auf parallelen Systemen“ des Berliner Centrums für genombasierte Bioinformatik (BCB) konzipiert Algorithmen zur Strukturvorhersage von Proteinen, die nicht nur verteilt im Grid ablaufen, sondern auch auf sehr umfangreiche Datenmengen zugreifen.

Lehre

Vorlesung

- IA-64 versus Power4: Architektur und effiziente Programmierung (PROF. DR. ALEXANDER REINEFELD, FELIX HUPFELD, THORSTEN SCHÜTT, WS 2003/04)
- Verteilte Algorithmen (DR. ARTUR ANDRZEJAK, WS 2003/04)

- Proteine: Klassifikation und Strukturvorhersage, Beitrag zur Vorlesung „Algorithmische Bioinformatik“, FU Berlin (DR. THOMAS STEINKE, WS 2002/03)

Seminare

- Verteilte Datenstrukturen in Peer-to-Peer-Systemen (DR. ARTUR ANDRZEJAK, PROF. DR. ALEXANDER REINEFELD, WS 2003/04)
- Parallele Algorithmen in der Bioinformatik (PROF. DR. ALEXANDER REINEFELD, DR. THOMAS STEINKE, SS 2003)
- Forschungsseminar Grid Computing (PROF. DR. ALEXANDER REINEFELD, SS 2003)
- Informationsverwaltung im Web (DR. ARTUR ANDRZEJAK, PROF. HEINZ F. SCHWEPPE (FU Berlin), SS 2003)

Softwarepraktikum

- Visualisierung, Modellierung und Simulation von Biomolekülen (DR. THOMAS STEINKE U.A., HUB/Charite, SS 2003)

Forschung

Projekt: GridLab – A Grid Application Toolkit and Testbed for a Grid Laboratory

Ansprechpartner: THORSTEN SCHÜTT

Beteiligte Mitarbeiter: MINOR GORDON, ANDREI HUTANU, FELIX HUPFELD, THILO KOCH, ANDRÉ MERZKY, PROF. DR. ALEXANDER REINEFELD, SIMON RIECHE, FLORIAN SCHINTKE, DR. BRYGG ULLMER, STEFAN ZWIERLEIN

Zusammenarbeit: Partner aus dem akad. Bereich: PSNC (PL), AEI (D), ZIB (D), Masaryk University (CZ), MTA SZTAKI (HU), VU (NL), ISUFI/HPCC (I), Cardiff University (GB), NTUA (GR), University of Chicago (USA), ISI (USA), University of Wisconsin (USA); Industriepartner: SUN, COMPAQ.

Forschungsförderung: Europäische Union

Um Anwendungsprogrammierern den Einstieg in das Grid Computing zu erleichtern, werden Werkzeuge benötigt, die ihnen bei der effizienten Einbettung ihrer komplexen Anwendungen in Grid-Umgebungen helfen. Im europäischen Projekt GridLab (www.gridlab.org) entwickeln wir eine „Middleware“, die als Zwischenschicht zwischen dem Anwendercode und der darunter liegenden Grid-Umgebung vermittelt. Diese Middleware soll stellvertretend für die Anwendung entscheiden, auf welche Dienste in den jeweiligen Grid-Umgebungen zugegriffen wird. Die Entscheidung soll vollständig autonom in der Middleware gefällt werden, damit die Anwendung ohne Modifikation – ja, sogar ohne genaue Kenntnis der Laufzeitumgebung – auf nahezu beliebigen Grid-Umgebungen ablaufen kann.

Das Funktionieren und die Effizienz der GridLab-Middleware soll beispielhaft anhand zweier Anwendungen demonstriert werden: CACTUS, ein Code zur Simulation kollidierender Neutronensterne, und TRIANA, ein Analysewerkzeug zur Auswertung der im GEO 600-Experiment gesammelten Daten, die der empirischen Verifikation der Einsteinschen Raumzeitkrümmung dienen sollen. Die in Hannover gewonnenen GEO 600-Daten werden über eine Weitverkehrsverbindung täglich ans ZIB geliefert und im Archiv für die spätere Offline-Auswertung gespeichert. Beide Anwendungen dienen auch der Erprobung komplexer Szenarien, wie z.B. der Migration einer laufenden GridLab-Anwendung von einem Rechner auf den nächsten, ohne dass die Anwendung davon Kenntnis nehmen muss oder gar der Quellcode geändert werden müsste.

Im GridLab-Projekt leiten wir das Arbeitspaket „Data Handling and Visualization“. Im Berichtsjahr haben wir einen webbasierten Visualisierungsserver, eine attributbasierte Dateiverwaltungskomponente und ein Tool zum effizienten Zugriff auf entfernte Daten entwickelt und implementiert. Der Visualisierungsserver erlaubt es Forschern, die Ergebnisse von entfernt laufenden Jobs komfortabel graphisch darzustellen. Typischerweise werden dafür nur Teilmengen der entfernten Daten verwendet. Unser Tool kann sehr effizient fast beliebige Teilmengen aus entfernten Dateien extrahieren. Die Dateiverwaltungskomponente setzt auf dem Replikatkatalog auf, den wir ebenfalls im Rahmen dieses Projekts entwickelt haben. Zusammen bilden diese eine Abstraktionsschicht, die vom Speicherort der Dateien abstrahiert und eine ausdrucksstarke Benennung der Dateien erlaubt.

Projekt: DataGrid – Research and Technological Development for an International Data Grid

Ansprechpartner: THOMAS RÖBLITZ

Beteiligte Mitarbeiter: JANIS BARDINS, JÖRG MELTZER, KARL PAULS, PROF. DR. ALEXANDER REINEFELD, FLORIAN SCHINTKE

Zusammenarbeit: CERN (CH), PPARC (UK), INFN (I), CNRS (F), NIKHEF (NL), ESA/ESRIN (I); IBM (UK), Compagnie des Signaux (F), Datamat (I); Helsinki Institute of Physics and CSC (SF), Swedish Natural Science Research Council (SE), Instituto Trentino di Cultura (I), KIP Heidelberg (D), CEA/DAPNIA (F), IFAE Barcelona, CNR (I), CESNET (CZ), KNMI (NL), SARA (NL), SZTAKI (HU)

Forschungsförderung: Europäische Union

Im DataGrid-Projekt (www.eu-datagrid.org) sind unter Federführung des CERN mehr als 170 Wissenschaftler aus 21 Institutionen beschäftigt. Ziel des Projektes ist es, für die nächste Generation von Kernteilchen-Experimenten auf dem Large Hadron Collider (LHC) des Cern die zur Datenauswertung erforderliche Rechnerinfrastruktur zu entwickeln und den Physikern zur Verfügung zu stellen. Die prognostizierte Datenmenge von mehreren Petabyte (10^{15} Byte) pro Jahr ist derart hoch, dass selbst die weltweit leistungsstärksten Supercomputer zur Auswertung der Daten nicht mehr ausreichen. Zur Lösung dieser internationalen Herausforderung werden einige tausend handelsübliche PCs zu Clustern verbunden und in sogenannten „Tier 1 Regionalzentren“ für die Datenauswertung betrieben. Zusammen mit den Clustern der kleineren „Tier 2 Zentren“ werden ab dem Jahr 2006/07 weltweit insgesamt ca. 50.000 PCs für die Datenanalyse zur Verfügung stehen. Alle Regionalzentren sind über eine Grid-Software miteinander verbunden, so dass die Jobs der Physiker an beliebigen Orten gerechnet werden können.

Unsere Aufgabe im DataGrid-Projekt liegt in der Integration der lokalen Cluster in das Grid. In unserer Architektur steuert ein leistungsfähiger Scheduler, der Maui-Scheduler, das Starten und Beenden von Benutzerjobs über eine neu entwickelte Softwareschicht. Diese bietet den Vorteil, dass verschiedene Ressourcen-Managementsysteme leicht integriert werden können. Darüber hinaus stellt sie eine einheitliche Schnittstelle für die Planung und Ausführung administrativer Aufgaben zur Verfügung, um z. B. das virtuelle An- und Abschalten von Rechenknoten für geplante Wartungsarbeiten in die Job-Abfolge einzubringen. Im Berichtszeitraum haben wir unsere Software um eine Schnittstelle für die Kopplung mit dem Accounting-System des DataGrid-Projektes erweitert. Unsere Ergebnisse haben wir in einem Beitrag auf der Konferenz Cluster 2003 vorgestellt.

Ein Problem besteht in dem automatischen Auffinden verfügbarer und passender Ressourcen für die Ausführung neuer Jobs. Die im DataGrid-Projekt entwickelten „Information Provider“ haben wir an das sog. GLUE-Schema angepasst. Dieses Schema wurde in

Kooperation verschiedener internationaler Grid-Projekte entwickelt und soll die Benutzung weltweiter Ressourcen vereinfachen.

Projekt: FlowGrid – Flow Simulations On-Demand Using Grid Computing

Ansprechpartner: DR. JAN WENDLER

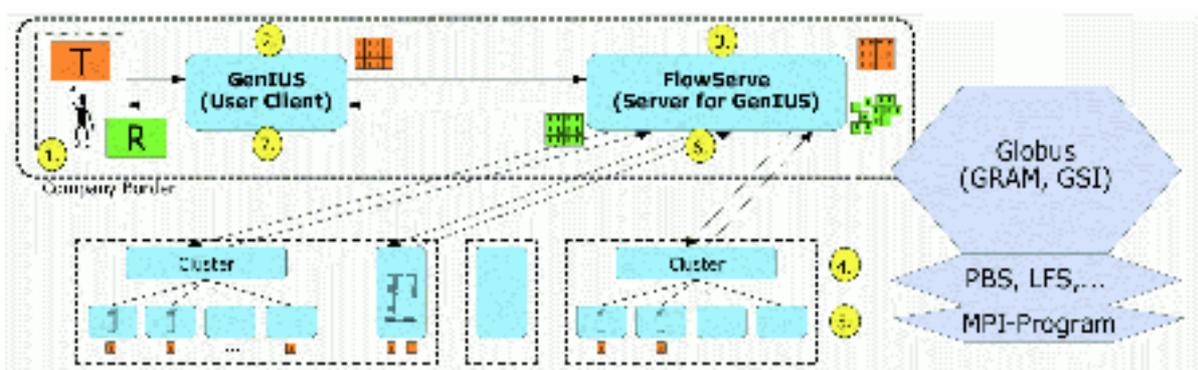
Beteiligte Mitarbeiter: DR. ARTUR ANDRZEJAK, GUILLERMO ARIAS DEL RÍO, PROF. DR. ALEXANDER REINEFELD, FLORIAN SCHINTKE

Zusammenarbeit: University of Zaragoza (E), CERTH/HIT (EL), SKODA (CZ), HSVA (D), SYMBAN (UK)

Forschungsförderung: Europäische Union

Im europäischen FlowGrid-Projekt (www.unizar.es/flowgrid) entsteht eine virtuelle Organisation für die Durchführung von Strömungssimulationsrechnungen, um damit eine gemeinsame Nutzung von Software, Computerressourcen und Wissen zu ermöglichen. Dazu wird mit Hilfe von Grid-Infrastruktur ein Netzwerk von Anbietern von Rechenkapazität, Programmeigentümern und Nutzern eingerichtet. Die Vorbereitung, Ausführung und Überwachung der Simulationen sind sowohl räumlich als auch organisatorisch verteilt. Das FlowGrid-System wird für die Durchführung anspruchsvoller Simulationen in den Bereichen Schiffbau, Autoabgase und Verbrennung fossiler Brennstoffe durch die Projektpartner eingesetzt.

Innerhalb des Projektes ist unsere Forschungsgruppe für die Entwicklung und Bereitstellung der Middleware, das sog. „FlowServe“, verantwortlich. Im Berichtszeitraum wurde ein voll funktionierender FlowServe-Prototyp entwickelt. FlowServe besitzt Schnittstellen zum Benutzerprogramm GenIUS, zu den Ressourcen-Managern und zur zentralen Datenbank DBMS. Das GenIUS-Benutzerinterface (oder andere Benutzerprogramme) übergibt FlowServe den Auftrag zur Durchführung einer Simulation. FlowServe verteilt die Simulation auf geeignete Cluster, überwacht dessen Ausführung und generiert Status- und Fortschrittsberichte. Die folgende Abbildung zeigt die FlowGrid-Software mit seinen Komponenten.



FlowGrid-Systemarchitektur

Unser Hauptanliegen in diesem Projekt ist die Erforschung und Entwicklung von Grid-Umgebungen, die einen benutzerfreundlichen und effizienten Zugriff auf geographisch verteilte Ressourcen ermöglichen. Als Spezialfall ist in FlowServe auch die gleichzeitige Ausführung gekoppelter Anwendungen auf verschiedenen Systemen möglich. Mit den Ergebnissen dieses Projektes werden Anwender (nicht nur der Strömungssimulation) in die

Lage versetzt, die modernen Techniken und Werkzeuge des Grid Computing ohne spezielle Kenntnisse anzuwenden.

Projekt: Alignment und Threading auf massiv parallelen Rechnern

Ansprechpartner: DR. THOMAS STEINKE

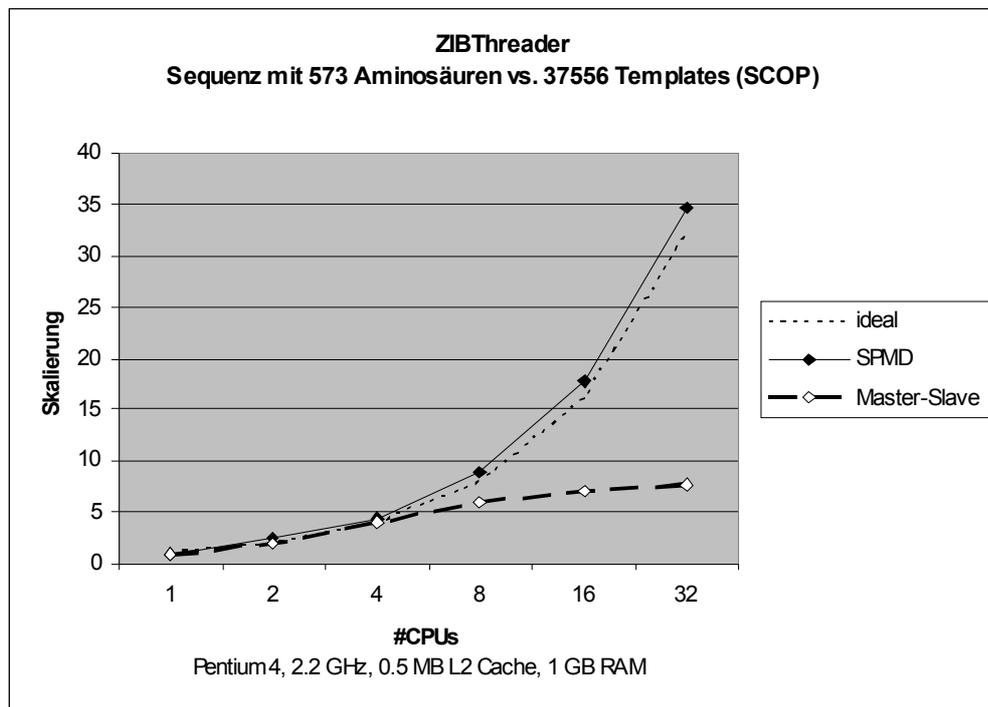
Beteiligte Mitarbeiter: RENÉ HEEK, PATRICK MAY, PROF. DR. ALEXANDER REINEFELD

Zusammenarbeit: Freie Universität Berlin, Humboldt-Universität zu Berlin, Max-Delbrück-Centrum für Molekulare Medizin Berlin-Buch, Max-Planck-Institut für Molekulare Genetik, Technische Fachhochschule Berlin.

Forschungsförderung: BMBF, BCB Forscher-Nachwuchsgruppe

Die Forschungsaufgaben der Nachwuchsgruppe „Alignment und Threading auf massiv parallelen Rechnern“ des „Berliner Centrum für genombasierte Bioinformatik“ (BCB) umfassen die beiden Schwerpunkte Strukturvorhersage mit der Threading-Methode sowie die Realisierung eines Applikationsportals für geeignete Bioinformatikanwendungen im BCB-Verbund. Im Threading wird der Faltungstyp von Proteinen berechnet, indem die Aminosäuresequenz des Proteins auf einen Satz von 3D-Struktur-Templates „gefädelt“ wird. Mit der anschließenden Bewertung mehrerer Threadings erhält man statistisch abgesicherte Vorschläge für den wahrscheinlichsten Faltungstyp.

Skalierungsverhalten des Threading-Core für die Master-Slave- und SPMD-



Implementierung für eine Anfragesequenz bestehend aus 573 Aminosäuren.

Im Berichtszeitraum haben wir das Threading parallelisiert indem mehrere Templates gleichzeitig gegen die Anfragesequenz aliniert werden. Die obige Abbildung zeigt die auf unserem Linux-Cluster erzielten Ergebnisse. Wie zu sehen ist, haben wir die Skalierbarkeit und die Einzel-CPU-Leistung unseres Threading-Cores deutlich verbessert, indem die bisherige Master-Slave-Implementierung durch ein SPMD-Konzept ersetzt wurde. Der durchgeführte FISCHER-Benchmark demonstrierte die erwartete Erkennung der Faltungstypen je nach Faltungsklasse bei 60-90%. Zur Verbesserung der Güte der Faltungserkennung müssen weitere Analysen der Sequenzdaten vorgenommen werden, wodurch letztendlich ein gekoppelter Datenfluss zwischen verschiedenen autonomen Programmen zustande kommt. Zu diesem Zweck haben wir uns mit der Evaluierung von Prozessmanagementsystemen im Kontext der Bioinformatik beschäftigt und speziell das Bio-Operasystem (ETHZ) auf unserem Linux-Cluster installiert.

In Kooperation mit den BCB-Arbeitsgruppen Frömmel/Preissner (Charite/HUB) und Leser (Informatik/HUB) wird die COLUMBA-Datenbank auf dem Frontend-Rechner unseres Cluster-Systems betrieben. Speziell haben wir am Entwurf und der Implementierung der Web-Schnittstelle von COLUMBA mitgearbeitet.

Projekt: Co-Reservierung von Ressourcen im Grid

Ansprechpartner: THOMAS RÖBLITZ

Beteiligte Mitarbeiter: JANIS BARDINS, JÖRG MELTZER, PROF. DR. ALEXANDER REINEFELD, FLORIAN SCHINTKE.

Zusammenarbeit: TONY GARGYA (IBM)

Forschungsförderung: IBM Deutschland Entwicklungs GmbH, Böblingen.

Ursprünglich gaben die Herausforderungen gegenwärtiger und zukünftiger wissenschaftlicher Projekte, zum Beispiel in der Hochenergiephysik, den Anlass zur Entwicklung des Grid Computing. Seit einiger Zeit engagiert sich nun aber auch die Industrie auf diesem Gebiet, um zum Beispiel im sog. Utility Computing den Nutzern flexibel Rechenleistung anbieten zu können. Dabei ist die Fähigkeit, einmal getroffene Vereinbarungen (zum Beispiel hinsichtlich der Erfüllung der Rechenzeitanforderungen) einzuhalten ein ganz zentraler Aspekt. Reservierungen von Ressourcen sind eine Möglichkeit, solche Vereinbarungen zu gewährleisten.

In diesem Projekt werden Methoden für die Reservierung von Ressourcen im Grid entwickelt und getestet. Eine besondere Eigenschaft der betrachteten Reservierungen ist ihre komplexe Gestalt. So sollen z.B. mehrere Ressourcen, die zu unterschiedlichen oder gleichen Zeitenräumen benötigt werden, in Form einer einzigen Reservierung behandelt werden können. Zusätzlich sollen Präferenzen der Nutzer, z.B. hinsichtlich der Reservierungszeit und der Reservierungskosten, berücksichtigt werden. Zunächst sollen effiziente Algorithmen für die Aushandlung einer Reservierung entwickelt werden. Anschließend werden diese in existierende Grid-Middleware zur Unterstützung komplexer Reservierungen integriert.

Projekt: Modellierung der Verfügbarkeit von Replikaten in großen Grid-Systemen

Ansprechpartner: FLORIAN SCHINTKE

Beteiligte Mitarbeiter: PROF. DR. ALEXANDER REINEFELD

Grid-Umgebungen werden derzeit hauptsächlich zur Verteilung von Rechenaufträgen auf geographisch verteilte Rechner genutzt, während Peer-to-Peer Umgebungen, wie Tauschbörsen, eher beim Lagern und Austauschen von Daten eingesetzt werden. Erstere werden eher im wissenschaftlichen, letztere im privaten Umfeld genutzt. Gründe für die bisher

geringe Akzeptanz von Peer-to-Peer Speicherlösungen im wissenschaftlichen Kontext könnten die ungewisse Dateiverfügbarkeit und die Datensicherheit bzw. Persistenz sein.

Um ein dem Peer-to-Peer-Ansatz folgendes System zu bauen, das vorgegebene Dateiverfügbarkeiten automatisch gewährleistet, wurden Modelle entwickelt, die die Dateiverfügbarkeit in solchen Systemen beschreiben. Kernpunkt ist dabei der zweistufig ablaufende Prozess des Dateizugriffs. Zunächst werden die Replika-Kataloge befragt, auf welchen Rechnern Kopien einer Datei abgelegt sind. Danach wird versucht, auf diese Kopien zuzugreifen. Die entwickelten Modelle berücksichtigen mehrere Architekturparameter, wie z.B. die Knotenzuverlässigkeit, die Katalogzuverlässigkeit, das Konsistenzmodell zwischen Katalogen und die Anzahl der Kataloge und Replikate im System.

Im Gegensatz zu anderen Modellen unterscheiden wir zwischen einer lokalen und einer globalen Sichtweise auf das System. Während die lokale Sicht in ein sich selbst überwachendes System eingebaut werden kann, um eine vorgegebene Mindestverfügbarkeit zu gewährleisten, dient die globale Sicht dem Systemadministrator, den Bedarf an Replikaten und somit der benötigten Speicherkapazität im Gesamtsystem abzuschätzen.

Das Modell und seine Einsatzmöglichkeiten sind detailliert in einem Zeitschriftenbeitrag im Journal of Grid Computing beschrieben.

Projekt: StorageBox – Attributbasiertes Management persönlicher Informationen

Ansprechpartner: FELIX HUPFELD

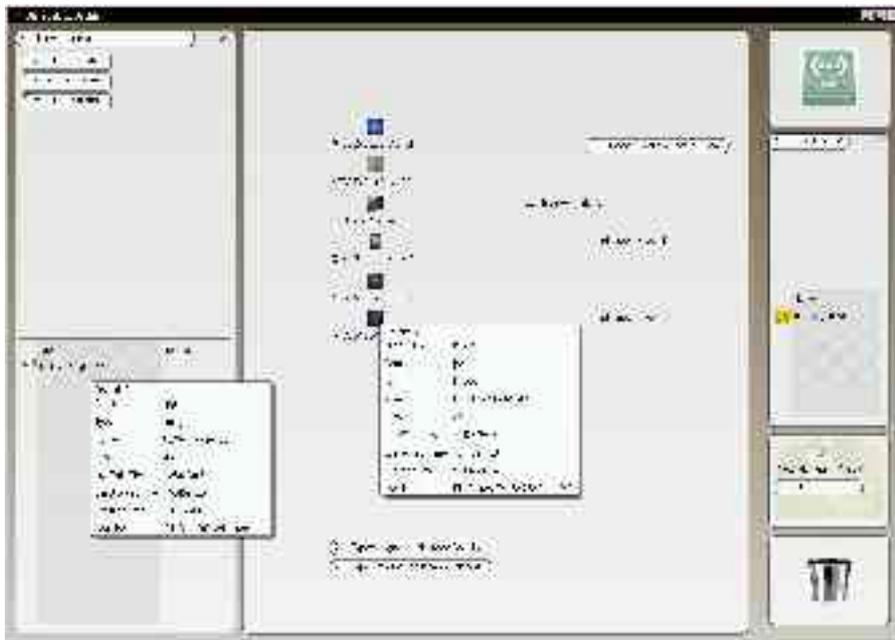
Beteiligte Mitarbeiter: MINOR GORDON, THILO KOCH

Persönliche Informationen wie Fotos, Mails, Briefe und Termine sind heute auf viele verschiedene Arten im Computer gespeichert und jeweils nur über spezielle Benutzerschnittstellen zugreifbar und verwaltbar. Das Dateisystem hält zwar die meisten der Daten, ist mit seiner hierarchischen Benennung aber sehr unflexibel, weil Daten meist nicht eindeutig in ein Verzeichnis passen, sondern unter mehreren Kategorien abgelegt werden müssen. Auch werden Daten wie z.B. Emails und Termine nur in speziellen Datenbasen gespeichert und sind deshalb nur über die jeweilige Anwendung erreichbar. Diese inhomogene Ablage macht es unmöglich, persönliche Informationen untereinander zu verknüpfen und auf sie über eine konsistente Benutzerschnittstelle zuzugreifen.

Screenshot von Storagebox

Mit StorageBox (*s. Abbildung*) wird ein System zur verteilten Speicherung von Information entwickelt, das alle Daten des Benutzers auf dessen persönlichen Geräten speichert und verwaltet. Die Daten können dabei flexibel mit Hilfe von Attribut-Wert Paaren in Kategorien einsortiert werden. Der Zugriff auf das System erfolgt über einen Desktop; ein eigens dafür entwickelter Data-Mining-Algorithmus erlaubt einen hierarchischen Zugriff auf die Daten, der sehr der Interaktion mit dem vertrauten Dateisystem ähnelt. Durch Mehrbenutzerfähigkeit und die Verwendung standardkonformer Schnittstellen ist StorageBox auch zur Datenorganisation im Grid einsetzbar.

Kern von StorageBox ist ein Datenbanksystem, das der mobilen Natur heutiger Geräte Rechnung trägt, und es deshalb erlaubt, die verwalteten Daten auch ohne verfügbares Netzwerk zu ändern und später die Änderungen mit beliebigen Partnern zu synchronisieren. Wegen der begrenzten Speicher- und Kommunikationsressourcen liegt hierbei ein Haupt-



augenmerk auf der Entwicklung eines effizienten Protokolls zur Kommunikation der Änderungen und einer sparsamen Speicherung der dazu nötigen Metadaten.

Projekt: Effiziente Synchronisation von Dateien in verteilten Systemen

Ansprechpartner: THORSTEN SCHÜTT

Beteiligte Mitarbeiter: FLORIAN SCHINTKE, PROF. DR. ALEXANDER REINEFELD

Das Management großer Datenmengen spielt eine immer wichtigere Rolle. Für das DataGrid-Projekt ist es beispielsweise notwendig, große Datenmengen auf mehrere Rechenzentren in Europa zu verteilen und die Daten untereinander zu synchronisieren. Auch innerhalb von Clustern gewinnen mit zunehmender Knotenanzahl Werkzeuge zur effizienten Synchronisation und Verteilung von Daten an Bedeutung.

Im Rahmen dieses Projekts wurde ein effizientes Verfahren zur Synchronisation verteilter Verzeichnisstrukturen entwickelt und implementiert. Das Verfahren benötigt zur Synchronisation keine zentrale Instanz, wodurch eine verbesserte Skalierbarkeit erreicht werden konnte. Dabei wurden Erkenntnisse aus der Graphentheorie eingesetzt und weiterentwickelt, um die Netzwerktopologie und Übertragungsbandbreite zwischen den Rechnern zu berücksichtigen (Gossip). Durch die Verwendung einer Offline-Synchronisation werden Änderungen erst dann an andere Rechner propagiert, wenn der Nutzer dies anstößt. Das kann zum Beispiel nach einer Transaktion mit Änderungen an mehreren Dateien erfolgen.

Die Projektergebnisse sind auf der ICCS-2003 und in der Diplomarbeit von Thorsten Schütt detailliert beschrieben.

Projekt: Grid-Systeme mit Peer-to-Peer Techniken

Ansprechpartner: FLORIAN SCHINTKE

Beteiligte Mitarbeiter: MONIKA MOSER, THORSTEN SCHÜTT, JOCHEN WITTE

Mit dem zunehmenden Einsatz von Grid-Systemen im produktiven Betrieb rücken zwei Aspekte immer mehr in den Vordergrund: die Skalierbarkeit und die Zuverlässigkeit. Um beides zu erhöhen, werden zentrale, leistungs- und ausfallkritische Komponenten, wie beispielsweise der Meta Directory Server (MDS) in Globus, durch hierarchisch vernetzte Implementierungen ersetzt.

Ein weitergehender Ansatz liegt in der Nutzung von Techniken aus dem Peer-to-Peer Computing (P2P). Die flachen, nicht-hierarchischen P2P-Systeme bieten sowohl eine gute Skalierbarkeit als auch eine verbesserte Ausfallsicherheit durch Redundanz. Jeder Peer führt denselben verteilten Algorithmus aus und handelt nach denselben Regeln. Die Gesamtheit der Peers erbringt den Dienst.

Grid-Systeme sind in der Regel sehr umfangreich und bieten eine Vielzahl von Diensten an. Wenn jeder Dienst durch ein einzelnes, speziell dafür ausgelegtes P2P-System erbracht wird, erhält man ein Komponentenmodell, in dem jede (verteilte) Komponente einem P2P-Dienst entspricht. Ähnlich dem erfolgreichen UNIX-Toolset-Ansatz wird dadurch die Komplexität der einzelnen Dienste reduziert, was eine erhöhte Robustheit zur Folge hat. Weitere Informationen sind auf der DEXA-03 vorgestellt worden.

Projekt: Modellierung des Applikationsbedarfs gemeinsam genutzter Ressourcen

Ansprechpartner: DR. ARTUR ANDRZEJAK

Beteiligte Mitarbeiter: MEHMET CEYRAN, PROF. DR. ALEXANDER REINEFELD

Zusammenarbeit: Hewlett-Packard Labs

Forschungsförderung: ZIB

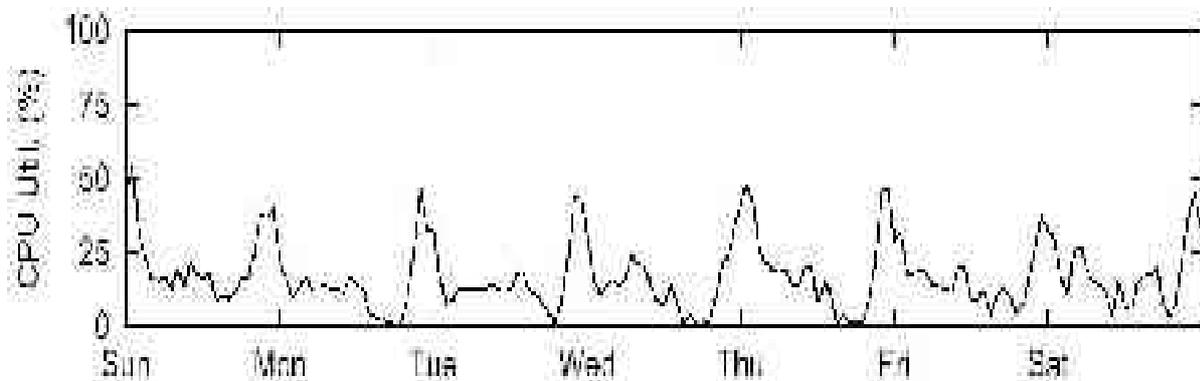
Die Erfahrung zeigt, dass die Ressourcen in industriellen wie auch wissenschaftlichen Rechenzentren nicht immer vollständig genutzt werden. Insbesondere in Rechenzentren großer Firmen, wie Federal Express oder Volkswagen, wird oftmals jeder Anwendung wie z.B. der Lohnbuchhaltung oder dem Webserver ein eigener Rechner zur Verfügung gestellt. Bei derartigen Anwendungen, die immer laufen müssen, wie auch bei periodisch wiederkehrenden Rechenjobs ergeben sich große Ersparnispotentiale durch die gemeinsame Nutzung von Ressourcen. Bei gleich bleibendem Bedarf kann auf diese Weise die Anzahl der benötigten Server reduziert werden. Dieser als *Serverkonsolidierung* bezeichnete Vorgang ist aus Gründen der Wirtschaftlichkeit besonders interessant im Bereich des Utility Computing, wo Rechenressourcen von Drittanbietern geleased werden.

Das Projekt verfolgt drei Ziele. Zunächst soll der Bedarf an Ressourcen geeignet modelliert werden, um Voraussagen über erwartetes künftiges Verhalten der Applikationen treffen zu können. Auf der Basis dieser Daten soll dann eine langfristige Planung von benötigten Kapazitäten ermöglicht werden. Für die kurzfristige Planung werden schließlich Algorithmen und Verfahren entwickelt, die die Zuteilung der Applikationen zu Ressourcen automatisieren.

Die Modellierung des Bedarfs erfolgt im Projekt auf Basis von sog. Kalender-Verfahren: eine feste Zeitperiode, z.B. ein Tag, wird in Abschnitte, z.B. jeweils eine Stunde, unterteilt. Anhand der historischen Daten wird dann der erwartete Bedarf jeder Applikation in jedem Abschnitt prognostiziert. Der aktuelle Forschungsgegenstand ist die automatische Klassifizierung von Applikationen in solche mit periodischem oder nichtperiodischem Verhalten, die Ermittlung von (ggf. hierarchischen) Periodenmustern im ersten Fall, und die Modellierung durch Verfahren aus Ökonometrie im anderen Fall.

Für die kurz- und langfristige Planung wurde ein effizienter Scheduler entwickelt, der auf Basis von genetischen Algorithmen arbeitet. Die Geschwindigkeit dieses Schedulers ermöglicht eine Online-Berechnung der Zuordnung von Applikationen zu Ressourcen, wodurch auch die kurzfristige Planung abgedeckt wird. Eine Fallstudie hat gezeigt, dass die Qualität unwesentlich von den optimalen Lösungen abweicht, wobei die letzteren durch zeitaufwendiges quadratisches Programmieren gefunden wurden. Die künftige Forschung in

diesem Teilbereich wird sich auf die Probleme der sog. pareto-effizienten Optimierung und Zuordnung unter mehreren Gütekriterien konzentrieren.



Periodischer Rechenbedarf einer Applikation mit starkem Potential für Server-Konsolidierung

Projekt: Virtual Supercomputer Berlin - Hannover

Ansprechpartner: HUBERT BUSCH

Beteiligte Mitarbeiter: SEBASTIAN HEIDL, MATTHIAS HEYDER, THOMAS RÖBLITZ

Forschungsförderung: DFN-Verein

Mit der Installation des neuen Hochleistungsrechners für die norddeutschen Länder (HLRN) steht den Wissenschaftlern ein außergewöhnlich leistungsfähiges System zur Verfügung. Durch die Verteilung der Rechenelemente auf zwei verschiedene Standorte in Berlin (ZIB) und Hannover (RRZN) entstehen jedoch auch neue Herausforderungen für den Betrieb und die effiziente Nutzung des Gesamtsystems. Inhalte dieses Projektes sind die Erforschung und Lösung der durch die Verteilung des Systems hervorgerufenen Probleme (z. B. Scheduling, Kommunikation, I/O). Es sollen effiziente Lösungen zur Bereitstellung eines virtuellen, hochperformanten und transparenten Systems entwickelt werden, die auf vergleichbare Installationen übertragbar sind.

Im Berichtszeitraum wurden Untersuchungen zum Datentransfer zwischen verschiedenen Dateisystemen durchgeführt. Die Ergebnisse sollen für die Entwicklung eines Tools verwendet werden, welches automatisch (abhängig von Quelle und Ziel) die effizienteste Übertragungsmethode benutzt. Eine parallele Anwendung kann gegebenenfalls auf beiden Standorten gleichzeitig ausgeführt werden. Für die Kommunikation zwischen den Prozessen stehen verschiedene MPI-Implementationen zur Verfügung. Diese wurden hinsichtlich ihrer Leistungsfähigkeit untersucht. Erste Ergebnisse wurden in einem DFN-Projektbericht zusammengefasst.

Veröffentlichungen

G. ALLEN, K. DAVIS, K. N. DOLKAS, N. D. DOULAMIS, T. GOODALE, T. KIELMANN, A. MERZKY, J. NABRZYSKI, J. PUKACKI, T. RADKE, M. RUSSELL, E. SEIDEL, J. SHALF, I. TAYLOR: *Enabling Applications on the Grid: A GridLab Overview*, Int. Journal of High Performance Computing Applications, November 2003, vol. 17, no. 4, pp. 449-466(18).

A. ANDRZEJAK, K. FUKUDA: *Debugging Distributed Computations by Reverse Search*. Parallel and Distributed Computing and Networks (PDCN'2003), Februar 2003.

- A. ANDRZEJAK: *Memory-Efficient and Fast Enumeration of Global States*. 7th IEEE International Conference on Information Visualization, IV03-AGT, London, Juli 2003.
- S. GRAUPNER, V. KOTOV, A. ANDRZEJAK, H. TRINKS: *Service-Centric Organization of Globally Distributed Computing*. IEEE Internet Computing, Juli/August 2003.
- S. GRAUPNER, A. ANDRZEJAK, V. KOTOV, H. TRINKS: *Adaptive Control Overlay for Service Management*. First Workshop on the Design of Self-Managing Systems (AASMS'2003), Juni 2003.
- F. HUPFELD: *Hierarchical Structures in Attribute-based Namespaces and their Application to Browsing*. ZIB Report 03-06, März 2003.
- A. REINEFELD, F. SCHINTKE: *Grid Computing in Deutschland*. Spektrum der Wissenschaft, Juli 2003, pp. 72-73.
- T. RÖBLITZ, F. SCHINTKE, A. REINEFELD: *From Clusters to the Fabric: The Job Management Perspective*. IEEE Intl. Conference on Cluster Computing (Cluster 2003), Hongkong, Dezember 2003, pp. 468 - 473.
- J. ROLIA, A. ANDRZEJAK, M. ARLITT: *Automating Enterprise Application Placement in Resource Utilities*. 14th IFIP/IEEE Workshop on Distributed Systems: Operations and Management (DSOM 2003), Oktober 2003.
- K. ROTHER, S. TRISSL, H. MÜLLER, P. MAY, R. HEEK, R. PREISSNER, TH. STEINKE, I. KOCH, U. LESER, C. FRÖMMEL: *COLUMBA - A Database of Annotations of Protein Structure*. ECCB 2003, September 2003, Paris, pp. 109 - 110.
- F. SCHINTKE, A. REINEFELD: *Modeling Replica Availability in Large Data Grids*. Journal of Grid Computing (2) 2003, Kluwer Acad. Publisher.
- F. SCHINTKE, T. SCHÜTT, A. REINEFELD: *A Framework for Self-Optimizing Grids Using P2P Components*. Intl. Workshop on Autonomic Computing Systems, 14th Intl. Workshop on Database and Expert Systems Applications (DEXA'03), September 2003, pp. 689 - 693.
- T. SCHÜTT, F. SCHINTKE, A. REINEFELD: *Efficient Synchronization of Replicated Data in Distributed Systems*. Intl. Conf. on Computational Science ICCS - 2003, St. Petersburg, Juni 2003, Springer LNCS 2657, pp. 274-283.
- J. WENDLER, J. BACH: *Case Based Behavior Assessment in Simulated Robotic Soccer*. Modellierung und Simulation menschlichen Verhaltens, Informatik-Bericht Nr. 163, Humboldt-Universität zu Berlin, März 2003.
- J. WENDLER, J. BACH: *Recognizing and Predicting Agent Behavior with Case Based Reasoning*. Proceedings of the 7th RoboCup Workshop, Padua, Juli 2003.

Vorträge

- ALEXANDER REINEFELD: *Managing Distributed Data in the Grid: Case Studies in Drug Design and Particle Physics*. 1st European Grid Symposium for Life Sciences, InfoTechPharma 2003, London, 10.02.2003.
- ALEXANDER REINEFELD: *The Grid - Status and Perspectives*. ENES Workshop, Antwerpen, 12.02.2003.
- JAN WENDLER: *Automatisches Modellieren von Agenten-Verhalten – Erkennen, Verstehen und Vorhersage*. Fachhochschule Zittau/Görlitz, Fachbereich Informatik, 26.3.2003.

FLORIAN SCHINTKE: *Grid Computing with Peer-to-Peer Techniques*. Tag der Informatik, Humboldt-Universität zu Berlin, Mai 2003.

ALEXANDER REINEFELD: *Grid Computing – Status and Perspectives*. Ringvorlesung "Schwerpunkte der Informatik", Humboldt-Universität zu Berlin, 05.06.2003.

ALEXANDER REINEFELD: *Towards a German Grid/eScience Program – Grid Activities in Germany*. Workshop on ERA and Grid Research, Brüssel, 17.07.2003.

JAN WENDLER: *Modellierung von Agenten-Verhalten – Erkennen, Verstehen und Vorhersagen von Verhalten in komplexen Multi-Agenten-Systemen*. Dissertations-Verteidigung, Humboldt-Universität zu Berlin, 26.8.2003.

ALEXANDER REINEFELD: *Distributed Data Management in the Grid*. VLDB 2003, Berlin, 11.09.2003.

ALEXANDER REINEFELD: *Aktuelle technische Fragestellungen bei der Grid-Entwicklung*. Konferenz „E-Science: Die technische Infrastruktur“, Wissenschaftszentrum Bonn, 25.09.2003.

JAN WENDLER: *FlowGrid – Flow Simulations On-Demand Using Grid-computing*. HP Labs, Bristol, 3.10.2003.

JAN WENDLER: *Agent Behaviour Modelling – Recognizing and Predicting Soccer Agent Behaviour using Case Based Reasoning*. HP Labs, Bristol, 3.10.2003.

Sonstige Aktivitäten

Advisory Boards

Prof. Dr. Alexander Reinefeld

- ALICE High-Level Trigger Review Board (Cern)
- apeNEXT Supercomputer Review Board (INFN, DESY)
- Autonomic Computing Advisory Board (IBM)
- BMBF Ref. 522 „Diskussionsforum Internet“
- D-Grid Lenkungsausschuss
- DFN Technischer Ausschuss
- Global Grid Forum Advisory Committee (GFAC)
- Global Grid Forum Research Oversight Committee (GROC)
- IEEE Task Force on Cluster Computing (TFCC)

Editorial Boards and Conference Advisory Boards

Prof. Dr. Alexander Reinefeld

- European Conference on Parallel Computing (Euro-Par)
- Journal Future Generation Computing Systems (FGCS), Elsevier
- IEEE/ACM Symposium on Cluster Computing and the Grid (CCGrid)
- Journal of Grid Computing (JoGC), Kluwer Academic Publisher

Veranstaltete Tagungen und Workshops

Prof. Dr. Alexander Reinefeld

- German Grid Day: Towards a German D-Grid, Heidelberg, 01.10.2003

- Workshop „Quantenchemie/Dichtefunktionaltheorie“ des Norddeutschen Verbundes für Hoch- und Höchstleistungsrechnen, ZIB, Berlin, 08.07.2003

Dr. Thomas Steinke

- First Berlin Molecular Dynamics Meeting, Berlin, 30.10.-1.11.2003

Mitarbeit in Programmkomitees

Prof. Dr. Alexander Reinefeld

- PDP2003: Conference on Parallel Distributed and Networking based Processing, Genoa, Italy, 05.-07.02.2003
- HIPS'03: 8th International Workshop on High-Level Parallel Programming Models and Supportive Environments, Nice, France, 22.04.2003
- CCGrid2003: 3rd IEEE/ACM International Symposium on Cluster Computing and the Grid, Tokyo, Japan, 12.-15.05.2003
- HPDC-12: High Performance Distributed Computing, Seattle, Washington, 22.-24.06.2003
- 1st International Summer School on Grid Computing, organized by Global Grid Forum, Vico Equense, Italy, 13 - 25.07. 2003
- Euro-Par 2003: International Conference on Parallel and Distributed Computing, Klagenfurt, Austria, 26.-29.08.2003
- AGridM 2003: Workshop on Adaptive Grid Middleware, in conjunction with PACT 2003, New Orleans, Louisiana, September 2003
- ParCo2003: Parallel Computing 2003, Dresden, Deutschland, 02.-05.09.2003
- EDG2003: European Datagrid Conference, German Grid Day, Heidelberg, 29.09.-03.10.2003
- ICPP 2003: International Conference on Parallel Processing, Kaohsiung, Taiwan, 06.-09.10.2003
- Grid2003: 4th International Workshop on Grid Computing, Phoenix, Arizona, 17.11.2003
- Cluster 2003: IEEE International Conference on Cluster Computing, Kowloon, Hongkong, 01.-04.12.2003

Dr. Jan Wendler

- 7th RoboCup International Symposium, Padua, Italy, 10.-11.7.2003

Advisory Committees

Dr. Thomas Steinke

- Vizevorsitzender des Management Committees der EU Concerted Research Action D23, METACHEM - Metalaboratories for Complex Computational Applications in Chemistry

Betreute Diplomarbeit

Prof. Dr. Alexander Reinefeld

SIMON RIECHE: *Lastbalancierung in Peer-to-Peer-Systemen*, Januar 2004.

Betreute Studienarbeiten

Prof. Dr. Alexander Reinefeld

JOCHEN WITTE: *Fallstudie: GDMP und SRB – Vergleich zweier Datenmanagementsysteme für das Grid*, Mai 2003.

Dr. Thomas Steinke

SEBASTIAN MOLL: *Die Berechnung metastabiler Konformationen von Biomolekülen am Beispiel des Bicyclo[3.3.1.]nonan*, Juni 2003.

Lehr- und Forschungseinheit

Spezifikation, Verifikation and Testtheorie

<http://www.informatik.hu-berlin.de/svt>

Leitung

PROF. DR. HOLGER SCHLINGLOFF

Tel.: 030 6392 1907

E-Mail: hs@informatik.hu-berlin.de

Die Forschungs- und Lehrereinheit SVT (Spezifikation, Verifikation und Testtheorie) beschäftigt sich mit grundlegenden Aspekten der Qualitätssicherung von Software. Assoziiert ist der Bereich „Synthese, Validierung und Test“ am Fraunhofer Institut für Rechnerarchitektur und Softwaretechnik (FIRST). Fragen, an denen wir arbeiten, sind unter anderem folgende:

- „Wie kann man eindeutig beschreiben, was ein Programm tun soll?“
- „Wie kann man die Korrektheit eines Programms beweisen?“
- „Wie kann man ein Programm in seiner Umgebung testen?“

Die Formalismen, die dabei erforscht werden, sind temporale und modale Logiken, Prozessalgebren, Transitionssysteme und Petrinetze, StateCharts, und andere. Für diese Sprachen entwickeln wir Methoden zur Modellprüfung, Konformanzanalyse, Verfeinerung, Deduktion und Transformation. Ein besonderer Schwerpunkt ist die automatisierte Erstellung von Testfällen und die automatische Durchführung von „Black-Box“-Tests für eingebettete Echtzeitsysteme. Untersuchte Forschungsthemen beinhalten dabei sowohl grundlagenorientierte Fragestellungen wie etwa nach der Expressivität und Komplexität bestimmter Spezifikationsprachen, als auch praktische Gesichtspunkte wie etwa die effiziente Repräsentation und Manipulation großer Datenstrukturen in speziellen Verifikationsverfahren. Ein wesentlicher Gesichtspunkt ist in jedem Fall die Anwendbarkeit der erzielten Ergebnisse in praktischen Systemen, etwa in eingebetteten Steuergeräten im Verkehr, Kommunikationswesen oder beim Chip-Design. Neue Herausforderungen ergeben sich derzeit bei formalen Methoden zur Angriffssicherung und in der multimedialen und dreidimensionalen Präsentation von Forschungsinhalten.

Die Gruppe wurde im Mai 2002 gegründet und fungiert als „Theorieunterstützung“ des Bereichs SVT beim FIRST. Dort werden die erzielten Ergebnisse in konkreten Anwendungsprojekten mit industriellen Partnern, etwa Alcatel, O₂ oder DaimlerChrysler, umgesetzt. Während frühere Projekte vor allem in den Bereichen Luft- und Raumfahrt (Satelliten- und Flugzeug-Steuergeräte) und in der Telekommunikation (UMTS Mobiltelefone) angesiedelt waren, sind neue Anwendungen derzeit hauptsächlich im Schienenverkehr (Zuverlässigkeit eines fehlertoleranter Stellwerkscomputers), Automobil (Entwicklungsmethodik eines Fahrgastkomfort-Türsteuergerätes), Zahlungsverkehr (Formalisierung einer Kartenterminal-Spezifikation) und bei Internetdiensten (Qualität eines großen Informationssystems).



Lehre

- Vorlesung WS 2002/2003: Management großer Informationssysteme (Halbkurs)
- Vorlesung WS 2003/2004: Praktische Informatik 1 (Grundvorlesung)

Forschung

Projekt: HyMoC

Ansprechpartner: PROF. DR. M DE RIJKE, PROF. DR. HOLGER SCHLINGLOFF

Beteiligter Mitarbeiter: MASSIMO FRANCESCHET

Zusammenarbeit: Universität Amsterdam, Universität Udine, Universität Pescara, Universität Swansea

Forschungsförderung: NWO Netherlands Organization for Scientific Research

HyMoC ist ein Kooperationsprojekt zwischen den beteiligten Universitäten zur Entwicklung von Modellprüfungsverfahren für hybride Logiken. In vielen Anwendungen modaler Spezifikationstechniken ist es vorteilhaft, einzelne Zustände mit Namen benennen zu können (z.B. „Start“ oder „Fehler“). Hybride Logik ist eine Erweiterung klassischer Modallogik um erststufige Konzepte wie Namen, Individuenvariablen und (eingeschränkte) Quantoren. Trotz ihrer größeren Ausdrucksmächtigkeit sind viele hybride Logiken „gutartig“, ihre Komplexität ist nicht wesentlich schlechter als die der zu Grunde liegenden Modallogik. Andererseits sind das Modellprüfungsproblem und die Anwendungsmöglichkeiten hybrider Logiken bislang noch nicht hinreichend untersucht. In diesem Projekt geht es darum, die algorithmische Komplexität des Modellprüfungsproblems für verschiedene Sprachen und Modellklassen zu untersuchen, Heuristiken und Techniken der klassischen Modellprüfung

auf den hybriden Fall zu übertragen und einige Fallstudien zur Spezifikation mit hybrider Logik durchzuführen.

Im Jahr 2003 wurden wichtige Ergebnisse zur Expressivität der hybriden Modallogik erarbeitet und auf einer internationalen Konferenz vorgestellt. Des Weiteren wurden Arbeiten zur Modellierung der Beispielapplikation mit abstrakten Datentypen begonnen und in diesem Zusammenhang eine Kooperation mit der Zühlke AG, Schweiz, sowie der Universität Swansea, Wales, vereinbart.

Projekt: MMiSS

Projektleitung: PROF. DR. WOLFGANG REISIG, PROF. DR. HOLGER SCHLINGLOFF

Mitarbeiter: DIRK FAHLAND

Mittelgeber: BMBF / Univ. Bremen

Ziel des Projektes MMiSS ist die Erstellung eines multimedialen, internetgestützten, adaptiven Lehr- und Lernsystems für die Informatikausbildung. Es wird das gesamte Curriculum des Fachgebiets „Sichere Systeme“ abdecken und soll im normalen Lehrbetrieb der beteiligten Universitäten eingesetzt werden. Das System bietet eine konsistente, hypermediale Aufbereitung von klassischen Vorlesungsmaterialien (Folien, Büchern, Skripten, Übungsaufgaben etc.) und eine Integration von Werkzeugen zur formalen Entwicklung korrekter Software und ist damit sowohl zur Unterstützung der Präsenzlehre, zum Fernstudium als auch zum interaktiven, betreuten und kooperativen Selbststudium geeignet. Besonderer Wert wird auf Kohärenz und Konsistenz der Inhalte durch umfassende semantische Vernetzung gelegt.

An der Humboldt-Universität wird für MMiSS ein Modul zum Thema „Modellbasierte Softwareentwicklung“ erstellt, welche das bereits in MMiSS vorhandene Material komplementär ergänzen. Inhaltlich umfassen die Materialien einen wesentlichen Aspekt der Modellierung sicherheitskritischer Systeme: Um die Korrektheit eines Softwaresystems nachzuweisen, muss zunächst ein Modell davon entworfen werden - die Spezifikation. Verschiedene Spezifikationsarten für verteilte Algorithmen werden in diesem Modul behandelt. Insbesondere werden dabei netzbasierende Formalismen berücksichtigt, die in der modellbasierten Entwicklung eine besondere Rolle spielen. Weiterhin kommen typische Eigenschaften von Systemen (Sicherheit, Lebendigkeit, Fairness) vor, die nachzuweisen sind, um die Korrektheit eines Systems zu zeigen.

Im Berichtsjahr wurden die Lehrmaterialien für die Integration in den gemeinsamen Rahmen aufbereitet sowie Untersuchungen zum Einsatz von Software-Engineering-Werkzeugen in multimedialen Lernumgebungen durchgeführt.

Veröffentlichungen

M. FRANCESCHET, M. DERIJKE, AND H. SCHLINGLOFF: *Hybrid Logics on Linear Structures: Expressivity and Complexity*. In: "TIME/ICTL 2003. 4th International Conference on Temporal Logic" (July 2003). IEEE Computer Society Press

M. FREY, H. SCHLINGLOFF: *Conformance of Distributed Systems*. In: "TestCom 2003: Next Generation Testing for Next Generation Networks". 15th IFIP International Conference on Testing of Communicating Systems, Sophia Antipolis, France, May 26-28, 2003, Springer LNCS 2644 (May 2003).

Vorträge

H. SCHLINGLOFF: ETSI, *Conformance of Distributed Systems* (zusammen mit M. Frey, o2 Germany GmbH). Sophia Antipolis, Frankreich (TestCom 2003), Mai 2003.

H. SCHLINGLOFF: *Modellbasiertes Testen von Software*. HU Berlin (Ringvorlesung), Juni 2003.

H. SCHLINGLOFF: *Systematisches Testen von Software*. FhG FIRST (ViSEK), Juni 2003.

H. SCHLINGLOFF: *A temporal logic extension of role-based access control covering dynamic separation of duties*. Cairns, Australien (TIME/ICTL 2003), Juli 2003.

H. SCHLINGLOFF: *Modellbasierte Testfallermittlung und Lasttests mit TTCN-3*. Berlin, DaimlerChrysler RIC, Berlin, September 2003.

H. SCHLINGLOFF: *Software-Qualitätssicherung – Testgenerierung aus UML, Fehlertoleranz und Realzeit-Modellprüfung*. Siemens CT, München, Oktober 2003.

H. SCHLINGLOFF: *Linear Time Hybrid Logics*. (Kolloquium) Swansea, November 2003.

Sonstige Aktivitäten

Prof. Dr. Holger Schlingloff

- Mitglied des Organisationskomitees für das zweijährlichen Kolloquium „Methods for Modalities“, Amsterdam
- Mitglied des Programmkomitees der Tagung „TIME-ICTL 2003, 10th International Symposium on Temporal Representation and Reasoning and 4th International Conference on Temporal Logic“, Cairns, Queensland, Juli 2003
- Lecturer bei der ESSLLI2004: in Nancy "Temporal Logic for Systems Validation"

Diplomarbeiten

RENE BARTHOLOMÄ: *Entwurf einer Applikationsbeschreibungssprache für Smart-Card Betriebssysteme* (in Zusammenarbeit mit der MidiTec GmbH, Bremen) Februar 2003, Univ. Bremen

MARKUS GESCHKE: *Vergleich des strukturellen Entwurf von Gerätetreibern in verschiedenen Betriebssystemen am Beispiel eines SmartCardLeser-Treibers* (in Zusammenarbeit mit der MidiTec GmbH, Bremen) Februar 2003, Univ. Bremen

KIM SCHÖN: *Sichere Kommunikation in sporadischen Geschäftsbeziehungen* Januar 2003, Univ. Bremen

Lehr- und Forschungseinheit
Informationsintegration

<http://www.informatik.hu-berlin.de/mac/>

Leiter

PROF. DR. FELIX NAUMANN, JUNIORPROFESSOR
Tel.: (030) 2093 3905
E-Mail: naumann@informatik.hu-berlin.de

Sekretariat

CORNELIA MISPELHORN
Tel.: (030) 2093 3901
E-Mail: mispel@informatik.hu-berlin.de

Wissenschaftliche Mitarbeiter

DIPL.-INF. JENS BLEIHOLDER
DIPL.-ING. MELANIE WEIS

Tutoren

CHRISTOPH BÖHM
N.N.

Die Professur wurde am 5.5.2003 eingerichtet. Der Berichtszeitraum erstreckt sich damit über 8 Monate. Neben dem räumlichen, technischen und personellen Aufbau des Lehrstuhls konnten bereits erste Forschungsarbeiten in Angriff genommen werden.

Lehre im Hauptstudium

Crashkurs: Informationssysteme (Prof. Naumann, WS 2003/2004)
Vorlesung: Informationsintegration I (Prof. Naumann, WS 2003/2004)
Seminar: Data Cleansing (Prof. Naumann, WS 2003/2004)

Forschung

Projekt: "MAC: Merging Autonomous Content"

Ansprechpartner: PROF. DR. FELIX NAUMANN

Beteiligte Mitarbeiter: Dipl.-Inf. JENS BLEIHOLDER, DIPL.-ING. MELANIE WEIS

Merging Autonomous Content (MAC) ist die Vereinigung von Daten und Informationen mehrerer, autonomer Quellen zu einer einheitlichen Informationsmenge. Trotz intensiver Entwicklung und Untersuchung integrierender Informationssysteme, angefangen mit verteilten und föderierten Datenbanksystemen, über Methoden der Schemaintegration, bis hin zu Mediator-basierten Integrationsarchitekturen und Anfragesystemen existieren überraschend wenig wissenschaftliche Ergebnisse und praktische Ansätze zur Integration der Daten und Informationen selbst. Nach dem Entwurf und Aufbau eines integrierten Informationssystems wird der nächste Schritt, die eigentliche Integration der heterogenen

Informationen, oft ignoriert. MAC ordnet sich in den Bereich Contentware Engineering ein und bietet ein großes Forschungspotenzial.

Teilprojekt: "HumMer: Der Humboldt Merger"

Ansprechpartner: PROF. DR. FELIX NAUMANN

Beteiligte Mitarbeiter: Dipl.-Inf. JENS BLEIHOLDER, DIPL.-ING. MELANIE WEIS, CHRISTOPH BÖHM

Der Humboldt Merger (HumMer) ist ein integriertes Informationssystem, welches relationale, XML-, und unstrukturierte Daten liest und in ein gemeinsames, strukturiertes Datenmodell vereint. HumMer dient der Forschungsgruppe als eine Plattform zur Implementierung und zum Testen der Machbarkeit, Effizienz und Effektivität neuer Ideen der Informationsintegration. HumMer befindet sich zurzeit in der Entwicklung.

Teilprojekt: "Objektidentifikation in XML"

Ansprechpartner: Dipl.-Ing. MELANIE WEIS

Ein Problem der Datenintegration ist das Auftreten von verschiedenen Datensätzen in der Zielrepräsentation, die das gleiche Objekt beschreiben. Ziel dieses Projektes ist ein Algorithmus, der solche Duplikate in XML Dokumenten identifiziert und eliminiert. Ein Teil der Arbeit besteht darin, existierende Verfahren zur Objektidentifikation für Relationen auf ihre Übertragbarkeit auf XML zu prüfen und gegebenenfalls zu erweitern.

Teilprojekt: "Data Merging - Zusammenführen von Daten"

Ansprechpartner: Dipl.-Inf. JENS BLEIHOLDER

Dieses Projekt befasst sich mit der Untersuchung von Verfahren, die heterogene Daten aus autonomen Datenquellen zusammenführen (*data merging*). Neben klassischen Datenbank-techniken und Techniken in Multi- und föderierten Datenbanksystemen werden auch Techniken und Algorithmen in existierenden integrierten Informationssystemen betrachtet. Ziel ist es, ein Verfahren zur Zusammenführung von Daten für das HumMer-Projekt zu entwickeln, das beim Zusammenführen auftretende Datenkonflikte intelligent auflöst.

Projekt: "BioFast: Efficient and seamless access to Internet accessible life sciences data sources"

Ansprechpartner: PROF. DR. FELIX NAUMANN

Beteiligte Mitarbeiter: Dipl.-Inf. JENS BLEIHOLDER

Kooperation: PROF. DR. LOUIQA RASCHID, University of Maryland, PROF. DR. ZOE LACROIX, Arizona State University

Informationssysteme der life sciences (Lebenswissenschaften) sind oft stark untereinander verlinkt: Daten über Proteine speichern Verknüpfungen zu verwandten Krankheiten, Daten über Gene verweisen auf entsprechende Fachpublikationen, usw. Im BioFast Projekt wird diese Linkstruktur qualitativ und quantitativ untersucht. Ziel ist es, Anfragen an solcherart verlinkte Informationsquellen effizienter (schnellere Ausführung) und effektiver (bessere Ergebnis) auszuführen.

Projekt: "Instanz-gesteuertes Schema-Matching"

Ansprechpartner: PROF. DR. FELIX NAUMANN

Zusammenarbeit: ALEXANDER BILKE, Technische Universität zu Berlin

Schema-Matching ist der Vorgang, Elemente zweier unterschiedlich strukturierter Schemata über die gleiche Domäne semantisch miteinander zu verknüpfen. Beispielsweise haben zwei Tabellen mit Adressdaten jeweils ein Feld, in dem der Wohnort gespeichert wird. Diese Felder können unterschiedlich bezeichnet sein (etwa, „Wohnort“ und „Stadt“), tragen jedoch die gleiche Semantik. Ziel des Projektes ist es, ein Verfahren zu finden, welches, geleitet durch vorhandene Datensätze, automatisch oder zumindest semi-automatisch solche Gemeinsamkeiten findet.

Veröffentlichungen

Z. LACROIX, F. NAUMANN, L. RASCHID, M. ESTHER: *Exploring Life Sciences Data Sources*. Vidal Proceedings of IJCAI-03 Workshop on Information Integration on the Web (IIWeb-03) Acapulco, Mexico.

H. MÜLLER, F. NAUMANN, J.-C. FREYTAG: *Data Quality in Genome Databases*. Proceedings of the International Conference on Information Quality (IQ 2003), Cambridge, MA.

A. LÖSER, F. NAUMANN, W. SIBERSKI, W. NEJDL, U. THADEN: *Semantic Overlay Clusters within Super-Peer Networks*. Proceedings of the International Workshop on Databases, Information Systems and Peer-to-Peer Computing, 2003 (P2P03), Berlin, Germany.

F. NAUMANN, V. JOSIFOVSKI, S. MASSMANN: *Super-Fast XML Wrapper Generation in DB2 (demo)*. Proceedings of the International Conference on Data Engineering (ICDE 2003), Bangalore, India.

F. NAUMANN, J.-C. FREYTAG, U. LESER: *Completeness of Information Sources*. Workshop on Data Quality in Cooperative Information Systems 2003 (DQCIS) 2003 Siena, Italy.

F. NAUMANN: *BRAIN-GAIN: Wie verlockend sind Juniorprofessuren?* Science Next Wave, 5. Sept. 2003.

Vorträge

F. NAUMANN: *Informationsintegration - Große Datenräume in Web-basierten Umgebungen*. Ringvorlesung "Schwerpunkte der Informatik", 3.Juli 2003, Humboldt Universität, Institut für Informatik.

F. NAUMANN: *Information Quality and its Impact on Information Integration*. Invited Talk at "1. Deutsche Information Quality Management Konferenz", 20.Oktober 2003, Frankfurt, Germany.

Sonstige Aktivitäten

Prof. Dr. Felix Naumann

- Gastwissenschaftler: Prof. Louiqa Raschid, University of Maryland, September
- Programmkomiteemitglied der International Conference on Information Quality (IQ) 2003

- Programmkomiteemitglied des International Workshop on Data Quality in Cooperative Information Systems (DQCIS) 2003
- Technical Co-Chair der International Conference on Very Large Databases (VLDB) 2003
- Teilnahme an der Konferenz über Datenbanken für Business, Technologie und Web (BTW) 2003
- Teilnahme am Dagstuhl Seminar 03162 „GIBU 2003“
- Teilnahme am Dagstuhl Seminar 03362 „Data Quality on the Web“
- Forschungsaufenthalt an der University of Maryland, Smith School of Business
- Teilnahme an der German Information Quality Management Conference 2003 (GIQM)

Dipl.-Inf. Jens Bleiholder

- Teilnahme am Workshop WebDB im Rahmen der Berliner XML-Tage, Oktober 2003, Berlin
- Unterstützung bei der Organisation der VLDB 2003, September 2003, Berlin

Lehr- und Forschungseinheit
Rechnerorganisation und Kommunikation

<http://www.informatik.hu-berlin.de/rok>

Leiter

PROF. DR. MIROSLAW MALEK
 Tel.: (030) 2093 3027
 E-Mail: malek@informatik.hu-berlin.de

Sekretariat

SABINE BECKER
 Tel.: (030) 2093 3028
 Fax: (030) 2093 3029
 E-Mail: sbecker@informatik.hu-berlin.de

Wissenschaftliche Mitarbeiter und Mitarbeiterinnen

DR. GÜNTER DOLLNY
 DIPL.-INF. PETER IBACH
 DIPL.-INF. NIKOLA MILANOVIC
 DIPL.-INF. JAN RICHLING
 DIPL.-ING. FELIX SALFNER
 DR. KATINKA WOLTER

Technische Mitarbeiter und Mitarbeiterinnen

DR. SIEGMAR SOMMER
 DIPL.-INF. STEFFEN TSCHIRPKE
 DIPL.-CHEM. BIRGIT SCHIEFNER
 CHRISTINE HENZE

Tutoren

STEFFEN BUHLE
 MARKUS MERGNER
 VLADIMIR STANTCHEV
 STEFAN WIANKE

Die Lehr- und Forschungseinheit Rechnerorganisation und Kommunikation hat sich auf verschiedene Aspekte netzwerkbasierter Computersysteme spezialisiert. Unser Interesse liegt auf dem Gebiet des verteilten und parallelen Rechnens mit den Schwerpunkten Fehlertoleranz, Echtzeitfähigkeit, IT-Sicherheit, Kommunikation und Modellierung.

Unsere Prioritäten sind Hochleistung und Fehlertoleranz im Internet sowie Echtzeit, Zuverlässigkeit und Sicherheit für eingebettete Systeme.

Zur Lehr- und Forschungseinheit gehört die Netzwerkbetriebsgruppe. Ihre Arbeit wird in **Abschnitt V** genauer beschrieben.

Dienstgütegarantien für Standardsysteme

Parallele Architekturen haben sich bisher nur in wenigen, speziellen Anwendungsfällen durchsetzen können. Gründe dafür sind die schwierige Programmierbarkeit und schlechte Softwareunterstützung für Parallelrechner sowie deren hohe Kosten. Zudem veraltet die spezielle Hardware paralleler Systeme schnell. Beachtliche Rechenleistungen können auch von vernetzten, kooperierenden Workstations erbracht werden, die heute schon in großer Zahl installiert sind. Wir beschäftigen uns mit netzwerkbasierten Parallelrechnern, die als Multicomputer-Systeme Workstations unter Benutzung schneller lokaler Netze integrieren.

Schwerpunkte sind die Untersuchung neuer Paradigmen für parallele Programmierung, neue Ansätze zur virtuellen verteilten Speicherverwaltung und die Benutzung und Evaluierung von Standard-Middleware-Architekturen wie .NET für responsives (also fehlertolerant und echtzeitfähig) Rechnen sowie internetbasierte verlässliche Fernsteuerungen.

Verlässliche eingebettete Systeme

Die derzeitige Entwicklung der Hardware ist im wesentlichen davon geprägt, dass die Leistungsfähigkeit sich nach dem Gesetz von Moore steigert, zugleich aber die Kosten und die physikalische Größe der nötigen Technik weiter sinkt. Das führt dazu, dass einstige High-End-Hardware inzwischen in großen Mengen zu extrem niedrigen Preisen verfügbar ist, und dass auf kleinstem Raum sehr komplexe Funktionen realisiert werden können.

Diese Entwicklung führt auf der einen Seite zu immer leistungsfähigeren Computersystemen, sie erlaubt es aber auch, leistungsfähige Hardware in eingebetteten Systemen einzusetzen.

Ergebnis ist eine ständig steigende Vielfalt derartiger eingebetteter Systeme, deren Einsatzgebiet von Steueranlagen in Fahr- und Flugzeugen bis hin zu persönlichen Geräten wie Organizern und Mobiltelefonen reicht.

Den meisten dieser Einsatzgebiete ist eine sehr hohe Anforderung an die Zuverlässigkeit gemein, in vielen Fällen führt das Versagen solcher Systeme zu lebensbedrohlichen Situationen für Menschen oder zu hohen wirtschaftlichen Schäden.

Auf dem Gebiet der eingebetteten Systeme konzentrieren sich die Forschungen unserer Gruppe auf Hardware- und Softwarelösungen, um Fragen der Fehlertoleranz, der Echtzeitfähigkeit, des Energieverbrauchs, der Kommunikationsmöglichkeiten und des Umganges mit beschränkten Ressourcen zu beantworten.

Responsive Hochleistungs-Rechnersysteme

Unter dem Titel "High-Performance Responsive Computing" führen wir Forschungsarbeiten auf dem Gebiet der Responsiven Systeme und dem Gebiet der Dienstgütegarantien für Standardsysteme durch. Unser Augenmerk richtet sich auf Fehlertoleranz, Echtzeitfähigkeit, Sicherheit und Performance als Dienstgütecharakteristika.

Responsive Systeme erreichen Fehlertoleranz und Echtzeitverhalten durch Redundanz in Raum und Zeit. Workstations in einem Netz werden als unabhängige Einheiten mit unabhängigem Fehlverhalten angenommen, die gleichzeitige oder wiederholte Ausführung eines Programms auf verschiedenen Maschinen kann also als Maßnahme zur Tolerierung von Fehlern vorgesehen werden.

Gleichzeitig können die miteinander kommunizierenden Workstations als paralleles System angesehen werden. Der durch Parallelverarbeitung in einer solchen Umgebung mögliche Performancegewinn kann dabei benutzt werden, um den *overhead* des responsiven Systems zu verbergen.

Projekt: Komponierbarkeit eingebetteter Echtzeitsysteme

Ansprechpartner: DIPL.-INF. JAN RICHLING

Zusammenarbeit: DR. LOUCHKA POPOVA-ZEUGMANN, LFE Logik in der Informatik
DR.-ING. MATTHIAS WERNER, Arbeitsgruppe Kommunikations- und Betriebssysteme, TU
Berlin, DaimlerChrysler

Durch den zunehmenden Einsatz sehr komplexer, komponentenbasierter Software gibt es einen großen Bedarf an Verfahren, die beschreiben, wie mit dieser Komplexität umgegangen werden kann. Konzepte der strukturierten Programmierung, der Objektorientierung und der komponentenbasierten Softwareentwicklung haben in diesem Zusammenhang sehr an Gewicht gewonnen. Ihnen ist gemeinsam, dass sie in erster Linie die funktionalen Aspekte von Schnittstellen zwischen verschiedenen Teilen oder Komponenten betrachten.

In den meisten Fällen ist das ausreichend, nicht jedoch, wenn der Einsatz von Computersystemen in eingebetteten Umgebungen betrachtet wird. Für solche Systeme sind begrenzte Ressourcen und strenge Anforderungen an das zeitliche Verhalten typisch. Es genügt in einem solchen Szenario für das Zusammenfügen zweier Komponenten nicht mehr, wenn ihre Schnittstellen funktional zusammen passen, es ist zusätzlich erforderlich, dass auch nicht-funktionale Aspekte berücksichtigt werden. Nichtfunktionale Aspekte gehen im Gegensatz zu funktionalen Aspekten nicht aus dem Code der Komponente hervor und sind von der Ausführungsumgebung und Interaktionen mit anderen Komponenten abhängig.

In der Praxis ist es von Bedeutung, Echtzeitsysteme aus Komponenten so zusammensetzen zu können, dass Komponenteneigenschaften auch im System verfügbar sind, und es möglich ist, neue Eigenschaften, die bei solchen Kompositionen auf Systemebene auftauchen (etwa End-zu-End-Zeiten) im voraus zu bestimmen.

Insbesondere in der Entwicklung eingebetteter Echtzeitsysteme stellen Test und Verifikation einen hohen Anteil der Entwicklungskosten dar, die durch die Möglichkeit einer solchen Komponierbarkeit erheblich reduziert werden können.

Dieses Problem wird als die "Komponierbarkeit von Echtzeitsystemen in Bezug auf das zeitliche Verhalten" identifiziert und untersucht.

Folgende Schwerpunkte werden dabei betrachtet:

- Bestimmung des Begriffes "Komponierbarkeit" als Eigenschaft einer Systemarchitektur
- Beschreibung von Komponenten in ihren funktionalen und nichtfunktionalen Eigenschaften (z.B. Beschreibungssprache)
- Entwicklung einer Architektur für eingebettete Echtzeitsysteme, die Komponierbarkeit unterstützt

Der Begriff Komponierbarkeit für die Komposition in Bezug auf nichtfunktionale Eigenschaften wurde in Zusammenarbeit mit Dr. Matthias Werner klarer gefasst und „Komponierbarkeit“ als Eigenschaft einer Systemarchitektur (unter der ein Satz von Regeln zum Bau von Systemen verstanden wird) definiert. Diese Betrachtungsweise hat den Vorteil, dass nichtfunktionale Eigenschaften wie das zeitliche Verhalten bereits beim Entwurf einer (komponierbaren) Architektur berücksichtigt und verifiziert werden können, so dass die spätere Konstruktion von Systemen (nach den Regeln der Architektur) die betreffenden Eigenschaften per Konstruktion hat. Auf diese Weise wird Entwicklungs- und Verifikationsaufwand vom Design eines einzelnen Systems hin zum Design einer Systemarchitektur (die für eine ganze Klasse von Systemen anwendbar ist) verschoben.

Unter dem Namen "Message Scheduled System" (MSS, s. Abb. 1) wurde eine gemischt ereignis- und zeitgesteuerte Architektur für eingebettete Systeme entworfen und in Teilen verifiziert. MSS unterstützt Komponierbarkeit in der beschriebenen Art und erfordert dabei ein beschränktes Vorauswissen über Details zeitlicher Abläufe im System.

Im aktuellen Berichtszeitraum fokussierten sich die Arbeiten auf zwei Schwerpunkte:

- Die in der Vergangenheit begonnene komponierbare Spezifikation von MSS und deren Analyse mit Hilfe von Timed Petrinetzen wurde in Zusammenarbeit mit Dr. Matthias Werner und Dr. Louchka Popova-Zeugmann weiterentwickelt. Im Rahmen dieser Arbeiten ist eine Analysetechnik für die Nichterreichbarkeit in prioritäts-behafteten Timed Petrinetzen entstanden und publiziert worden. Diese Technik wurde nicht nur auf konkrete endliche Netze angewandt, sondern auch auf eine Klasse beliebig großer Petrinetze, die durch Parameter beschrieben wird und Teilaspekte der Architektur MSS allgemein beschreibt. Diese Methode soll im weiteren Verlauf des Projektes auf die gesamte Architektur MSS angewandt werden, womit die Verifikation von MSS auf rein formalem Wege erreicht wird.
- Das oben vorgestellte allgemeine Konzept der „Komponierbarkeit in Bezug auf nichtfunktionale Eigenschaften“ wurde in Zusammenarbeit mit Dr. Matthias Werner weiterentwickelt und formalisiert.

Das Projekt wird fortgesetzt. Dabei sollen die entwickelten Methoden sowohl auf MSS, aber auch auf andere Systemarchitekturen angewandt werden.

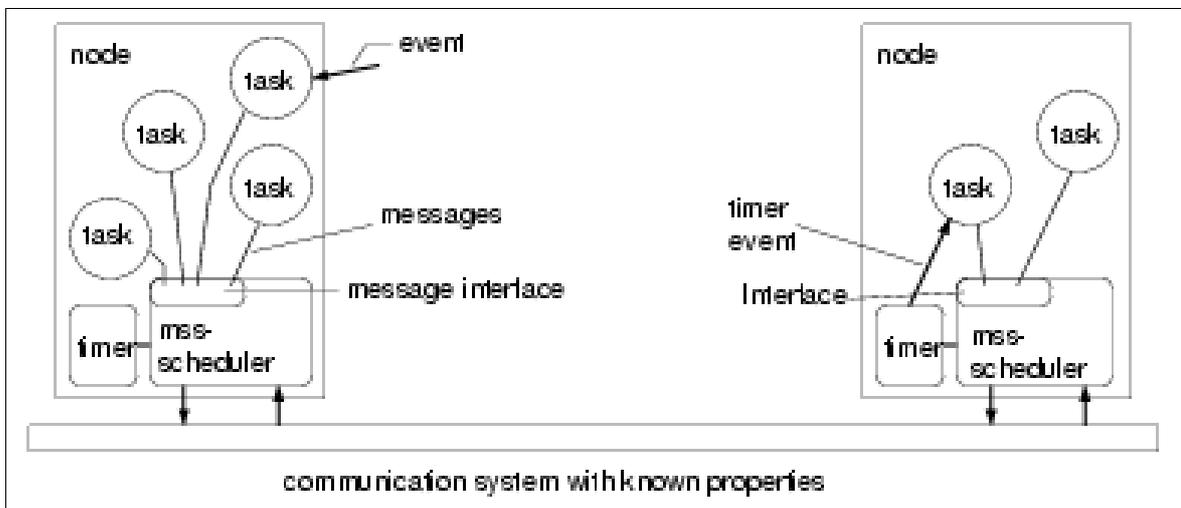


Abbildung 1: Message Scheduled System

Projekt: DISCOURSE – <http://www.discourse.de>**Ansprechpartner:** JAN RICHLING, VLADIMIR STANTCHEV**Zusammenarbeit:** Sechs Arbeitsgruppen an drei anderen Universitäten von Berlin/Brandenburg**Forschungsförderung:** Microsoft Corporation

Technische Universität Berlin

- Prof. Kurt Geihs (Intelligent Networks and Management of Distributed Systems, TU-iVS)
- Prof. Hans-Ulrich Heiß (Communication and Operating Systems Group, TU-KBS)
- Prof. Adam Wolisz / Dr. Holger Karl / Dr. Günther Schäfer (Telecommunication Networks Group, TU-TKN)

Freie Universität Berlin

- Prof. Klaus-Peter Löhr (Software Engineering and Systems Software Group, FU-SESS)
- Prof. Jochen Schiller (Computer Systems and Telematics Group, FU-CST)

Universität Potsdam, HPI (Hasso-Plattner-Institut)

- Prof. Andreas Polze (Operating Systems and Middleware Group, HPI-OSM)

DISCOURSE (*DIStributed & COLlaborative University Research & Study Environment*) ist ein verteiltes Laboratorium für verteiltes Rechnen unter Benutzung fortgeschrittener Middleware-Technologie. Das Labor erstreckt sich über die vier Universitäten der Region Berlin/Potsdam und wird gemeinsam von sieben Forschungsgruppen an sechs geographischen Standorten betrieben. Diese Gruppen nutzen es als Testbed für die Forschung auf den Gebieten der Interoperabilität, der Zuverlässigkeit, der Mobilität und der Dienstgüte, sowie als eine realitätsnahe Referenzplattform für die Lehre.

Im Rahmen von DISCOURSE liegt der Fokus der Arbeiten der LFE „Rechnerorganisation und Kommunikation“ auf nichtfunktionalen Eigenschaften wie Performance, Verfügbarkeit, Fehlertoleranz, Komponierbarkeit, Echtzeit und Sicherheit (sowohl Security, als auch Safety). Zu diesem Zwecke ist der Aufbau eines „Virtual Presence and Experience Laboratory“ (ViPEL) als Teil des DISCOURSE-Netzwerkes geplant, in dem reale Experimente (z.B. Robotik) den Kooperationspartnern über das Netz verfügbar gemacht werden.

Im Berichtszeitraum gab es folgende Aktivitäten in Lehre und Forschung:

- Weiterer Ausbau des VPN
- Vom 18.2. bis 21.2. 2003 wurde von allen beteiligten Forschungsgruppen eine gemeinsame Blocklehrveranstaltung für Studierende aller vier Universitäten zum Thema .NET durchgeführt
- Im Sommer 2003 wurde von allen beteiligten Gruppen in Kooperation mit Microsoft ein Programmierwettbewerb auf Basis der DISCOURSE-Infrastruktur und .NET-Technologie für Studierende aller vier Universitäten durchgeführt
- Ab September 2003 ist das DISCOURSE-Netz in den Lehrbetrieb (Praktikum) der Kurse „Zuverlässige Systeme“ und „Eigenschaften Mobiler und Eingebetteter Systeme“ integriert
- Versuche zur entfernten Steuerung von Experimenten
- Integration in Studien- und Diplomarbeiten

Projekt: CERO - CE Robots Community**Ansprechpartner:** JAN RICHLING**Beteiligte Mitarbeiter:** VLADIMIR STANTCHEV, NIKOLA MILANOVIC**Forschungsförderung:** Microsoft Corporation

Roboter sind ein klassisches Beispiel für eingebettete Systeme mit sehr konkreten nichtfunktionalen Anforderungen. Das Problem bei universitärer Robotik-Forschung sind die erheblichen Kosten, die bereits für einfache Ausrüstung erforderlich sind. Unser besonderes Interesse gilt jedoch Gruppen von Robotern, die miteinander kommunizieren, was den Aufwand zusätzlich in die Höhe treibt.

Idee bei CERO ist es darum, Robotik auf Basis von Standard-Komponenten zu betreiben, um auf diese Weise sowohl bereits existierende Ansätze übertragen zu können, als auch neuartige Ansätze mit vertretbarem Aufwand untersuchen zu können.

CERO nutzt dazu den Umstand, dass es Betriebssysteme wie Microsoft Windows CE gibt, die sowohl auf PDAs, als auch auf eingebetteten Boards lauffähig sind. Diese kann man mit flexibler modularer Robotik-Hardware (LEGO-Mindstorm) verbinden.

Ziel der Forschung in CERO ist es, unter Benutzung von Adhoc-Routingverfahren eine Gruppenkommunikation unter Robotern zu etablieren, die sowohl die Möglichkeiten der Mobilität, als auch des adaptiven Verhaltens ausnutzt.

Projekt: MONARCH – Modelle und Architekturen für Zuverlässige Systeme**Ansprechpartner:** VLADIMIR STANTCHEV**Kooperationspartner:** Microsoft Research, Cambridge, UK.

Forschung im Bereich von service-basierten Umgebungen gewinnt immer mehr an Bedeutung. Am Lehrstuhl ROK (Rechnerorganisation und Kommunikation) wird im Rahmen des NOMADS (Network of Mobile Adaptive Dependable Systems) Projekts die Architektur einer neuartigen service-basierten Umgebung gestaltet. Dabei wird auch die Integration vorhandener Technologien behandelt, insbesondere Component Frameworks wie .NET und EJB.

Das MONARCH Projekt wird die Integration solcher Frameworks (speziell .NET) untersuchen. Insbesondere wird die Frage untersucht, ob sich das .NET Framework als Kommunikations- und Interaktionsnetzwerk für die NOMADS Umgebung eignet. Weiterhin werden Interoperabilitätskonzepte erprobt, die die Zusammenarbeit zwischen .NET Komponenten und NOMADS Services ermöglichen sollen.

RES.NET – Responsive .NET

Unter dem Titel RES.NET untersuchen wir im Rahmen des MONARCH Projekts neue Ansätze zur Anbindung von responsiven virtuellen Experimenten in Enterprise Komponenten-Frameworks wie .NET und EJB für den Einsatz in verteilten Szenarien. Diese verwalteten Umgebungen bieten viele Vorteile für die Entwicklung und Erweiterung der Systemlogik und ermöglichen eine standard-basierte Kommunikation zwischen den einzelnen Stellen mit Hilfe von Web-Services. Die Responsivität des Gesamtsystems verlangt von den einzelnen Komponenten Fehlertoleranz und Echtzeitverhalten. Während Komponenten-Frameworks Fehlertoleranz durch Redundanz im Raum gewährleisten können, sind sie wegen ihres verwalteten Speichermodells nicht in der Lage ohne weiteres Echtzeitgarantien einzuhalten. Ziel unseren Forschungsaktivitäten ist eine .NET basierte Softwareumgebung für responsive Dienste zu entwickeln, deren Konzepte sich auch auf EJB

anwenden lassen. Dabei wenden wir Konzepte wie Synchronisation mittels Konsensalgorithmen, Interprozeß-kommunikation via Remoting und Web-Services, sowie auch aspektorientierte Programmierung an. Im aktuellen Berichtszeitraum erweiterten wir unsere vorhandenen Prototypen mit den Funktionalitäten eines contract negotiating Frameworks. So sind die einzelnen Komponenten in der Lage gewisse funktionale und nichtfunktionale Eigenschaften untereinander auszuhandeln. Derzeit arbeiten wir an einer Architektur-erweiterung, die es ermöglichen soll, gewisse funktionale und nichtfunktionale Eigenschaften von .NET Komponenten automatisch auszuwerten und in dem von uns spezifizierten Format (CDL – Contract Definition Language) abzuspeichern.

Responsive Remote Control of Virtual Presence Experiments

Innerhalb des MONARCH Projekts verwenden wir die RES.NET Umgebung für den speziellen Fall der Remote-Control-Experimente (durch Zustandsmaschinen modelliert). Sie gewährleistet die Kohärenz der Präsenz am Remote Site mit der physikalischen Präsenz. Die Kontrollinfrastruktur wurde auch um die Funktionalitäten erweitert, die die Erstellung und Verarbeitung von CDL Spezifikationen erlaubt. Somit können unsere Remote Control Experimente in einer service-basierten Umgebung wie NOMADS integriert werden.

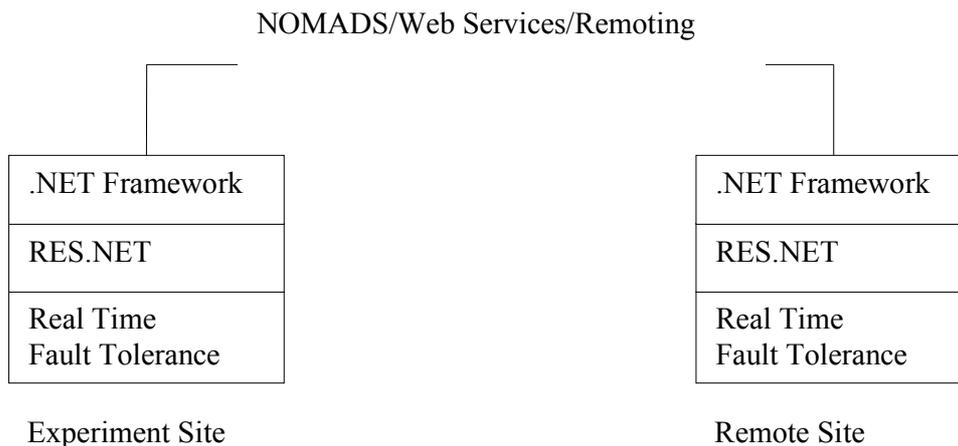


Abbildung 2: RES.NET als zusätzliche Schicht, die Responsivität gewährleistet

Projekt: Quality-of-Service Management in Kommunikationssystemen

Ansprechpartner: DR. KATINKA WOLTER

Zusammenarbeit: Graduiertenkolleg „Stochastische Modellierung und quantitative Analyse großer Systeme in den Ingenieurwissenschaften“, TU Berlin, FU Berlin, HU Berlin

Forschungsförderung: Deutsche Forschungsgemeinschaft unter dem Kennzeichen Wo 989/1-1. Bewilligt im Juli 2003, Projektbeginn 2004.

Service-Level-Agreements (SLAs) sind eine Form, in der vertragliche Bedingungen bezüglich Preis und Qualität eines erbrachten Dienstes festgehalten werden. In Netzwerken ist es schon seit einiger Zeit üblich, solche Verträge abzuschließen. Mit der Verbreitung des e-business in den letzten Jahren wurde auch im Bereich des e-Business Quality-of-Service ein zunehmend wichtiges Thema. Hier geht es dann um Qualitätsgarantien, die viel mehr als nur eine Netzwerkverbindung betreffen.

In diesem Projekt sollen Service-Level-Agreements für Netzwerke mit teils mobilen Nutzern definiert werden. Es sollen Verfahren entwickelt werden, mit denen es möglich ist für verschiedene Teilnehmer zu verschiedenen Zeiten Quality-of-Service Garantien zu vergeben und einzuhalten.

Eine modellbasierte Leistungsbewertung soll anschließend Aufschluss darüber geben, welche Formulierung eines Service-Level-Agreements den größten Nutzen für alle Vertragsparteien bringt. Dabei geht man davon aus, dass Verletzungen eines SLAs mit Kosten (Strafen) verbunden sind, die möglichst vermieden werden sollten.

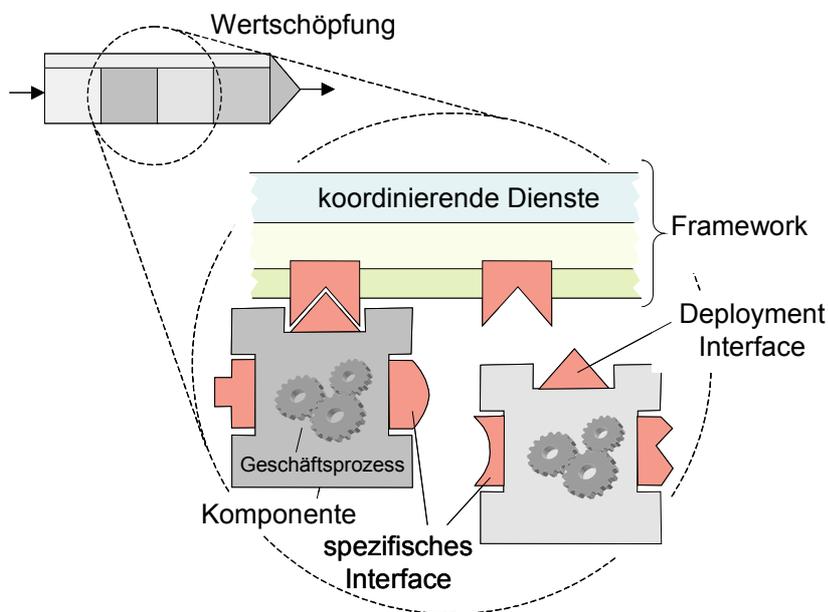
Projekt: Optimierung des Antwortzeitverhaltens durch Restart

Ansprechpartner: DR. KATINKA WOLTER

Zusammenarbeit: Forschungsvorhaben „Quality-of-Service in rekonfigurierbaren, teilweise drahtlosen Rechensystemen: Definition und Bewertung“, gefördert von der DFG unter dem Kennzeichen Wo 989/1-1. Graduiertenkolleg „Stochastische Modellierung und quantitative Analyse großer Systeme in den Ingenieurwissenschaften“, TU Berlin, FU Berlin, HU Berlin.

Die Situation des downloads von Seiten aus dem Internet steht exemplarisch für eine ganze Klasse von Problemen in denen Antwortzeitverhalten von Interesse ist und von unüberschaubaren Faktoren abhängt. Dies sind sowohl Themen in Zusammenhang mit Kommunikation über das Internet, wie allgemeine Dienste im Internet, oder Agenten, als auch Fragestellungen der Terminierung von Algorithmen. In diesem Projekt wird das Thema einerseits sehr allgemein und theoretisch und andererseits konkret für die einfache Anwendung des downloads von Seiten im Internet behandelt.

Es werden zahlreiche, teils wahrscheinlichkeitstheoretische, teils netzwerkpraktische Fragen rund um das oben dargestellte Szenario untersucht. Dies sind auf der praktischen Seite u.a. Warum treten die langen



Wartezeiten auf? An welcher Stelle des Internet, oder beteiligter Server liegt das Problem. Darauf folgt sofort die Frage, wie sich dies vermeiden lässt, oder mehr ingenieurwissenschaftlich, wie das Problem möglicherweise umgangen werden kann. Weiterhin beschäftigt uns, ob sich reproduzierbare Szenarien finden lassen.

Da man nun weiß, dass 'reload' hilft, stellt sich sofort die Frage, wann man den 'reload' Knopf drücken sollte, um die gewünschte Seite möglichst schnell zu erhalten und des

weiteren, ob sich dieses Vorgehen nicht automatisieren lässt. Zu diesem Zweck haben wir einen Proxy-Server implementiert, der automatische reloads durchführt und mit dem wir Experimente zur Ladezeit unter Mehrfachausführung einer Anfrage durchführen.

Als Stochastiker fragen wir nach statistischen Eigenschaften der Ladezeiten, die sicherlich stochastischen Einflüssen unterliegen und sich nur als stochastische Größen adäquat beschreiben lassen. Wir interessieren uns für die Verteilung der Ladezeiten ohne und mit 'reload', für erwartete Ladezeiten ohne und mit 'reload' und für eine optimale Wahl des 'reload' Zeitpunkts. Wir haben einen einfachen Algorithmus zur Bestimmung des optimalen 'reload' Zeitpunktes entwickelt, der in dem bereits genannten Proxy implementiert ist.

Graduiertenkolleg "Stochastische Modellierung und quantitative Analyse großer Systeme

Sprecher: PROF. HOMMEL, TU Berlin

Stipendiat: DIPL.-ING. FELIX SALFNER

Die Zuverlässigkeit von großen Softwaresystemen gewinnt mit fortschreitender Abhängigkeit von solchen Systemen an Bedeutung. Ein Ansatz zur Verbesserung der Zuverlässigkeit ist, Ausfälle des Systems oder Teilen des Systems vorherzusagen, um präventive Maßnahmen ergreifen zu können. Im Rahmen dieses Forschungsvorhabens wird eine Methode entwickelt werden, die eine solche Vorhersage für große Softwaresysteme ermöglicht. Die verwendete stochastische Modellierung verfolgt den bislang wenig beachteten Ansatz, das zeitliche Fehlverhalten des Systems zu modellieren. Grundlage des Modells ist eine zeitkontinuierliche, zustandsdiskrete Markov-Kette, die Muster von Ereignissen, die in der Vergangenheit zu Systemausfällen geführt haben, modelliert. Die Arbeit wird im Zusammenhang mit dem Siemens Projekt durchgeführt.

Projekt: Early Fault Detection for @vantage Platform

Ansprechpartner: DIPL.-ING. FELIX SALFNER

Beteiligte Mitarbeiter: DIPL.-ING. STEFFEN TSCHIRPKE, DIPL.-INF. GÜNTHER HOFFMANN

Forschungsförderung: Siemens ICM

Für den Bereich der Telekommunikation wird von der Fa. Siemens ein komplexes komponentenbasiertes Softwaresystem entwickelt. Zu den wichtigsten Zielen bei der Entwicklung dieses Systems gehören Eigenschaften wie Fehlertoleranz, Verfügbarkeit und die Einhaltung von Zeitanforderungen. Als Mittel zur Verbesserung dieser Eigenschaften werden Verfahren entwickelt um Fehler frühzeitig zu erkennen und Ausfälle vorherzusagen. Damit sollen Ausfall- und Wartungszeiten minimiert werden. Es werden zeitgleich drei Verfahren entwickelt, die auf unterschiedlichen Wegen dieses Ziel zu erreichen suchen: Eine Modellierung des Systems anhand von analytisch gewonnenem Systemwissen. Hierbei werden die Kommunikationsstrukturen des zugrundeliegenden Komponentensystems dynamisch ermittelt und aus Fehlermeldungen wird auf globale Systemzustände geschlossen. Das zweite Verfahren, das auf Markov-Ketten basiert, modelliert das zeitliche Auftreten von Fehlermustern. Die Muster werden algorithmisch aus Fehlermeldungen des Systems gewonnen. Eine Vorverarbeitung der Eingangsdaten erfolgt ähnlich zum ersten Ansatz. Das dritte Verfahren basiert auf Funktionsapproximation mit radialen Basisfunktionen. Im ersten Schritt wird eine Bewertung der aus den Systemmeldungen gewonnenen Parameter vorgenommen. Die wichtigsten werden anschließend dazu verwendet, die Wahrscheinlichkeitsfunktion des Auftretens eines Ausfalls zu approximieren. All diesen Arbeiten liegen unter

anderem Fehlermodelle und Fehlerverzeichnisse zugrunde, die im Rahmen des Projekts entwickelt wurden.

Projekt: Architekturen und Geschäftsmodelle für die Internet-Ökonomie

Ansprechpartner: PETER K. IBACH (ibach@informatik.hu-berlin.de)

Zusammenarbeit: Freie Universität Berlin, Technische Universität Berlin

Forschungsförderung: BMBF

Der Berliner Forschungsverbund „InterVal – Internet and Value Chains“ (<http://interval.hu-berlin.de>) untersucht in einer Synthese technologischer und ökonomischer Gesichtspunkte Grundlagen, Ursachen und Wirkungen des Internets. Im Rahmen des Forschungsverbunds konzentriert sich das Teilprojekt am Lehrstuhl Rechnerorganisation und Kommunikation auf dienstbasierte Architekturen. Auf der Grundlage komponierbarer Dienste versprechen solche Architekturen effizienter auf wechselnde unternehmensinterne und -externe Anforderungen reagieren zu können. Unterschiedliche Ansätze dienstbasierter Architekturen sollen analysiert, Verbesserungen entwickelt und in praxisnahen Szenarien ihre Eignung als Infrastruktur einer Internet-Ökonomie überprüft werden. Spezieller Fokus des Teilprojektes gilt dabei:

- Entwurf von Methoden zur Modellierung und Bewertung von dienst- und komponentenbasierten Architekturen hinsichtlich Handhabbarkeit, Skalierbarkeit, Sicherheit, Verfügbarkeit, Echtzeitfähigkeit und effektivem Ressourcenmanagement
- Integration und Abbildung nichtfunktionaler Eigenschaften
- Entwurf bzw. Erweiterung von Schnittstellen und Protokollen, um die Ad-Hoc Kombination von Diensten zu erleichtern

Systemintegration bzw. -migration: Wie können bestehende Systeme in neue Architekturkonzepte integriert bzw. migriert werden?

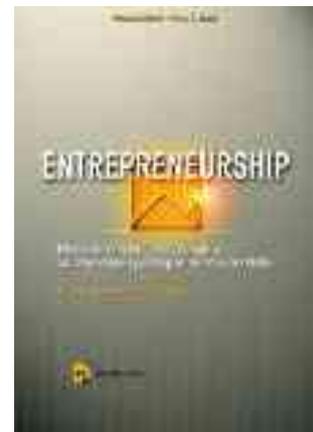
Geschäftsprozesse werden in einer Software-Hülle verpackt (Wrapping). Über weitgehend standardisierte Schnittstellen erhalten die Prozesse eine einheitliche Erscheinungsform und können in einer Rahmenarchitektur (Framework), die von der spezifischen Einsatzumgebung abstrahiert, als Dienst aufgerufen werden. Dadurch werden Voraussetzungen geschaffen, Softwarekomponenten unternehmensübergreifend, interoperabel und komponierbar zu machen und plattform- bzw. umgebungsunabhängig wiederzuverwenden.

Buch: „Entrepreneurship - Prinzipien, Ideen und Geschäftsmodelle zur Unternehmensgründung im Informationszeitalter“, dpunkt.verlag, Januar 2004

Seit 1996 unterrichten wir am Lehrstuhl neben den Themen der technischen Informatik unter dem Titel „High-Tech Entrepreneurship“ auch Prinzipien, Ideen und Geschäftsmodelle zur technologieorientierten Unternehmensgründung. Das Buch „Entrepreneurship“ fasst unsere Erfahrungen dieser Lehrveranstaltungen zusammen. Es vermittelt nicht nur die klassischen Gründerthemen wie Geschäftsplanung, Finanzierung und Marketing, sondern auch die Erneuerungsprozesse einer Innovationsökonomie, die Förderung und Steuerung von Technologietransfer und Innovationsdiffusion, Internet- bzw. E-Business-Geschäftsmodelle, sowie durch Informationstechnologien induzierte gesellschaftliche und ökonomische Veränderungen.

Es ist als Lehrbuch zur inter fakultativen Entrepreneurship-Ausbildung ausgelegt. Wir freuen uns, wenn es den „Entrepreneurial Spirit“ in Deutschland und die Synthese von Humboldt und Schumpeter an den Hochschulen fördert. Denn fundierte Entrepreneurship-Ausbildung kann nicht nur Erfolgsaussichten von Unternehmensgründungen sowie Wissenstransfer und wirtschaftliche Entwicklung verbessern. Der Rückkopplungseffekt zwischen Forschung und Praxis war schon immer auch ein Faktor für den herausragender Erfolg einiger Universitäten wie Stanford, Cambridge oder des MIT.

Weitere Informationen zum Buch mit Inhaltsangabe, Inhaltsverzeichnis und Leseproben haben wir bereitgestellt unter: www.entrepreneurship-par-excellence.de



Veröffentlichungen

N. MILANOVIC, V. STANTCHEV, J. RICHLING, M. MALEK: *Towards Adaptive and Composable Services*. Proceedings of IPSI2003, Sveti Stefan, Montenegro 2003.

M. MALEK: *Towards Dependable Networks of Mobile Arbitrary Devices – Diagnosis and Scalability*. Future Directions in Distributed Computing, Vol 2548, Lecture Notes in Computer Science, A. Schiper, A. Shvartsman, H. Weatherspoon and B.Y. Zhao, Eds., Springer 2003, 191-196.

M. MALEK: *The NOMADS Republic*. Proceedings of International Conferences on Advances in Infrastructure for Electronic Business, Education, Science, Medicine, and Mobile Technologies on the Internet, Scuola Superiore G Reiss Romoli (SSGRR), Telecom Italia, L'Aquila, Italy, 2003.

M. WERNER, J. RICHLING, N. MILANOVIC, V. STANTCHEV: *Applying Composability to Dependable Embedded Systems*. Proceedings of the International Workshop on Dependable Embedded Systems at the 22nd Symposium on Reliable Distributed Systems (SRDS 2003).

M. WERNER, L. POPOVA-ZEUGMANN, J. RICHLING: *Non-Reachability in Priority Duration Petrinets*. Proceedings of CS&P 2003.

L. POPOVA-ZEUGMANN, M. WERNER, J. RICHLING: *Using State Equation to Prove Non-reachability*. Timed Petrinets Fundamentae Informatica, 55(2003)2, 187-202.

FELIX SALFNER, MIROSLAW MALEK: *Early Fault Detection and Failure Prediction in Large Software Systems*. IBM Academy Conference on Proactive Problem Prediction, Avoidance, and Diagnosis; Yorktown Heights, USA; April 2003.

K. WOLTER, K. KASPROWICZ: *WebAppLoader: A Simulation Tool Set for Evaluating Web Application Performance*. Proc. Computer Performance Evaluation -- Modelling Techniques and Tools; P. Kemper, W. H. Sanders (Eds.); pp. 47—62; Urbana, IL, USA, LNCS 2794, Springer, Sept. 2003.

A. VAN MOORSEL, K. WOLTER: *Optimization of Failure Detection Retry Times*. Proc. Performability Workshop", Monticello, IL, USA, Sept. 2003.

A. VAN MOORSEL, K. WOLTER: *Analysis and Algorithms for Restart*. (submitted for publication, Dec. 03).

P. REINECKE, A. VAN MOORSEL, K. WOLTER: *A Measurement Study of the Interplay between Application Level Restart and Transport Protocol*. (submitted for publication, Jan. 04).

N. MILANOVIC, M. MALEK, A. DAVIDSON, V. MILUTINOVIC : *Routing and Security in Mobile Ad Hoc Networks*, IEEE Computer, February 2004.

F. SALFNER, S. TSCHIRPKE, M. MALEK: *Comprehensive Logfiles for Autonomic Systems*. Proc. 9th IEEE Workshop on Fault-Tolerant Parallel, Distributed and Network-Centric Systems, Santa Fe, New Mexico, USA, April 2004.

N. MILANOVIC, J. RICHLING, M. MALEK: *Lightweight Services for Embedded Systems*. Proc. of 2nd IEEE Workshop on Software Technologies for Embedded and Ubiquitous Computing Systems, Vienna, Austria, May, 2004.

M. MALEK, P. IBACH: *Entrepreneurship, Prinzipien, Ideen und Geschäftsmodelle zur Unternehmensgründung im Informationszeitalter*. dpunkt.verlag, Januar 2004.

Vorträge

M. MALEK: *The NOMADs Republic*. 43rd IFIP WG 10.4 Meeting, Sal, Kapverden, 7. Januar 2003

M. MALEK: *Eingebettete Systeme - die Herausforderung für Kreativität*. Berlin, 28. April 2003

F. SALFNER: *Early Fault Detection and Failure Prediction in Large Software Systems*. IBM Academy, IBM Yorktown Heights, NY, USA, 28. April 2003

M. MALEK: *Erhöhung der Verfügbarkeit und Dienstgüte der @vantage Plattform*. Siemens, Berlin, 12. Mai 2003

M. MALEK: *Models and Tools for Dependability Evaluation*. Technische Universität Berlin, 15. Mai 2003.

M. MALEK: *The NOMADS Republic (Networks of Mobile, Adaptive, Dependable Systems)*. SSGRR, Telekom-Italia, L'Aquila, Italien, 26. Juli 2003

M. MALEK: *Fünf Thesen für Softwarestandort Berlin aus der akademischen Perspektive*. Innovationsdialog, Investitionsbank Berlin, 24. September 2003

M. MALEK: *The NOMADS Republic (Networks of Mobile, Adaptive, Dependable Systems)*. Ringvorlesung, Freie Universität Berlin, 11. November 2003

M. MALEK: *Berlin - the City of Innovation, Science and Entrepreneurship*. Hosei University, Japan, 3. Dezember 2003 (remote)

M. MALEK: *The NOMADS Republic (Networks of Mobile, Adaptive, Dependable Systems)*. Ringvorlesung, Humboldt-Universität zu Berlin, 4. Dezember 2003

Kooperationen

- Graduiertenkolleg *Stochastische Modellierung und quantitative Analyse großer Systeme in den Ingenieurwissenschaften* (gemeinsame Einrichtung des Fachbereichs Informatik der Freien Universität Berlin, des Fachbereichs Informatik und des Fachbereichs Ingenieurwissenschaften der Technischen Universität Berlin, des Instituts für Mathematik und des Instituts für Informatik der Humboldt-Universität Berlin)
- Technische Universität Berlin, Prof. Kurt Geihs (Intelligent Networks and Management of Distributed Systems, TU- iVS)
- Technische Universität Berlin, Prof. Hans-Ulrich Heiß (Communication and Operating Systems Group, TU-KBS)
- Technische Universität Berlin, Prof. Adam Wolisz / Dr. Holger Karl / Dr. Günther Schäfer (Telecommunication Networks Group, TU-TKN)
- Freie Universität Berlin, Prof. Klaus-Peter Lühr (Software Engineering and Systems Software Group, FU-SESS)
- Freie Universität Berlin, Prof. Jochen Schiller (Computer Systems and Telematics Group, FU-CST)
- Universität Potsdam, HPI (Hasso-Plattner-Institut), Prof. Andreas Polze (Operating Systems and Middleware Group, HPI-OSM)
- Hewlett-Packard Laboratories, Palo Alto, USA
- DaimlerChrysler Forschung Berlin
- Università di Pisa, Italien
- University of Illinois at Urbana-Champaign, USA
- T. J. Watson Laboratory, IBM, Yorktown Heights, USA
- Lucent Technologies, Bell Laboratories, Murray Hill, USA
- Microsoft Research, Cambridge, UK
- Consiglio Nazionale di Ricerca, Pisa, Italien
- New York University, USA
- Stanford University, USA
- Siemens AG

Sonstige Aktivitäten

Prof. Miroslaw Malek

- Program Co-Chair: 22nd Symposium on Reliable Distributed Systems, Florenz, Italien, 6. – 8. Oktober 2003
- Mitglied der Steering Committee Dependable Systems and Networks
- Mitherausgeber Journal of Real-Time Systems

- Chair, Dependable Systems, AlaRI, Lugano, March 2003
- Cochair, Entrepreneurship and High-Tech Workshop, Wolfsburg, Juli 2003
- Cochair, Autumn School on Entrepreneurship for High-Tech Startups, Università della Svizzera italiana, Lugano, October 2003

Gäste am Lehrstuhl

PROF. DR. ST. OLARIU, Old Dominion University

Diplomarbeiten

SVEN HERSCHEL: *Web Content Annotation Using Aspect-Oriented Programming Techniques*, April 2003.

STEPHAN ECKART: *Generierung feingranularer Zugriffskontrolle für verteilte Komponentensysteme*, Mai 2003.

Lehr- und Forschungseinheit

Signalverarbeitung und Mustererkennung

<http://www.informatik.hu-berlin.de/Institut/struktur/signalverarbeitung/>

Leiterin

PROF. DR. BEATE MEFFERT
E-Mail: meffert@informatik.hu-berlin.de

Sekretariat

SABINE DZIWI SZ
Tel.: (030) 2093 3044

Mitarbeiter

DR.-ING. MANFRED GÜNTHER
LOTHAR HEESE
DR.-ING. OLAF HOCHMUTH
DIPL.-INF. UWE KNAUER
DIPL.-INF. AXEL WEISS
DR.-ING. FRANK WINKLER

DIPL.-ING. KARL-HEINRICH HAUPTVOGEL
DR.-ING. THOMAS MORGENSTERN
DR.-ING. MICHAEL RITZSCHKE
PRIV. DOZ. DR.-ING. DIETRICH SCHILDER

Doktoranden

DIPL.-ING. XIAOYUN CHEN
DIPL.-ING. MATHIAS HAASE
DIPL.-INF. DAVID KRUTZ
DIPL.-ING. THOMAS LOEWEL
DIPL.-ING. HOLGER MAASS
DIPL.-INF. STEFAN MAYER
DIPL.-INF. KATJA WORM

Tutoren

ROMAN BLASCHEK
SASCHA DESCHLER
MARCUS EHRIG
ERIK FISCHER
SEBASTIAN HINZ
NICKY HOCHMUTH
MATTHIAS PREISS
MARCUS RÄTZEL
MICHAEL SCHAEFFER
THOMAS SCHERSCHMIDT
RAINER SCHNABEL
ROBERT STEIN

Das Fachgebiet „Signalverarbeitung und Mustererkennung“, vertreten durch die gleichnamige Professur innerhalb der Technischen Informatik, befasst sich in Lehre und Forschung mit der Erfassung, Verarbeitung und Auswertung von Signalen unterschiedlicher Dimension. Ein Schwerpunkt der Forschungsaktivitäten des Jahres 2003 war die Entwicklung von Hardwarelösungen für die Signalverarbeitung unter Echtzeitbedingungen mit Hilfe von FPGA (field programmable gate arrays). Um die Vorteile der FPGA für die Signalverarbeitung effektiv nutzen zu können, sind Algorithmen erforderlich, die gute Möglichkeiten der Parallelisierung und eine Flexibilität der Dimensionierung bieten.

Ein weiterer Forschungsschwerpunkt war die Entwicklung von Algorithmen zur Bildauswertung unter Einbeziehung biometrischer Merkmale.

Auch die Analyse von Biosignalen zur Unterstützung der Therapiekontrolle ist - in Zusammenarbeit mit dem Institut für Rehabilitationswissenschaften - Gegenstand der Forschung.

In der Lehre werden neben den Grundlagen der Signalverarbeitung Lehrveranstaltungen zur Bildverarbeitung, Mustererkennung und zum Entwurf von Hard- und Software für die digitale Signalverarbeitung angeboten.

Lehre

Grundstudium

- Technische Informatik 1 (F. WINKLER, K.-H. HAUPTVOGEL, D. SCHILDER, A. WEISS)
- Einführung in die Informatik (K.-H. HAUPTVOGEL, M. RITZSCHKE)
- Einführung in die Photonik (D. SCHILDER)
- Praktische Informatik I – Praktikum (M. RITZSCHKE)

Hauptstudium (Halbkurse)

- Grundlagen der Signalverarbeitung (B. MEFFERT, O. HOCHMUTH, D. SCHILDER)
- Signalverarbeitung (B. MEFFERT, O. HOCHMUTH)
- Bildverarbeitung - Spezialgebiete (B. MEFFERT, O. HOCHMUTH)
- Schaltkreisentwurf (F. WINKLER)
- Prozessinformatik (K.-H. HAUPTVOGEL, M. RITZSCHKE)
- Übertragungstechnik (D. SCHILDER)

Mathematisches Ergänzungsfach

- Grundlagen der Signalverarbeitung (B. MEFFERT, O. HOCHMUTH)

Seminare

- Umweltmonitoring (B. MEFFERT)
- Photonische Netze (D. SCHILDER)

Projekte

- Biosignalverarbeitung im Snoezelen-Labor (B. MEFFERT, K. MERTENS)
- Programmierung von digitalen Signalprozessoren (A. WEISS)
- Parallele Signalverarbeitung mit FPGA (F. WINKLER)

Ringvorlesung „Eingebettete Systeme“

- Systemkomponenten für ein Bluetooth-Positioniersystem (F. WINKLER)

Forschung

Projekt: Algorithmen und Implementation von Bildverarbeitungsfunktionen auf FPGA

Ansprechpartner: DR.-ING. FRANK WINKLER

Beteiligte Mitarbeiter: DR.-ING. OLAF HOCHMUTH, PROF. DR.-ING. BEATE MEFFERT

Zusammenarbeit: Deutsches Zentrum für Luft- und Raumfahrt DLR, Berlin-Adlershof

Für die Steuerung und Kopplung von mehreren Bilddatenquellen und für die Signalverarbeitung unter Echtzeitbedingungen kommen zunehmend Hardwarelösungen auf der Basis von FPGA in Betracht. Die traditionellen Nachteile, wie geringe Gatterkomplexität, hoher Entwicklungsaufwand und uneinheitliche Schnittstellengestaltung, werden durch die Fortschritte der Halbleitertechnik und der Entwurfssoftware zunehmend reduziert, wobei die Vorteile, wie der hohe Grad an Parallelität, die Flexibilität der Struktur und die geringe Taktfrequenz/Verlustleistung, immer mehr zum Tragen kommen. Ziel des Projekts ist es, parallelisierbare Algorithmen zu finden und sie so zu modifizieren, dass eine flexible Umsetzung in FPGA-Strukturen möglich ist. Die Algorithmen werden mit Hardwarebeschreibungssprachen formuliert. An beispielhaften Implementierungen wird die Effizienz der Lösungen für verschiedene Sprach- und Synthesetools bewertet. Die Funktion des Entwurfes wird durch Simulation, Funktionstests im Labor und im Feldversuch nachgewiesen.

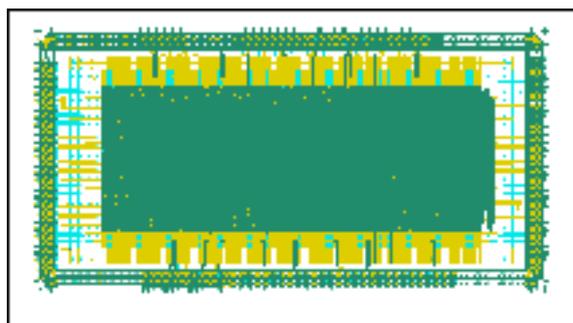
Projekt: Entwicklung eines digitalen Moduls zur Korrelation von Bluetooth-Signalen mit einer für Positionsbestimmungen geeigneten Zeitauflösung

Ansprechpartner: DR.-ING. FRANK WINKLER

Beteiligte Mitarbeiter: LOTHAR HEESE, DR.-ING. THOMAS MORGENSTERN, PRIV. DOZ. DR.-ING. DIETRICH SCHILDER

Zusammenarbeit: IHP GmbH - Innovations for High Performance Microelectronics, Frankfurt/Oder

Zur Positionsbestimmung mobiler Bluetooth-Netzteilnehmer ist eine genaue zeitliche Lokalisation der eintreffenden Funksignale erforderlich. Dazu wird im Projekt als Kernbaugruppe ein schneller digitaler Korrelator entwickelt, der die Zeitdifferenz zwischen zwei zeitverschobenen, aber sonst gleichen Signalen bestimmt. Die Realisierung erfolgt mit Hardwarebeschreibungssprachen so, dass eine Implementation auf unterschiedlichen Technologien möglich ist. Prototypisch wurde eine FPGA-Lösung entwickelt, die im Zusammenspiel mit einem PC erste Messungen zur Beurteilung der Leistungsfähigkeit ermöglichen. Aufbauend auf diesen Ergebnissen erfolgte eine Integration in 0,25- μ m-CMOS-Technologie und die Realisierung eines Funktionsmusters einer Basisstation zur Positionsbestimmung.



Layout des Bluetooth-Korrelator-Chips

Projekt: Einsatz der Biometrik in der Bildverarbeitung**Ansprechpartner:** PROF. DR. BEATE MEFFERT**Beteiligte Mitarbeiter:** DIPL.-INF. UWE KNAUER, SASCHA DESCHLER, SEBASTIAN HINZ, MATTHIAS PREISS, MICHAEL SCHAEFFER, DR.-ING. OLAF HOCHMUTH, DR.-ING. THOMAS MORGENSTERN, DR.-ING. MICHAEL RITZSCHKE**Zusammenarbeit:** Lüth & Dümchen Automatisierungsprojekt GmbH, Berlin**Forschungsförderung:** Investitionsbank Berlin (IBB)

Das Projekt ist Teil eines längerfristigen Vorhabens „Sicherheitsmanagement mit Bildauswertung“. In der ersten Projektphase wurden Verfahren zur Identifikation und Verifikation einzelner Personen anhand biometrischer Merkmale untersucht. Es wurden Algorithmen zur Erkennung des Gesichts und der Gesichtsmerkmale entwickelt und ihre Leistungsfähigkeit verglichen. Derzeit erfolgt die Integration der Algorithmen in das BDE-System des Kooperationspartners. Ziel der nächsten Projektphase ist die Entwicklung einer Bildverarbeitungskomponente zur Erkennung sicherheitsrelevanter Ereignisse in kameraüberwachten Räumen und ihre Integration in ein modulares Gesamtsystem.

Projekt: Integration einer 3D-Herzdarstellung in die Darstellung des elektrischen Herzfeldes**Ansprechpartner:** DR.-ING. OLAF HOCHMUTH**Beteiligte Mitarbeiter:** PROF. DR. BEATE MEFFERT, DIPL.-PHYS. M. WEIHRAUCH, A. TRIANTAFYLLOU**Zusammenarbeit:** Deutsches Herzzentrum Berlin**Forschungsförderung:** DFG-Forschungsprojekt „Verfahren zur Darstellung von EKG-Maps auf der mit Ultraschallmessungen ermittelten Herzoberfläche“ am Deutschen Herzzentrum Berlin (DR.-ING. G. KRENZKE)

Die Untersuchungen unterstützen die Forschungsarbeiten im Rahmen des DFG-Forschungsprojekts, das von einem interdisziplinären Wissenschaftlerteam bearbeitet wird. Mit Hilfe eines EKG-Mapping-Systems wird das elektrische Feld des Herzens ermittelt und auf einer Kugeloberfläche dargestellt. In diese Darstellung des elektrischen Herzfeldes soll lagerichtig das Herz des Patienten projiziert werden, dessen Lage und Form mit Hilfe einer Ultraschallmessung ermittelt wurde.

Die Aufgabenstellung ist für die programmtechnische Lösung in mehrere Teilschritte zerlegt worden. Diese sind im einzelnen:

- interaktive Messwertinterpolation für ausgefallene Ableitelektroden
- Korrektur von Basislinienschwankungen
- 50-Hz-Tiefpaß mit besonderem Einschwingverhalten
- Korrektur der Signalschwächung durch die Lunge.

Projekt: Bildhafte Identifizierung von postalischen Objekten**Ansprechpartner:** PROF. BEATE MEFFERT**Beteiligte Mitarbeiter:** DR.-ING. OLAF HOCHMUTH**Zusammenarbeit:** Siemens Postautomation Berlin-Adlershof

Bearbeiterin: DIPL.-INF. KATJA WORM (Siemens Postautomation Berlin-Adlershof)

Schwerpunkte bei der Bearbeitung des Projekts sind derzeit ein Vergleich der Leistungsfähigkeit verschiedener Orthogonaltransformationen und Segmentierungsverfahren für die

Identifizierung postalisch relevanter Objekte. Bei der Entwicklung der Erkennungsalgorithmen stellt die Echtzeitfähigkeit eine besondere Herausforderung dar.

Projekt: Biosignalverarbeitung im Snoezelenlabor

Ansprechpartner: PROF. BEATE MEFFERT

Beteiligte Mitarbeiter: DIPL.-ING. KARL-HEINZ HAUPTVOGEL, DR.-ING. OLAF HOCHMUTH

Zusammenarbeit: Institut für Rehabilitationswissenschaften, Abteilung Körperbehindertenpädagogik (PROF. KRISTA MERTENS)

Biosignale können über den Zustand eines Menschen Auskunft geben. Die Gewinnung von Informationen aus Biosignalen bedarf der Anwendung von Werkzeugen der digitalen Signalverarbeitung. Das Projekt hatte zum Ziel, diese Werkzeuge anzuwenden und an einem praktischen Beispiel die Probleme der Signalerfassung, -verarbeitung und -auswertung zu lösen. Die zu verarbeitenden Biosignale sind an hyperaktiven Kindern gewonnen worden, die sich einmal wöchentlich einem speziellen Entspannungs- und Aufmerksamkeitstraining im Snoezelenlabor unterzogen. Insbesondere ging es darum, aus den Biosignalen Kenngrößen zu extrahieren, die es ermöglichen, den Entspannungszustand eines Probanden zu quantifizieren.

Projekt: Vergleich der Extraktion von Objekten aus Landsat ETM- und Quickbird-daten

Ansprechpartner: PROF. BEATE MEFFERT, PROF. PATRICK HOSTERT (Geographisches Institut der HU Berlin)

Beteiligte Mitarbeiter: DR.-ING. THOMAS MORGENSTERN, RONNY SCHULZ

Zusammenarbeit: Geographisches Institut der HU Berlin, Abteilung Geofernerkundung

Ziel des Projekts ist ein Vergleich von Methoden, die bei der Extraktion von Objekten aus Satellitenbildern unterschiedlicher Auflösung zur Anwendung kommen. Die zu verwendenden Klassifikationsverfahren sind nicht unabhängig von der Auflösung der Satellitenbilder. Die Untersuchungen umfassen die Erarbeitung von Kriterien für die Wahl der Klassifikationsverfahren und die Festlegung von Auflösungsgrenzen für die Erkennung unterschiedlicher Objektkategorien (Häuser, Straßen, Flüsse usw.).

Projekt: Entwurf von Teststrukturen zur Inbetriebnahme hochintegrierter Schaltkreise

Ansprechpartner: DR.-ING. FRANK WINKLER

Beteiligte Mitarbeiter: LOTHAR HEESE, DR.-ING. THOMAS MORGENSTERN

Zusammenarbeit: IHP GmbH - Innovations for High Performance Microelectronics, Frankfurt/Oder

Der Test hochintegrierter Schaltkreise und der Vergleich ihrer Parameter mit Simulationsergebnissen oder FPGA-Prototypen erfordert eine übergreifende Entwicklungsmethodik und daran angepasste Teststrukturen. Während der Entwurfsprozess weitgehend automatisiert ist, erfordert die Zusammenfügung aller Komponenten zu ersten technologischen Mustern in der Regel kostenintensive Mehrfachentwürfe des Systems.

Am Beispiel eines I²C-Busses und eines WLAN-Prozessors werden Teststrukturen entwickelt, die mit heterogener Konfiguration, z. B. wahlweiser FPGA- und ASIC-Bestückung und variabler Rechner- oder Mikrocontrollerunterstützung arbeiten können.

Projekt: Plattform-unabhängige Lösungen für die Online-Signalverarbeitung**Ansprechpartner:** DIPL.-INF. AXEL WEISS**Beteiligte Mitarbeiter:** DR.-ING. FRANK WINKLER, LOTHAR HEESE**Zusammenarbeit:** SINUS Messtechnik GmbH, Leipzig; Infineon AG, München; Altium GmbH, Karlsruhe; DResearch GmbH, Berlin

Die Online-Verarbeitung von Sensorsignalen ist eine wichtige Teilaufgabe von eingebetteten Systemen und digitalen Mess- und Übertragungssystemen. Für deren Implementierung werden in zunehmendem Maße verteilte und heterogene Systeme entwickelt, um den wachsenden Performance- und Kostenanforderungen gerecht zu werden.

Zur Aufbereitung von Sensorsignalen und deren Weiterverarbeitung zum Zweck der Komprimierung, Mustererkennung und Ereignisklassifikation sind viele Standard-Algorithmen bekannt. Einerseits werden für die schnelle Abarbeitung von Kernoperationen (parallele Arithmetik, akkumulierte Multiplikationen, bitreverse Indizierung u.a.) spezielle Prozessoren (DSPs) entwickelt, die eine prozessor-gebundene Programmierung erfordern. Andererseits ist die generische Formulierung der Standard-Algorithmen in Hochsprachen möglich, die die Wiederverwendbarkeit und Zuverlässigkeit garantieren. In der Regel sind aber die Compiler nicht in der Lage, die generische Formulierung auf die Prozessor-Fähigkeiten abzubilden, und die Performance der in einer Hochsprache formulierten Algorithmen ist um Größenordnungen schlechter als das Optimum.

Ziel des Projekts ist die Entwicklung einer Infrastruktur zur Programmierung signalverarbeitender Systeme, deren Lösungen unabhängig von der Plattform sind, auf der sie umgesetzt werden. Elementare signalverarbeitende Algorithmen sollen generisch formuliert werden, um Wiederverwendbarkeit und Zuverlässigkeit zu garantieren. Dabei sollen die besonderen Eigenschaften der verwendeten Plattform zur Maximierung der Performance ausgenutzt werden. Die freie, aber konsistente Verknüpfung elementarer signalverarbeitender Algorithmen zu einem signalverarbeitenden System soll mit einfachen sprachlichen Mitteln umgesetzt werden. Multiprozessor-Plattformen sollen ebenso unterstützt werden wie verteilte, heterogene Systeme.

Für den Nachweis stehen verschiedene Plattformen zur Verfügung:

- Desktop-PCs und Workstations mit verschiedenen Betriebssystemen
- Singleboard-DSP-Karte von Integrated Innovations Inc. (SBC32)
- Entwicklungsboard für den TriCore-Prozessor von Infineon (TC1775B)
- Entwicklungsboard für den TriMedia-Prozessor von Philips (TriMeleon von DResearch)
- High-End Multiprozessor-Messkarte von Sinus-Messtechnik (Harmonie-PCI)
- PCI-DSP-Karte von Integrated Innovations Inc. (ADC64)

Wissenschaftliche Kooperationen

- Altium GmbH, Karlsruhe
- Avicom GmbH, Berlin
- Deutsches Herzzentrum Berlin
- Deutsches Zentrum für Luft- und Raumfahrt, Berlin-Adlershof und Neustrelitz
- DResearch GmbH, Berlin
- ESYS GmbH, Berlin
- Geographisches Institut der HU Berlin
- IHP GmbH - Innovations for High Performance Microelectronics, Frankfurt/Oder
- Infineon AG, München
- Institut für Rehabilitationswissenschaften der HU Berlin

- Lüth & Dümchen Automatisierungsprojekt GmbH, Berlin
- PicoQuant GmbH, Berlin
- Stahlwerk Gröditz GmbH
- Siemens Dematic Postdienst - Automatisierung GmbH, Berlin
- SINUS Messtechnik GmbH, Leipzig
- Universität Potsdam, Hasso-Plattner-Institut für Softwaresystemtechnik

Veröffentlichungen

HARMUTH, H.; MEFFERT, B.: *Calculus of Finite Differences in Quantum Electrodynamics*. Series: Advances in Imaging and Electron Physics, Volume 129, ISBN: 0-12-0147718.

MEFFERT, B.; DESCHLER, S.; HINZ, S.; LÜTH, G.: *Biometrik. Ein aktueller Forschungsgegenstand - Exponat auf der CeBIT 2003*. Berlin, Humboldt-Spektrum 10 (2003) Heft 1, S. 42 - 45.

WEISS, A.: *Linux-Treiber API HARMONIE*. Dokumentation, SINUS Messtechnik GmbH, April 2003.

WINKLER, F.: *Interface description of the HU Berlin/IHP correlator chip*. Technical Documentation, IHP Frankfurt/Oder, 11.6.2003.

WINKLER, F.; FISCHER, E.: *Korrelatorchip-Testumgebung auf Basis des 32-bit-Prozessors LEON*. Technical Documentation, IHP Frankfurt/Oder, 29.10.2003.

Poster

KRUTZ, D.; WINKLER, F.: *Konzept eines auf VHDL basierenden Betriebssystems für FPGA-Plattformen*. Poster, 11. E.I.S.-Workshop (Entwurf Integrierter Schaltungen und Systeme), Berlin, 31. März - 1. April 2003.

Vorträge

DÜFFERT, U.; WINKLER, F.: *Untersuchungen zur Klassifikation multispektraler Bilddaten mit FPGA's in Handel-C*. Wissenschaftliches Seminar, DLR, Berlin-Adlershof, 5.12.2003.

KNAUER, U.: *Automatische Vermessung gewalzter Stahlringe für die Qualitätssicherung*. Tag der Informatik, 8. Mai 2003.

KRUTZ, D.: *FPGA Stereo Matching Algorithmus mit Betriebssystem*. Vortrag, Workshop OIS-Software-Review, Deutsches Zentrum für Luft- und Raumfahrt, Berlin, 10. Februar 2003.

MEFFERT, B.: *Snoezelen und seine Wirkung auf Biosignale*. Eingeladener Vortrag, 2. Internationales Symposium der International Snoezelen Association (ISNA), Ede, Niederlande, 9. bis 11. Oktober 2003.

MEFFERT, B.: *Biosignalanalyse*. Workshop, 2. Internationales Symposium der International Snoezelen Association (ISNA), Ede, Niederlande, 9. bis 11. Oktober 2003.

WEISS, A.: *Konzepte für die Verarbeitung fließender Daten*. Tag der Informatik, 8. Mai 2003.

WEISS, A.: *Beiträge des Instituts für Informatik der HU Berlin zum Verifikationsprojekt*. Kolloquium anlässlich des 15-jährigen Jubiläums des Bochumer Verifikationsprojektes (BVP), Bochum, 29. Mai 2003.

WEISS, A.: *Harmonie-Gerätetreiber für LINUX*. Erweitertes ingenieur-wissenschaftliches Seminar der SINUS Messtechnik GmbH, Leipzig, 14. November 2003.

WEISS, A.: *FreeSP - plattformunabhängige Lösungen für die digitale Signalverarbeitung*. Erweitertes ingenieur-wissenschaftliches Seminar der SINUS Messtechnik GmbH, Leipzig, 14. November 2003.

WINKLER, F.: *Echtzeitverarbeitung, Ansätze und Ergebnisse der VHDL-Implementation*. Workshop OIS-Software-Review, Deutsches Zentrum für Luft- und Raumfahrt, Berlin, 10. Februar 2003.

WINKLER, F.: *FPGA-Einsatz in der Bildverarbeitung*. Wissenschaftliches Seminar, DLR, Berlin-Adlershof, 24. 2. 2003.

WINKLER, F.: *Post Processing Requirements for Bluetooth In-Door Positioning Systems*. IPS-Workshop, IHP, Frankfurt/Oder, 6.8.2004.

Sonstige Aktivitäten

Prof. Dr. Beate Meffert

- Organisation des „*Workshop für Stipendiatinnen*“ im Rahmen des Berliner Programms zur Förderung der Chancengleichheit für Frauen in Forschung und Lehre
- Mitherausgeberin der Zeitschrift „Das Hochschulwesen“ des Universitätsverlages Webler
- Mitglied des Kuratoriums der Stiftung „Konsul Karl und Dr. Gabriele Sandmann“ (KKGS-Stiftung)
- Vorsitzende der Gutachterkommission Natur- und Technikwissenschaften des Berliner Programms zur Förderung der Chancengleichheit für Frauen in Forschung und Lehre
- Mitglied der Auswahlkommission für das Berliner Programm zur Förderung der Chancengleichheit für Frauen in Forschung und Lehre
- Mitglied der Kommission für Frauenförderung des Akademischen Senats
- Mitglied der Deutschen Gesellschaft für Biomedizinische Technik
- Mitglied des Deutschen Hochschullehrerinnenbundes

Dr.-Ing. Thomas Morgenstern

- Mitglied der Medienkommission des Akademischen Senats der HU Berlin
- Mitglied des Konzils der HU Berlin

Dr.-Ing. Michael Ritzschke

- Aufsichtsrat der actina Aktiengesellschaft

Diplomarbeiten

BRÜGGERT, STEFFEN: *Klassifikation militärischer Fahrzeuge mit Orthogonaltransformationen*, Dezember 2003.

KNAUER, UWE: *Automatische Vermessung gewalzter Stahlringe für die Qualitätssicherung*, April 2003.

IV. Lehre

SOMMERSEMESTER 2003

GRUNDSTUDIUM

Diplomstudiengang (D)

2. Semester

Praktische Informatik 2 (D, L, M, N)

Vorlesung	4 Std.	H.-D. BURKHARD
Übung	8 Std.	G. LINDEMANN-V. TRZEBIATOWSKI
Übung	6 Std.	M. MINOR
Übung 4 Std.	K. SCHRÖTER	
Übung	2 Std.	A. OSHERENKO
Praktikum		K. AHRENS, B. HOHBERG

Technische Informatik 1 (D)

Vorlesung	2 Std.	F. WINKLER
Übung	2 Std.	F. WINKLER, A. WEISS
Praktikum		K.-H. HAUPTVOGEL, D. SCHILDER
		A. WEISS, F. WINKLER

Mathematik für Informatiker 2 (D)

Vorlesung	4 Std.	L. RECKE
Übung	6 Std.	R. THRUM
Übung	4 Std.	M. ROCZEN
Übung	2 Std.	G. FUCHS

4. Semester

Informatik und Gesellschaft (D)

Vorlesung	2 Std.	W. COY
-----------	--------	--------

Technische Informatik 2 (D, M)

Vorlesung	4 Std.	M. MALEK
Übung	4 Std.	J. RICHLING
Übung	2 Std.	M. Mergner, M. Werner

Theoretische Informatik 3 (D)

Vorlesung	2 Std.	T. NIERHOFF
Übung	4 Std.	D. OSTHUS
Übung	2 Std.	T. NIERHOFF

Proseminare

Wissenschaftliches Publizieren mit LaTeX

Proseminar	2 Std.	M. PIEFEL
------------	--------	-----------

Compilergeneratoren

Proseminar/Seminar	2 Std.	A. KUNERT
--------------------	--------	-----------

Beauty is our Business

Proseminar	2 Std.	W. REISIG
------------	--------	-----------

3-D Computergrafik

Proseminar	2 Std.	R. KUBICA
------------	--------	-----------

Programmverifikation

Proseminar	2 Std.	K. SCHMIDT
------------	--------	------------

Formale Sprachen

Proseminar	2 Std.	L. POPOVA-ZEUGMANN
------------	--------	--------------------

Approximationsalgorithmen

Proseminar	2 Std.	T. NIERHOFF
------------	--------	-------------

Temporale Logik und Model Checking

Proseminar	2 Std.	S. ROCH
------------	--------	---------

Petrinetze

Proseminar/Seminar	2 Std.	S. ROCH
--------------------	--------	---------

Markov-Ketten und Monte-Carlo-Algorithmen

Proseminar	2 Std.	A. TARAZ
------------	--------	----------

Mikroprozessorentwurf

Proseminar	2 Std.	F. WINKLER
------------	--------	------------

Einführung in die Photonik

Proseminar	2 Std.	D. SCHILDER
------------	--------	-------------

Magister – 2. Hauptfach (M) und Magister-Nebenfach (N)

Praktische Informatik 2 (D, L, M, N)

Vorlesung	4 Std.	H.-D. BURKHARD
Übung	8 Std.	G. L INDEMANNV-V. TRZEBIATOWSKI
Übung	6 Std.	M. MINOR
Übung	4 Std.	K. SCHRÖTER
Übung	2 Std.	A. OSHERENKO
Praktikum		K. AHRENS, B. HOHBERG

Technische Informatik 2 (D, M)

Vorlesung	4 Std.	M. MALEK
Übung	4 Std.	J. RICHLING
Übung	2 Std.	M. MERGNER, M. WERNER

Einführung in die Informatik (N)

Vorlesung	2 Std.	M. RITZSCHKE
Übung	1 Std.	M. RITZSCHKE
Übung	2 Std.	K.-H. HAUPTVOGEL
Praktikum	2 Std.	M. RITZSCHKE
Praktikum	3 Std.	K.-H. HAUPTVOGEL

Lehramt (L)

Praktische Informatik 2 (D, L, M, N)

Vorlesung	4 Std.	H.-D. BURKHARD
Übung	8 Std.	G. L INDEMANNV-V. TRZEBIATOWSKI
Übung	6 Std.	M. MINOR
Übung	4 Std.	K. SCHRÖTER
Übung	2 Std.	A. OSHERENKO
Praktikum		K. AHRENS, B. HOHBERG

Analyse, Planung und Beurteilung von Informatikunterricht

Seminar	2 Std.	J. PENON
---------	--------	----------

Unterrichtspraktikum/Blockpraktikum

Praktikum	Block	CH. DAHME
-----------	-------	-----------

Fakultative Veranstaltung

UNIX für Anfänger

Vorlesung	2 Std.	E. HOLZ
-----------	--------	---------

HAUPTSTUDIUM

Kurse/Halbkurse (D, N, L, M)Praktische und angewandte Informatik**Objektorientierte Modellierung in C++ (HK)**

Vorlesung	2 Std.	J. FISCHER
Vorlesung	2 Std.	K. AHRENS
Praktikum		K. AHRENS

UML/SDL (HK)

Vorlesung	2 Std.	J. FISCHER
Vorlesung	2 Std.	M. BORN
Praktikum		M. PIEFEL

Heterogene und komplexe Systemarchitekturen (HK)

Vorlesung	4 Std.	E. HOLZ
Praktikum/Übung		E. HOLZ

Software Engineering (HK)

Vorlesung	4 Std.	K. BOTHE
Vorlesung/Übung	2 Std.	K. BOTHE, K. SCHÜTZLER
Übung	4 Std.	K. SCHÜTZLER

Parallelisierende Compiler (HK)

Vorlesung	4 Std.	R. HÄNISCH / K. BOTHE
-----------	--------	-----------------------

Informatik & Informationsgesellschaft II: Technik, Geschichte und Kontext der Informatik (HK)

Vorlesung	4 Std.	W. COY
Übung/Praktikum	2 Std.	J. KOUBEK

Konzepte für die frühen Phasen der Softwareentwicklung (HK)

Vorlesung	4 Std.	CH. DAHME
-----------	--------	-----------

Moderne Methoden der Künstlichen Intelligenz (HK)

Vorlesung	4 Std.	H.-D. BURKHARD / H.-M. VOIGT
Übung/Praktikum	2 Std.	J. BACH, J. HOFFMANN

Projekt: Kognitive Robotik (HK)

Vorlesung	4 Std.	H.-D. BURKHARD / J. BACH / M. HILD, J. HOFFMANN
Übung/Praktikum	2 Std.	J. BACH / M. HILD / J. HOFFMANN

DBS II: Implementierung von Datenbanksystemen (HK)

Vorlesung	4 Std.	J.-C. FREYTAG
-----------	--------	---------------

Praktikum	4 Std.	C. BEN NECIB
Praktikum	2 Std.	D. SCHEFFNER

Werkzeuge der empirischen Forschung (HK)

Vorlesung	4 Std.	W. KÖSSLER
-----------	--------	------------

Theoretische Informatik

Verteilte Algorithmen (HK – auch Praktische Informatik)

Vorlesung	4 Std.	W. REISIG
Übung	2 Std.	A. ALEXANDER

Einführung in die Informations- und Kodierungstheorie (HK – auch Math. Erg.fach)

Vorlesung	4 Std.	E. RÖDEL
Übung	2 Std.	E. RÖDEL

Lineare Optimierung (HK – auch Math. Erg.fach)

Vorlesung	4 Std.	L. POPOVA-ZEUGMANN
-----------	--------	--------------------

Computergestützte Verifikation (HK)

Vorlesung	4 Std.	K. SCHMIDT
-----------	--------	------------

Graphen und Algorithmen (K – 2. Teil)

Vorlesung	4 Std.	S. HOUGARDY
Übung	2 Std.	A. TARAZ

Komplexitätstheorie (HK)

Vorlesung	4 Std.	O. BEYERSDORFF
-----------	--------	----------------

Technische Informatik

Grundlagen der Signalverarbeitung (HK – auch Math. Erg.fach)

Vorlesung	2 Std.	B. MEFFERT
Übung	2 Std.	O. HOCHMUTH
Praktikum	2 Std.	D. SCHILDER

Prozessinformatik (HK – auch Praktische Informatik)

Vorlesung	2 Std.	M. RITZSCHKE
Übung	2 Std.	M. RITZSCHKE, K.-H. HAUPTVOGEL
Praktikum	2 Std.	M. RITZSCHKE, K.-H. HAUPTVOGEL

Übertragungstechnik (HK)

Vorlesung	3 Std.	D. SCHILDER
Übung	1 Std.	D. SCHILDER
Praktikum	2 Std.	D. SCHILDER

Rechnerkommunikation in IP-Netzen (HK)

Vorlesung	2 Std.	G. DOLLNY
Praktikum	2 Std.	G. DOLLNY

Mathematisches Ergänzungsfach (D)

Einführung in die Informations- und Kodierungstheorie

Vorlesung	4 Std.	E. RÖDEL
Übung	2 Std.	E. RÖDEL

Lineare Optimierung

Vorlesung	4 Std.	L. POPOVA-ZEUGMANN
-----------	--------	--------------------

Grundlagen der Signalverarbeitung

Vorlesung	2 Std.	B. MEFFERT
Übung	2 Std.	O. HOCHMUTH
Praktikum	2 Std.	D. SCHILDER

Pflichtfach (L)

Rechnereinsatz in der Schule

Seminar	2 Std.	R. ERBRECHT
---------	--------	-------------

Informatik und Gesellschaft

Vorlesung	2 Std.	W. COY
Projekt	n. V.	W. COY

Seminare (D, M, L)

Praktische Informatik

XML für Fortgeschrittene

Seminar	2 Std.	O. BECKER
---------	--------	-----------

Multimedia-Middleware, Plattformen und Systeme

Seminar	2 Std.	J. FISCHER
---------	--------	------------

Objektorientierter Entwurf und Analyse

Projekt/Forschungsseminar	2 Std.	J. FISCHER
---------------------------	--------	------------

Compilergeneratoren

Seminar/Proseminar 2 Std. A. KUNERT

Software Engineering

Projekt 2 Std. K. BOTHE

Parallele Algorithmen in der Bioinformatik

Seminar 2 Std. A. REINEFELD / T. STEINKE

Beweis verstehen: Schriftlogik & bildliches Argument

Seminar BLOCK W. COY, J. BRÜNING

Internet-Governance – Wer regiert das Internet?

Seminar BLOCK P. BITTNER

Webdesign

Seminar BLOCK J. KOUBEK

Bewusste evolutionäre Softwareentwicklung

Seminar 2 Std. CH. DAHME

Bioinformatik

Projekt 2 Std. J.-C. FREYTAG

Evolution und Robotik II

Seminar 2 Std. M. HILD

Systementwurf

Seminar 2 Std. W. REISIG

E - Commerce

Seminar 2 Std. A. MARTENS

Komposition verteilter Systeme

Seminar 2 Std. A. ALEXANDER

*Theoretische Informatik***E-Learning – Gegenwart und Zukunft**

Seminar 2 Std. K.-P. NEUENDORF

Sicherheit im Internet

Seminar 2 Std. E.-G. GIESSMANN

Prozessalgebra

Seminar 2 Std. K. SCHMIDT

Petrinetze

Seminar/Proseminar 2 Std. S. ROCH

Algorithmen und Komplexität

Seminar 2 Std. T. NIERHOFF

Semidefinite Programmierung

Seminar 2 Std. A. COJA-OGHLAN

The Strange Logic of Random Graphs

Seminar 2 Std. M. KANG, M. BODIRSKY

Technische Informatik

Biosignalverarbeitung im Snoezelen-Labor (Fortsetzung)

Projekt 2 Std. B. MEFFERT / K. MERTENS

Parallele Signalverarbeitung mit FPGA

Projekt 2 Std. F. WINKLER

Programmierung von digitalen Signalprozessoren

Projekt 2 Std. A. WEISS

Electronic Business Forum

Seminar 2 Std. M. MALEK

Spezielle Probleme des Echtzeitrechnens

Seminar 2 Std. J. RICHLING

Betriebssystemarchitektur an Beispiel von .NET

Seminar 2 Std. V. STANTCHEV

Spezialvorlesungen (D, M)

XML und Datenmodellierung

Vorlesung 2 Std. R. ECKSTEIN

Bioinformatik

Vorlesung 2 Std. J.-C. FREYTAG

Data Warehousing

Vorlesung 2 Std. U. LESER

Übung 2 Std. U. LESER

Molekularbiologische Datenbanken

Vorlesung	2 Std.	U. LESER
Übung	2 Std.	U. LESER

Web Services

Vorlesung	2 Std.	A. MARTENS
-----------	--------	------------

UNIX – Systemadministration

Vorlesung	2 Std.	J.-P. BELL
Praktikum	2 Std.	J.-P. BELL

Studium generale**Schwerpunkte der Informatik**

Vorlesung	2 Std.	J.-C. FREYTAG
-----------	--------	---------------

Olle Kamellen. Ausgewählte Beispiele aus der Technikgeschichte

Vorlesung	2 Std.	W. ORTHBANDT
-----------	--------	--------------

Forschungsseminare/Kolloquium**Forschungsseminare, 2stündig**

Geleitet von: K. BOTHE, H.-D. BURKHARD, J.-C. FREYTAG, J. KÖBLER, M. MALEK, B. MEFFERT, H.-J. PRÖMEL, T. NIERHOFF, W. REISIG, E. RÖDEL, A. REINEFELD

Kolloquium

Kolloquium	2 Std.	N.N.
------------	--------	------

WINTERSEMESTER 2003/2004

GRUNDSTUDIUM

Diplomstudiengang (D)

1. Semester

Praktische Informatik 1 (D, L, M, N)

Vorlesung	4 Std.	H. SCHLINGLOFF
Übung	4 Std.	G. LINDEMANN-V. TRZEBIATOWSKI, M. RITZSCHKE, K. SCHÜTZLER
Übung	2 Std.	B. ALTMEYER, R. KUBICA, K. SCHMIDT
Praktikum		K. AHRENS, B. HOHBERG

Theoretische Informatik 1 (D, M)

Vorlesung	4 Std.	M. GROHE
Übung	10 Std.	L. POPOVA-ZEUGMANN
Übung	4 Std.	S. KREUTZER, N. SCHWEIKHARDT

Mathematik 1 (D, M)

Vorlesung	4 Std.	E. HERRMANN
Übung	4 Std.	E. HERRMANN, M. GRABITZ, I. NOWAK

3. Semester

Praktische Informatik 3 (D)

Vorlesung	4 Std.	J.-C. FREYTAG
Praktikum	3 Std.	A. KUNERT
Praktikum	1 Std.	R. HEESE, E. ECKSTEIN
Übung	2 Std.	E. ECKSTEIN

Theoretische Informatik 2 (D)

Vorlesung	4 Std.	J. KÖBLER
Übung	4 Std.	C. SCHWARZ
Übung	2 Std.	O. BEYERSDORFF, M. FÜSSEL, P. LISKE

Technische Informatik 1 (D)

Vorlesung	2 Std.	F. WINKLER
Übung	2 Std.	F. WINKLER, A. WEISS
Praktikum		M. GÜNTHER, K.-H. HAUPTVOGEL, F. WINKLER

Mathematik 3 (D)

Vorlesung	3 Std.	R. THRUM
Übung	2 Std.	R. LAMOUR, F. NICOLAE, B. GEIGER

Proseminare**Verarbeitung natürlicher Sprache**

Proseminar	2 Std.	K. SCHRÖTER
------------	--------	-------------

Maschinelles Lernen

Proseminar	2 Std.	T. SCHEFFER
------------	--------	-------------

Beauty is our Business

Proseminar	2 Std.	W. REISIG
------------	--------	-----------

Funktionales Programmieren

Proseminar	2 Std.	M. GROHE
------------	--------	----------

Magister – 2. Hauptfach (M) und Magister – Nebenfach (N)**Praktische Informatik 1 (D, L, M, N)**

Vorlesung	4 Std.	H. SCHLINGLOFF
Übung	4 Std.	G. LINDEMANN-V. TRZEBIATOWSKI, M. RITZSCHKE, K. SCHÜTZLER
Übung	2 Std.	B. ALTMAYER, R. KUBICA, K. SCHMIDT
Praktikum		K. AHRENS, B. HOHBERG

Theoretische Informatik 1 (D, M)

Vorlesung	4 Std.	M. GROHE
Übung	10 Std.	L. POPOVA-ZEUGMANN
Übung	4 Std.	S. KREUTZER
Übung	4 Std.	N. SCHWEIKHARDT

Mathematik 1 (D, M)

Vorlesung	4 Std.	E. HERRMANN
Übung	4 Std.	E. HERRMANN, M. GRABITZ, I. NOWAK

Einführung in die Informatik (auch für Diplomnebenfächer)

Vorlesung	2 Std.	M. RITZSCHKE
Übung	1 Std.	M. RITZSCHKE
Übung	2 Std.	K.-H. HAUPTVOGEL
Praktikum	2 Std.	M. RITZSCHKE
Praktikum	3 Std.	K.-H. HAUPTVOGEL

**Einführung in die technische Informatik
(Rechnerorganisation/Betriebssysteme) (N, L)**

Vorlesung	2 Std.	K. WOLTER
Übung	2 Std.	K. WOLTER

Lehramt (L)

Praktische Informatik 1 (D, L, M, N)

Vorlesung	4 Std.	H. SCHLINGLOFF
Übung	4 Std	G. LINDEMANN-V. TRZEBIATOWSKI, M. RITZSCHKE, K. SCHÜTZLER
Übung	2 Std	B. ALTMEYER, R. KUBICA, K. SCHMIDT
Praktikum		K. AHRENS, B. HOHBERG

Einführung in die technische Informatik (Rechnerorganisation/Betriebssysteme) (N, L)

Vorlesung	2 Std.	K. WOLTER
Übung	2 Std.	K. WOLTER

Einführung in die Fachdidaktik (L)

Seminar	2 Std.	J. KOUBEK
---------	--------	-----------

Unterrichtspraktikum/Blockpraktikum (L)

Praktikum	BLOCK	CH. DAHME
-----------	-------	-----------

Fakultative Veranstaltungen

UNIX für Anfänger

Vorlesung	2 Std.	J. BELL
-----------	--------	---------

Einführung in C

Vorlesung/Praktikum	BLOCK	J.-C. FREYTAG
---------------------	-------	---------------

HAUPTSTUDIUM

Kurse/Halbkurse (D, N, L, M)

Praktische und angewandte Informatik

UNIX-Werkzeuge

Vorlesung	2 Std.	J. BELL
Praktikum	2 Std.	J. BELL

UNIX-Architektur

Vorlesung	2 Std.	J. BELL
Praktikum	2 Std.	J. BELL

Objektorientierte Analyse, Design und Implementation [OADI]

Vorlesung	4 Std.	J. FISCHER, K. AHRENS
Praktikum	2 Std.	M. PIEFEL

Software Engineering (HK)

Vorlesung	4 Std.	K. BOTHE
Übung	4 Std.	K. SCHÜTZLER

IA-64 versus Power4: Architektur und effiziente Programmierung (HK)

Vorlesung	2 Std.	A. REINEFELD
Übung	2 Std.	A. REINEFELD / F. HUPFELD / T. SCHÜTT
Praktikum		F. HUPFELD / T. SCHÜTT

Informatik und Informationsgesellschaft I: Digitale Medien (HK)

Vorlesung	4 Std.	W. COY
Übung/Praktikum	2 Std.	J. KOUBEK

Zwischen „Information Sharing“ und „Information Control“ (HK)

Vorlesung	4 Std.	P. BITTNER
Übung/Praktikum	2 Std.	P. BITTNER

Kooperatives Prototyping (HK)

Vorlesung	4 Std.	CH. DAHME
-----------	--------	-----------

Einführung in die Datenbanken/DBS I (HK)

Vorlesung	4 Std.	J.-C. FREYTAG
Praktikum	4 Std.	C. BEN NECIB
Praktikum	2 Std.	S. HERSCHEL, K. THAM

Bioinformatik (HK)

Vorlesung	4 Std.	U. LESER
-----------	--------	----------

Crashkurs Informationssysteme

Vorlesung BLOCK F. NAUMANN

Informationsintegration I (HK-1. Teil)

Vorlesung 2 Std. F. NAUMANN

Einführung in die Künstliche Intelligenz (HK)

Vorlesung 4 Std. H.-D. BURKHARD

Übung 4 Std. M. MINOR

Projekt: Kognitive Robotik II (HK)

Vorlesung 4 Std. H.-D. BURKHARD / J. BACH / M. HILD /
J. HOFFMANN

Übung/Praktikum 2 Std. H.-D. BURKHARD / J. BACH / M. HILD /
J. HOFFMANN

Maschinelle Sprachverarbeitung (HK)

Vorlesung 4 Std. T. SCHEFFER

Übung 2 Std. T. SCHEFFER

Methoden und Modelle des Systementwurfs (HK)

Vorlesung 4 Std. W. REISIG

Geschäftsprozessmodellierung (HK)

Vorlesung 4 Std. A. MARTENS

Theoretische Informatik

Graphen und Algorithmen (K – 1. Teil)

Vorlesung 4 Std. A. COJA-OGHLAN

Übung 2 Std. A. COJA-OGHLAN

Kombinatorik und ihre Anwendungen (HK)

Vorlesung 4 Std. D. OSTHUS

Algorithmen in der Bioinformatik (HK - 1. Teil)

Vorlesung 2 Std. S. HOUGARDY

Übung 2 Std. S. HOUGARDY

Randomisierte Algorithmen und Probabilistische Analyse (HK)

Vorlesung 4 Std. A. TARAZ

Übung 2 Std. A. TARAZ

Kryptologie 1 (HK)

Vorlesung 4 Std. J. KÖBLER

Übung 2 Std. O. BEYERSDORFF

Chipkartentechnologie (HK – 1. Teil)

Vorlesung	2 Std.	E.-G. GIESSMANN
-----------	--------	-----------------

Technische Informatik**Grundlagen der Signalverarbeitung (HK)**

Vorlesung	2 Std.	B. MEFFERT
Übung	2 Std.	O. HOCHMUTH
Praktikum	2 Std.	U. KNAUER

Signalverarbeitung (HK)

Vorlesung	2 Std.	B. MEFFERT
Übung	2 Std.	O. HOCHMUTH
Praktikum	2 Std.	O. HOCHMUTH

Bildverarbeitung/Grundlagen (HK)

Vorlesung	2 Std.	B. MEFFERT
Übung	2 Std.	O. HOCHMUTH
Praktikum	2 Std.	O. HOCHMUTH

Schaltkreisentwurf (HK)

Vorlesung	3 Std.	F. WINKLER
Praktikum	2 Std.	F. WINKLER

Zuverlässige Systeme (HK)

Vorlesung	4 Std.	M. MALEK
-----------	--------	----------

Eigenschaften von mobilen und eingebetteten Systemen (HK)

Vorlesung	4 Std.	J. RICHLING
-----------	--------	-------------

Kommunikation in IP-Netzwerken (HK)

Vorlesung	2 Std.	G. DOLLNY
Praktikum	2 Std.	G. DOLLNY

Spezialvorlesungen (D, M)

Modellierung von Kommunikationssystemen

Vorlesung	2 Std.	K. WOLTER
Projekt	2 Std.	K. WOLTER

Pflichtfach (L)

Hauptseminar Fachdidaktik

Seminar	2 Std.	S. SPOLWIG
---------	--------	------------

Seminare (D, M, L)

Praktische Informatik

Multimedia-Middleware, Plattformen und Systeme

Seminar/Forschungsseminar	2 Std.	J. FISCHER
---------------------------	--------	------------

Objektorientierter Entwurf und Analyse

Seminar	2 Std.	J. FISCHER
---------	--------	------------

Projekt: Softwaresanierung

Projekt	2 Std.	K. BOTHE
---------	--------	----------

Projekt: Softwaresanierung für Neueinsteiger

Projekt	2 Std.	K. BOTHE
---------	--------	----------

Verteilte Datenstrukturen in Peer-to-Peer-Systemen

Seminar		A. ANDRZEJAK / A. REINEFELD
---------	--	-----------------------------

3D-Grafik

Seminar	2 Std.	R. KUBICA
---------	--------	-----------

Algorithmen in der Proteomforschung

Seminar	2 Std.	LFG DBIS
---------	--------	----------

Advanced Data Warehousing

Seminar	2 Std.	U. LESER
---------	--------	----------

Data Cleansing

Seminar	2 Std.	F. NAUMANN
---------	--------	------------

Evolution und Robotik

Seminar 2 Std. M. HILD

Simulation sozialer Systeme

Seminar BLOCK G. LINDEMANN – V. TRZEBIATOWSKI / D. URBIG

Systementwurf (auch Th I)

Seminar 2 Std. W. REISIG

Spezifikation und Verifikation verteilter Systeme (auch Th I)

Seminar 2 Std. A. ALEXANDER

*Theoretische Informatik***Logik und Komplexität**

Seminar 2 Std. M. GROHE

IT-Sicherheit

Seminar 2 Std. J. KÖBLER / M. SCHWAN

Discrete Random Walks: Theory and Applications

Seminar 2 Std. M. KANG

Algorithmen und Komplexität

Seminar 2 Std. A. COJA-OGHLAN

*Technische Informatik***Bildanalyse und Objekterkennung**

Seminar 2 Std. B. MEFFERT / U. KNAUER

Programmierung von digitalen Signalprozessoren

Projekt 2 Std. A. WEISS

High-Tech Entrepreneurship

Seminar 2 Std. M. MALEK

Studium generale**Schwerpunkte der Informatik**

Vorlesung 2 Std. J. FISCHER

Die Epoche der sinusförmigen Signale: Ausgewählte Beispiele aus der Technikgeschichte

Vorlesung 2 Std. W. ORTHBANDT

Forschungsseminare/Kolloquium

Forschungsseminare, 2stündig

Geleitet von: K. BOTHE, H.-D. BURKHARD, W. COY, J.-C. FREYTAG, M. GROHE, U. LESER, J. KÖBLER, M. MALEK, B. MEFFERT, F. NAUMANN, H.-J. PRÖMEL, W. REISIG, T. SCHEFFER

Kolloquium

Kolloquium 2 Std. N.N.

V. Informationstechnik des Instituts für Informatik

Rechnerbetriebsgruppe

<http://www.informatik.hu-berlin.de/rbg>

Leiter

DR. JAN-PETER BELL

Tel.: (030) 2093 3131

e-mail: bell@informatik.hu-berlin.de

Mitarbeiter

DIPL.-ING. (FH) GABRIELE BÄRWOLFF

DIPL.-ING. WOLFGANG GANDRE

DIPL.-ING. PETRA KÄMPFER

FRANK WOZOBULE

Tutoren

N.N.

N.N.

Die Rechnerbetriebsgruppe gehörte zum Lehr- und Forschungsgebiet Systemanalyse.

Im Jahre 2003 vergrößerten sich die Zahl, der durch die Rechnerbetriebsgruppe betreuten Nutzer, nicht wesentlich. Zum Jahresende sind mehr als 2050 Nutzer bei der Informatik registriert. Die seit 1998 auftretenden Disproportionen bei der Auslastung der studentischen Rechnerpools in Adlershof verringerten sich im Wintersemester 2003/2004 trotz weiterhin einer hohen Zahl von Studenten in den ersten Studienjahren. Es traten kaum noch Wartezeiten für Arbeitsplätze in Pools auf. Dies ist vor allem auf die Durchführung der Lehrveranstaltungen (Vorlesungen und Übungen) in Adlershof zurückzuführen. Die Studenten sind fast jeden Tag in Adlershof und nutzen die Pools damit an allen Tagen der Woche und nicht nur wie in den vergangenen Jahren an zwei oder drei Tagen.

Entwicklung der rechentechnischen Ausstattung des Instituts 2003

Wesentlichste Neuerung im zentralen Bereich im Jahre 2003 war die Ablösung des SUN-Sparc5-Pools durch einen SUN-Ray1-Pool mit 20 Arbeitsplätzen. Als Server dafür konnten zwei Server SUN V480 mit je 4 Prozessoren und 16 GByte Hauptspeicher in Betrieb genommen werden. Damit stehen den Studenten jetzt 20 moderne Arbeitsplätze mit hoher Prozessorleistung und großem Hauptspeicher zur Verfügung.

Die rechentechnische Ausstattung in einigen LFGs konnte verbessert werden. Wesentliche Neuerungen sind:

- Erweiterung des PC-Clusters um 6 Knoten im LFG Algorithmen und Komplexität
- Neuausstattung des LFG Logik in der Informatik mit einem SUN-Ray1-Pool (10 Arbeitsplätze und zwei Server SUN V280) und 4 PCs
- Neuausstattung des LFG Wissensmanagement in der Bioinformatik mit 7 PCs und 2 PC-Servern mit je 1 TB Plattenkapazität
- Beschaffung eines neuen PC-Pools mit 15 Arbeitsplätzen und zwei PC-Servern im LFG Signalverarbeitung und Mustererkennung

- Bereitstellung von 6 PCs im LFG Theorie der Programmierung
- Bereitstellung von 12 SUN-Ray1-Arbeitsplätze in diversen LFGs

Erbrachte Dienstleistungen

Die grundlegenden Dienste, wie E-Mail, WWW, FTP, Remote-Einwahl, News, NIS, DNS und zentraler Backup-Service wurden das ganze Jahr über stabil zur Verfügung gestellt. Auf den zentralen Mailservern wurden neue Virens Scanner und neue SPAM-Filter installiert. Die Software des News-Servers wurde aktualisiert. Die Umrüstung der zentralen Server führte nicht zu längeren Unterbrechungen bzw. Datenverlusten.

Auf Grund von Kapazitätsproblemen sind die Zugriffszahlen des FTP-Servers rückläufig. Die Zugriffszahlen auf den NEWS-Server sind bei gleichem Angebot stabil geblieben. Zum Jahresende traten erhebliche Kapazitätsprobleme auf, die erst 2004 beseitigt werden können. Weiterhin rückläufig sind die Zugriffszahlen für die Remote-Einwahl. Hier sind die Ursachen sicher in der preiswerten Einwahlmöglichkeit des RZs über DFN@HOME zu sehen. Der Terminalserver für die analoge Einwahl sollte 2004 abgeschaltet werden.

Das Mailaufkommen ist auf eine Größe von täglich 20.000 bis 40.000 Mails gestiegen. Dieser Wert hat sich, wie schon im Jahre 2002, in diesem Jahr im Vergleich zum Vorjahr fast verdoppelt. Dies ist nicht nur auf SPAM-Mail zurückzuführen, wie die stark gestiegenen Zahlen bei lokaler und extern versendeter Mail belegen.

Die WWW-Zugriffe auf den Institutserver liegen semesterabhängig zwischen 75.000 und 145.000 täglich. Damit haben sich die Zugriffszahlen im Vergleich zum Vorjahr leicht erhöht. Das Gleiche gilt auch für die transportierten Datenmengen. Problematisch erweist sich der Mail-Zugriff über der WWW-Server. Hier verlangsamten sich die Zugriffszeiten wesentlich, da der Dienst sehr häufig genutzt wird und die Prozessorkapazität des WWW-Servers erschöpft ist.

Die Account-Beantragung mittels WWW hat sich bewährt. Dadurch konnten die Einrichtung von Accounts weitgehend automatisiert werden. In der Regel kann auch in den Spitzenzeiten während des Semesteranfangs eine Bearbeitungszeit von maximal 24 Stunden garantiert werden. Die Verlängerung der Accounts für Studenten der Informatik konnte ebenfalls automatisiert werden. Dieses Verfahren hat sich bewährt. Die Kopplung des Account-Management mit dem System Goya wurde realisiert. Alle anfänglichen Probleme bei der Abgleichung beider System konnten gelöst werden.

Die Datenspeicherungsmöglichkeiten der beiden SAN-Server wurden durch die LFGs „Datenbanken und Informationssysteme“, „Rechnerorganisation und -kommunikation“, „Künstliche Intelligenz“ und „Komplexität und Kryptographie“ angenommen.

Mailverkehr 2003:

Mailserver 1 (mail)

<i>Monat</i>		<i>Empfangen extern</i>	<i>Versendet extern</i>	<i>lokale E-Mail</i>
Januar	236.880	123.317	21.768	91.795
Februar	215.343	117.834	16.597	80.912
März	205.598	132.600	23.073	49.925
April	215.210	139.704	23.566	51.940
Mai	289.165	176.758	30.683	81.724
Juni	289.914	171.029	33.530	85.355
Juli	418.888	193.199	45.148	180.541
August	410.649	213.064	54.227	143.358
September	478.354	254.781	65.721	157.852
Oktober	519.139	268.277	53.415	197.447
November	574.477	303.615	55.850	215.012
Dezember	606.807	333.830	76.647	196.330

Mailserver 2 (mailslv1)				
<i>Monat</i>		<i>Empfangen extern</i>	<i>Versendet extern</i>	<i>lokale E-Mail</i>
Januar	214.157	97.726	49.483	66.948
Februar	188.266	88.508	36.348	63.410
März	152.460	98.916	32.125	21.419
April	203.500	111.142	41.898	50.460
Mai	251.299	143.976	51.036	56.287
Juni	261.395	147.134	54.129	60.132
Juli	377.554	162.437	64.232	150.885
August	497.039	199.379	100.214	197.446
September	421.952	206.490	87.997	127.465
Oktober	455.834	218.635	87.135	150.064
November	528.482	245.331	128.259	154.892
Dezember	569.364	282.306	130.657	156.401

Zugriffe zum WWW-Server im Jahr 2003

<i>Monat</i>	<i>Transfer in MByte</i>	<i>Zugriffe</i>
Januar	72.408	2.865.164
Februar	70.834	2.513.986
März	68.522	2.346.237
April	76.754	3.087.322
Mai	77.469	3.203.313
Juni	77.793	3.132.740
Juli	77.832	2.933.482
August	73.576	2.354.940
September	68.760	2.503.287
Oktober	100.268	3.852.416
November	121.492	4.358.618
Dezember	92.244	3.363.538

Zugriffe zum FTP-Server im Jahr 2003

<i>Monat</i>	<i>Dateien</i>	<i>MByte</i>	<i>Systeme</i>	<i>Dateien tägl.</i>	<i>MByte tägl.</i>
Januar	145.566	1.005.710	7.489	4.549	31.428
Februar	132.232	806.070	5.488	4.723	28.788
März	75.218	393.233	5.381	2.507	13.107
April	34.630	106.684	5.195	1.117	3.441
Mai	96.210	167.451	6.351	3.007	5.232
Juni	74.046	93.393	4.893	2.389	3.012
Juli	25.033	72.206	4.180	782	2.256
August	51.899	100.328	5.056	1.622	3.135
September	34.755	80.815	5.596	1.121	2.606
Oktober	72.898	126.104	6.098	2.278	3.940
November	67.171	177.128	6.115	2.167	5.713
Dezember	43.185	85.815	5.378	1.350	2.681

Zugriffe zum News-Server im Jahr 2003

<i>Monat</i>	<i>Empfangene Artikel</i>	<i>Gesendete Artikel</i>	<i>Gelesene Artikel</i>
Januar	2.819.112	1.438.744	10.808
Februar	3.581.897	2.231.466	9.278
März	3.882.690	2.398.624	6.863
April	3.558.389	2.12.210	7.845
Mai	3.590.425	2.085.323	9.791
Juni	3.496.664	2.077.819	15.062
Juli	-----	-----	-----
August	-----	-----	-----
September	3.411.130	1.742.762	11.063
Oktober	3.263.157	1.688.845	11.262
November	3.091.984	1.627.509	10.829
Dezember	3.550.645	1.289.994	6.821

Zugriffe zu den Terminal-Servern im Jahr 2003

<i>Monat</i>	<i>Nutzer analog</i>	<i>Verbind.</i>	<i>Dauer in Stunden</i>	<i>Nutzer ISDN</i>	<i>Verbind.</i>	<i>Dauer in Stunden</i>
Januar	2	2	0:47	143	2384	584:32
Februar	12	107	7:35	111	1964	532:40
März	15	108	8:51	109	1539	429:04
April	17	121	7:12	117	1422	374:41
Mai	18	110	4:16	105	1292	376:35
Juni	4	6	0:23	79	1152	310:53
Juli	4	11	0:27	82	1126	286:37
August	3	7	0:06	65	848	235:33
September	5	8	0:25	70	1097	252:15
Oktober	11	29	0:51	133	1367	356:09
November	8	71	2:42	124	1478	465:31
Dezember	9	72	2:55	126	1292	404:17

Netzwerkbetriebsgruppe

<http://www.informatik.hu-berlin.de/rok>

Leiter

DR. GÜNTER DOLLNY
Tel.: (030) 2093 3032
E-Mail: dollny@informatik.hu-berlin.de

Mitarbeiter

DIPL.-CHEM. BIRGIT SCHIEFNER
DIPL.-INF. STEFFEN TSCHIRPKE
DR. SIEGMAR SOMMER

Die Lehr- und Forschungseinheit „Rechnerorganisation und Kommunikation“ ist seit 1994 mit den Arbeitsanteilen der Netzwerkbetriebsgruppe verantwortlich für die Bereitstellung der Netzwerk-Kommunikations-Dienste unseres Instituts.

Durch die Netzwerkbetriebsgruppe wurden alle arbeitsorganisatorisch notwendigen Aufgaben zur Aufrechterhaltung und Weiterentwicklung der Funktionsfähigkeit auf der Netzwerkebene erfolgreich bearbeitet. In dieser Verantwortlichkeit konnten langjährig grundlegende Voraussetzungen zur Bereitstellung von Kommunikations- Dienstinfrastrukturen im Institut erbracht werden.

Im Berichtsjahr 2003 wurden die physikalischen Netzwerkbereiche VLAN, WLAN und SAN des Instituts-Netzes in ihre veränderte Backbone-Umgebung eingebunden. Die für den Einsatz vorgesehenen neuen Netzwerkkomponenten wurden im Netzwerk-Labor unseres Lehrstuhls vorab getestet. Die Testergebnisse hatten entscheidenden Einfluß auf die Zeit und Kosten effektive Ausführung der Migrationsphasen für Switching (Etagenbereiche) und Routing (Backbone). Damit konnten längere Netzwerkausfälle im Institut und Anforderungen von Installationsleistungen der Vertriebspartner, verbunden mit zusätzlichen Kosten, ausgeschlossen werden.

Bei der perspektivischen Entwicklung von Netzwerk-Kommunikations-Diensten des Instituts (IT-Planung) werden von der Netzwerkbetriebsgruppe folgende Vorschläge unterbreitet, die absehbar geplant und realisiert werden müssen :

im Backbone-Switch-Bereich

Erweiterung der GigaBit Port Kapazität im ausgewogenen Verhältnis zu den Etagen-Bereichen

in den Etagen-Switch-Bereichen

Migration der existenten Super Stack Technologie
auf GigaBit Port Verfügbarkeit (Entlastung Backbone)
auf höhere Portdichten
auf den Einsatz optischer Komponenten(CWDM)

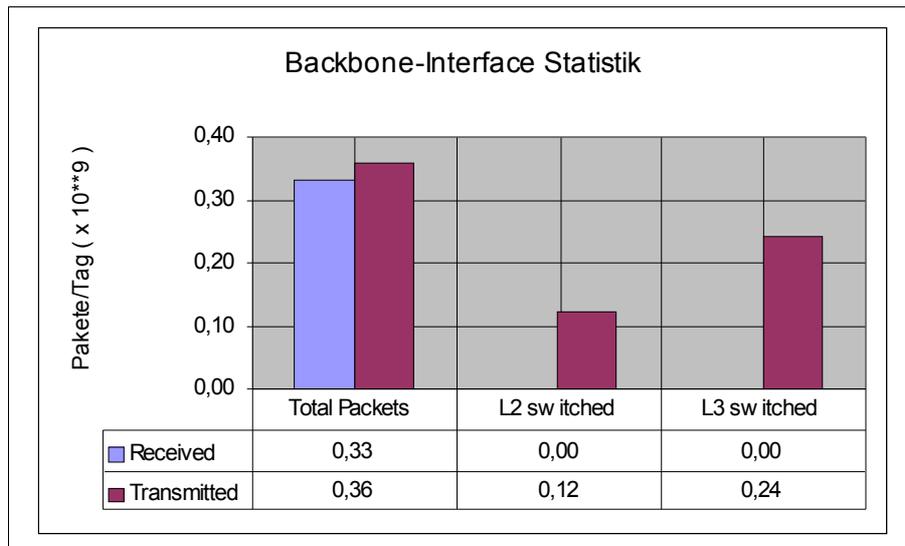
Management-Software

Einsatz eines komplexen Netzwerk-Management-Systems (z.B. HP-Open-View).

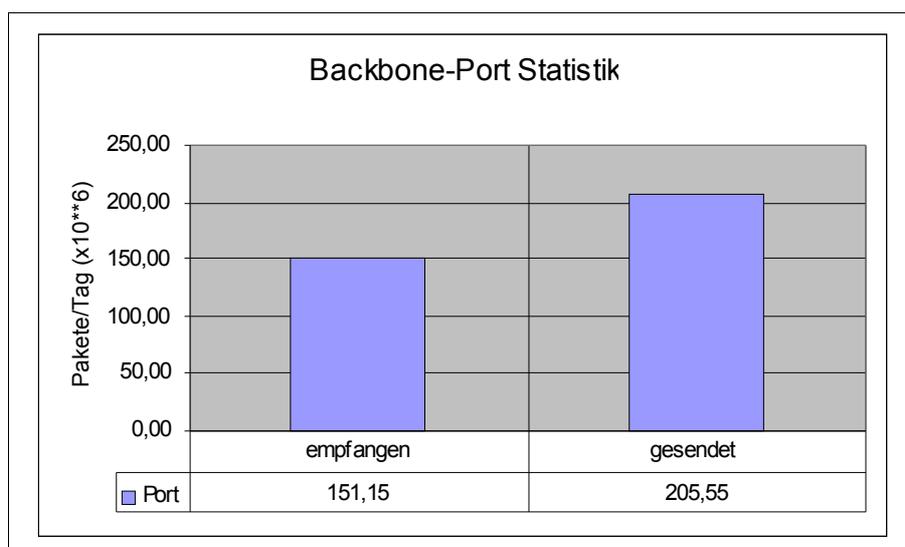
Nach einer Phase von produkt spezifischen „Frühausfällen“ des Backbone-Switch wird im vorliegenden Jahresbericht eine erste Statistik über die Netzwerk-Performance vorgelegt.

Dazu wurden die Durchsatzzahlen an Unicast-Paketen für das Backbone und die VLAN- und Server-Ports zu repräsentativen Produktionszeiträumen des Netzwerks erfasst. Die Statistik-Diagramme zeigen für jeden Bereich die normierte Anzahl der empfangenen und gesendeten Pakete, gemittelt auf einen Tag.

Aus der Backbone-Interface Statistik ist das Verhältnis zwischen Switching (Layer-2) und Routing (Layer-3) Anteilen im Verhältnis von etwa 1 zu 2 festzustellen. Damit ist der interne und externe Datentransfer im Institutsnetzwerk statistisch bewertet.

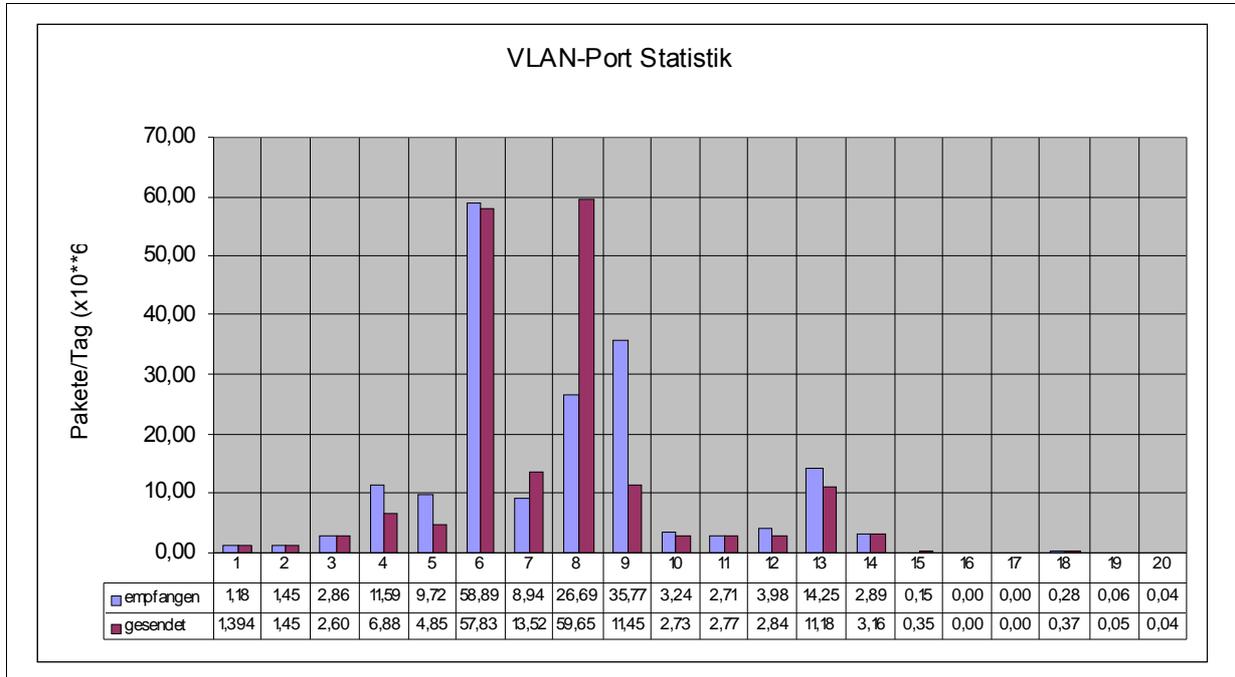


Das Diagramm über die Backbone-Port Statistik zeigt die gemessenen Werte der Unicast-Paket Performance auf dem entsprechenden physikalischen Port.

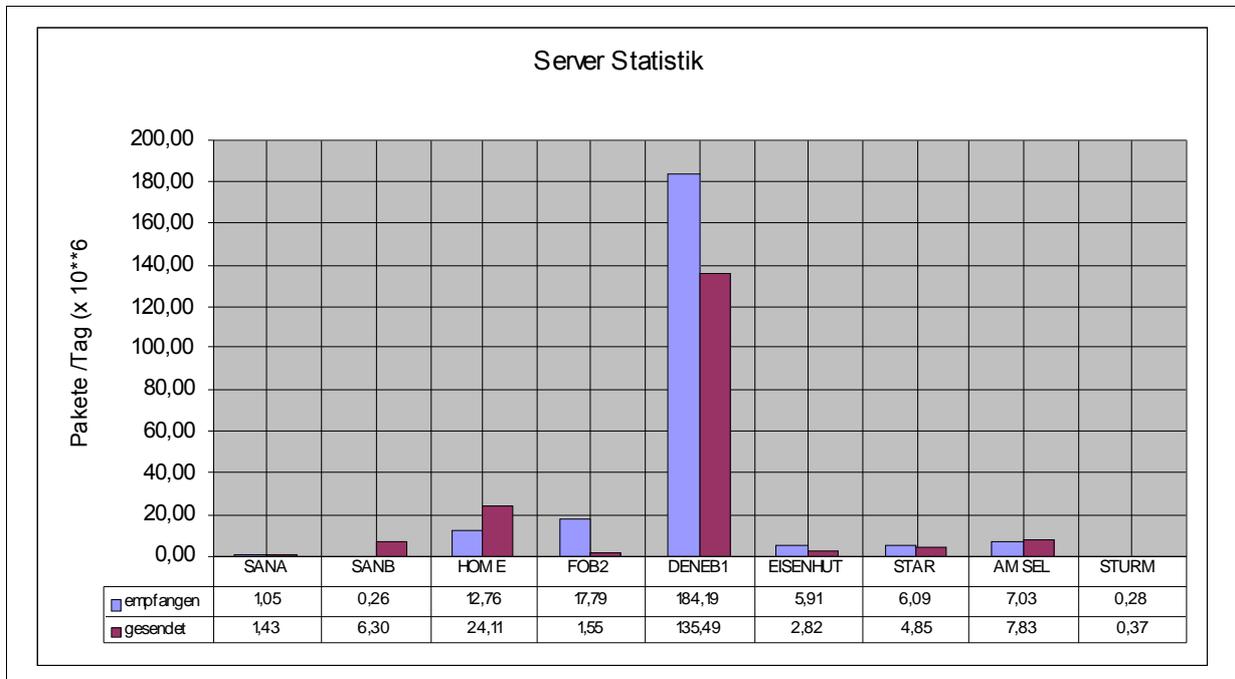


Der Anteil des internen Unicast-Paket Durchsatzes im Instituts-LAN für die VLAN-Ports ist aus dem folgenden Statistik-Diagramm ersichtlich. Die Verbindung zwischen den logischen VLAN-Interfaces und den verwendeten physikalischen VLAN-Ports werden durch den

Monitor des Backbone-Switch nicht ausgewiesen. Die statistische Durchsatz Bewertung eines konkreten VLANs kann dadurch nur über seine Port Verbindung ermittelt werden.



Das Diagramm der Server Statistik weist den Durchsatz an Unicast-Paketen für die zugeordneten Ports des Backbone Switch aus. Die Port Zugehörigkeit der Server ist eindeutig und damit können ihre SNMP-System-Namen in der Tabelle verwendet werden.



"Mehr Frauen in die Informatik!"

Ideen-Werkstatt

zur Studienwahl, zum Studium und zum Beruf am Institut für Informatik der Humboldt-Universität zu Berlin

LEITERIN: DR. MÄRTA GUTSCHE

Das Institut für Informatik hat sich mit dem Projekt „Mehr Frauen in die Informatik!“ zum Ziel gesetzt, bei der Förderung der Chancengleichheit den Schwerpunkt auf die Steigerung des Anteils der Frauen in den Studiengängen der Informatik zu legen. Übergreifendes Anliegen des Projekts ist die Absicht, im Bereich der Informatik die notwendige Teilhabe von Frauen zu bewirken und in Schulen, Instituten und bei Praxispartnern die Frage der Geschlechtergerechtigkeit - die Chancengleichheit von Frauen und Männern - allgemein zum Thema zu machen.

Um dies effektiv und nachhaltig zu erreichen, wurde durch die Projektleiterin der Ideen-Werkstatt eine Reihe von Maßnahmen initiiert und betreut. Dabei war die Kooperation mit der Institutsleitung, den Lehrkräften und Studierenden selbstverständlich und hilfreich.

Von den bisher erbrachten Leistungen des Projektes seien hier genannt:

- Organisation und Pflege kontinuierlicher Kontakte, vorrangig zu den Gymnasien des Stadtbezirks Treptow-Köpenick und Neukölln
- enge Zusammenarbeit mit Informatiklehrerinnen und -lehrern der Schulen
- Angebot von halbjährlich wechselnden Arbeitsgemeinschaften nur für Schülerinnen
- Durchführung von Projekttagen in den Schulen und einer Projektwoche „Neue Medien“ mit der Anna-Seghers-Oberschule
- Aufbau eines an den Interessen und Bedürfnissen der Schülerinnen orientierten speziellen Beratungsangebots, koordiniert durch die Projektleiterin, durchgeführt von Studentinnen
- Präsentationen zum Informatikstudium in den Schulen
- Angebot eines „Probier-Studiums“
- Erarbeitung und Veröffentlichung von Selbstportraits von Studierenden
- Präsentation der Ideen-Werkstatt im Internet durch die Gymnasiastin, Friderike Rogner, während ihres Betriebspraktikums
- Organisation und Betreuung von Betriebspraktika für Schülerinnen am Institut
- Veränderung der Außendarstellung des Instituts im Sinne einer frauenfreundlichen „Corporate Identity“
- Kontaktaufpflege zur Gleichstellungsbeauftragten des Stadtbezirks
- Kooperation mit den anderen Instituten der Humboldt-Universität und außeruniversitären Forschungs- und Wirtschaftseinrichtungen in Adlershof
- Mitarbeit bei der Organisation der Tage der Forschung in Adlershof
- Kooperation mit der Stiftung der Deutschen Wirtschaft (sdw) bei der Durchführung des MINT-Schülerinnen-Camps und Koordinierung der Aktivitäten der naturwissenschaftlichen Institute in Adlershof durch die Leiterin der Ideen-Werkstatt.

Hervorzuheben ist das erste MINT-Schülerinnen-CAMP, das am 25. September 2003 an der Humboldt-Universität in Berlin-Adlershof startete, mit dem sich die Stiftung der Deutschen Wirtschaft und die Humboldt-Universität zu Berlin gemeinsam für den weiblichen Nachwuchs in den naturwissenschaftlich-technischen Berufen engagierten. Sechzig Mädchen haben sich vier Tage lang mit praktischen Aufgaben aus den Bereichen Mathematik, Informatik, Naturwissenschaften und Technik auseinandergesetzt und erneut den Beweis angetreten, dass diese Fächer keine klassisch männlichen Begabungsfelder sind. Das MINT-

Schülerinnen-CAMP ist Teil des Projektes "Mit MINT zum Beruf", das die Stiftung der Deutschen Wirtschaft im vergangenen Jahr in sechs Bundesländern auf den Weg gebracht hat. Gefördert wird das Projekt aus Mitteln des Bundesministeriums für Bildung und Forschung im Rahmen des Programms "Schule-Wirtschaft/Arbeitsleben" (SWA). Studentinnen der Institute für Informatik, Physik und Chemie haben die Projektgruppen geleitet. Die Projekte des Instituts für Informatik wurden durch unsere Studentinnen Nannette Kase (Informatikprojekt), Eileen Hilges (Kryptologieprojekt), Jessica Ebrahimi (PC-Projekt), Jeanine Umlang (Erkundungsprojekt) und Daniela Rantzuch (Homepageprojekt) betreut. Die Koordinierung dieser Aktivitäten erfolgte durch die Ideen-Werkstatt.

(http://www.informatik.hu-berlin.de/~gutsche/ideen-werkstatt/mint-camp_2003.html)



Die Homepage-Bau-Gruppe der Ideen-Werkstatt

Alle Maßnahmen sind auf sehr positive Resonanz gestoßen, was sich in erster Linie in der Anzahl der Teilnehmerinnen und ihr wachsendes Interesse zeigt. Gegenwärtig ist das Interesse geweckt. Schülerinnen und Informatiklehrer gewinnen zunehmend die Einsicht, dass Mädchen sehr wohl für ein Informatik-Studium geeignet sind, sich aber häufig von falschen Vorstellungen abschrecken lassen. Es ist deshalb unbedingt erforderlich, den Sensibilisierungsprozess in Schulen, im Elternhaus, am Institut für Informatik und am Standort Adlershof weiter zu befördern. Solange noch Mädchen in Informatikkursen mit den Worten wie „Mädchen wollen sowieso nichts mit Informatik zu tun haben“ übergangen

werden, muss es ein Projekt wie die Ideen-Werkstatt geben. Eine anhaltende Wirkung der laufenden Aktivitäten kann nur durch Kontinuität des Angebots erreicht werden.

Dr. Márta Gutsche
Projektleiterin
Sitz: Rudower Chaussee 25,
Haus IV, Raum 108
D - 12489 Berlin
2093-5468

gutsche@informatik.hu-berlin.de

www.informatik.hu-berlin.de/~gutsche/ideen-werkstatt/index.htm