

Humboldt-Universität zu Berlin

Institut für Informatik

www.informatik.hu-berlin.de

Jahresbericht 2005



© Humboldt-Universität zu Berlin, 2005
Institut für Informatik
Unter den Linden 6
10099 Berlin

Besuchsanschrift: Rudower Chaussee 25, 12489 Berlin-Adlershof

Druck: Papyrus-Druck GmbH, Bismarckstraße 19, 10625 Berlin

Redaktion: Frau Prof. Beate Meffert, Christine Henze

Redaktionsschluss: 31.12.2005

Vorwort

Das Jahr 2005 verzeichnet eine Reihe erfreulicher Ereignisse. An erster Stelle stehen sicher die Berufungen von Frau Dr. Nicole Schweikardt auf die Juniorprofessur „Logik und Datenbanktheorie“ und von Herrn Dr. Stephan Kreuzer auf die Juniorprofessur „Logik und diskrete Systeme“. Die beiden Nachwuchsgruppen werden von der Deutschen Forschungsgemeinschaft im Rahmen der Initiative "Aktionsplan Informatik" gefördert. Mit der Berufung von Frau Prof. Schweikardt ist der Anteil der Professorinnen am Institut auf nunmehr 10% gestiegen, hoffentlich bleibt es nicht dabei.

Die Förderung des Nachwuchses beginnt bekanntlich in der Schule. Auch hier hat das Institut eine erfreuliche Nachricht zu vermelden: Die Arbeit der Ideenwerkstatt des Instituts unter Leitung von Frau Dr. Gutsche, einer Einrichtung zur Gewinnung von Informatikstudentinnen, hat erste Erfolge aufzuweisen. Die enge Kooperation mit Schulen, die Durchführung zahlreicher Arbeitsgemeinschaften für Schülerinnen, die individuelle Beratung und Betreuung interessierter Mädchen haben den Anteil der Studentinnen im Diplomstudiengang von 6,5% im Jahre 2003 auf 19% im Jahre 2005 ansteigen lassen.

Zu den guten Nachrichten zählt auch die Auszeichnung von Prof. Wolfgang Reisig mit dem *Faculty Award* von IBM. Prof. Reisig ist damit für seine Forschungsarbeit im Bereich der Modellierung von öffentlichen Business-to-Business-Geschäftsprozessen ausgezeichnet worden. Mit dem *Faculty Award* würdigt IBM weltweit herausragende wissenschaftliche Arbeiten, die besonders richtungsweisend für die Informationstechnologie sind.

Der Bericht über erfreuliche Ereignisse ließe sich noch fortsetzen. Die im AIBO-Team Fußball spielenden Roboter haben im German Team den Weltmeistertitel beim RoboCup 2005 verteidigen können, die Ringvorlesung „Themen der Informatik im historischen Kontext“ ist auf ein besonders großes Interesse gestoßen, so dass eine Veröffentlichung der Beiträge in Vorbereitung ist.

Die Forschungsaktivitäten des Jahres 2005 waren ganz auf die weitere Stärkung der drei Forschungsschwerpunkte des Instituts ausgerichtet. So wird der Schwerpunkt „Modellbasierte Systementwicklung“ in den nächsten Jahren vor allem durch ein im Jahre 2005 beantragtes und kürzlich bewilligtes interdisziplinäres Graduiertenkolleg geprägt werden. Es trägt den Titel "Modellbasierte Entwicklung von Technologien für selbstorganisierende dezentrale Informationssysteme - zur Anwendung im Katastrophenmanagement" und wird im Sommersemester 2006 mit 10 Doktoranden - und sicher auch Doktorandinnen - seine Arbeit aufnehmen. Die Vergabe der Stipendien an begabte junge Wissenschaftler und Wissenschaftlerinnen wird auch einen wichtigen Beitrag zur dringend notwendigen Nachwuchsförderung auf dem Gebiet der Informatik leisten.

Die Berichte der einzelnen Lehr- und Forschungsgebiete auf den folgenden Seiten geben Auskunft über die Forschung und Lehre, die am Institut auch im Jahre 2005 dazu beigetragen hat, die Wissenschaftsdisziplin Informatik weiter zu entwickeln, die interdisziplinären Verknüpfungen inner- und außerhalb der Universität zu festigen und zu erweitern und in der Ausbildung der Studierenden für eine hohe Qualität Sorge zu tragen.



Prof. Dr. Beate Meffert
Geschäftsführende Direktorin
Berlin, Februar 2006

Inhaltsverzeichnis

I. Institutsorganisation	7
II. Akademische Angelegenheiten	8
III. Lehrkörper am Institut	11
IV. Lehr- und Forschungseinheiten	22
<i>Theoretische Informatik</i>	
Logik in der Informatik	22
<i>Leiter: Prof. Dr. Martin Grohe</i>	
Logik und diskrete Systeme	22
<i>Leiter: Prof. Dr. Stephan Kreutzer</i>	
Logik und Datenbanktheorie	22
<i>Leiterin: Prof. Dr. Nicole Schweikardt</i>	
Algorithmen und Komplexität	33
<i>Leiter: Prof. Dr. Hans Jürgen Prömel</i>	
Komplexität und Kryptografie	48
<i>Leiter: Prof. Dr. Johannes Köbler</i>	
<i>Praktische Informatik</i>	
Systemanalyse, Modellierung und Computersimulation	55
<i>Leiter: Prof. Dr. Joachim Fischer</i>	
Systemarchitektur	70
<i>Leiter: Prof. Dr. Jens-Peter Redlich</i>	
Theorie der Programmierung	85
<i>Leiter: Prof. Dr. Wolfgang Reisig</i>	
Softwaretechnik	100
<i>Leiter: Prof. Dr. Klaus Bothe</i>	
Datenbanken und Informationssysteme	106
<i>Leiter: Prof. Johann-Christoph Freytag, Ph. D.</i>	
Wissensmanagement in der Bioinformatik	129
<i>Leiter: Prof. Dr. Ulf Leser</i>	
Künstliche Intelligenz	147
<i>Leiter: Prof. Dr. Hans-Dieter Burkhard</i>	
Wissensmanagement	168
<i>Leiter: Prof. Dr. Tobias Scheffer</i>	
Datenanalyse	176
<i>Leiter: Prof. Dr. Egmar Rödel</i>	

Informatik in Bildung & Gesellschaft	182
<i>Leiter: Prof. Dr. Wolfgang Coy</i>	
Parallele und Verteilte Systeme	195
<i>Leiter: Prof. Dr. Alexander Reinefeld</i>	
Spezifikation, Verifikation und Testtheorie	211
<i>Leiter: Prof. Dr. Bernd-Holger Schlingloff</i>	
Informationsintegration	215
<i>Leiter: Prof. Dr. Felix Naumann</i>	
Technische Informatik	
Rechnerorganisation und Kommunikation	222
<i>Leiter: Prof. Dr. Mirosław Malek</i>	
Signalverarbeitung und Mustererkennung	243
<i>Leiterin: Prof. Dr. Beate Meffert</i>	
Computer Vision	254
<i>Leiter: Prof. Dr. Ralf Reulke</i>	
V. Ideenwerkstatt und Studienberatung für Studentinnen und Schülerinnen	259
<i>Leiterin: Dr. Martà Gutsche</i>	
VI. Informationstechnik des Instituts für Informatik	262
VII. Lehrveranstaltungen	268

Institutsorganisation

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

Sitz: Rudower Chaussee 25 / Ecke Magnusstraße
12489 Berlin-Adlershof

Geschäftsführende Direktorin: PROF. BEATE MEFFERT
Sekretariat: SABINE DZIWISZ, Tel.: 2093 3044, Raum IV 319

Prüfungsausschuss

Vorsitzender: PROF. 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 – 15.00 Uhr
dienstags, donnerstags 09.00 – 11.30 Uhr
mittwochs 13.00 – 15.00 Uhr
und nach Vereinbarung*

Studienfachberatung

PROF. JOHANNES KÖBLER, Tel.: 2093 3189, Raum IV 001
Sprechzeit: dienstags 15.00 – 17.00 Uhr

Studentische Studienberatung

Fachschaft: CHRISTIAN BECKER, Tel.: 2093 3923, Raum II 321
Sprechzeit: donnerstags 11.00 – 12.30 Uhr und per Mail

Studienberatung für Studentinnen und Schülerinnen

DR. MÀRTA GUTSCHE, Tel.: 2093 5468, Raum IV 108
Sprechzeit: mittwochs 10.00 – 17.00 Uhr und nach Vereinbarung

Ideenwerkstatt „Mehr Frauen in die Informatik“

Leiterin: DR. MÀRTA GUTSCHE, Tel.: 2093 5468, Raum IV 108

Verwaltung

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

Akademische Angelegenheiten

Rufe und Berufungen

Neue Mitglieder des Lehrkörpers:

HERR DR. STEPHAN KREUTZER

wurde im April 2005 zum Juniorprofessor für das Gebiet „Logik und diskrete Systeme“ ernannt.

FRAU DR. NICOLE SCHWEIKARDT

wurde im April 2005 zur Juniorprofessorin für das Gebiet „Logik und Datenbanktheorie“ ernannt.

Einen Ruf an eine andere Universität hat erhalten:

Herr Privatdozent DR. KARSTEN SCHMIDT hat einen Ruf an die Universität Rostock auf die Professur Theorie der Programmiersprachen und Programmierung angenommen.

Ehrungen und Preise

IBM Faculty Award

für Prof. Reisig. Im Rahmen eines weltweiten Hochschulprogramms fördert IBM mit dem Faculty Award herausragende wissenschaftliche Arbeiten im Bereich der Informationstechnologie. Der IBM Faculty Award ist mit 40.000 US-Dollar dotiert.

KDD Cup 2005

für David Vogel und Steve Bridges (AI Insight), Steffen Bickel, Peter Haider, Rolf Schimpfky, Peter Siemen, Tobias Scheffer (Humboldt-Universität).

Der KDD Cup ist ein Wettbewerb, in dem jährlich wechselnde Data-Mining-Aufgaben gelöst werden müssen. Der Preis wurde für einen Algorithmus vergeben, der die Bedeutung von Suchmaschinenanfragen analysiert und Anfragen auf Knoten einer Taxonomie abbildet. Der Algorithmus erkennt etwa, dass sich die Suchanfrage „die Fliege“ entweder auf Kinofilme, auf Köder für den Angelsport oder auf Insekten bezieht und bildet damit eine Grundlage für bessere zukünftige Suchmaschinen.

ECML 2005 Best Paper Award

für Ulf Brefeld, Christoph Büscher, Tobias Scheffer, European Conference on Machine Learning 2005.

In der Arbeit „Multi-view discriminative sequential learning“ ist ein diskriminatives Verfahren zur Annotierung von Sequenzen entwickelt worden, das nur wenige annotierte Beispiele benötigt und darüber hinaus unannotierte Sequenzen zum Trainieren ausnutzen kann. Das Verfahren eignet sich z.B. zur Eigennamenerkennung und Informationsextraktion.

Best Paper Award

Für M. Grohe, C. Koch und N. Schweikardt. Received the ICALP 2005 (Track B)

Best Student Paper Award

für F. Salfner, er erhielt den Preis von der European Dependable Computing Conference (EDCC-5), Budapest, Ungarn, April 2005.

Young Investigator Award

für Alexander Suhrbier und Heiko Scharff, sie erhielten jeweils den Preis auf dem 29. Wissenschaftlichen Kongress Hypertonie in Berlin im November 2005. Ausgezeichnet wurden sie für ihre Ergebnisse, die sie im Rahmen ihrer Studienarbeiten zur Herzratenvariabilität erzielten.

Zum Mitglied der Jungen Akademie

an der Berlin-Brandenburgischen Akademie der Wissenschaften und der Deutschen Akademie der Naturforscher Leopoldina ist Frau Prof. Schweikardt im Juni 2005 berufen worden

Habilitationen***Dr. rer. nat. Eckhardt Holz***

derzeit wissenschaftlicher Mitarbeiter am Hasso-Plattner-Institut Potsdam, Vertretung des Fachgebiets Softwaretechnik

Titel: „*Kombination von Modellierungstechniken für den Softwareentwurf*“

Gutachter: Prof. Dr. Joachim Fischer, Prof. Dr. Andreas Prinz, Prof. Dr. Adam Woliz

Zuerkennung der Lehrbefähigung am 18. April 2005

Dr. rer. nat. Amin Coja-Oghlan

derzeit wissenschaftlicher Mitarbeiter am Lehrstuhl Algorithmen und Komplexität des Instituts für Informatik

Titel: „*Spectral Techniques, Semidefinite Programs, and Random Graphs*“

Gutachter: Prof. Dr. Hans Jürgen Prömel, Prof. Dr. Alan Frieze, Prof. Dr. Günter Ziegler

Zuerkennung der Lehrbefähigung am 11.7.2005

Erteilung der Lehrbefugnis: 24.10.2005

Prof. Dr. rer. nat. Tobias Scheffer

Juniorprofessor für Wissensmanagement am Institut für Informatik

Titel: „*Knowledge Discovery and Machine Learning: Theory, Methods, and Applications*“

Gutachter: Prof. Dr. Hans-Dieter Burkhard, Prof. Dr. Thomas Hofmann, Prof. Dr. Stefan Wrobel

Zuerkennung der Lehrbefähigung am 28.11.2005

Promotionen

M. Sc. Dagmar Monett Diaz

Titel: „Agent-Based Configuration of (Metaheuristic) Algorithms“

Gutachter: Prof. Dr. Hans-Dieter Burkhard, Prof. Dr. Egmar Rödel, Prof. Dr. Pedro Larranaga

Dipl.-Geograph Adrian Schischmanow

Titel: „Analyse des Straßenverkehrs mit verteilten opto-elektronischen Sensoren“

Gutachter: Prof. Dr. Ralf Reulke, Prof. Dr. Herbert Jahn, Prof. Dr. Barbara Lenz

Dipl.-Inf. Günther Hoffmann

Titel: „Failure Prediction in Complex Computer Systems: A Probabilistic Approach“

Gutachter: Prof. Dr. Mirosław Malek, Prof. Dr. Günter Hommel, Prof. Dr. Kishor Trivedi, PD Dr. Aad van Moorsel, Prof. Dr. Erik Maehle

Dipl.-Inf. Adrianna Alexander

Titel: „Komposition temporallogischer Spezifikationen“

Gutachter: Prof. Dr. Wolfgang Reisig, Prof. Dr. Martin Grohe, Prof. Dr. Jörg Siekmann

Beendigung von Arbeitsverhältnissen

Herr Prof. Egmar Rödel hat im September 2005 planmäßig seine Arbeit am Institut beendet. Die Angehörigen des Instituts danken ihm dafür, dass er mehr als zehn Jahre lang erfolgreich das Lehr- und Forschungsgebiet Datenanalyse vertreten hat. Dank gilt ihm vor allem auch für sein großes Engagement bei der Beratung und Betreuung von Studierenden des Lehramts Informatik.

Lehrkörper am Institut für Informatik



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. 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 Automatentheorie) 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 Automatentheorie, Schaltkreis-Diagnose, Petrinetze, Verteilte Systeme und Künstliche Intelligenz. Die aktuellen Interessengebiete sind Verteilte Künstliche Intelligenz, Agentenorientierte 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. 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). Mit-herausgeber 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 Herausbergremium der Zeitschrift „Informatik-Spektrum“ und Sprecher der DFG-Forschergruppe „Bild-Schrift-Zahl“.

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-Beschreibungs-techniken 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 und der ITU. Prof. Fischer selbst leitete als Rapporteur in der Studiengruppe 17 der ITU-T (Sprachen und allgemeine Software-Aspekte für Telekommunikationssysteme) der-

zeitig 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).

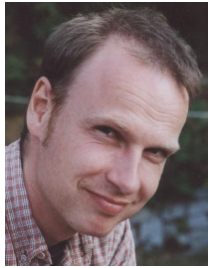
Von 1997 bis 1998 leitete Prof. Fischer als Geschäftsführender Direktor die Verlagerung des Instituts von Berlin-Mitte nach Berlin-Adlershof.



PROF. 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 ist. 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 Besonderen dem Bereich Bioinformatik/Life Science.

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



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. 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 Dialogcompilation - Arbeitsweise und Realisierung des MS-Algol Dialogcompilers“. 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 standen Probleme der Modellierung und Verifikation verteilter Systeme im Mittelpunkt seiner Forschungsarbeiten.



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, Automatentheorie und formale Sprachen, Berechenbarkeitstheorie und Logik.



PROF. DR. STEPHAN KREUTZER

Studierte von 1994 bis 1999 Informatik an der Rheinisch-Westfälischen Technischen Hochschule (RWTH) Aachen und promovierte dort im Jahre 2002 bei Erich Grädel am Institut für Informatik. Im Anschluss an die Promotion verbrachte er im Jahre 2003 einen sechsmonatigen Forschungsaufenthalt an der University of Edinburgh. Von dort wechselte er an den Lehrstuhl Logik in der Informatik der Humboldt-Universität zu Berlin, bevor er 2005 auf eine durch die DFG im Rahmen des Aktionsplans Informatik geförderte Juniorprofessur am selben Institut berufen wurde.

Die Forschungsinteressen von Professor Kreutzer liegen im Bereich der Logik, besonders ihrer Anwendungen innerhalb der Informatik. Schwerpunkte seiner aktuellen Forschungsarbeit liegen im Bereich der Verifikation, der Theorie von Auswertungsspielen, der Graphentheorie und Algorithmik. 2003 wurde er mit dem GI Dissertationspreis für das Jahr 2002 ausgezeichnet, ebenso mit der Borchers-Plakette der RWTH Aachen.



PROF. DR. ULF LESER

Studierte Informatik an der Technischen Universität München und arbeitete danach ca. 1,5 Jahre am Max-Planck-Institut für molekulare Genetik in Berlin an der Entwicklung von integrierten Datenbanken im Rahmen des Human Genome Projekts. Von 1997 bis 2000 promovierte er am Graduiertenkolleg "Verteilte Informationssysteme" über Anfragealgorithmen in heterogenen Informationssystemen. In dieser Zeit arbeitete er außerdem mit an Projekten zur automatischen Annotation von Ergebnissen aus Hochdurchsatzexperimenten in der molekularbiologischen Forschung, zur Standardisierung von Datenmodellen und Zugriffsmethoden in den Lebenswissenschaften auf Basis von CORBA, und zur qualitätsbasierten Anfrageoptimierung in der Informationsintegration. Nach der Promotion ging er in die Industrie und leitete bei der UBIS AG Softwareentwicklungsprojekte im Bereich Data Warehousing, eCommerce und Wissensmanagement. Seit 2002 ist er Professor für Wissensmanagement in der Bioinformatik an der Humboldt-Universität. Die Forschungsarbeiten von Prof. Leser und seinen Mitarbeitern beschäftigen sich mit allen Aspekten der Integration heterogener, verteilter Datenbanken, beispielsweise im Bereich der Proteomics und der Proteinstrukturforschung, der Modellierung und Implementierung komplexer Datenbanken zum Wissensmanagement und der automatischen Analyse von natürlichsprachlichen Fachpublikationen. Weitere Themen der Gruppe sind Wissensrepräsentationssprachen, Data Mining und Algorithmen zum Zeichenkettenvergleich, speziell zur Berechnung von Abstammungsverhältnissen.



PROF. DR. MIROSLAW MALEK

Erhielt 1970 das Diplom für Elektronik und promovierte 1975 auf dem Gebiet der Technischen Informatik an der Technischen Universität Wrocław (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 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

Studierte nach dem Abitur und einer gleichzeitigen Ausbildung als Funkmechanikerin 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 Walshttransformation, 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 und 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. Meffert hat gemeinsam mit Olaf Hochmuth das Lehrbuch „Werkzeuge der Signalverarbeitung“ geschrieben.

**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 im Jahre 2000 und erhielt für seine Arbeit den Dissertationspreis der Gesellschaft für Informatik. In den Jahren 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 semiautomatischen Transformation von Daten verschiedener Strukturen. Seine Ergebnisse vollzogen erfolgreich den Technologietransfer in IBMs Produktabteilung, honoriert durch dem IBM Research Division Award.

**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. JENS-PETER REDLICH

Begann 1988 ein Informatikstudium an der Humboldt-Universität zu Berlin, welches er 1992, nach einem Gastaufenthalt am City College New York, mit dem Diplom abschloss. Als wissenschaftlicher Mitarbeiter arbeitete er anschließend 5 Jahre am Lehrstuhl Systemarchitektur, wo er sich der Erforschung objektorientierter Telekommunikationssysteme widmete und 1995 zum Dr. rer. nat. promovierte. 1996 schrieb er das erste deutschsprachige Buch über CORBA – eine moderne objektorientierte Middleware-Plattform. 1997 wechselte Herr Redlich zur Industrie, zunächst als Gastwissenschaftler am C&C Labor der Firma NEC in Princeton, NJ, USA. Ab 1998 arbeitete er dort als unbefristeter Mitarbeiter (Research Staff Member) bis er 2000 als Department Head die Leitung der Abteilung ‚Mobile Internet‘ übernahm. In dieser Zeit entstand Point-M, ein System für den sicheren drahtlosen Zugang zu Firmennetzwerken, für welches mehrere Patente in den USA, Europa und Japan angemeldet wurden. Im Sommer 2004 kehrte Herr Redlich nach Deutschland zurück, wo er nun an der Humboldt-Universität als Professor für Systemarchitektur (C4) und für NEC Europe als Senior Research Advisor tätig ist. Seine Forschungsinteressen umfassen Betriebssysteme und Middleware, Sicherheit und Mobilkommunikation. Derzeitiger Schwerpunkt sind selbstorganisierende Netzwerke und Dienstplattformen, wie z.B. Ad-Hoc-Netzwerke für 802.11-basierte Community-Netzwerke. Als Gutachter für Zeitschriften sowie als Mitglied von Programm-Komitees ist Herr Redlich international seit vielen Jahren tätig.



PROF. DR. 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 1987 arbeitete Herr Reinefeld als wissenschaftlicher Mitarbeiter und von 1989 bis 1992 als Hochschulassistent an der Universität Hamburg. In den dazwischenliegenden Jahren 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. 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 zur Analyse kooperierender sequentieller Prozesse. 1983 vertrat er eine Professur an der Universität Hamburg und leitete anschließend bei der Gesellschaft für Mathematik und Datenverarbeitung Projekte 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 für Softwaretechnik und Theorie der Programmierung am Institut für Informatik der Humboldt-Universität zu Berlin. Prof. Reisig war Geschäftsführender Direktor des Institutes für Informatik 1994-1996 und 2002-2004 sowie Dekan der Mathematisch-Naturwissenschaftlichen Fakultät II der HU Berlin 1996-1998.

Die Forschungsarbeiten von Prof. Reisig sind in zwei Bereiche gegliedert: In den Projekten des ersten Bereichs werden Methoden und Modellierungstechniken für den Systementwurf (weiter-) entwickelt, insbesondere Petrinetze, Abstract State Machines und die Temporal Logic of Actions. Sie sind Grundlage für den zweiten Bereich, in dem diese Methoden und Techniken praktisch eingesetzt werden.

Prof. Reisig hat mehrmonatige Forschungsaufenthalte im ICSI, Berkeley, als „Lady Davis Visiting Professor“ am Technion, Haifa, und bei Microsoft Research (Redmond) verbracht. Für seine grundlegenden Projekte zur Modellierung und Analyse von Geschäftsprozessen zusammen mit dem IBM-Labor in Böblingen hat Prof. Reisig 2003 und 2005 einen „IBM Faculty Award“ erhalten.

**PROF. DR. RALF REULKE**

Studierte von 1975 bis 1980 Physik an der Humboldt-Universität zu Berlin, wo er auch 1984 promovierte. Seit 1983 arbeitete er am Institut für Kosmosforschung der Akademie der Wissenschaften der DDR in Berlin-Adlershof auf den Gebieten der multispektralen Fernerkundung, Signal- und Bildverarbeitung, sowie der optischen Sensorik.

Seit 1992 ist Ralf Reulke Mitarbeiter am Deutschen Zentrum für Luft- und Raumfahrt und arbeitete dort in verschiedenen Instituten und Einrichtungen. Er war an einer Vielzahl von Projekten beteiligt. Dazu gehörten Sensoren, die z.B. zur Erforschung des Mars und dem Saturn dienten, aber auch hochauflösende Kamerasysteme für die Erdfernerkundung, so zum Beispiel zusammen mit der Firma Leica, die Flugzeugkamera ADS40 und eine hochauflösende terrestrische Panoramakamera.

Im Jahre 2002 erhielt er als Professor einen Ruf an die Universität Stuttgart. Seit August 2004 ist er Professor für Computer Vision am Institut für Informatik der Humboldt-Universität. Gleichzeitig leitet er am deutschen Zentrum für Luft- und Raumfahrt eine Abteilung für Bild- und Signalverarbeitung.

Seine Forschungsinteressen liegen auf dem Gebiet der Signal- und Bildverarbeitung, der Sensor- und Datenfusion und der Visualisierung.



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 Hilbert-raum“. Von 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“. Von 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“. Im Jahre 1992 wurde Edgar Rödel zum Professor für Datenanalyse an das Institut für Informatik berufen. Seit Oktober 2005 ist er im Ruhestand.



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, N.J. und an der University of New South Wales in Sydney, Australien, 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. Seit 2003 ist er Juniorprofessor an der Humboldt-Universität zu Berlin, er wird im Emmy-Noether-Programm der Deutschen Forschungsgemeinschaft gefördert. 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 Information Retrieval.



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 Dr. 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 Holger 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. NICOLE SCHWEIKARDT

Studierte Mathematik mit Nebenfach Informatik an der Johannes Gutenberg-Universität Mainz. Nach dem Abschluss des Diploms war sie von 1998 bis 2002 als wissenschaftliche Mitarbeiterin am dortigen Institut für Informatik beschäftigt, wo sie im Juni 2002 ihre Promotion über die Ausdrucksstärke der Logik erster Stufe mit eingebauten Prädikaten abschloss. Nach einem Postdoc-Jahr an der University of Edinburgh war sie ab September 2003 als wissenschaftliche Assistentin am Lehrstuhl für Logik in der Informatik der HU Berlin tätig. Im Jahr 2005 wurde sie im Aktionsplan Informatik in das Emmy-Noether-Programm der Deutschen Forschungsgemeinschaft aufgenommen und zur Juniorprofessorin für Logik und Datenbanktheorie an die Humboldt-Universität zu Berlin berufen.

Die Forschungsinteressen von Frau Prof. Schweikardt liegen in den Bereichen Logik, Datenbanktheorie und Komplexitätstheorie, mit einem Schwerpunkt auf der Erforschung der Grundlagen der Verarbeitung von großen Datenmengen und Datenströmen.

Ihre Arbeit wurde mit dem Dissertationspreis 2002 der Gesellschaft für Informatik und mit dem Forschungsförderpreis 2003 der Freunde der Universität Mainz ausgezeichnet. Im Juni 2005 wurde sie zum Mitglied der Jungen Akademie an der Berlin-Brandenburgischen Akademie der Wissenschaften und der Deutschen Akademie der Naturforscher Leopoldina berufen.

IV. Lehr- und Forschungseinheiten

Lehr- und Forschungseinheiten

Logik in der Informatik

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

Leiter

PROF. DR. MARTIN GROHE

Tel.: (030) 2093 3078

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

Logik und diskrete Systeme

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

Leiter

PROF. DR. STEPHAN KREUTZER

Tel.: (030) 2093 3075

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

Logik und Datenbanktheorie

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

Leiterin

PROF. DR. NICOLE SCHWEIKARDT

Tel.: (030) 2093 3086

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

Sekretariat

BIRGIT EISENMANN

Tel.: (030) 2093 3080

Fax: (030) 2093 3081

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

Wissenschaftliche Mitarbeiterinnen und Mitarbeiter

DR. YIJIA CHEN

DR. PANAGIOTIS GIANNOPOULOS

DIPL.-INF. MAGDALENA GRÜBER

DIPL.-INF. ANDRÉ HERNICH

PAUL HUNTER, B.Sc.

DR. DÁNIEL MARX

DR. LOUCHKA POPOVA-ZEUGMANN

DIPL.-MATH. MARK WEYER

Gäste

DR. HUBIE CHEN (BARCELONA)
PROF. FEDOR FOMIN (HUMBOLDT-STIPENDIAT, BERGEN)

Doktoranden

DIPL.-INF. DANIEL ROLF
DIPL.-MATH. GÖTZ SCHWANDTNER

Technikerin

DIPL.-ING. PETRA KÄMPFER

Tutoren

DZIFA AMETOWOBLA
KORNELIUS KALNBACH
THOMAS KUNZE
SEBASTIAN ORDYNIK
FLORIAN SEFFLER
MARK THURLEY
WOJCIECH WOJCIKIEWICZ

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 an diesen Lehr- und Forschungseinheiten 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 grundlegende Fragen der Komplexitätstheorie oft mittels logischer Methoden analysieren. Ein nützliches Paradigma besagt, dass die algorithmische Komplexität eines Problems sehr eng mit der sprachlichlogischen 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.

Lehre**Vorlesungen**

- Theoretische Informatik 1 (M. GROHE, WiSe 2004/2005, WiSe 2005/2006)
- Berechenbarkeit (M. GROHE, WiSe 2004/2005)
- Parametrische Algorithmen und Komplexitätstheorie (M. GROHE, SoSe 2005)
- Logik in der Informatik (M. GROHE, WiSe 2005/2006)
- Lineare Optimierung (L. POPOVA-ZEUGMANN, SoSe 2005)

- Logik und Komplexität (S. KREUTZER, N. SCHWEIKARDT, SoSe 2005)
- SAT-Solving und Constraint Satisfaction Probleme (S. Kreutzer, WiSe 2005/2006)

Seminare und Proseminare

- Automaten und semistrukturierte Daten (XML) (M. GROHE, WiSe 2004/2005)
- Algorithmische Spieltheorie und Grundlagen der Internetökonomie (M. GROHE, SoSe 2005)
- Proseminar Die Grenzen der Berechenbarkeit (M. GROHE, SoSe 2005)
- Proseminar Sprachen und Automaten (L. POPOVA-ZEUGMANN, SoSe 2005)
- Endliche Modelltheorie (S. KREUTZER, N. SCHWEIKARDT, WiSe 2005/2006)
- Kommunikationskomplexität (N. SCHWEIKARDT, WiSe 2005/2006)
- Forschungsseminar Logik in der Informatik (M. GROHE, S. KREUTZER, N. SCHWEIKARDT, WiSe 2004/2005, SoSe 2005, WiSe 2005/2006)
- Proseminar Mathematica (K.-P. NEUENDORF, WiSe 2004/2005)
- Proseminar Mit Mathematica spielend durch die Informatik (K.-P. NEUENDORF, WiSe 2005/2006)

Übungen

- Theoretische Informatik 1 (M. GROHE, M. GRÜBER, S. KREUTZER, L. POPOVA-ZEUGMANN, N. SCHWEIKARDT, M. WEYER, WiSe 2004/2005, WiSe 2005/2006)
- Parametrische Algorithmen und Komplexitätstheorie (M. GRÜBER, SoSe 2005)
- Logik in der Informatik (M. WEYER, WiSe 2005/2006)
- Lineare Optimierung (L. POPOVA-ZEUGMANN, SoSe 2005)
- Logik und Komplexität (S. KREUTZER, N. SCHWEIKARDT, SoSe 2005)
- SAT-Solving und Constraint Satisfaction Probleme (S. KREUTZER, WiSe 2005/2006)

Forschung

Projekt: Grundlagen der Verarbeitung von großen Datenmengen und Datenströmen

Ansprechpartnerin: PROF. DR. NICOLE SCHWEIKARDT

Beteiligter Mitarbeiter: DIPL.-INF. ANDRÉ HERNICH

Forschungsförderung: Deutsche Forschungsgemeinschaft (DFG Aktionsplan Informatik)

Die effiziente Verarbeitung extrem großer Datenmengen wird - nicht zuletzt weil Sekundär- und Tertiär-Speicherplatz in den letzten Jahren zu sehr preiswerten Ressourcen geworden sind - zu einer immer wichtigeren Herausforderung für die Informatik. Solch große Datenmengen treten in vielen Anwendungsbereichen auf, etwa als Sammlung wissenschaftlicher Ergebnisse (z.B. Medizindatenbanken wie Medline und Swiss-Prot), in Form von Sensordaten oder als Börsenticker.

Häufig liegen die Daten dabei nicht in einer klassischen, effizient bearbeitbaren Datenbank, sondern nur in semistrukturierter Form vor, z.B. als XML-Dokument. Solche semistrukturierten Daten können auf natürliche Weise durch Bäume repräsentiert werden. Wegen der großen Datenmenge kann in der Regel nicht die Baumrepräsentation der gesamten Daten im Hauptspeicher eines Rechners vorgehalten werden, sondern nur ein gewisser Ausschnitt. In vielen Anwendungen sind die Daten sogar nur nach und nach, als Datenstrom zugänglich, etwa beim Börsenticker, der mit der Zeit immer wieder neue Informationen sendet.

Zur effizienten Verarbeitung solcher Daten sind daher neue, über die aus der klassischen Datenbankverarbeitung bekannten hinausgehende Techniken erforderlich. Schwerpunkt dieses Projekts ist die Erforschung der theoretischen Grundlagen der Verarbeitung solcher großer, semistrukturierter Datenmengen hinsichtlich Anfrageoptimierung, Eigenschaften von Anfragesprachen und Komplexitätstheorie für die Verarbeitung großer Datenmengen.

Projekt: Strukturelle und spielbasierte Analyse von Auswertungs- und Erfüllbarkeitsproblemen

Ansprechpartner: PROF. DR. STEPHAN KREUTZER

Beteiligter Mitarbeiter: PAUL HUNTER, B.SC.

Studentische Hilfskraft: SEBASTIAN ORDYNYAK

Forschungsförderung: Deutsche Forschungsgemeinschaft (DFG Aktionsplan Informatik)

Algorithmische Probleme der Logik haben aufgrund ihrer Allgemeinheit vielfältige Einsatzgebiete in der Informatik. Zu den wichtigsten dieser Probleme gehört das Erfüllbarkeitsproblem der Aussagenlogik (SAT), mit wichtigen Anwendungen unter anderem in der Verifikation und der künstlichen Intelligenz. Daneben sind Auswertungsprobleme temporaler Logiken von großer praktischer Bedeutung, vor allem im Bereich der Verifikation, wo temporale Logiken die Basis von Spezifikationssprachen bilden. Zur Lösung solcher Auswertungsprobleme hat sich ein Ansatz als sehr erfolgreich erwiesen, der auf Verfahren aus der Automaten- und Spieltheorie basiert. Prominentestes Beispiel dieses Ansatzes ist die Charakterisierung des modalen μ -Kalküls durch Paritätsspiele. Trotz intensiver Forschung sind hier noch zentrale Aspekte weitgehend unverstanden, etwa die genaue Komplexität des Paritätsspielproblems.

Im Rahmen des Projektes sollen das SAT-Problem sowie spieltheoretische Verfahren zur Lösung von Auswertungsproblemen in der Verifikation untersucht werden. Kern des Projektes ist die Untersuchung der Zusammenhänge zwischen strukturellen Eigenschaften der Eingabeinstanzen und der Komplexität der betrachteten Probleme und Algorithmen. Dabei wird zum einen auf bekannte Strukturmaße zurückgegriffen, wie sie etwa in der Graphentheorie entwickelt wurden. Da diese sich jedoch vor allem auf die Modellierung der Eingaben durch ungerichtete Graphen beziehen, in der Verifikation auftretende Systeme jedoch typischerweise eher durch gerichtete Graphen modelliert werden, wird im Rahmen dieses Projektes an einer Strukturtheorie für gerichtete Graphen gearbeitet. Daneben soll eine Theorie von Spielen entwickelt werden, mit denen komplexere Eigenschaften von Systemen modelliert werden können, wie sie etwa aus der Verifikation nebenläufiger Prozesse erwachsen.

Projekt: Die Struktur Parametrischer Komplexitätsklassen

Ansprechpartner: PROF. DR. MARTIN GROHE

Beteiligte Mitarbeiter: DR. YIJIA CHEN, DR. PANAGIOTIS GIANNOPOULOS

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 be-

rücksichtigt 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

Beteiligter Mitarbeiter: DR. DÁNIEL MARX

Studentische Hilfskraft: MARK THURLEY

Forschungsförderung: Deutsche Forschungsgemeinschaft

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: Endliche Modelltheorie und ihre Verbindungen zur Informatik

Ansprechpartner: PROF. DR. MARTIN GROHE

Zusammenarbeit: PROF. DR. ANDREAS BAUDISCH (HU Berlin) und 13 weitere Logikgruppen in Europa

Forschungsförderung: EU

URL: <http://www.logique.jussieu.fr/modnet/index.html>

Das Projekt ist Teil des von der EU finanzierten *MODNET Research Training Networks in Model Theory*. Dieses Netzwerk von 13 europäischen Arbeitsgruppen beschäftigt sich mit der Modelltheorie, einem Teilbereich der mathematischen Logik, der in gewisser Weise

ein Bindeglied zwischen Logik und Algebra darstellt. Beim in unserer Arbeitsgruppe bearbeiteten Teilprojekt geht es um die Modelltheorie endlicher Strukturen, die zahlreiche Verbindungen zur Komplexitätstheorie und anderen Bereichen der Informatik aufweist.

Projekt: Zeitabhängige Systeme

Ansprechpartnerin: 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 zentrale Problem 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. 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 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 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 ist auf die gesamte Architektur MSS angewandt worden und damit die zeitlichen Garantien von MSS größtenteils bewiesen.

Projekt: Funktionalität und Zeit in biochemischen Netzwerken

Ansprechpartnerin: DR. LOUCHKA POPOVA-ZEUGMANN

Zusammenarbeit: BTU Cottbus, Institut für Informatik, PROF. DR. MONIKA HEINER und TFH Berlin, FB V, Bioinformatik, PROF. DR. INA KOCH

Gegenstand dieses Projekts ist Modellierungsmittel für biochemische Netzwerke zu entwerfen und Algorithmen zu entwickeln, die quantitative und qualitative Analyse ermöglichen ohne den gesamten Zustandsraum zu kennen. Diese Bedingung kann wegen der Größe der biochemischen Netzwerke im Allgemeinen nicht vernachlässigt werden.

Veröffentlichungen

I. ADLER, G. GOTTLOB, M. GROHE: *Hypertree-width and related hypergraph invariants*. In S. Felsner, editor, Proceedings of the 3rd European Conference on Combinatorics, Graph Theory and Applications, volume AE of DMTCS Proceedings Series, pages 5–10, 2005.

- A. ATSERIAS, A. DAWAR, M. GROHE: *Preservation under extensions on well-behaved finite structures*. In L. Caires, G.F. Italiano, L. Monteiro, C. Palamidessi and M. Yung, editors, Proceedings of the 32nd International Colloquium on Automata, Languages and Programming, volume 3580 of Lecture Notes in Computer Science, pages 1437–1450. Springer-Verlag, 2005.
- D. A. M. BARRINGTON, N. IMMERMANN, C. LAUTEMANN, N. SCHWEIKARDT, D. THÉRIEN: *First-Order Expressibility of Languages with Neutral Letters Or: The Crane Beach Conjecture*. Journal of Computer and System Sciences, volume 70, pages 101-127, 2005.
- A. BLUMENSATH, S. KREUTZER: *An Extension of Muchnik's Theorem*. Journal of Logic and Computation, 15(1), p. 59-74, 2005.
- J. BRADFIELD, S. KREUTZER: *The Complexity of Independence-friendly Fixpoint Logic*. In Proceedings of the 14th Annual Conference of the European Association for Computer Science Logic (CSL), Lecture Notes in Computer Science 3634, Springer-Verlag, 2005.
- A. BULATOV, M. GROHE: *The complexity of partition functions*. Theoretical Computer Science. 348: 148–186, 2005. ICALP 2004 (Track A) Special Issue.
- S. CABELLO, P. GIANNOPOULOS, C. KNAUER, G. ROTE: *Matching Point Sets with Respect to the Earth Mover's Distance*. 13th Annual European Symposium on Algorithms (ESA 2005): 520-531.
- Y. CHEN, J. FLUM, M. GROHE: *Machine-based methods in parameterized complexity theory*. Theoretical Computer Science, 339: 167–199, 2005.
- A. DAWAR, E. GRÄDEL, S. KREUTZER: *Backtracking Games and Inflationary Fixed Points*. ICALP 2004 Selected Paper Issue, Theoretical Computer Science, 2005.
- M. FARSHI, P. GIANNOPOULOS, J. GUDMUNDSSON: *Finding the best shortcut in a geometric network*. Symposium on Computational Geometry 2005: 327-335.
- J. FLUM, M. GROHE: *Model-checking problems as a basis for parameterized intractability*. Logical Methods in Computer Science, 1(1), 2005.
- G. GOTTLOB, M. GROHE, N. MUSLIU, M. SAMER, F. SCARCELLO: *Hypertree decompositions: Structure, algorithms, and applications*. In D. Kratsch, editor, Proceedings of the 31st International Workshop on Graph-Theoretic Concepts in Computer Science, volume 3787 of Lecture Notes in Computer Science, pages 1–15. Springer-Verlag, 2005.
- M. GROHE, C. KOCH, N. SCHWEIKARDT: *The complexity of querying external memory and streaming data*. In M. Liskiewicz and R. Reischuk, editors, Proceedings of the 15th International Symposium on Fundamentals of Computation Theory, volume 3623 of Lecture Notes in Computer Science, pages 1–16. Springer-Verlag, 2005.
- M. GROHE, C. KOCH, N. SCHWEIKARDT: *Tight lower bounds for query processing on streaming and external memory data*. In L. Caires, G.F. Italiano, L. Monteiro, C. Palamidessi, and M. Yung, editors, Proceedings of the 32nd International Colloquium on Automata, Languages and Programming, volume 3580 of Lecture Notes in Computer Science, pages 1076–1088. Springer-Verlag, 2005. Received the ICALP 2005 (Track B) Best Paper Award.
- M. GROHE, S. KREUTZER, N. SCHWEIKARDT: *The expressive power of two-variable least fixed-point logics*. In J. Jedrzejowicz and A. Szepietowski, editors, Proceedings of the 30th

International Symposium on Mathematical Foundations of Computer Science, volume 3618 of Lecture Notes in Computer Science, pages 422–434. Springer-Verlag, 2005.

M. GROHE, N. SCHWEIKARDT: *Lower bounds for sorting with few random accesses to external memory*. In Proceedings of the 24th ACM Symposium on Principles of Database Systems, pages 238–249, 2005.

M. GROHE, N. SCHWEIKARDT: *The succinctness of first-order logic on linear orders*. Logical Methods in Computer Science, Volume 1 (1:6) pp. 1-25, 2005,.

A. HERNICH, A. NICKELSEN: *Combining Self-Reducibility and Partial Information Algorithms*. In Proceedings of the 30th International Symposium on Mathematical Foundations of Computer Science. Lecture Notes in Computer Science, Band 3618, Seiten 483-494. Springer-Verlag, 2005.

D. MARX: *A short proof of the NP-completeness of minimum sum interval coloring*. Operations Research Letters, 33(4), 382-384, 2005.

D. MARX: *Efficient approximation schemes for geometric problems?* 13th Annual European Symposium on Algorithms (ESA 2005), Lecture Notes in Computer Science, 3669, Springer, Berlin, 448-459, 2005.

D. MARX: *Minimum sum multicoloring on the edges of planar graphs and partial k-trees*. 2nd Workshop on Approximation and Online Algorithms (WAOA), 2004 (Bergen), 9-22, Lecture Notes in Computer Science, 3351, Springer, Berlin, 2005.

D. MARX: *NP-completeness of list coloring and precoloring extension on the edges of planar graphs*. Journal of Graph Theory. 49(4): 313-324, 2005.

D. MARX: *The Closest Substring problem with small distances*. In Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science, 63-72, 2005.

L. POPOVA-ZEUGMANN, M. HEINER, I. KOCH: *Time Petri Nets for Modelling and Analysis of Biochemical Networks*. Fundamenta Informaticae (FI), 67(2005), pp 149-162, IOS-Press, Amsterdam.

L. POPOVA-ZEUGMANN, M. WERNER: *Extreme Runtimes of Schedules Modelled by Time Petri Nets.*, Fundamenta Informaticae (FI), 67(2005), pp 163-174, IOS-Press, Amsterdam.

D. ROLF: *Derandomization of PPSZ for Unique- k-SAT*. In Proceedings of the 8th International Conference on Theory and Applications of Satisfiability Testing. Lecture Notes in Computer Science 3569, Springer-Verlag, 216-225, 2005.

N. SCHWEIKARDT: *Arithmetic, First-Order Logic, and Counting Quantifiers*. ACM Transactions on Computational Logic (ACM TOCL), Volume 6, Issue 3, July 2005, pages 634-671.

Vorträge

P. GIANNOPOULOS: *Matching Point Sets with Respect to the Earth Mover's Distance*. 13th Annual European Symposium on Algorithms (ESA 2005).

M. GROHE: *Hypertree width and related hypergraph invariants*. 3rd European Conference on Combinatorics, Graph Theory and Applications (EUROCOMB'05), Berlin, September 2005.

M. GROHE: *Law Enforcement on Hypergraphs*. Eingeladener Vortrag, Workshop on "Perspectives in Verification" in Honour of Prof. Thomas, Paris, November 2005.

M. GROHE: *Preservation Theorems on Well-Behaved Finite Structures*. Workshop on Algorithmic Model Theory Darmstadt, Februar 2005.

M. GROHE: *Preservation under Extensions on Well-Behaved Finite Structures*. 32nd International Colloquium on Automata, Languages and Programming (ICALP'05), Lissabon, Juli 2005.

M. GROHE: *The Complexity of Querying External Memory and Streaming Data*. Eingeladener Vortrag, 15th International Symposium on Fundamentals of Computation Theory (FCT'05), Lübeck, August 2005.

M. GROHE: *Tight Lower Bounds for Query Processing on Streaming and External Memory Data*. Eingeladener Seminarvortrag Edinburgh, Mai 2005.

A. HERNICH: *Combining Self-Reducibility and Partial Information Algorithms*. Mathematical Foundations of Theoretical Computer Science (MFCS 2005), Danzig, Polen, 2. September 2005.

S. KREUTZER: *Approximation through Logic*. Gastvortrag auf dem Logic Colloquium 2005, 25.7.2005 – 3.8.2005, Athen.

S. KREUTZER: *Backtracking Spiele*. Vortrag auf dem Jahrestreffen des GI Beirats der Universitätsprofessoren (GIBU), 21.3.2005 – 23.3.2005, Schloß Dagstuhl.

S. KREUTZER: *DAG-Decompositions and Parity Games*. Gastvortrag am mathematischen Institut der Universität Warschau, 18.11.2005.

S. KREUTZER: *DAG-Decompositions and Parity Games*. Gastvortrag im Graduiertenkolleg „Mathematische Logik und Anwendungen“ der Albert-Ludwigs-Universität Freiburg, 5.12.2005.

S. KREUTZER: *Spiele und Model-Checking*. Antrittsvorlesung am Tag der Informatik der Humboldt-Universität zu Berlin, 26.5.2005.

D. MARX: *Efficient approximation schemes for geometric problems?* European Symposium on Algorithms (ESA) 2005, Palma de Mallorca, Spain, October 4, 2005.

D. MARX: *The Closest Substring problem with small distances*. Dagstuhl Seminar 05301: Exact Algorithms and Fixed-Parameter Tractability, July 25, 2005.

D. MARX: *The Closest Substring problem with small distances*. Foundations of Computer Science (FOCS) 2005, Pittsburgh, Pennsylvania, October 23, 2005.

D. MARX: *Trees, tree width, hypertree width, fractional hypertree width*. 33. Berliner Algorithmen-Tag, Freie Universität Berlin, July 15, 2005.

L. POPOVA-ZEUGMANN: *A Memo on Computability in Time Petri Nets*. Workshop CS&P 2005, Ruciane-Nida, Poland, September 28-30, 2005.

L. POPOVA-ZEUGMANN: *Quantitative Analysis of Time Petri Nets Used for Modelling Biochemical Networks*. Max-Planck-Institut für Molekulare Pflanzenphysiologie, Golm, May 04, 2005.

L. POPOVA-ZEUGMANN: *Time Petri Nets State Space Reduction Using Dynamic Programming and Time Paths*. 17th Triennial Conference of the IFORS, Honolulu, Hawaii USA, July 11-15, 2005.

N. SCHWEIKARDT: *A Logical Characterisation of Linear Time on Nondeterministic Turing Machines*. Vortrag beim Symposium zum Gedenken an Clemens Lautemann, Johannes Gutenberg-Universität Mainz, Oktober 2005.

N. SCHWEIKARDT: *Logik und Datenbanktheorie*. Antrittsvorlesung am Tag der Informatik des Instituts für Informatik der Humboldt-Universität zu Berlin, Mai 2005.

N. SCHWEIKARDT: *Lower Bounds for Sorting with Few Random Accesses to External Memory*. Gastvortrag im Graduiertenkolleg Combinatorics, Geometry and Computation der ETH Zürich, der 3 Berliner Universitäten sowie des Konrad-Zuse-Forschungszentrums Berlin, Januar 2005.

N. SCHWEIKARDT: *Lower Bounds for Sorting with Few Random Accesses to External Memory*. 24th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (PODS'05), Baltimore, Maryland, U.S.A., Juni 2005.

N. SCHWEIKARDT: *The Complexity of Querying External Memory and Streaming Data*. Gastvortrag im Informatikkolloquium der Universität des Saarlandes, November 2005.

N. SCHWEIKARDT: *The Expressive Power of Two-Variable Least Fixed-Point Logics*. 30th International Symposium on Mathematical Foundations of Computer Science (MFCS'05), Gdansk, Polen, August 2005.

N. SCHWEIKARDT: *Tight Lower Bounds for Query Processing on Streaming and External Memory Data*. Vortrag beim Dagstuhl-Seminar Foundations of Semistructured Data. Schloss Dagstuhl, Februar 2005.

N. SCHWEIKARDT: *Tight Lower Bounds for Query Processing on Streaming and External Memory Data*. 32nd International Colloquium on Automata, Languages and Programming (ICALP'05), Lissabon, Portugal, Juli 2005.

N. SCHWEIKARDT: *XQuery-Anfrageoptimierung mittels Schema-Informationen*. Vortrag beim Jahrestreffen 2005 des GI Beirats der Universitätsprofessoren (GIBU). Schloss Dagstuhl, März 2005.

N. SCHWEIKARDT: *Zur Komplexität von Anfragen an Datenströme und große Datenmengen*. Gastvortrag im Graduiertenkolleg Mathematische Logik und Anwendungen, Albert-Ludwigs-Universität Freiburg, Januar 2005.

Sonstige Aktivitäten

Prof. Dr. Martin Grohe

- Mitherausgeber der Zeitschriften Journal of Symbolic Logic und Journal of Discrete Algorithms
- Mitglied folgender Programmkomitees:
- 20th Annual IEEE Symposium on Logic in Computer Science (LICS 2005)
- 19th International Workshop on Computer Science Logic (CSL '05)
- Mitglied des Councils der Association for Symbolic Logic
- Diverse Gutachtertätigkeiten
- Direktor für Lehre und Studium am Institut für Informatik

Prof. Dr. Stephan Kreutzer

- Gutachtertätigkeiten für verschiedene Fachzeitschriften und Konferenzen
- Mitglied der Prüfungskommission des Instituts für Informatik der HU-Berlin (seit Oktober 2005)
- Publicity Co-Chair des IEEE Symposium on Logic in Computer Science (LICS)
- Publicity Co-Chair der Federated Logic Conference 2006 (FLoC'06)

Dr. Louchka Popova-Zeugmann

- Mitglied des Institutsrats
- Mitglied des Prüfungsausschusses
- Mitarbeit in der Kommission für Lehre und Studium
- Gutachten für CS&P 2005
- Gutachten für die Zeitschrift Fundamenta Informaticae

Prof. Dr. Nicole Schweikardt

- Gutachtertätigkeiten für verschiedene Fachzeitschriften und Konferenzen
- Mitglied der Frauenförderungskommission des Instituts für Informatik
- Berufung in die Junge Akademie an der Berlin-Brandenburgischen Akademie der Wissenschaften und der Deutschen Akademie der Naturforscher Leopoldina
- Best Paper Award bei ICALP 2005 (Track B) für die Arbeit "Tight Lower Bounds for Query Processing on Streaming and External Memory Data" (mit Martin Grohe und Christoph Koch)
- Mitglied im Programmkomitee des 10th International Symposium on Databases and Programming Languages (DBPL'05)
- Publicity Co-Chair des IEEE Symposium on Logic in Computer Science (LICS)
- Publicity Co-Chair der Federated Logic Conference 2006 (FLoC'06)
- Jurymitglied bei der Endrunde des 23. Bundeswettbewerb Informatik

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

Sekretariat

EVA SANDIG/MARGRIT HOPPE
Tel.: (030) 2093 3190
Fax: (030) 2093 3191
E-Mail: sandig@informatik.hu-berlin.de
mhoppe@informatik.hu-berlin.de

Wissenschaftliche Mitarbeiterinnen und Mitarbeiter

DIPL.-INF. MICHAEL BEHRISCH
DIPL.-INF. MANUEL BODIRSKY
DR. AMIN COJA-OGHLAN
PD DR. STEFAN HOUGARDY
DR. MIHYUN KANG
DIPL.-INF. STEFAN KIRCHNER
DIPL.-INF. DIRK SCHLATTER
DR. MATHIAS SCHACHT
DIPL.-MATH. MARTIN THIMM

Stipendiaten

MGR. JAN KARA
CAND. SCIENT. TARAL SEIERSTAD
DR. OLEG VERBITSKY

Techniker

DIPL.-MATH. RALF OELSCHLÄGEL

Tutorinnen und Tutoren

DORATHA DRAKE
GENEVIÈVE GRUNERT
HIEP HAN
MATTHIAS KILLAT
PETER LISKE
NICOLAS ROCCA
CHRISTIAN ROTHE
PER VIOLET
MARIANO ZELKE
VALENTIN ZIEGLER

Zentrale Lehr- und Forschungsgegenstände sind der Entwurf und die Analyse effizienter Algorithmen. In der Komplexitätstheorie werden Probleme hinsichtlich verschiedener Komplexitätsmaße wie Laufzeit oder Speicherplatz klassifiziert. Algorithmen, die sich durch eine besonders kurze Laufzeit bzw. einen besonders geringen Speicherplatzbedarf 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 der Anwendung der dabei erzielten probabilistischen und asymptotischen Resultate bei dem Entwurf und der Analyse von Graphenalgorithmen. Diese Resultate sind u.a. von Bedeutung bei der Untersuchung randomisierter Algorithmen, welche vom Zufall Gebrauch machen. Ein weiterer Schwerpunkt ist das Studium von Approximationsalgorithmen, 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 (MATHEON).

Innerhalb des Instituts besteht eine Zusammenarbeit mit den Lehrstühlen Komplexität und Kryptografie und Logik in der Informatik im Rahmen des Schwerpunktes „Modelle und Algorithmen“.

Lehre

Die Grundlagen der Gebiete Algorithmen und Komplexität werden im Grundstudium in den Vorlesungen *Theoretische Informatik 2* (Schwerpunkte: Berechenbarkeit und Komplexitätstheorie) und *Theoretische Informatik 3* (Schwerpunkt: Algorithmen und Datenstrukturen) 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 im Drug-Design* und durch weitere vertiefende Vorlesungen. In den Seminaren werden spezielle Klassen von Algorithmen wie approximative 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 „Modelle und Algorithmen“.

Veranstaltungen im Grundstudium

- Theoretische Informatik 2 (A. COJA-OGHLAN, M. BODIRSKY, G. GRUNERT, H. HAN, M. KILLAT, P. VIOLET, WiSe 05/06)
- Theoretische Informatik 3 (H. HAN, SoSe 05)

Kernveranstaltungen (Halbkurse)

- Graphen und Algorithmen, Teil 2 (A. COJA-OGHLAN, SoSe 05)
- Graphen und Algorithmen, Teil 1 (S. HOUGARDY, M. BEHRISCH, WiSe 05/06)
- Randomized Algorithms and Probabilistic Analysis (M. KANG, WiSe 05/06)

Spezialveranstaltungen

- Approximationsalgorithmen (A. COJA-OGHLAN, SoSe 05)
- Graphen in der Ebene (M. SCHACHT, D. SCHLATTER, SoSe 05)
- Combinatorics and its applications (M. BODIRSKY, M. KANG, SoSe 05)
- Algorithmen im Drug Design (S. HOUGARDY, SoSe 05)

Seminare und Proseminare

- Proseminar: ACM Programming Contest (M. BEHRISCH, SoSe 05)
- Proseminar: Einführung in Approximationsalgorithmen (S. KIRCHNER, SoSe 05)
- Seminar: Extremale Kombinatorik (M. SCHACHT, WiSe 05/06)

Forschung

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

Ansprechpartner: DR. AMIN COJA-OGHLAN, PD DR. STEFAN HOUGARDY, PROF. DR. HANS JÜRGEN PRÖMEL

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

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

Forschungsförderung: Deutsche Forschungsgemeinschaft

Netzwerke in den Lebenswissenschaften und Anwendungen

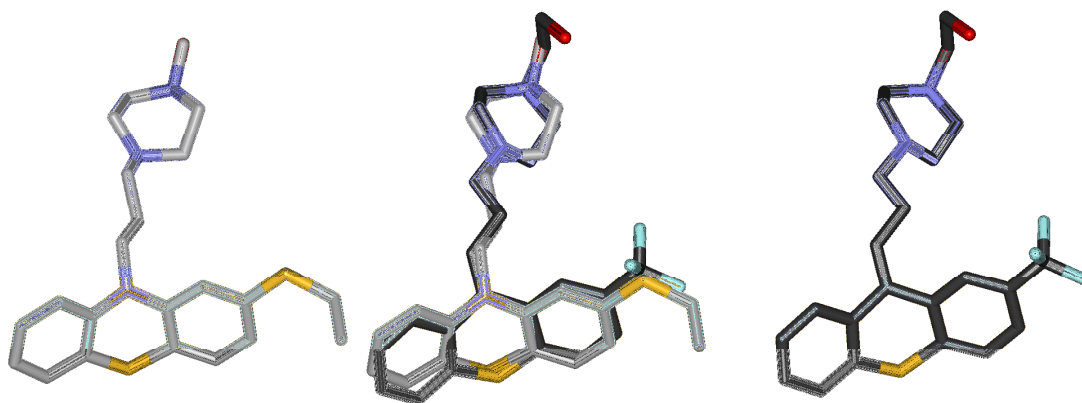
Ein Schwerpunkt unserer derzeitigen Forschung ist die möglichst exakte Berechnung der räumlichen Ähnlichkeit von Molekülen. Von uns entwickelte und implementierte Algorithmen zeigen, dass 3D-Ähnlichkeit von Molekülen stark mit Ähnlichkeit bezüglich Wirkung/Nebenwirkung korreliert. Des Weiteren zeigt sich, dass der von uns verfolgte rein geometrisch definierte Ähnlichkeitsbegriff in der Lage ist, Gemeinsamkeiten von Molekülen zu entdecken, die mit bisher benutzten fingerprint-basierten Techniken verborgen bleiben.

Wir haben einen weiteren Algorithmus zur Ähnlichkeitsbestimmung von Molekülen entwickelt, der eine optimale Lösung etwa 1000 mal schneller findet als bisherige Verfahren. Dieser Algorithmus ist 2005 zum Patent angemeldet worden. Er ermöglicht es, sehr große Mengen von Molekülen paarweise auf 3D-Ähnlichkeit zu testen, was im Moment am Beispiel einer Datenbank von potenziellen Krebsmedikamenten unternommen wird. Erste Auswertungen zeigen erneut, dass Wirkung (in diesem Beispiel: Hemmung von Tumorstadium) und 3D-Ähnlichkeit gut korrelieren.

Stochastische Modelle und Netzwerkparameter

Zur Modellierung der auftretenden Ähnlichkeitsnetzwerke ist das bisherige Standardmodell zufälliger Graphen von Erdős und Renyi mangels Abbildung wichtiger Eigenschaften wie Transitivität und Gradverteilung wenig geeignet. Wir untersuchen deshalb die Evolution zufälliger Schnittgraphen (random intersection graphs) in Bezug auf wesentliche Parameter wie Komponentengröße, Durchmesser und Gradverteilung. Strukturelle Informationen dieser Art erlauben die Entwicklung von Algorithmen und Beweise über die asymptotische Optimalität derselben.

So haben wir zum Beispiel einfache Greedy-Strategien zum Clustering von Schnittgraphen analysiert. Diese erlauben es, die Suche in Daten dieser Art zu beschleunigen und die Daten effizient zu organisieren. Außerdem gewinnt man dadurch weitere Einblicke in die Ähnlichkeitsstruktur des Netzwerks.



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: DIPL.-INF. MANUEL BODIRSKY, DR. AMIN COJA-OGHLAN, PD DR. STEFAN HOUGARDY, DR. MIHYUN KANG, DR. MATHIAS SCHACHT, DIPL.-INF. DIRK SCHLATTER

Zusammenarbeit: Technische Universität Berlin, Konrad-Zuse-Zentrum für Informatik 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: Europäisches Graduiertenkolleg „Combinatorics, Geometry, and Computation“

Ansprechpartner: PROF. DR. HANS JÜRGEN PRÖMEL

Beteiligte Stipendiaten: MGR. JAN KARA (MARIE CURIE-STIPENDIAT), DIPL.-INF. DIRK SCHLATTER, CAND. SCIENT. TARAL SEIERSTAD, DR. MATHIAS SCHACHT

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.

Projekt: Eigene Stelle „The Regularity Method for Sparse Graphs and Hypergraphs“

Beteiligter Mitarbeiter: DR. MATHIAS SCHACHT

Forschungsförderung: Deutsche Forschungsgemeinschaft

Das zentrale Thema dieses Projektes ist die Weiterentwicklung der *Regularitätsmethode*, welche bereits viele wichtige Anwendungen in der diskreten Mathematik und der theoretischen Informatik hatte. Szemerédi's *Regularitätslemma* für Graphen besagt, dass die Knotenmenge jedes dichten Graphen so partitioniert werden kann, dass die meisten induzierten bipartiten Graphen viele Eigenschaften mit dem zufälligen bipartiten Graphen derselben Dichte gemeinsam haben. In gewissem Sinne kann also die Kantenmenge jedes Graphen durch wenige, zufallsartige bipartite Graphen sehr gut approximiert werden. Da zufällige Graphen oft einfacher zu analysieren sind als allgemeine Graphen, erlaubt das Regularitätslemma, Methoden und Resultate von zufälligen Graphen auf die Klasse aller Graphen zu übertragen, und hatte viele solche Anwendungen in der extremalen Graphentheorie.

In den letzten Jahren wurde das Regularitätslemma auf neue diskrete Strukturen erweitert: auf *dünne Graphen* und *k-uniforme Hypergraphen*. Im Besonderen die Generalisierung auf Hypergraphen war von Interesse, da sie zum Beispiel einen neuen Beweis Szemerédi's berühmten Satzes über die obere Dichte von Untermengen der ganzen Zahlen ohne arith-

metische Progression und dessen multidimensionalen Erweiterungen lieferte. In diesem Projekt untersuchen wir die zufallsartigen Eigenschaften, welche diese Generalisierungen des Regularitätslemmas garantieren, und wenden diese neuen Techniken auf Probleme der extremalen Graphentheorie und der theoretischen Informatik an.

Im Rahmen dieses Projektes kooperieren wir mit verschiedenen anerkannten Experten des Forschungsgebiets, unter anderem mit Professor Vojtěch Rödl (Atlanta, USA), Professor Endre Szemerédi (Budapest, Ungarn und New Brunswick, USA), Professor Yoshiharu Kohayakawa (São Paulo, Brasilien), Professor Andrzej Ruciński (Poznań, Polen), und Professor Anusch Taraz (München).

Veröffentlichungen

Patent

ST. HOUGARDY, M. THIMM, V. ZIEGLER: *Verfahren und Vorrichtung zum computergestützten Auffinden von ähnlichen Molekülen*. 2005. Patent.

Artikel

CH. AVART, V. RÖDL, M. SCHACHT: *Every Monotone 3-Graph Property is Testable*. In: 7th International Colloquium on Graph Theory, Band 22 der Reihe *Electronic Notes in Discrete Mathematics*, Seiten 539-542, 2005. (Extended abstract).

M. BODIRSKY: *The core of a countably categorical structure*. In: Proceedings of the 22nd International Symposium on Theoretical Aspects of Computer Science (STACS05), Stuttgart, LNCS 3404, Seiten 110-120, Springer Verlag 2005.

M. BODIRSKY, C. DANGELMAYR, J. KÁRA: *Representing Series-parallel Graphs as Intersection Graphs of Line Segments in Three Directions*. in: Innovative Applications of Information Technology for Developing World, 2005.

M. BODIRSKY, O. GIMENEZ, M. KANG, M. NOY: *On the number of series-parallel and outerplanar graphs*. In: Proceedings of European Conference on Combinatorics, Graph Theory, and Applications (EuroComb05), DMTCS Proceedings Series, Seiten 383-388, 2005.

M. BODIRSKY, C. GRÖPL, M. KANG: *Sampling unlabeled biconnected planar graphs*. In: Proceedings of the 16th Annual International Symposium on Algorithms and Computation (ISAAC05), LNCS 3827, Seiten 593-603, 2005. Springer Verlag.

M. BODIRSKY, D. KRÁL: *Locally Consistent Constraint Satisfaction Problems with Binary Constraints*. In: Proceedings of the 31st International Workshop on Graph Theoretic Concepts in Computer Science (WG05), LNCS 3787, Seiten 295-306, 2005. Springer Verlag.

M. BODIRSKY, M. KUHLMANN, M. MÖHL: *Well-nested Drawings as Models of Syntactic Structure*. In: Proceedings of the 10th conference on Formal Grammar and the 9th Meeting on Mathematics of Language (FG-MOL05), Edinburgh, 2005.

A. COJA-OGHLAN: *A spectral heuristic for bisecting random graphs*. In: Proceedings of the 16th ACM-SIAM Symposium on Discrete Algorithms, Seiten 850-859, 2005.

A. COJA-OGHLAN: *The Lovasz number of random graphs*. *Combinatorics, Probability and Computing*, 14: 439-465, 2005.

- A. COJA-OGHLAN, S. O. KRUMKE, T. NIERHOFF: *A hard dial-a-ride problem that is easy on average*. Journal of Scheduling, 8: 197-210, 2005.
- A. COJA-OGHLAN, J. MICHALICEK: *Limit theorems for non-commutative random variables*. Journal of Theoretical Probability, 18: 595-614, 2005.
- D. DRAKE, ST. HOUGARDY: *A linear time approximation algorithm for weighted matchings in graphs*. ACM Transactions on Algorithms, 1(1), 2005.
- F. HARARY, W. SLANY, O. VERBITSKY: *On the computational complexity of the forcing chromatic number*. In: V. Diekert und B. Durand (Herausgeber): Proc. of the 22nd Int. Symp. on Theoretical Aspects of the Computer Science (STACS 2005), Lecture Notes in Computer Science, 2005.
- J.-H. KIM, O. PIKHURKO, J. SPENCER, O. VERBITSKY: *How complex are random graphs in first order logic?* Random Structures and Algorithms, 26: 119-145, 2005.
- Y. KOHAYAKAWA, B. NAGLE, V. RÖDL, M. SCHACHT, J. SKOKAN: *The hypergraph regularity method and its applications*. Proceedings of the National Academy of Sciences USA, 102(23): 8109-8113, 2005.
- D. KÜHN, D. OSTHUS: *Spanning triangulations in graphs*. Journal of Graph Theory, 49: 205-233, 2005.
- D. KÜHN, D. OSTHUS: *Packings in dense regular graphs*. Combinatorics, Probability and Computing, 14: 325-337, 2005.
- D. KÜHN, D. OSTHUS: *4-cycles in graphs without a forbidden even cycle*. Journal of Graph Theory, 48: 147-156, 2005.
- O. PIKHURKO, J. SPENCER, O. VERBITSKY: *Decomposable graphs and definitions with no quantifier alternation*. In: Proc. of the 3rd European Conference on Combinatorics, Graph Theory and Applications EuroComb'2005, Berlin, Seiten 25-30, 2005.
- O. PIKHURKO, O. VERBITSKY: *Descriptive complexity of finite structures: saving the quantifier rank*. Journal of Symbolic Logic, 70(2): 419-450, 2005.
- H.-J. PRÖMEL: *Complete disorder is impossible. The mathematical work of Walter Deuber*. Combinatorics, Probability and Computing, 14: 3-16, 2005.
- V. RÖDL, M. SCHACHT, E. TENGAN, N. TOKUSHIGE: *Density theorems and extremal hypergraph problems*, erscheint im Israel Journal of Mathematics.
- M. SCHACHT, A. TARAZ: *Globally bounded local edge colourings of hypergraphs*. In: Proceedings of GRACO 2005, Band 19 der Reihe Electronic Notes in Discrete Mathematics, Seiten 179-185, 2005.
- O. VERBITSKY: *The first order definability of graphs with separators via the Ehrenfeucht game*. Theoretical Computer Science, 343(1-2): 158-176, 2005.

Preprints

- M. BEHRISCH, A. TARAZ: *Efficiently covering complex networks with cliques of similar vertices*. erscheint in Theoretical Computer Science.
- M. BEHRISCH: *Component evolution in random intersection graphs*. submitted.
- M. BEHRISCH, A. TARAZ, M. UECKERDT: *Colouring random intersection graphs and complex networks*. submitted.

- M. BODIRSKY, V. DALMAU: *Datalog and Constraint Satisfaction with Infinite Templates*. erscheint: Proceedings of the 23rd International Symposium on Theoretical Aspects of Computer Science (STACS06).
- M. BODIRSKY, J. NESETRIL: *Constraint Satisfaction with Countable Homogeneous Templates*. erscheint im Journal of Logic and Computation.
- MANUEL BODIRSKY, CLEMENS GRÖPL, MIHYUN KANG: *Generating Labeled Planar Graphs Uniformly at Random*. erscheint in Theoretical Computer Science.
- M. BODIRSKY, M. KANG: *Generating outerplanar graphs uniformly at random*, erscheint in Combinatorics, Probability and Computing.
- M. BODIRSKY, H. CHEN: *Oligomorphic Clones*. submitted.
- M. BODIRSKY, J. KÁRA: *The Complexity of Equality Constraint Languages*. submitted.
- M. BODIRSKY, A. BONATO: *The Infinite Locally Random Graph*. submitted.
- M. BODIRSKY, C. GRÖPL, D. JOHANNSEN, M. KANG: *A direct decomposition of 3-connected planar graphs*. 2005. Presented at the 17th Annual International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC05).
- M. BODIRSKY, M. KANG, M. LÖFFLER, C. MCDIARMID: *Random cubic planar graphs*. submitted.
- M. BODIRSKY, M. KANG, ST. VIGERSKE: *Enumeration of Unlabeled Outerplanar Graphs*. submitted.
- M. BODIRSKY, M. KUTZ: *Determining the Consistency of Partial Tree Descriptions*. submitted.
- T. BOHMAN, A. FRIEZE, T. LUCZAK, O. PIKHURKO, C. SMYTH, J. SPENCER, O. VERBITSKY: *The first order definability of trees and sparse random graphs*. erscheint in Combinatorics, Probability and Computing.
- A. COJA-OGHLAN: *A spectral heuristic for bisecting random graphs*. erscheint in Random Structures and Algorithms.
- A. COJA-OGHLAN: *Finding large independent sets in polynomial expected time.*, erscheint in Combinatorics, Probability and Computing.
- A. COJA-OGHLAN: *Solving NP-hard semirandom graph problems in polynomial expected time*. erscheint im Journal of Algorithms.
- A. COJA-OGHLAN, ANDREAS GOERDT, ANDRE LANKA: *Strong refutation heuristics for random k -SAT*. erscheint in Combinatorics, Probability and Computing.
- A. COJA-OGHLAN, SVEN O. KRUMKE, TILL NIERHOFF: *A heuristic for the stacker crane problem on trees which is almost surely exact*. erscheint im Journal of Algorithms.
- A. COJA-OGHLAN, CRISTOPHER MOORE, VISHAL SANWALANI: *MAX k -CUT and approximating the chromatic number of random graphs*. erscheint in Random Structures and Algorithms.
- A. COJA-OGHLAN: *Coloring semirandom graphs*.
- A. COJA-OGHLAN: *On the Laplacian eigenvalues of $G(n,p)$* .
- A. COJA-OGHLAN, M. KANG: *The evolution of the min-min random graph process*.

A. COJA-OGHLAN, L. KUHTZ: *An improved algorithm for approximating the chromatic number of $G(n,p)$.*

ST. HOUGARDY: *Classes of Perfect Graphs.* erscheint in Discrete Mathematics.

ST. HOUGARDY: *On a conjecture of Hoang and Tu concerning perfectly orderable graphs.* erscheint in Discrete Mathematics.

ST. HOUGARDY, ST. KIRCHNER: *Lower Bounds for the Realtime Greedy Algorithm for Approximating Steiner Trees.* erscheint in Networks.

ST. HOUGARDY, D. VINKEMEIER: *Approximating weighted matchings in parallel.* erscheint in Information Processing Letters.

ST. HOUGARDY, I. KÖTHNIG: *Enumeration of best possible low degree expanders.*

ST. HOUGARDY, F. LUTZ, M. ZELKE: *Realizations of polyhedra of genus 3.*

ST. HOUGARDY, F. LUTZ, M. ZELKE: *Polyhedra of genus 2 with 10 vertices and minimal coordinates.*

ST. HOUGARDY, F. LUTZ, M. ZELKE: *Polyhedral Tori with minimal coordinates.*

M. KANG, Y. KOH, T. LUCZAK, S. REE: *The connectivity threshold for the min-degree random graph process.* erscheint in Random Structures and Algorithms.

M. KANG, O. PIKHURKO: *Maximum K_{r+1} -free graphs which are not r -partite.* erscheint in Mat. Studii.

S. KIRCHNER: *An FPTAS for computing the similarity of three-dimensional sets.* eingereicht.

Y. KOHAYAKAWA, V. RÖDL, M. SCHACHT, P. SISSOKHO, J. SKOKAN: *Turán's theorem for pseudo-random graphs.* submitted.

B. NAGLE, V. RÖDL, M. SCHACHT: *The counting lemma for regular k -uniform hypergraphs.* Erscheint in Random Structures and Algorithms.

B. NAGLE, V. RÖDL, M. SCHACHT: *Extremal hypergraph problems and the regularity method.* submitted.

B. NAGLE, V. RÖDL, M. SCHACHT: *A short proof of the 3-graph counting lemma.* submitted.

V. RÖDL, M. SCHACHT: *Regular partitions of hypergraphs.* submitted.

V. RÖDL, M. SCHACHT, M. SIGGERS, N. TOKUSHIGE: *Integer and fractional packings of hypergraphs.* submitted.

Vorträge

M. BEHRISCH: *Euler und das Haus des Nikolaus.* GirlsDay 2005. Institut für Informatik der Humboldt-Universität zu Berlin, April 2005.

M. BEHRISCH: *Probabilistische Modelle biologischer Netzwerk.* Forschungsseminar Neue Entwicklungen in der Bioinformatik und Informationsintegration, Institut für Informatik der Humboldt-Universität zu Berlin, Dezember 2005.

M. BODIRSKY: *The Core of a Countably Categorical Structure.* 21th Annual Symposium on Theoretical Aspects of Computer Science (STACS'05), Stuttgart, Februar 2005.

M. BODIRSKY: *Constraint Satisfaction with Infinite Domains*. Research colloquium of the Departament de Tecnologia, Universitat Pompeu Fabra, Barcelona, Spanien, März 2005.

M. BODIRSKY: *Constraint Satisfaction with Infinite Domains*. Research colloquium of the Department of Computer Science, University of Durham, Großbritannien, März 2005.

M. BODIRSKY: *A model-theoretic approach to the complexity of constraint satisfaction problems*. Programming Systems Lab, Universität des Saarlandes, Saarbrücken, Juni 2005.

M. BODIRSKY: *Warum der Fußball aus 32 Flächen besteht*. Lange Nacht der Wissenschaften, Humboldt-Universität zu Berlin, Juni 2005.

M. BODIRSKY: *Constraint Satisfaction Problems for the Infinite Linear Order*. Midsummer Combinatorial Workshop, Department of Applied Mathematics (KAM), Karls-Universität Prag, Tschech. Republik, Juli 2005.

M. BODIRSKY: *Maximal Infinite-Valued Constraint Languages*. Workshop on Graphs, Morphismus and Applications am Centre de Recerca Matemàtica, Campus de la Universitat Autònoma de Barcelona, Bellaterra, Spanien, September 2005.

A. COJA-OGHLAN: *A spectral heuristic for bisecting random graphs*. 16. ACM-SIAM Symposium on discrete algorithms, Vancouver, Kanada, Januar 2005.

A. COJA-OGHLAN: *Heuristiken für NP-schwere Probleme*. Institut für Informatik, Universität Dortmund, April 2005.

A. COJA-OGHLAN: *A spectral heuristic for bisecting random graphs*. Seminar Design and Analysis of Randomized and Approximation Algorithms, IBFI Schloss Dagstuhl, Mai 2005.

A. COJA-OGHLAN: *A spectral heuristic for bisecting random graphs*. Seminar Exact Algorithms and Fixed-Parameter Tractability, IBFI Schloss Dagstuhl, Juli 2005.

A. COJA-OGHLAN: *Counting connected (hyper)graphs without counting*. 12th International Conference on Random Structures and Algorithms, Poznań, Polen, August 2005.

A. COJA-OGHLAN: *Stochastische Analyse von Algorithmen*. Universität Passau, Dezember 2005.

S. HOUGARDY: *Partitioning and Matching Algorithms*. Zentrum Mathematik, Technische Universität München, Februar 2005.

S. HOUGARDY: *Approximation Algorithms for Combinatorial Optimization Problems*. Fachbereich Informatik, Universität Dortmund, April 2005.

S. HOUGARDY: *Analysis and modelling of complex networks*. Workshop des DFG-Forschungszentrums (MATHEON) Mathematik für Schlüsseltechnologien, Berlin, April 2005.

S. HOUGARDY: *Effiziente Matchingalgorithmen*. Institut für Informatik, Heinrich-Heine-Universität Düsseldorf, Juli 2005.

S. HOUGARDY: *Greedy Algorithmen*. Institut für Mathematik, Technische Universität Ilmenau, Oktober 2005.

S. HOUGARDY: *Das Matchingproblem in Graphen*. Institut für Mathematik, Technische Universität Ilmenau, Oktober 2005.

S. HOUGARDY: *Effiziente Matchingalgorithmen*. Institut für Mathematische Optimierung, Otto-von-Guericke-Universität Magdeburg, Oktober 2005.

S. HOUGARDY: *Effiziente Matchingalgorithmen*. Institut für Computergraphik und Algorithmen, Technische Universität Wien, November 2005.

S. HOUGARDY: *Computing Exact Solutions for Hard Problems*. Waterman-Seminar, Institut für Pflanzengenetik und Kulturpflanzenforschung, Gatersleben, Dezember 2005.

M. KANG: *Random Planar Structures*. Karls-Universität Prag, Tschech. Republik, März 2005.

M. KANG: *Random Cubic Planar Graphs*. Midsummer Combinatorial Workshop XII, Institute for Theoretical Computer Science, Karls-Universität, Prag, Tschech. Republik, Juli 2005.

M. KANG: *Random Cubic Planar Graphs*. The 12th International Conference on Random Structures and Algorithms, Poznań, Polen, August 2005.

M. KANG: *On the Number of Series-parallel and Outerplanar Graphs*. European Conference on Combinatorics, Graph Theory, and Applications (EuroComb2005), TU Berlin, September 2005.

M. KANG: *Random Planar Structures*. Combinatorics Group Seminar, School of Mathematical Sciences Queen Mary, University of London, Großbritannien, Oktober 2005.

M. KANG: *Random Graph Processes*. Seminar on Discrete and Applicable Mathematics, Centre for Discrete and Applicable Mathematics, London School of Economics & Political Science, Großbritannien, Oktober 2005.

M. KANG: *Random Planar Structures*. Combinatorial theory Seminar, Mathematical Institute, University of Oxford, Großbritannien, Oktober 2005.

M. KANG: *Random Planar Structures*. Oberseminar, Institut für Informatik, HU Berlin, November 2005.

M. KANG: *Sampling unlabeled biconnected planar graphs*. 16. Int. Symposium Algorithms and Computation (ISAAC 2005), Hainan, China, Dez. 2005.

M. SCHACHT: *Discrepancy and Eigenvalues of Cayley Graphs*. ETH Zürich, Schweiz, Januar 2005.

M. SCHACHT: *Turán's theorem for sparse random graphs*. Karls-Universität Prag, Tschech. Republik, Februar 2005.

M. SCHACHT: *Ramsey properties of random hypergraph*. Kolloquiumsvortrag, Emory University, Atlanta, USA, April 2005.

M. SCHACHT: *Globally bounded local edge colourings of hypergraphs*. GRACO 2005 - 2nd Brazilian Symposium on Graphs, Algorithms, and Combinatorics, Angra dos Reis, Brasilien, April 2005.

M. SCHACHT: *Ramsey properties of random hypergraphs*. Universidade de São Paulo, Brasilien, Mai 2005.

M. SCHACHT: *On the regularity method for hypergraphs*. Alfréd Rényi Institut für Mathematik, Budapest, Ungarn, August 2005.

M. SCHACHT: *On the regularity method for hypergraphs*. Bristol-LMS Workshop on Additive Combinatorics, Bristol, Großbritannien, September 2005.

M. SCHACHT: *On the regularity method for hypergraphs*. Kolloquiumsvortrag, Emory University: Atlanta, USA, Oktober 2005.

M. SCHACHT: *On the regularity method for hypergraph.*, Oberseminar Theoretische Informatik, Humboldt-Universität zu Berlin, Dezember 2005.

DIRK SCHLATTER: *The random planar graph process*. Forschungsseminar, TU München, Juni 2005.

DIRK SCHLATTER: *The random planar graph process*. Konferenz RS&A Poznań, August 2005.

DIRK SCHLATTER: *Restricted random graph processes*. CGC-Colloquium, TU Berlin, November 2005.

DIRK SCHLATTER: *Restricted random graph processes*. Forschungsseminar, ETH Zürich, November 2005.

TARAL GULDAHL SEIERSTAD: *The minimum degree multi-graph process*. Learn- & Workshop "Random Graphs and Probabilistic Methods", Humboldt-Universität zu Berlin, März 2005.

TARAL GULDAHL SEIERSTAD: *The phase transition in the minimum degree random graph process*. The 12th International Conference on Random Structures and Algorithms in Poznań, Polen August 2005.

TARAL GULDAHL SEIERSTAD: *The minimum degree graph process*. Annual Workshop of the Graduiertenkolleg "Combinatorics, Geometry, and Computation", Hiddensee, September 2005.

TARAL GULDAHL SEIERSTAD: *The critical phase for random graphs with given degree sequence*. Department for Discrete Mathematics, Adam Mickiewicz Universität Poznań, Polen, November 2005.

OLEG VERBITSKY: *On the computational complexity of the forcing chromatic number*. The 22nd International Symposium on Theoretical Aspects of the Computer Science (STACS'05), Stuttgart, Februar 2005.

OLEG VERBITSKY: *Descriptive complexity of graphs*. Learn- & Workshop "Random Graphs and Probabilistic Methods". Humboldt Universität zu Berlin, März 2005.

OLEG VERBITSKY: *Descriptive complexity of finite graphs*. Institut für Informatik, TU München, Mai 2005.

OLEG VERBITSKY: *On the computational complexity of combinatorial forcing numbers*. Institut für Softwaretechnologie, TU Graz, Österreich, Juli 2005.

OLEG VERBITSKY: *Decomposable graphs and definitions with no quantifier alternation*. The 3rd European Conference on Combinatorics, Graph Theory and Applications (EuroComb'05), Berlin, September 2005.

Tagungen / Workshops

Learn&Workshop "Random Graphs and Probabilistic Methods"

Veranstalter: A. COJA-OGHLAN, M. KANG

Im Rahmen der DFG-Forschergruppe "Algorithmen, Struktur, Zufall" fand dieser Workshop vom 7.-9. März 2005 an der Humboldt-Universität mit ca. 40 Teilnehmern statt. Die Hauptvortragenden waren Joel Spencer vom Courant Institut der New York University und Oleg Verbitsky, der Alexander-von-Humboldt-Stipendiat am Institut für Informatik der Humboldt-Universität ist. Weitere Vorträge wurden gehalten von Joshua Cooper (New York University), Benjamin Dörr (MPI Saarbrücken), Volker Kaibel (ZIB), Mathias Schacht (HU Berlin) und Taral Seierstad (HU Berlin).

Berlin-Poznań Workshop on Random Structures

Dies ist ein gemeinsames Seminar mit Prof. M. Karoński von der Adam Mickiewicz Universität Poznań, Polen. Es werden ausgewählte Themen der algorithmischen Diskreten Mathematik und der Theorie zufälliger Graphen behandelt. Das Seminar findet als Blockseminar abwechselnd in Berlin und in Poznań statt. Im Jahr 2005 haben wir das Seminar am 4. Juni an der Humboldt-Universität ausgerichtet.

Die Vortragenden waren:

- Coja-Oghlan: Counting connected (hyper)graphs via the probabilistic method
- E. Szymanska: Distributed approximation algorithms in planar graphs
- J. Grytczuk: The game of arboricity
- T. Seierstad: The giant component in the minimum degree graph process

Sonstige Aktivitäten

Michael Behrisch

- Mitglied der Haushaltskommission
- Organisation des Studentenprogrammierwettbewerbs

Amin Coja-Oghlan

- Stellvertreter von Prof. Prömel als Leiter der Lehr- und Forschungseinheit Algorithmen und Komplexität

Stefan Hougardy

- Gastprofessor an der TU Berlin (bis April 2005)

Stefan Kirchner

- Nachrücker im Institutsrat für die wissenschaftlichen Mitarbeiter
- Aufgabe für den MATHEON-Adventskalender (erschieden am 23.12.2005)

Ralf Oelschlägel

- Mitglied des Institutsrates

Hans Jürgen Prömel

- Vizepräsident für Forschung der Humboldt-Universität zu Berlin
- Präsident der Humboldt-Universität zu Berlin, in Vertretung (Sept. - Dez. 2005)

- Mitglied des Präsidiums der Deutschen Mathematiker-Vereinigung
- Sprecher der DFG-Forschergruppe „Algorithmen, Struktur, Zufall“
- Mitglied des Rates des DFG-Forschungszentrums „Mathematik für Schlüsseltechnologien“ (MATHEON)
- 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
- Mitglied des Aufsichtsrats der Innovations-Zentrum Berlin Management GmbH
- Aufsichtsratsvorsitzender der Humboldt-Innovation GmbH
- Mitherausgeber der Zeitschriften Random Structures and Algorithms; Electronic Journal of Combinatorial Number Theory; Combinatorics, Probability and Computing; SIAM Journal on Discrete Mathematics

Mathias Schacht

- Vorlesung „Modern Methods in Ramsey Theory“ mit E. FRIEDGUT und V. RÖDL im Doccourse Prague 2005, Karls-Universität Prag, Tschech. Republik, Januar/Februar 2005.

Gäste am Lehrstuhl

PROF. DR. JOEL SPENCER, New York University, Courant Institute, März 2005.

DR. BENJAMIN DOERR, Max-Planck-Institut Saarbrücken, März 2005.

DR. JOSHUA COOPER, New York University, Courant Institute, März 2005.

DR. COLIN COOPER, King's College, Department of Computer Science, Mai 2005.

PROF. DR. BERTHOLD VÖCKING, RWTH Aachen, Department of Computer Science, Juni 2005.

PROF. DR. VOJTECH RÖDL, Emory University Atlanta, Department of Mathematics and Computer science, USA, Juni 2005.

A. LANKA, Technische Universität Chemnitz, Institut für Informatik, August 2005.

PROF. DR. ANDREAS GOERDT, Technische Universität Chemnitz, Institut für Informatik, August 2005.

HUBIE CHEN, Universität Pompeu Fabra, Barcelona, Spanien, Sept. bis Nov. 2005.

PROF. DR. MARTIN LOEBL, Department of Applied, Prag, November 2005.

ERIC FUSY, Projet Algo, INRIA (Institut National de Recherche en Informatique et en Automatique) Rocquencourt, Frankreich, November 2005.

DR. TIBOR SZABO, Institute of Theoretical Computer Science, ETH-Zentrum Zürich, Schweiz, Nov./Dezember 2005.

Diplomarbeiten

J. BÖTTCHER: *Coloring sparse random k -colorable graphs in polynomial expected time.* Januar 2005.

L. DÖLLE: *Untersuchung der Laufzeit des Random-Walk-Algorithmus auf zufälligen 2-CNF.* September 2005.

M. UECKERDT: *Färben von zufälligen Schnittgraphen,* Sept. 2005.

ST. VIGERSKE: *Asymptotic enumeration of unlabeled outerplanar graphs.* Oktober 2005.

M. LÖFFLER: *Counting and uniform generation of labeled planar structures.* November 2005.

M. ZELKE: *Ein Approximationsalgorithmus zur Berechnung eines 2-Spanners in planaren Triangulationen.* November 2005.

Lehr- und Forschungseinheit

Komplexität und Kryptografie

<http://www.informatik.hu-berlin.de/institut/struktur/algorithmenII>
(<http://www.informatik.hu-berlin.de/forschung/gebiete/algorithmenII>)

Leiter

PROF. DR. JOHANNES KÖBLER
Tel.: (030) 2093 3189
E-Mail: koebler@informatik.hu-berlin.de

Sekretariat

EVA SANDIG/MARGRIT HOPPE
Tel.: (030) 2093 3190
Fax.: (030) 2093 3191
E-Mail: sandig@informatik.hu-berlin.de
mhoppe@informatik.hu-berlin.de

Wissenschaftliche Mitarbeiter

DIPL.-MATH. OLAF BEYERSDORFF
PROF. DR. ERNST GÜNTHER GIESSMANN (APL. PROFESSOR)
PRIV.-DOZ. DR. WOLFGANG KÖSSLER
DIPL.-INFORM. MATTHIAS SCHWAN
DIPL.-INFORM. CARSTEN SCHWARZ

Techniker

DIPL.-ING. NORBERT HEROLD

Studentischer Mitarbeiter

MARTIN APEL

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 Beweissysteme 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 Querbezüge 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 3 (J. Köbler, O. Beyersdorff, C. Schwarz, H. Han, SoSe 05)
- Übungen zur Praktischen Informatik (W. KÖSSLER, WiSe 05/06)

Kernveranstaltungen (Halbkurse)

- Schaltkreiskomplexität (J. KÖBLER, SoSe 05)
- Kryptologie 1 (J. KÖBLER, O. BEYERSDORFF, WiSe 05/06)
- Stochastik für InformatikerInnen (W. KÖSSLER, WiSe 05/06)

- OpenSSL – Kryptologie in C (E.G. GIESSMANN, SoSe 05)
- Werkzeuge der empirischen Forschung (W. KÖSSLER, SoSe 05)

Seminare und Proseminare

- Proseminar: Kryptografische Algorithmen und Protokolle (J. KOEBLER, O. BEYERSDORF, WiSe 05/06)
- Seminar: Komplexität und Kryptologie (J. KÖBLER, SoSe 2005)
- Seminar: Elektronische Bezahlssysteme (J. KÖBLER, M.SCHWAN, SoSe 2005)
- Seminar: Komplexität und Kryptologie (J. KÖBLER, O. BEYERSDORFF, WiSe 05/06)
- Seminar: Interoperabilität und Sicherheit (E. GIESSMANN, WiSe 05/06)

Weitere Veranstaltungen

- Zirkel: Mathematische Schülergesellschaft (W. KÖSSLER, SoSe 2005)
- Zirkel: Mathematische Schülergesellschaft (W. KÖSSLER, WiSe 05/06)

Forschung

Projekt: Aussagenlogische Beweiskomplexität und disjunkte NP-Paare

Ansprechpartner: PROF. DR. JOHANNES KÖBLER

Beteiligte Mitarbeiter: OLAF BEYERSDORFF

Forschungsförderung: Deutsche Forschungsgemeinschaft

Die aussagenlogische Beweiskomplexität ist ein aktives Forschungsgebiet im Schnittpunkt von Komplexitätstheorie und Logik mit wichtigen Anwendungen in der künstlichen Intelligenz. Bislang hat sich die Forschung dabei vor allem auf schwache Beweissysteme wie Resolution konzentriert, während für starke Beweissysteme wie Frege-Systeme relativ wenig bekannt ist. Ein zentrales Anliegen dieses Projekts ist daher die Entwicklung einer allgemeinen Theorie zu starken Beweissystemen. Dazu wollen wir neue Ansätze aus der Komplexitätstheorie, Kryptografie und Logik kombinieren, bezüglich derer in den letzten Jahren beachtliche Fortschritte erzielt wurden. Im Einzelnen sind dies:

- die Verwendung von Pseudozufallsgeneratoren in der Beweistheorie,
- die Untersuchung des Verbands disjunkter NP-Paare und
- die Beziehung von Beweissystemen zur beschränkten Arithmetik.

Hauptziele des Projekts sind der Nachweis unterer Schranken für die Beweislänge in starken Beweissystemen unter Benutzung kryptografischer und komplexitätstheoretischer Härte voraussetzungen und die Charakterisierung von Beweissystemen im Verband disjunkter NP-Paare, den wir mit Methoden der Logik untersuchen wollen.

Projekt: Erstellung und Verifizierung eines Sicherheitsmodells für eine Signaturerstellungseinheit mit biometrischer Nutzerauthentifizierung

Ansprechpartner: PROF. DR. JOHANNES KÖBLER

Beteiligte Mitarbeiter: DIPL.-INF. MATTHIAS SCHWAN, PROF. DR. ERNST GÜNTHER GIESSMANN

Zusammenarbeit: Business Unit ITC Security der T-Systems International GmbH

Forschungsförderung: Deutsche Telekom AG

Das Projekt dient der Entwicklung sicherer, nach allgemein anerkannten Sicherheitskriterien evaluierbarer IT-Sicherheitssysteme, wie sie in immer stärkerem Maße z.B. für elektronische Bezahlsysteme, Systeme biometrischer Merkmalserkennung sowie Anwendungen elektronischer Signaturen notwendig werden.

In dem Forschungsvorhaben soll eine Sicherheitsstrategie für eine IT-Sicherheitsanwendung aufgestellt und formal modelliert sowie verifiziert werden. Die Anwendung umfasst das Erstellen einer elektronischen Signatur mit Hilfe eines geheimen kryptographischen Schlüssels mit vorheriger biometrischer Authentifizierung des Inhabers auf einer Chipkarte. Für die Entwicklung des Sicherheitsmodells wird auf einen generischen Ansatz Wert gelegt, so dass das Modell für verschiedene Implementationen nutzbar ist. Weiterhin werden Möglichkeiten der Werkzeugunterstützung genutzt, da Entwicklungszeiten durch Automatisierung verkürzt werden können. Es werden das Werkzeug "Verification Support Environment (VSE)" des DFKI sowie das "Simple Homomorphism Verification Tool (SHVT)" des SIT FhG gewählt.

Das Forschungsvorhaben wird gemeinschaftlich von der ITC Security der T-Systems GmbH und der Humboldt-Universität zu Berlin bearbeitet und ist die Fortführung des Vorgängerprojektes "Sicherheitsmodelle".

Projekt: Classical and Quantum Complexity of Graph Isomorphism and Related Problems

Ansprechpartner: PROF. DR. JOHANNES KÖBLER

Beteiligte Mitarbeiter: DIPL.-MATH. OLAF BEYERSDORFF

Zusammenarbeit: Institute of Mathematical Sciences, Chennai, Indien

Forschungsförderung: Deutscher Akademischer Austauschdienst

Das Graphenisomorphieproblem, bestehend in der Aufgabe, zwei Graphen auf Isomorphie zu testen, ist eines der faszinierendsten algorithmischen Probleme, da für dieses Problem bislang weder effiziente Algorithmen noch befriedigende Vollständigkeitsresultate bekannt sind.

Ziel dieses Projektes ist zum einen die genauere Einordnung des Graphenisomorphieproblems für eingeschränkte Graphklassen, wie etwa Graphen mit beschränktem Grad oder beschränkter Farbklasse, in Komplexitätsklassen unterhalb von P. Hierfür soll das Graphenisomorphieproblem im größeren Kontext gruppentheoretischer Probleme untersucht werden, da viele der bekannten Algorithmen auf gruppentheoretischen Prinzipien basieren.

Ein weiteres offenes Problem ist die Frage nach der Existenz effizienter Quantenalgorithmen für das Graphenisomorphieproblem. Auch hier spielen gruppentheoretische Probleme (hidden subgroup problem) eine zentrale Rolle. In Bezug auf das Graphenisomorphieproblem ist die wichtigste Frage, ob die bekannten Quantenalgorithmen für Abelsche Gruppen auf Permutationsgruppen übertragen werden können.

Projekt: Konsultationszentrum Statistik

Ansprechpartner: DR. WOLFGANG KÖSSLER

Zusammenarbeit: Zentraleinrichtung Computer- und Medienservice (Rechenzentrum)

Im vergangenen Jahr wurden unter anderem folgende Projekte unterstützt:

- Analyse von Konzentration und Leistung im Tagesverlauf (Institut für Psychologie)

- Soziologische Untersuchungen zur Schuleintrittszeit (Institut für Erziehungswissenschaften)
- Analyse der Zeitdauer ausgesprochener Vokale (Institut für Linguistik)
- Einkaufsverhalten und Verkehrsmittelnutzung (Institut für Geographie)
- Analyse von Kompetenz und Aggression von Grundschulern (Institut für Psychologie)

Projekt: Unabhängigkeitstests

Ansprechpartner: PROF. DR. EGMAR RÖDEL, DR. WOLFGANG KÖSSLER

Eine grundlegende statistische Fragestellung ist die nach der Unabhängigkeit von zwei Populationen. Wir untersuchen die wichtigsten in der statistischen Theorie und Praxis verwendeten Modelle, und berechnen asymptotische absolute und relative Effizienzen sowie die Gütefunktion verschiedener linearer Rangtests. Rangtests sind ein beliebtes Werkzeug in der Statistik, da sie nicht auf der Normalverteilungsannahme beruhen und somit recht robust sind. Weiterhin sind die Güteeigenschaften von Rangtests meistens besser als die der klassischen parametrischen Tests. Die Güteeigenschaften hängen stark von dem (angenommenen) Modell ab. Wie auch in vielen anderen Problemstellungen erweisen sich die einfachen, mittlerweile klassischen Rangtests, hier insbesondere der Spearman-Test, als eine gute Wahl.

Projekt: Lokationstests unter Verwendung von U-Statistiken

Ansprechpartner: DR. WOLFGANG KÖSSLER

Zusammenarbeit: Universität Chandigarh, Panjeb, Indien

U-Statistiken sind eine interessante Klasse von Statistiken, bei denen alle möglichen Teilstichproben von bestimmtem vorgegebenen Umfang aus der zugrunde liegenden Stichprobe gezogen werden. Für das Zweistichproben-Lageproblem ist der Mann-Whitney-Wilcoxon Test der klassische Vertreter dieser Klasse. Wir betrachten einige Verallgemeinerungen und untersuchen die asymptotischen Eigenschaften. Weiterhin konstruieren wir adaptive Tests, bei denen zunächst die Verteilung anhand einer Selektorstatistik bezüglich Tailstärke (und Schiefe) geschätzt und dann ein geeigneter, auf U-Statistiken beruhender Test, ausgewählt wird. Der vorgeschlagene adaptive Test hat gute asymptotische und finite Güteeigenschaften.

Projekt: SAS-Bediensystem

Ansprechpartner: DR. WOLFGANG KÖSSLER

Beteiligte Studenten und Mitarbeiter: DIRK-JOHANN DETERS, WOLF F. LESENER

Für den Vergleich zweier unabhängiger Messreihen gibt es eine Reihe von parametrischen und nichtparametrischen Methoden (z.B. t-Test, Wilcoxon-Test), aber auch erst in jüngerer Zeit entwickelte adaptive Verfahren. Der Anwender steht u.U. vor der Frage, welche der vielen Methoden anzuwenden ist. Wir haben ein in der SAS-Makrosprache geschriebenes Bediensystem entwickelt, das den Anwender durch seine Problemstellung führt, und, je nach seinem Kenntnisstand über die Daten, eine geeignete Methode auswählt. Es werden beliebige (stetige) Verteilungen zugelassen, also sowohl symmetrische als auch Verteilungen mit kurzen oder langen Tails. Für den Fall unbekannter Verteilungen sind adaptive Tests implementiert worden, bei denen zunächst die Form der Verteilung anhand einer Selektorstatistik klassifiziert und dann ein geeigneter Rangtest ausgewählt wird.

Veröffentlichungen

O. BEYERSDORFF: *Disjoint NP-pairs from propositional proof systems*. Technical report TR05-083, Electronic Colloquium on Computational Complexity, 2005. Erscheint in: Theory and Applications of Models of Computation (TAMC).

O. BEYERSDORFF: *Tuples of disjoint NP-sets*. Technical Report TR05-123, Electronic Colloquium on Computational Complexity, 2005.

J. KÖBLER, V. ARVIND: *On Hypergraph and Graph Isomorphism with Bounded Color Classes*. Erscheint in: Symposium on Theoretical Aspects of Computer Science (STACS).

J. KÖBLER, W. LINDNER: *The complexity of learning concept classes with polynomial general dimension*. Erscheint in: Theoretical Computer Science.

W. KÖSSLER, E. RÖDEL: *The asymptotic efficacies and relative efficiencies of various linear rank tests for independence*. Metrika, angenommen.

W. KÖSSLER: *Some c -Sample Rank Tests of Homogeneity Against Ordered Alternatives Based on U -Statistics*. Journal of Nonparametric Statistics 17 (2005), 777-795.

W. KÖSSLER: *Some c -Sample Rank Tests of Homogeneity Against Umbrella Alternatives With Unknown Peak*. Journal of Statistical Computation and Simulation, 76 (2006), 57-74.

M. SCHWAN, G. LASSMANN, G. ROCK, L. CHEIKHROUHOU: *Secure Biometric Identification System*. T-Systems International University Conference, Düsseldorf, Oktober 2005, to be published.

M. SCHWAN, G. LASSMANN: *Vertrauenswürdige Chipkartenbasierte Biometrische Authentifikation*. Jahrestagung der Gesellschaft für Informatik, Fachbereich Sicherheit, Magdeburg, Februar 2006, angenommen.

Vorträge

O. BEYERSDORFF: *Disjoint NP-pairs and propositional proof systems*. Mathematisches Institut der Akademie der Wissenschaften, Prag, Juni 2005.

O. BEYERSDORFF: *Complexity classes of disjoint NP-pairs*. Fall School of the Logic Seminar, Pec pod Snezkou, Tschechien, September 2005.

O. BEYERSDORFF: *Disjunkte NP-Paare und disjunkte NP-Tupel*. Komplexitätstheoretisches Colloquium, Universität Jena, Oktober 2005.

J. KÖBLER: *Complexity of testing isomorphism for restricted graph classes*. Institute of Mathematical Sciences, Chennai, Indien, März 2005.

J. KÖBLER: *Complexity of testing isomorphism for restricted graph classes*. UPC Barcelona, Spanien, April 2005.

J. KÖBLER: *Parameterisierte Erlernbarkeit von k -Juntas*. Universität Jena, Oktober 2005.

J. KÖBLER: *Graphisomorphie für spezielle Graphklassen*. Universität Würzburg, Dezember 2005.

W. KÖSSLER: *Ein SAS-Bediensystem für Lokationstests*. 9. Konferenz der SAS-Anwender in Forschung und Entwicklung, Humboldt-Universität zu Berlin, März 2005.

W. KÖSSLER: *Ein auf U-Statistiken beruhender adaptiver Test für das Zwei-Stichproben-Lokationsproblem (Poster)*. 9. Konferenz der SAS-Anwender in Forschung und Entwicklung, Humboldt-Universität zu Berlin, März 2005.

E.G. GIESSMANN: *IT-Sicherheit, Biometrie und Zugangsschutz*. 6. Berliner Katastrophenschutzforum des Landesamts für Arbeitsschutz, Gesundheitsschutz und technische Sicherheit, Berlin, September 2005.

Sonstige Aktivitäten

Prof. Dr. Johannes Köbler

- Studienberatung am Institut für Informatik
- Mitglied im Prüfungsausschuss
- Mitglied der Haushaltskommission
- Mitglied der Wahlleitung des Instituts für Informatik
- Juror beim Bundeswettbewerb Informatik

Dr. Wolfgang Kössler

- Eine Reihe von Gutachten für die Zeitschriften *Statistics*, *Statistical Papers*, *Allgemeines Statistisches Archiv*
- Mitglied des erweiterten Fakultätsrates
- Mitglied der Kommission Lehre und Studium
- Verantwortlicher für die Evaluierung der Lehre im Institut für Informatik
- Aktive Mitarbeit im Mentoring Programm
- Korrektor bei der Mathematik-Olympiade

Matthias Schwan

- Freier Mitarbeiter der T-Systems International GmbH im BMBF geförderten Projekt
- „Verisoft – Beweisen als Ingenieurwissenschaft“
- 6-monatiges Praktikum in Singapur bei der T-Systems Singapore Pte Ltd.

Gäste am Lehrstuhl

PROF. DR. V. ARVIND, Mathematical Institute C.I.T. Campus, Chennai, Indien, Juni/Juli 2005.

PROF. DR. OSAMU WATANABE, Tokyo, Institute of Technology, Japan, Februar 2005.

Diplomarbeiten

CARSTEN SCHWARZ, *Undeniable Signatures*, April 2005.

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. ANDREAS KUNERT

DIPL.-INF. MICHAEL PIEFEL

Projektmitarbeiter

DIPL.-INF. ALEXANDER AUERBACH

DIPL.-INF. IVO KÖTHNIG

DIPL.-INF. TOBY NEUMANN

Technischer Mitarbeiter

DIPL.-MATH. MANFRED HAGEN

Habilitanden

DR. RER. NAT. HABIL. ECKHARDT HOLZ

MARIO KOLBERG, PH. D.

Promovenden

DIPL.-INF. HARALD BÖHME

DIPL.-INF. TOM RITTER

DIPL.-INF. MARKUS SCHEIDGEN

DIPL.-INF. MICHAEL SODEN

DIPL.-INF. FRANK STOINSKI

EMERITUS

PROF. DR. HABIL. GUNTER SCHWARZE

TUTOREN

GLENN SCHÜTZE

KONRAD VOIGT (BIS 11/05)

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 des Lehrstuhls.

Im Berichtszeitraum wurden folgende strategische Zielstellungen in der Forschung umgesetzt:

1. **Weitere Stabilisierung von Konzepten, Modellen und Technologien** einer komponentenorientierten Softwareentwicklung bei Bereitstellung geeigneter Entwicklungsumgebungen und Werkzeuge. Im Rahmen eines BMBF-Projektes wurde die Ausgründung einer Firma vorbereitet, die eine neuartige Komponentenplattform mit Patentabsicherung als Produkt anbietet. Eine Vielzahl von Transformationswerkzeugen sichert eine wesentlich höhere Flexibilität dieser Plattform als die am Markt verfügbaren. Diese drückt sich in einer sowohl server- als auch clientseitigen Unterstützung eines Komponentenansatzes aus. Zudem ist die Wahl der Implementationssprache nicht vom Konzept oder Framework der Komponenteninteraktion abhängig. Des Weiteren kann dieses Framework durch CORBA, Web-Services oder .net ohne Modifikation der Anwendung bestimmt sein, wobei zusätzlich leicht von einer Einzelplatzanwendung auf eine verteilte Anwendung umgestellt werden kann.

2. **Konsolidierung der ULF-Ansatzes**

ULF (Unified Language Family) ist eine metamodellbasierte Plattform, für die von uns Konzepte und Technologien zur Definition und Anwendung domänenspezifischer Modellierungssprachen entwickelt werden. Die Motivation für diesen Teil unserer Forschung ergab sich ursprünglich aus der sehr praxisrelevanten Fragestellung nach der Kombination von Modellierungssprachen im Kontext einer modellgetriebenen Softwareentwicklung, wofür eODL, SDL, MSC und UML als relevant angesehen worden sind.

Inzwischen steht eine Modelldatenbank als Kern der Plattform zur Verfügung, über die eine Reihe von Transformationen identifiziert wurden, von der momentan die SDL-nach-C++/Java-Transformation (als Codegenerator) bearbeitet wird. Voraussetzung dafür ist zum einen ein Metamodell für SDL und ein hoffentlich gemeinsames Metamodell von C++ und Java. Daraus wiederum wurde die allgemeine Fragestellung abgeleitet, wie man überhaupt aus Grammatik-basierten Sprachdefinitionen metamodellbasierte Entsprechungen ableiten kann, die sich für Analysen und Transformationen eignen. Hier wurden Erkenntnisse gewonnen, die insbesondere eine automatische Generierung ermöglichen.

Die Entscheidung, derartige Metamodelle konform zum neuesten MOF-Standard (MOF-2.0) zu entwickeln, erwies sich dabei als strategisch wertvoll, weil sie sich bereits im Hinblick auf die erreichte Ausdrucksmächtigkeit der geschaffenen Modelldatenbank auszahlt. Für eine Teilsprache von SDL ist zudem eine Übertragung klassischer Compiler-Aufgaben, wie die der Syntax- und Semantikanalyse auf die metamodellbasierte Repräsentation gelungen, wobei entscheidende Vorteile gegenüber den klassischen Ansätzen herausgearbeitet werden konnten. Die Anbindung von Simulationswerkzeugen zum Ausführen und Testen von Modellen befindet sich dagegen noch in den Anfängen.

3. Erfolgreiche Projektvorbereitung

Der Berichtszeitraum war durch eine Reihe intensiver Projektvorbereitungen geprägt. Als erfolgreich gelten die Zusagen der DFG zur Förderung eines interdisziplinären Graduiertenkollegs für die modellbasierte Entwicklung selbstorganisierender Systeme für den Einsatz im Katastrophenmanagement, das Prof. J. Fischer als Sprecher vertreten wird, und weitere Kooperationsinitiativen zwischen dem Geoforschungszentrum Potsdam und den Instituten für Informatik und Geographie der Humboldt-Universität. Die Beteiligung der Lehr- und Forschungseinheiten *Systemarchitektur* (Prof. J.-P. Redlich) und *Systemanalyse* (Prof. J. Fischer) an dem EU-Projekt SAFER zum Aufbau eines Europäischen Erdbebenfrühwarnsystems (in dem auch die Türkei und Ägypten einbezogen sind) ist ein konkretes Beispiel dafür.

Lehre

Veranstaltungen im Grundstudium

- Praktikum zur Vorlesung „Praktische Informatik I“ (KLAUS AHRENS, WiSe 2004/2005 und WiSe 2005/2006)
- Praktikum zur Vorlesung „Praktische Informatik II“ (KLAUS AHRENS, SoSe 2005)
- Vorlesung „Praktische Informatik III“ (JOACHIM FISCHER, WiSe 2004/2005)
- Praktikum zur Vorlesung „Praktische Informatik III“ (KLAUS AHRENS, ANDREAS KUNERT, MICHAEL PIEFEL, WiSe 2004/2005)
- Praktikum zur Vorlesung „Praktische Informatik III“ (ANDREAS KUNERT, WiSe 2005/2006)
- Proseminar „Softwarespezifikation mit UML“ (JOACHIM FISCHER, SoSe 2005)
- Proseminar „Wissenschaftliches Publizieren mit L^AT_EX“ (MICHAEL PIEFEL, SoSe 2005)

Veranstaltungen im Hauptstudium

- Kurs „Objektorientierte Analyse, Modellierung, Spezifikation und Implementierung [OMSI]“ (JOACHIM FISCHER, KLAUS AHRENS, SoSe 2005)
- Praktikum zum Kurs „Objektorientierte Analyse, Modellierung, Spezifikation und Implementierung [OMSI]“ (MICHAEL PIEFEL, SoSe 2005)
- Vorlesung (Halbkurs) „Objektorientierte Analyse, Modellierung, Spezifikation und Implementierung [OMSI 1]“ (JOACHIM FISCHER, KLAUS AHRENS, WiSe 2005/2006)
- Praktikum zum Halbkurs „Objektorientierte Analyse, Modellierung, Spezifikation und Implementierung [OMSI 1]“ (MICHAEL PIEFEL, WiSe 2005/2006)
- Vorlesung „Softwareentwicklung mit UML 2“ (JOACHIM FISCHER, WiSe 2005/2006)
- Seminar „Objektorientierter Entwurf und Analyse verteilter Systeme“ (JOACHIM FISCHER, WiSe 2004/2005)
- Seminar „Komponenten-Orientierte Softwareentwicklung und Plattformen“ (JOACHIM FISCHER, WiSe 2004/2005)
- Seminar „Simulation von Rechnernetzen mit ns-2“ (JOACHIM FISCHER, SoSe 2005)

- Projektseminar “ULF – *Unified Language Family*”
(JOACHIM FISCHER, SoSe 2005)
- Seminar „Komponenten-Orientierte Softwareentwicklung“
(JOACHIM FISCHER, SoSe 2005)
- Seminar „Compilergeneratoren“
(ANDREAS KUNERT, SoSe 2005)
- Seminar „Metamodellierung für modellgetriebene Softwareentwicklung mit MDA und UML“
(MICHAEL PIEFEL, MARKUS SCHEIDGEN, WiSe 2005/2006)

Forschung

Schwerpunkt: Plattform-Entwicklung für Komponentenbasierte Systeme

Ansprechpartner: PROF. DR. JOACHIM FISCHER

Beteiligte Mitarbeiter: ALEXANDER AUERBACH, HARALD BÖHME, IVO KÖTHNIG, TOBY NEUMANN, TOM RITTER, GLENN SCHÜTZE, FRANK STOINSKI, KONRAD VOIGT

Zusammenarbeit: Siemens AG

Forschungsförderung: BMBF

Im Berichtszeitraum konnten weitere entscheidende Beiträge zur Ausgestaltung einer modellgetriebenen Software-Entwicklung komponentenbasierter Systeme erbracht werden.

Zum einen wurde mit der Entwicklung eines e-ODL-nach-CIDL-Compilers (etoc) der Nachweis erbracht, dass sich abstrakte, strukturell ausgerichtete Komponentenmodelle (als technologieunabhängige Beschreibungen in eODL) automatisch in technologieabhängige Beschreibungen überführen lassen, wobei die CORBA-Komponententechnologie (beispielhaft) als konkrete Technologie ausgewählt worden ist. Genau für diese wurde ja in den vergangenen Jahren an der LFE Systemanalyse mit der Open-Source-Entwicklung einer CCM-Implementierung überhaupt erst eine praktische Erprobungsmöglichkeit geschaffen.

Zum anderen konnten die Ergebnisse der Diplomarbeit von Alexander Auerbach erfolgreich für den Aufbau einer flexibleren Komponententechnologie als die von CCM im Rahmen eines BMBF-Projektes umgesetzt werden. Mit der im Projekt konzipierten, toolgestützten netCCM-Technologie werden außer der Bereitstellung einer Life-Cycle-Management-Funktionalität für Komponenten insbesondere eine Reihe neuer Paradigmen für eine komponentenbasierten Systementwicklung ermöglicht:

- Neben einer serverseitigen Komponentenarchitektur wird auch eine clientseitige Komponentengestaltung erlaubt.
- Der zweite Aspekt betrifft die prinzipielle Austauschbarkeit des Remoting-Frameworks der Plattform zur Komponenteninteraktion. Hier kann zwischen CORBA, .net und Web Services gewählt werden, wobei Einzelplatzanwendungen problemlos in Mehrplatzanwendungen (und umgekehrt) überführbar sind.
- Ein dritter Aspekt betrifft die Unabhängigkeit in der Wahl der Implementierungssprache der Komponenten vom jeweils gewählten Remoting-Framework.

Erreicht wird diese Flexibilität durch die Bereitstellung vielfältiger Transformationswerkzeuge. Ein weiterer Ansatz (bislang nur konzeptionell angedacht) soll die Konzepte einer Verhaltensmodellierung von Komponenten und einer automatischen Codegenerie-

rung auf die gleiche Art und Weise auf netCCM übertragen, so wie sie für die @vantage-Plattform für die Siemens AG im Rahmen der SDL-SITE- Entwicklung an der LFE Systemanalyse realisiert worden ist.

Leider ist es den externen Doktoranden, die an diesem Schwerpunktthema arbeiten, im Berichtszeitraum nicht gelungen, ihre vorliegenden und z.T. bereits publizierten Forschungsergebnisse zu CCM und @vantage betreffenden Technologien wie geplant als zusammenfassende Dissertationsarbeiten fertig zu stellen.

etoc

(Harald Böhme, Glenn Schütze, Konrad Voigt)

etoc ist ein Transformator von eODL nach CIDL. Mit seiner Hilfe können eODL-Modelle in CIDL-Modelle übersetzt werden. eODL kann zur Softwarekomponentenentwicklung von verteilten Systemen eingesetzt werden. Durch die Unterstützung von verschiedenen Sichten auf das Softwaremodell, können unterschiedliche Ziele bei der Modellierung verfolgt werden. Die in eODL verfügbaren Sichten sind:

- die rechenbezogene Sicht (computational view),
- die implementierungsbezogene Sicht (implementational view),
- die verteilungsbeschreibende Sicht (deployment view) und
- die Sicht auf die Zielumgebungen (target environment view).

Mit Hilfe der Sprache CIDL können die Struktur und Teile des Verhaltens von Komponentenimplementierungen im CORBA Komponentenmodell (CCM) beschrieben werden. Entsprechende ORB-Produkte generieren anhand der CIDL-Spezifikation Implementierungsschablonen. Diese können von weiteren Werkzeugen ausgefüllt werden. Letztendliches Ziel dieser Vorgehensweise ist eine vollständige Implementierung von Softwarekomponenten. Im Rahmen des etoc-Projektes ist ein Werkzeug entstanden, das zur modellbasierten Softwareentwicklung (MDA: Model Driven Architecture) im Kontext der Softwarekomponentenentwicklung beiträgt. Das Werkzeug etoc besteht aus drei Teilen: Eingabeteil, Transformationsteil und Ausgabeteil. Der Eingabeteil und der Ausgabeteil realisieren das Lesen bzw. Schreiben der Modelle in XMI oder Text-Format in bzw. aus einer MOF-Modelldatenbank (MOF-Repository). Der Transformationsteil implementiert die Regeln zur Transformation, die als typbasierte Abbildungsregeln definiert wurden. Das Projekt etoc wurde im Juni 2005 abgeschlossen und auf dem SDL-Forum 2005 in Grimstad (Norwegen) präsentiert.

netCCM

(Alexander Auerbach, Ivo Köthnig, Toby Neumann)

netCCM ist eine Sammlung eng aufeinander abgestimmter Werkzeuge und verschiedener Laufzeitumgebungen für die Entwicklung betriebswirtschaftlicher Standard- und Branchensoftware. Basierend auf einem neuen einheitlichen Komponentenmodell ermöglicht netCCM die Entwicklung komplexer verteilter Anwendungen wie Einzelplatzanwendungen. Damit wird der Entwicklungsaufwand für Softwarehäuser um 30 % bis 50 % reduziert.

Das Projekt zur Vorbereitung der Unternehmensgründung wurde im Rahmen des BMBF-Programmes **EXIST-SEED** gefördert. So konnte ein Businessplan geschrieben werden, der Anfang 2006 fertig gestellt werden wird. Der Entwurf des Businessplanes wurde im Juli 2005 im Rahmen des bundesweiten IT-Gründungswettbewerbs **start2grow**

(<http://www.start2grow.de/>) als einer der zehn besten Businesspläne prämiert. Er dient als Basis für die Finanzierung des Unternehmens, das netCCM vermarkten und weiterentwickeln wird. Daneben konnte ein Netzwerk aufgebaut werden, das die Unternehmensgründung unterstützt. Außerdem wurde das Produkt so vervollkommen, dass im Februar 2006 mit der Akquisition von Pilotkunden begonnen werden kann.

***Vergleich des FUS-Managements in @vantage –
Proprietäres Konzept versus Relationales Datenbanksystem***
(Frank Kühnlenz, im Auftrag der Siemens AG)

In dieser Diplomarbeit wurde die verteilte Datenhaltung in zwei Persistenzmedien innerhalb der Telekommunikationsplattform @vantage der Siemens AG untersucht. Telekommunikationsbetreiber stellen besondere Anforderungen (carrier-grade Aspekte), z.B. hinsichtlich Verfügbarkeit und Antwortzeiten. In diesem Kontext wurde die proprietäre Datenhaltung in der Anwendung Prepaid@vantage bezüglich architektonischer Vor- und Nachteile analysiert. Auf Basis von Oracles Relationalem Datenbanksystem entstand ein Prototyp, der die proprietäre Datenhaltung vollständig ablöst. Beide Implementierungen sind umfangreichen Messungen unter realitätsnahen Bedingungen unterzogen worden. Die gewonnenen empirischen Daten ermöglichten einen fundierten Vergleich beider Konzepte.

Schwerpunkt: Metamodellbasierte Entwicklung von Sprachwerkzeugen – ULF Ware

Ansprechpartner: PROF. DR. JOACHIM FISCHER

Beteiligte Mitarbeiter: ANDREAS KUNERT, MICHAEL PIEFEL, MARKUS SCHEIDGEN

Zusammenarbeit: IKV⁺⁺ AG (Berlin), Høgskolen i Agder (Grimstad, Norwegen)

Forschungsförderung: Land Berlin

ULF-Ware ist die technologische Umsetzung des konzeptuellen Modells von ULF (Unified Language Family) in Form einer metamodellbasierten Plattform. Kern dieser Plattform bildet eine Modelldatenbank, implementiert als Java MDR (Java Meta Data Repository). Diese Plattform wird in Abhängigkeit von ihrem Einsatzgebiet mit unterschiedlichen Werkzeugkomponenten zur Erfüllung semantischer Anforderungen einer Transformation angereichert. Folgende Einsatzfälle sind langfristig vorgesehen:

- SDL-nach-C++/Java
- eODL/SDL-nach-C#/.Net
- eODL/SDL/UML-nach-C++/Java/C#

Gegenwärtig wird die Realisierung von ULF-Ware: SDL-nach-C++/Java bearbeitet. Als SDL wird dabei zunächst „nur“ ein Sprachkern von SDL-2000 (SDL/R) untersucht, der von Andreas Prinz (Høgskolen i Agder) vorgeschlagen worden ist. Für diese Teilsprache gibt es bereits ein Metamodell (entwickelt von Markus Scheidgen) in der ULF-Ware-Modelldatenbank.

Semiautomatische Metamodellgenerierung
(Andreas Kunert)

Auch wenn in zahlreichen Publikationen gezeigt wurde, dass die modellbasierte Definition von (Programmier-) Sprachen der klassischen Definition (kontextfreie Grammatik und natürlichsprachlicher Begleittext) überlegen ist, existieren nach wie vor für die wenigsten der heute gebräuchlichen Sprachen Metamodelle. Ein Hauptgrund dafür liegt sicherlich

darin, dass der bisherige Ansatz zur Generierung von Metamodellen zu bereits existierenden Sprachen auf einem Grossteil Handarbeit beruht. Ein aktuelles Forschungsprojekt am Lehrstuhl beschäftigt sich mit einem semi-automatischen Lösungsansatz des Problems. Ziel ist es, aus der grammatikbasierten Definition der Programmiersprache automatisch ein Metamodell zu generieren. Zusätzlich soll es möglich sein, durch von Hand vorzunehmende Annotationen in der Grammatik die Metamodellgenerierung zu beeinflussen. Parallel zur Metamodellgenerierung soll automatisch ein Compiler generiert werden, der Programme der Grammatik einlesen kann und aus diesen Modelle erstellt, die dem vorher generierten Metamodell genügen. Ein Erfolg dieses Projektes hätte zur Folge, dass die im ULF-Ware-Projekt erstellte Werkzeugkette um ein Frontend erweitert werden könnte, welches die Verarbeitung verschiedenster Programmiersprachen erlaubt.

Metamodelle für objektorientierte Programmiersprachen

(Michael Piefel)

Das Projekt ULF-Ware beschäftigt sich in seiner ersten Ausbaustufe mit einem metamodellbasierten Compiler für SDL/R. Metamodellbasiert bedeutet in diesem Zusammenhang, dass nicht nur am Anfang ein Modell einer SDL/R-Spezifikation entsprechend des SDL/R-Metamodells erstellt wird, sondern dass auch auf der Ausgabeseite Modelle stehen. Das ermöglicht die Verwendung von Modelltransformationen zum Übergang zwischen abstrakter Spezifikation und konkreter Ausführung. Ähnlich wie für die Eingangsseite ein SDL/R-Metamodell geschaffen werden musste, wird ein Metamodell auch auf der Ausgangsseite benötigt. An der LFE haben wir dabei die Idee entwickelt, ein Metamodell nicht für eine konkrete Ausgabesprache wie etwa C++ zu entwickeln, sondern stattdessen die großen Gemeinsamkeiten der Mehrheit objektorientierter Programmiersprachen auszunutzen und ein Metamodell für diese verallgemeinerte Sicht zu entwerfen. Erste Entwürfe dieses Metamodells wurden zur Publikation eingereicht und sind auf großes Interesse gestoßen.

SDL/R: Syntax- und Semantikanalyse

(Markus Scheidgen)

Syntax- und Semantikanalyse sind Bestandteile herkömmlicher Compiler, also Werkzeuge zur Übersetzung von Spezifikationen und Programmen einer Sprache. Für die Syntaxanalyse werden kontextfreie Grammatiken als Modelle der Sprachsyntax verwendet, die Semantikanalyse wird herkömmlicherweise ohne weitere Hilfe manuell implementiert und am Syntaxbaum, also dem Ergebnis der Syntaxanalyse, durchgeführt. Bereits im letzten Jahr konnten wir zeigen, wie sich aus kontextfreien Grammatik, also Syntaxmodellen, objekt-orientierte Metamodelle erstellen lassen. Aus der Grammatik für SDL konnte ein SDL-Metamodell erstellt werden. Die dabei entstandene Abbildung von SDL-Grammatik nach Metamodell konnte nun dafür verwendet werden, eine Syntaxanalyse zu entwerfen, welche eingegebene SDL-Spezifikationen verarbeitet und im Ergebnis ein Modell erzeugt, welches dem SDL-Metamodell entspricht. Im Vergleich zum herkömmlichen Compilerbau ist also kein Syntaxbaum entstanden, sondern ein objekt-orientiertes Modell, welches die ursprüngliche SDL-Spezifikation widerspiegelt.

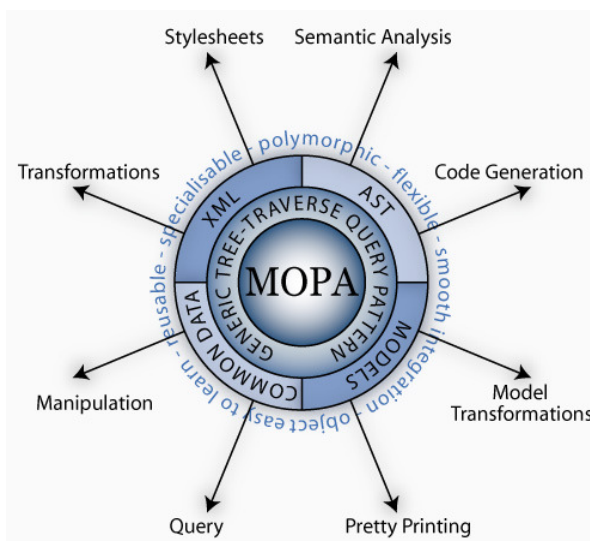
Der Vorteil dieser Vorgehensweise wird bei der Semantikanalyse deutlich. Während Syntaxbäume lediglich durch Programme analysiert werden, die in allgemeinen Programmiersprachen geschrieben sind, existieren für die Analyse von Modellen Sprachen, welche speziell auf die Modellierung statischer Zwangsbedingungen, also statischer Semantik, zugeschnitten sind. Dies bietet die folgenden Vorteile: Die Semantikanalyse erfolgt nach formalen Semantikmodellen mit Hilfe prädikatenlogischer Ausdrücke, welche sich mit der Ob-

ject Constraint Language (OCL) als Zwangsbedingungen formulieren lassen; des Weiteren können diese Zwangsbedingungen als Teil des Metamodells dargestellt werden, wodurch sich der Aufwand zur Entwicklung einer Semantikanalyse auf das Modellieren der Semantik beschränkt; des Weiteren beschreibt eine Zwangsbedingung nur die Semantik eines einzelnen Konzepts. Abstrakte Konzepte bleiben also isoliert und können weiterhin wieder verwendet werden, was bei den üblichen monolithischen Semantikanalysen herkömmlicher Compiler nicht möglich ist.

Neben der Prüfung von Zwangsbedingungen existieren noch weitere Aspekte der Semantikanalyse. Hierzu gehören das Auflösen von namentlichen Referenzen und das Erzeugen implizierter Sprachkonstrukte, die für die weitere maschinelle Verarbeitung eines Modells nötig sind. Allgemein lässt sich sagen, dass eine Semantikanalyse auf der einen Seite ein gegebenes Modell auf Korrektheit prüft und zum anderen das Modell aus einer für den Benutzer geeigneten Darstellung in eine für die maschinelle Verarbeitung günstige Darstellung umwandelt. Um diese Überführung zu realisieren, wurden die nötigen Umwandlungen als Modelltransformation dargestellt. Dabei werden zwei Metamodelle verwendet; ein Modell welches die Sprache des Benutzers beschreibt, die Eingabe der Semantikanalyse, und ein Modell für die Maschinenseite, die Ausgabe der Semantikanalyse. Im Falle von SDL konnte das Eingabemodell als Erweiterung des Ausgabemodells modelliert werden.

MOPA – Modelltransformationen mit Model Pattern (Markus Scheidgen¹)

Für die Implementierungen solcher Modelltransformationen haben wir das Werkzeug MOPA (Model Pattern) entwickelt. Im Gegensatz zu existierenden Transformationswerkzeugen, welche jeweils sehr speziell und auf eine bestimmte Art von Transformationsaufgabe zugeschnitten sind, ist MOPA zwar oft nicht so ausdrucksstark wie diese Konkurrenz, bietet jedoch eine höhere Flexibilität, welche MOPA als Allzweckwerkzeug für die verschiedensten Transformationsaufgaben interessant macht. Des Weiteren arbeitet MOPA auf einer Abstraktion von Modellen, welche es erlaubt, MOPA nicht nur für die verschiedensten MOF- oder MOF-ähnlichen Modelldatenbanken, sondern auch für viele andere Arten von Datenrepräsentationen, wie zum Beispiel Syntaxbäume oder XML-Dokumente, zu verwenden.



Eine Modelldatenbank für MOF 2.0 (Markus Scheidgen)

Bei der Arbeit an den Modellen zu SDL und ihrer statischen Semantik haben sich einige Mängel in den zur Verfügung stehenden Konzepten zur Metamodellierung, also Schwächen im Meta-Metamodell, gezeigt. Der bislang verwendete Standard zur Metamodellie-

¹ Markus Scheidgen erhält im Rahmen der Berliner Graduiertenförderung (NaFöG) ein Stipendium vom Land Berlin.

rung, MOF 1.4, welcher in seinen Grundzügen die Fähigkeiten der meisten heute existierenden Modelldatenbanken zur Verarbeitung von Modellen entspricht, erlaubt es nicht, einzelne Eigenschaften von Konzepten bei der Spezialisierung eines allgemeineren Sprachkonstrukts zu erweitern, wodurch das Formulieren und die Wiederverwendung von abstrakten Metamodellelementen erheblich eingeschränkt wird.

Um erweiterte Sprachkonzepte als Erweiterungen und Kombinationen von einfachen Sprachkonzepten darzustellen und damit die geplante Erweiterung von SDL/R zu ermöglichen, musste genau an dieser Schwachstelle in der Metamodellierung angesetzt werden. Nun wurde bereits für die Entwicklung von UML 2.0 der nachfolgende Metamodellierungsstandard MOF 2.0 so erweitert, dass sich auch Konzepteneigenschaften bei der Spezialisierung näher beschreiben lassen. Es ist gelungen, die entsprechenden neuen Metamodellierungskonzepte in ihrer Semantik so zu erweitern, dass sie in einer MOF-2.0-Implementierung realisiert werden konnten. Somit steht jetzt ein Metamodellierungswerkzeug zur Verfügung, mit dem die geplanten Schritte zur Erweiterung von Sprachen um zusätzliche Konzepte auch durchgeführt werden können. In Hinblick auf die künftig geplante Modellierung von gemeinsamen SDL- und UML-Konzepten ist damit bereits der erste Schritt getan, da man nun über die nötige Infrastruktur (genau genommen die UML-Infrastruktur) verfügt, um das offizielle UML-2.0-Metamodell zu verarbeiten, welches ein MOF-2.0-Metamodell ist.

Nach unserem Kenntnisstand ist diese an der LFE entstandene MOF-2.0-Modelldatenbank die einzige derzeit verfügbare Modelldatenbank für diesen neuen Standard.

MDA in practice

(Joachim Hößler, Michael Soden, Hajo Eichler)

Bei einer initialen Diskussion von MDA-bezogenen Promotionsthemen für J. Hößler, M. Soden und H. Eichler wurden ihre gewonnenen praktischen Erfahrungen aus einem Projekt mit einem großen japanischen Systemintegrationshaus vorgestellt, welches die IKV++ Technologies AG zusammen mit dem Fraunhofer Institut FOKUS realisiert hat. Projektgegenstand war die Einführung der Model Driven Architecture (MDA) der Object Management Group (OMG) in einen bestehenden Software Entwicklungsprozess. Ziel des Projektes war es, einen hohen Grad an Automation, Nachverfolgbarkeit und Konsistenzprüfung aller Modellierungsartefakte zu ermöglichen. Erreicht wurde dieses Ziel durch den Bau einer integrierten Werkzeugkette, bei der alle Systemmodellierungselemente in einem zentralen Repository abgelegt werden, wobei Transformationen, Constraint-Überwachung, Code-Generierung, Modellkombination, etc. für die nötige Unterstützung beim Entwickeln dienen. Die Technologiebasis lieferte die IKV++ Technologies AG mit einem eigens entwickelten Metatool namens "medini meta modeller", welches auf offenen Standards wie MOF, UML, OCL, XMI, CORBA, JET, etc. basiert und für diese Anwendungsfälle entwickelt wurde. Als eine entscheidende Lücke in der Werkzeugkette sind bei einer Bewertung sowohl Requirement-Engineering-Werkzeuge als auch Simulatoren oder Tester identifi-



ziert worden, mit denen ausführbare Modelle zu entwickelnder Software-Bausteine im Zusammenwirken mit bereits in Binärcode vorliegenden Bausteinen simuliert und getestet werden können.

Als eine große Herausforderung wurde die die Präzisierung der dynamischen Semantik von Modellelementen der Meta-Meta-Ebene mit einer entsprechenden Implementierung angesehen, die die Bereitstellung von Simulatoren/Tester für metamodellbasierte, ausführbare Modellierungssprachen ermöglichen.

Schwerpunkt: Modellierung von eingebetteten und Echtzeit-Systemen

Ansprechpartner: PROF. DR. JOACHIM FISCHER

Beteiligte Mitarbeiter: ANDREAS KUNERT, TOBY NEUMANN

Zusammenarbeit: Cinderella ApS (Kopenhagen), Schmiedewerke Gröditz GmbH

In diesem Forschungsbereich werden Arbeiten an der LFE zusammengefasst, die sich mit konkreten Modellierungsaufgaben aus unterschiedlichen Anwendungsbereichen beschäftigen. Ihre Gemeinsamkeit besteht darin, dass es sich dabei um eingebettete Systeme mit spezifischen Echtzeitanforderungen handelt. Die Forschungsergebnisse im Berichtszeitraum wurden in zwei verschiedenen Projekten erarbeitet:

- Die am Lehrstuhl entwickelte SDL-Entwicklungsumgebung SITE wurde um Ziellaufzeitsysteme erweitert, die eine Ausführung von verteilten, aus SDL heraus generierten Komponenten bei Einbettung in Standard-Middleware-Architekturen ermöglicht, wozu CORBA- und Web-Service-Architekturen zählen.
- Die Tragfähigkeit der als ODEM-Bibliotheksfamilie entwickelten Konzepte zur Modellierung und experimentellen Simulationen paralleler, zeitdiskreter und zeitkontinuierlicher Prozesse konnten im Rahmen eines komplexen Praxisprojektes zur Modellierung und Simulation von Vergütungsvorgängen in einem automatisierten Schmiede- und Walzwerk sowohl von ODEM als auch ODEMX unter Beweis gestellt werden. Es wurde damit begonnen, die Leistungsfähigkeit der C++ -basierten Simulatoren mit der eines Java-basierten Simulators zu vergleichen. Aus dieser Stärken- Schwächen-Analyse sollen Reserven einer Effizienzverbesserung für Simulatoren abgeleitet werden.

Code-Generierung für offene Systeme

(Toby Neumann)

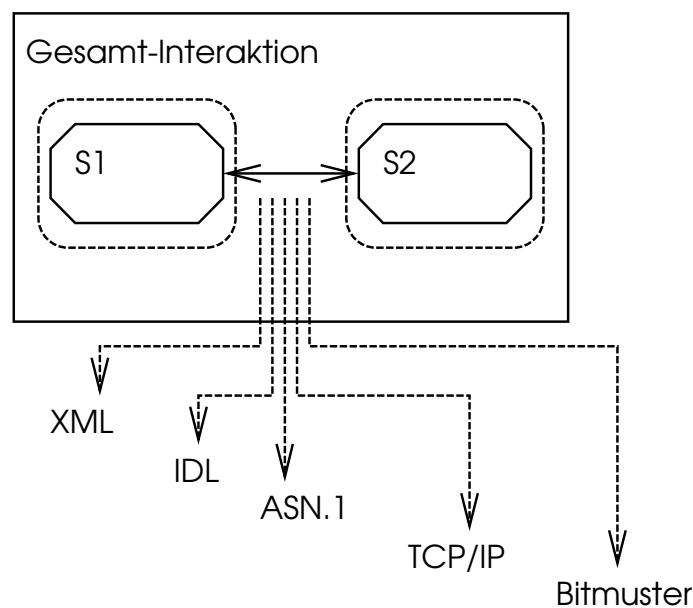
Es ist inzwischen gängige Praxis in vielen Anwendungsgebieten, beim Software-Entwurf aus der Beschreibung eines Systems automatisch eine Implementierung zu erzeugen. Dazu muss ein System in einer formalen Sprache beschrieben werden. Eine solche Beschreibungssprache, die vielfach in der Telekommunikation eingesetzt wird, ist SDL. An der LFE gibt es langjährige Erfahrungen, Code-Generatoren zu bauen, die SDL-Beschreibungen in Code einer Programmiersprache abbilden.

Die Klasse von Systemen, die mit ihrer Umgebung interagieren (offene Systeme), wirft im Vergleich zu autonomen Systemen ein zusätzliches Problem auf. Ein offenes System ist nur Teil eines größeren Systems, dessen andere Teile als gegeben angenommen werden müssen, weil sie in anderen Sprachen spezifiziert sind, von anderen Herstellern stammen oder bereits existieren. Das macht Schnittstellen zwischen diesen Teilen notwendig. Informationen, die zur Bereitstellung von Schnittstellen für unser offenes System und ihre Konfigurierung notwendig sind, fehlen üblicherweise in einer abstrakten Systemspezifikation.

Da sie aber zur automatischen Codegenerierung gebraucht werden, müssen diese Informationen ihren Weg auf andere Weise in den Ziel-Code finden. Herkömmliche Ansätze sind unflexibel, wenn sie sich nachträglicher Code-Manipulation bedienen, oder nicht portabel, wenn sie sich auf ein spezielles Werkzeug stützen.

Es wäre also wünschenswert, über eine standardisierte Methode zu verfügen, um Spezifikationen mit notwendigen Informationen anzureichern. Es gibt immer wieder Bemühungen, die Spezifikationsprachen um Aspekte des späteren praktischen Einsatzes zu erweitern (realtime, encoding, distribution). Allerdings sinkt damit der Grad der Abstraktion der Spezifikation und damit die Zahl der möglichen Umsetzungen.

Die an der LFE angestrebte Variante strebt an, Zusatzinformationen außerhalb der System-Beschreibung zu spezifizieren. Dafür könnten sich Ansätze eignen, die das Problem schon erkannt und berücksichtigt haben, wie die Sprache eODL oder das CORBA-Komponenten-Modell (CCM), die Möglichkeiten vorsehen, Zusatzinformationen zur Implementierung, Installation und Verteilung von Software-Teilen (deployment) zu notieren. Die Grafik zeigt (beispielhaft) zwei interagierende offene Systeme und einige Möglichkeiten zur Konfiguration der Kommunikation, die nicht in der Spezifikation enthalten ist und auch nicht sein sollte.



SimRing

(Ingmar Eveslage, Chris Helbing, Andreas Kunert)

Im Rahmen mehrerer Drittmittelprojekte ist das Simulationssystem SimRing für die Vergüterei der Schmiedewerke Gröditz GmbH (<http://www.stahl-groeditz.de>) entstanden. SimRing besteht aus einem mit Hilfe der Simulationsbibliothek ODEM implementierten Simulator, der die logistischen Abläufe in der Vergüterei analysiert, und einem Animator, der diese visuell aufbereitet. Dabei ist die zu bewältigende Aufgabenstellung seit dem ersten Projekt (Anfang 2001) kontinuierlich gewachsen. Ursprünglich wurde der Lehrstuhl mit der Simulation der logistischen Abläufe in der Vergüterei inkl. verschiedener Modifikationen beauftragt. Im Laufe der Zeit wurde die "echte" Vergüterei (u. a. als Folge der durchgeführten Simulationen) mehrfach weiterentwickelt, so dass zum einen die virtuelle

Vergüterei immer wieder an die reale angepasst werden musste und zum anderen neue Modifikationen möglich wurden.

Aktuell wird die Simulation im Rahmen zweier Studienarbeiten unter Verwendung der Simulationsbibliotheken ODEMX und DESMO-J neu implementiert. ODEMX entstand im Rahmen einer Diplomarbeit als Weiterentwicklung der am Lehrstuhl entstandenen Simulationsbibliothek ODEM und wird derzeit verstärkt in der Lehre angewendet. Mit der Portierung des SimRing-Projektes soll erstmals ein umfangreiches Projekt mit Hilfe von ODEMX implementiert werden. Bei DESMO-J handelt es sich um eine in Java implementierte Bibliothek zur Entwicklung ereignisorientierter Simulationen, die an der Universität Hamburg entwickelt wurde.

In einem abschließenden Vergleich soll ermittelt werden, inwieweit die verschiedenen Simulationsbibliotheken den Entwickler bei dem Entwurf und der Implementierung von Simulationen unterstützen und wie effizient sich letztere zur Laufzeit verhalten.

Schwerpunkt: Studentische Projekte

Ansprechpartner: PROF. DR. JOACHIM FISCHER

XML-DOM-Verarbeitung mit Kimwitu++

(Ralf Schuchardt)

Diese Studienarbeit beschäftigt sich mit der Frage, ob und wie mit Kimwitu++ XML verarbeitet werden könnte. XML (eXtensible Markup Language) ist ein Datenformat. Es definiert eine hierarchische Struktur, in der textuelle Daten und zugehörige Metainformationen abgelegt werden können. Das DOM (Document Object Model) ist eine programmiersprachenunabhängige Schnittstelle zur Verarbeitung von Daten, die im XML-Format vorliegen. Kimwitu++ ist ein Werkzeug, mit dem Operationen auf Baumstrukturen beschrieben und programmiert werden können. Es wird am Lehrstuhl hauptsächlich für das SITE-Projekt eingesetzt. Da sowohl DOM als auch Kimwitu++ auf Baumstrukturen operieren, wurde auch nach Möglichkeiten gesucht, beide Ansätze zusammenzubringen. Im Endergebnis wurden mehrere Möglichkeiten gefunden und teilweise implementiert. Dazu gehören eine Nachbildung der XML-Strukturen mit Kimwitu++ und ein zugehöriger Parser für XML, eine Untersuchung über Möglichkeiten weitergehende Metainformationen (DTDs, XML-Schema) über die Art und den Aufbau der in XML vorliegenden Daten zu nutzen, eine Zusammenführung von Kimwitu++- und DOM-Bäumen und ein XML-Speicherformat für Kimwitu++-Bäume.

Systementwicklung mit SDL-96

(Lars Münzberg)

Die im Rahmen dieser Diplomarbeit zu untersuchende Fragestellung ist, wie sich die Spezifikationstechnik SDL-96 anhand eines stufenweise weiterentwickelten Beispiels in ihren konzeptionellen Grundzügen umreißen lässt. Die formal definierte, werkzeuguunterstützte Sprache SDL-96 hat eine vollständige statische und dynamische Semantik und wird vorwiegend für die Spezifikation von Kommunikationssystemen verwendet. Das hier gewählte Beispiel eines Replikationssystems ist ein kommunizierendes System, welches Daten zum Beispiel für den Zweck der Sicherung in andere Datenspeicher spiegelt. Die Arbeit richtet sich vorwiegend an Leser mit geringen SDL-96 Kenntnissen, kann als komplementäres Begleitmaterial zur Vorlesung OMSI am Lehrstuhl Systemanalyse von Studenten

verwendet werden, aber auch als weiterführende beziehungsweise ergänzende Literatur zu den am Markt befindlichen Büchern herangezogen werden.

Veröffentlichungen

JOACHIM FISCHER, ECKHARDT HOLZ, ANDREAS PRINZ, MARKUS SCHEIDGEN: *Toolbased Language Development*. Computer Networks 49 (2005) 676 – 688, Elsevier B.V.

JOACHIM FISCHER, ANDREAS KUNERT, MICHAEL PIEFEL, MARKUS SCHEIDGEN: *ULF-Ware – An Open Framework for Integrated Tools for ITU-T Languages*. In *SDL 2005: Model Driven Systems Design: 12th International SDL Forum*, Grimstad, Norwegen, Juni 20-23, 2005. Proceedings, LNCS Vol. 3530 / 2005, Springer-Verlag GmbH.

JOACHIM FISCHER, TOBY NEUMANN, ANDERS OLSEN: *SDL Code Generation for Open Systems*, in *SDL 2005: Model Driven Systems Design: 12th International SDL Forum*, Grimstad, Norwegen, Juni 20-23, 2005. Proceedings, LNCS Vol. 3530 / 2005, Springer-Verlag GmbH.

HARALD BÖHME, GLENN SCHÜTZE, KONRAD VOIGT: *Component Development: MDA Based Transformation from eODL to CIDL*, in *SDL 2005: Model Driven Systems Design: 12th International SDL Forum*, Grimstad, Norwegen, Juni 20-23, 2005. Proceedings, LNCS Vol. 3530 / 2005, Springer-Verlag GmbH.

ANDREAS KUNERT: *Semi-automatic generation of metamodels and models from grammars and programs*, eingereicht beim 5th International Workshop on Graph Transformation and Visual Modelling Techniques, Wien, Österreich, April 2006.

MARKUS SCHEIDGEN: *CMOF-Model Semantics and Language Mapping for MOF 2.0 Implementation*, eingereicht beim Joint Meeting of the 4th Workshop on Model-Based Development of Computer Based Systems (MBD) und 3rd International Workshop on Model-based Methodologies for Pervasive and Embedded Software (MOMPES 2006).

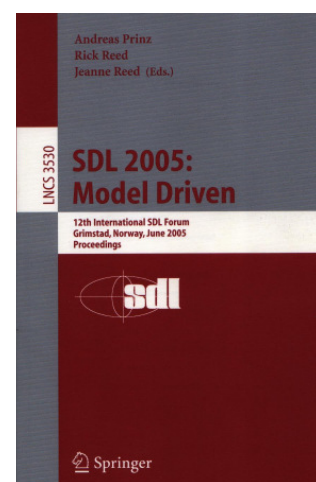
MARKUS SCHEIDGEN: *Model Patterns for Model Transformations in Model Driven Development*, Potsdam, eingereicht beim Joint Meeting of the 4th Workshop on Model-Based Development of Computer Based Systems (MBD) und 3rd International Workshop on Model-based Methodologies for Pervasive and Embedded Software (MOMPES 2006).

Vorträge

TOBY NEUMANN: *SDL Code Generation for Open Systems*, *SDL 2005: Model Driven Systems Design: 12th International SDL Forum*, Grimstad, Norwegen, Juni 2005.

MICHAEL PIEFEL: *ULF-Ware – An Open Framework for Integrated Tools for ITU-T Languages*, *SDL 2005: Model Driven Systems Design: 12th International SDL Forum*, Grimstad, Norwegen, Juni 2005.

KONRAD VOIGT: *Component Development: MDA Based Transformation from eODL to CIDL*, *SDL 2005: Model Driven Systems Design: 12th International SDL Forum*, Grimstad, Norwegen, Juni 2005.



Gastvorträge

MARIO KOLBERG (UNIVERSITY OF STIRLING, SCHOTTLAND): *Intelligent Devices*. Juli 2005.

ECKHARDT AUGENSTEIN (RWTH AACHEN): *Simulationssystem zur schnellen Bewertung von Energieversorgungssystemen*. Oktober 2005.

RICHARD SANDERS (SINTEF ICT, NORWEGEN): *Service Discovery and Component Reuse with Semantic Interfaces*. August 2005.

MICHAEL SODEN, J. HÖBLER, HAJO EICHLER (IKV⁺⁺ AG, BERLIN): *MDA in practice – Introducing Model Driven Development to an existing Software development process*. Juli 2005.

Graduierungsarbeiten

ECKHARDT HOLZ: *Kombination von Modellierungstechniken für den Softwareentwurf*, Habilitation, Februar 2005.

Diplomarbeiten

NICOLAS MICHAEL: *Vergleich von Enterprise JavaBeans mit dem @vantage-Komponentenmodell unter carrier-grade Aspekten*, Februar 2005.

FRANK KÜHNLENZ: *Vergleich des FUS-Managements in @vantage unter carrier-grade Aspekten – Proprietäres Konzept versus Relationales Datenbanksystem*, (eingereicht) Dezember 2005.

Sonstige Aktivitäten

12th SDL Forum

Vom 21. - 23. Juni fand in Grimstad, Norwegen, an der dortigen *Høgskolen i Agder* das zwölfte SDL Forum (<http://ikt.hia.no/sdl05/>) statt, organisiert von der SDL Forum Society (<http://www.sdl-forum.org>). Der Lehrstuhl war stark vertreten mit drei Konferenzbeiträgen und fünf Teilnehmern (sowie weiteren ehemaligen Mitarbeitern des Lehrstuhls).

ITU-T Study Group Meeting

Am Rande der SDL-Forum-Konferenz fand ein Treffen der SG17 der ITU-T statt, an dem die Vertreter der LFE teilnahmen. Inhalt war der weitere Standardisierungsprozess von SDL und das UML-Profil von SDL.

ITU-T Meeting der SG17 in Genf

Joachim Fischer und Harald Böhme nahmen im Oktober 2005 am Meeting der Study Group in Genf teil.

Graduiertenkolleg

„Modellbasierte Entwicklung von Technologien für selbstorganisierende dezentrale Informationssysteme – zur Anwendung im Katastrophenmanagement“

Unter Leitung von Joachim Fischer wurde ein Antrag für ein interdisziplinäres Graduiertenkolleg (<http://casablanca.informatik.hu-berlin.de/grk-wiki/>) bei der DFG erarbeitet und eingereicht. Im Dezember 2005 wurde der Antrag von der DFG befürwortet. Das Kolleg

wird zum 1. April 2006 für 4,5 Jahre eingerichtet und verfügt über insgesamt 15 Promotionsstellen.

Joachim Fischer

- Programmkomitee *SDL-Forum Konferenz, Grimstad 2005*
- diverse Gutachtertätigkeiten
- Instituts- und Fakultätsratsmitglied
- Vorsitzender der Strukturkommission des Instituts
- Vorsitzender der Haushaltskommission

Michael Piefel

- Institutsratsmitglied
- Mitglied der Strukturkommission des Instituts

Marita Albrecht

- Beauftragte des Instituts für Promotions- und Habilitationsarbeiten

Manfred Hagen

- Beauftragter des Instituts für die Publikation der Informatik-Berichte
- Mitglied der Raumkommission des Instituts

Lehr- und Forschungseinheit

Systemarchitektur (SAR)

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

Leiter

PROF. DR. JENS-PETER REDLICH
E-Mail: jpr@informatik.hu-berlin.de

Sekretariat

MARITA ALBRECHT
Tel.: (030) 2093 3400
Fax: (030) 2093 3112
E-Mail: albrecht@informatik.hu-berlin.de

Wissenschaftliche Mitarbeiter

DR. WOLF MÜLLER
DIPL.-INF. ANATOLIJ ZUBOW
DIPL.-INF. MATTHIAS KURTH (AB 05/05)

Gastwissenschaftler

DR. GANG LI (VR. CHINA, AVH STIPENDIAT)

Technischer Mitarbeiter

DIPL.-ING. WOLFGANG GANDRE

Rechnerbetriebsgruppe

DR. JAN-PETER BELL
DIPL.-ING. (FH) GABRIELE BÄRWOLFF
FRANK WOZOBULE

Tutoren

MATTHIAS JESCHKE
JENS MÜLLER (AB 10/05)
MARTIN STIGGE (AB 10/05)
HENRYK PLÖTZ (AB 11/05)

Der Lehrstuhl Systemarchitektur befasst sich im weitesten Sinne mit Betriebssystemen, Middleware-Plattformen und drahtlosen Kommunikationsnetzen, sowie mit Sicherheitsfragen in diesen Bereichen. Bei den Betriebssystemen interessieren besonders kleine (embedded) und besonders große (Grid) Systeme. Bei der drahtlosen Kommunikation liegt der Schwerpunkt auf selbst-organisierenden IP-basierten Ad-Hoc Maschen-Netzwerken die im unlizensierten Frequenzspektrum arbeiten, wie dem ISM Band bei 2.4 GHz (z.B. IEEE 802.11 – auch bekannt als WiFi oder WLAN). Beim Schwerpunkt Sicherheit liegt der Fokus auf dem Management von Informationen zur Reputation anonymer Nutzer – zur Unterstützung von Social Network Services (SNS), z.B. bei der Zuweisung von begrenzt ver-

fügbaren Ressourcen an Nutzer oder Prozeßabläufe, ohne die Identität der Teilnehmer offenlegen zu müssen.

Im Rahmen dieses breiten Interessengebietes werden am Lehrstuhl nacheinander vertiefende Projekte durchgeführt. Derzeitige Schwerpunkte sind das Berlin Roof Net – ein sich selbst-organisierendes drahtloses Community-Netzwerk, sowie dessen Anwendungsmöglichkeiten im Gebiet der Geo-Wissenschaften, z.B. als flexible, robuste und preiswerte Kommunikationsinfrastruktur für Monitoringsysteme und darauf aufbauende Frühwarnsysteme.

Lehre

Veranstaltungen im Grundstudium

- Unix Crash Kurs (Dr. Bell, WiSe 2004/05)
- Übungen zur VL Praktischen Informatik I (Dipl.-Inf. Zubow, WiSe 2004/05)
- Übungen und Praktika zur VL Praktischen Informatik I (Dipl.-Inf. Zubow, Dipl.-Inf. Kurth, WiSe 2005/06)
- Unix Crash Kurs (Dr. Bell, WiSe 2005/06)

Kernveranstaltungen (Halbkurse)

- Halbkurs „Unix Werkzeuge“ (Dr. Bell, WiSe 2004/05)
- Halbkurs „Unix Systemarchitektur“ (Prof. Redlich, WiSe 2004/05)
- Halbkurs „Prinzipien moderner Betriebssysteme“ (Prof. Redlich, SoSe 2005)
- Halbkurs „Mobilkommunikation“ (Prof. Redlich, SoSe 2005)
- Halbkurs „Security Engineering“ (Dr. Müller, SoSe 2005)
- Halbkurs „Systemadministration in Unix“ (Dr. Bell, SoSe 2005)
- Halbkurs „Middleware Plattformen“ (Prof. Redlich, WiSe 2005/06)
- Sezialvorlesung „Sichere Systemadministration“ (Dr. Bell, WiSe 2005/06)

Seminare

- „Enterprise Java Beans“ (Dipl.-Inf. Zubow, WiSe 2004/05)
- „Security Engineering“ (Dr. Müller, WiSe 2004/05)
- „Selbst-organisierende Middleware Plattformen“ (Prof. Redlich, WiSe 2004/05)
- „Ad-Hoc Netzwerke“ (Prof. Redlich, WiSe 2004/05)
- „Projektseminar Berlin Roof Net“ (Prof. Redlich, SoSe 2005)
- „Perr-to-Peer Netzwerke“ (Prof. Redlich, WiSe 2005/06)
- „Sicherheit in drahtlosen Netzen“ (Dr. Müller, WiSe 2005/06)
- „Projektseminar Berlin Roof Net“ (Prof. Redlich, WiSe 2005/06)

Workshops

- 5. Workshop „IT-Sicherheit“, September 2005.
- Kolloquium „Berlin Roof Net“, November 2005, (ausgerichtet von NEC Europe, Heidelberg).

Innovation in der Lehre

Ziel ist die Erhöhung der Qualität von Seminaren. Die (mit viel Fleiß erstellten) Ausarbeitungen der Studenten sollen permanent aufbewahrt und nachfolgenden Durchgängen als „Ausgangspunkt auf höherem Niveau“ bereitgestellt werden. Zu diesem Zweck wurde ein

Wiki bereitgestellt (<http://sarwiki.informatik.hu-berlin.de>), in das die Vortragenden ihre Ausarbeitungen in aufgearbeiteter Form eintragen müssen. Da das Wiki für alle Seminarteilnehmer offen ist, können diese ihre ergänzende oder abweichende Sichtweise sowie zusätzliches Wissen beitragen. In mehreren Fällen sind auf diese Weise bereits Themendarstellungen entstanden, die von externen Personen regelmäßig gelesen und referenziert werden (erkennbar über den von Google angebotenen Dienst „Analytics“).

Ein weiteres Anliegen ist, den Studenten zu helfen, ihre Vortragstechnik zu verbessern. Zu diesem Zweck wird, mit Zustimmung der Studenten, der Vortrag mit einer Videokamera aufgezeichnet. Am Ende jedes Vortrags erfolgt eine ausführliche Auswertung (Inhalt und Form) durch alle Seminarteilnehmer. Zusätzlich erhalten die Vortragenden die Videoaufzeichnung als DVD, um selbst einen Eindruck von Ihrer Leistung zu erhalten.

Forschung

Projekt: Berlin Roof Net

Ansprechpartner: Prof. Jens-Peter Redlich, Dipl.-Inf. Anatolij Zubow

Zusammenarbeit: Berliner Freifunk Projekt / OLSR-Experiment: Cornelius Keller; MIT, Computer Science and Artificial Intelligence Laboratory: John Bicket, Sanjit Biswas; Cornell University, School of Electrical Engineering: Prof. Zygmunt Haas

Durch das MIT Roof Net Projekt (<http://www.pdos.lcs.mit.edu/roofnet/>) wurde demonstriert, dass es prinzipiell möglich ist, große Teile einer Stadt wie Boston über drahtlose 802.11-basierte Netzwerke mit Internet-Zugang zu versorgen. Das Berlin Roof Net Projekt versucht herauszufinden, ob ähnliches in Berlin erreicht werden kann und ob diese Systeme vollständig selbst-organisierend gestaltet werden können. Es gibt weltweit nur sehr wenige reale Systeme dieser Art, an denen neue Protokolle durch Experimente evaluiert werden können.

Das Bostoner Netzwerk besteht aus ca. 40 Knoten, die auf Häuserdächern montiert werden (daher der Name Roof Net; Roof = Dach). Roof-Net Knoten in Radioreichweite entdecken sich automatisch und bilden ein Maschen-Netzwerk (Mesh Network), welches mit Hilfe geeigneter Routing-Protokolle in der Lage ist, IP-Pakete über mehrere Zwischenschritte zwischen beliebigen Knoten zu übertragen.

Die effiziente Umsetzung dieses viel versprechenden Konzepts stellt sich jedoch als schwierig heraus. So können traditionelle Routing-Protokolle nicht dem Umstand Rechnung tragen, dass die Verbindungsstrecken in drahtlosen Ad-Hoc Netzwerken eine variable Qualität haben, die ständig und schnell zwischen ‚sehr gut‘ und ‚sehr schlecht‘ schwanken kann. Auch beeinflussen sich die Verbindungsstrecken (Links) untereinander. Die bisher verwendeten Abstraktionen für Netzwerke sind in diesem Umfeld also nicht mehr sinnvoll. Es müssen neue Routing-Protokolle entwickelt werden; bzw. es muss sogar hinterfragt werden, ob nicht neue Forwarding-Mechanismen entwickelt werden müssen, die den inhärenten Eigenschaften dieser Netzwerkkategorie besser entsprechen.

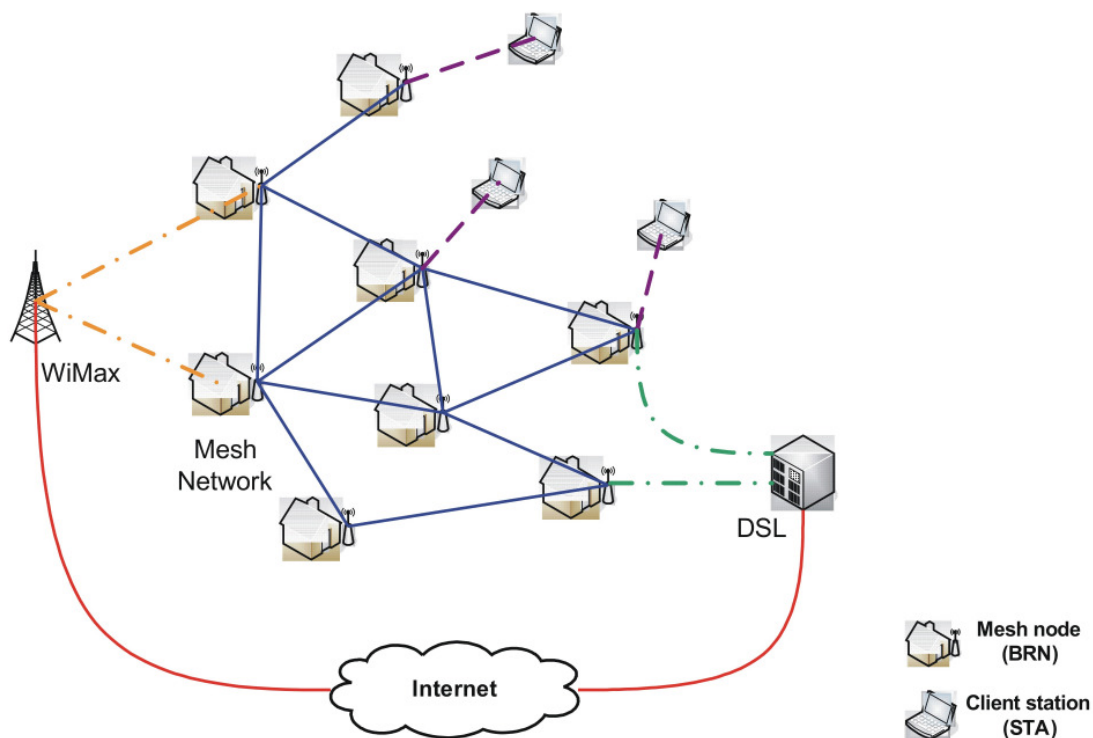


Abbildung 1: Die BRN-Knoten (mesh nodes) stellen die Infrastruktur des BRN-Netzes dar. End-Nutzer (client stations) können sich an einem BRN-Knoten in gewohnter Weise (wie bei einem Access Point) als 802.11 Klienten anmelden.

Technologischer Steckbrief für das Berlin Roof Net:

- **Ad-hoc**
 - Das Netz entwickelt sich spontan, d.h. in nicht detailliert vorausgeplanter Weise. Es benötigt keine zentrale Administrationsautorität, wie etwa einen Netzbetreiber. Die Netztopologie kann sich ungeplant, durch plötzliche und unkoordinierte Entscheidungen einzelner Knoten ändern.
- **Drahtlos (wireless)**
 - Die Geräte kommunizieren drahtlos miteinander. Es wird kostengünstige Standardhardware verwendet (z.B. IEEE 802.11g WLAN), die im lizenzfreien Radiospektrum arbeitet, z.B. 2.4 GHz ISM-Band.
- **Multi-hop**
 - Die Funkreichweite eines einzelnen Gerätes ist zu gering um weit entfernte Geräte direkt zu erreichen. Eine Nachricht wird in mehreren Schritten über mehrere Geräte hinweg übertragen. Bei der Weiterleitung können Datenpakete unterschiedlich priorisiert werden.
- **Maschen-Netz (mesh network)**
 - Zwischen Daten-Quelle und der verarbeitenden Station (Daten-Senke) existieren im Netz meist mehrere Pfade. Diese müssen so genutzt werden, dass die gegenseitige Behinderung gleichzeitig übertragener Datenströme möglichst minimal ist, aber auch so, dass der Ausfall einzelner Pfade automatisch (selbst-organisierend) durch alternative Pfade kompensiert wird.

- **Dezentral (de-centralized)**
 - Dienste werden durch die Geräte gemeinsam erbracht. Jedes Einzelgerät trägt einen kleinen Anteil dazu bei, ist aber individuell entbehrlich – d.h. es kann jederzeit und unangekündigt abgeschaltet werden, ohne dass dadurch das Gesamtsystem in seiner Leistungsfähigkeit beeinträchtigt wird (wenn zu viele Geräte gleichzeitig ausfallen, kann es jedoch zu Informationsverlust kommen).
- **Selbst-organisierend (self-organizing)**
 - Das Netz kann sich nicht auf das Vorhandensein menschlicher Intelligenz, z. B. in Form eines Netzbetreibers/Operators, verlassen, der einen ordnungsgemäßen Aufbau und Betrieb überwacht. Das Netz passt sich selbstständig an neue Gegebenheiten optimal an.
- **Nicht hochmobil (not highly mobile)**
 - Hier ist die Begrenzung der Dynamik bei der Änderung der Struktur (Topologie) des Netzes gemeint. Netzknoten dürfen sich langsam bewegen; sie dürfen auch mit „mäßiger Häufigkeit“ dem System entnommen oder hinzugefügt werden. Das System wird sich automatisch und schnell an die Änderungen adaptieren. Diese Fähigkeit erreicht ihre Grenzen, z.B. wenn sich die Netzknoten in schnell fahrenden Autos befinden oder in großen Zahlen und ständig, zum Beispiel im Sekundentakt, dem System entnommen bzw. hinzugefügt werden.

Teilprojekt: Indoor-Testbed für selbst-organisierende drahtlose Maschennetzwerke

Ansprechpartner: Dipl.-Inf. Anatolij Zubow

Mitarbeiter: Dipl.-Inf. Mathias Kurth, Mathias Jeschke, Jens Müller, Robert Sombrutzki

Zusammenarbeit: Lehrstühle und Verwaltung der Institute für Informatik und Mathematik

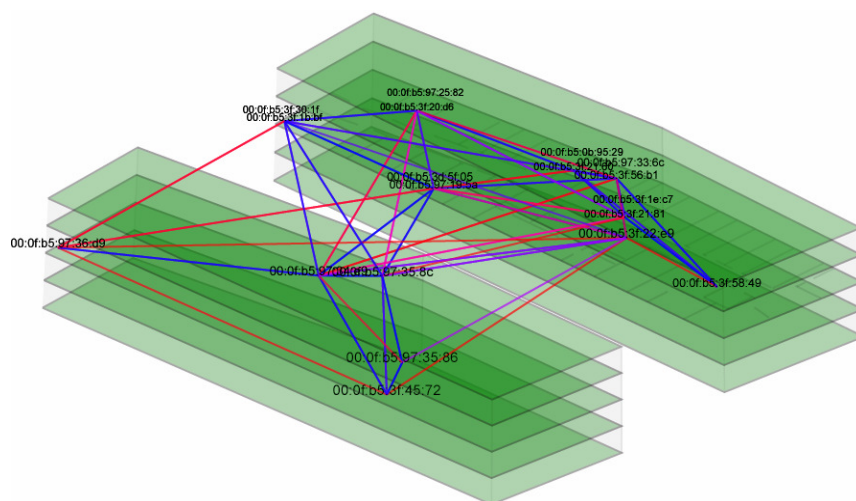


Abbildung 2: Online Statusbericht für das Indoor-Testbed in den Häusern 3 und 4 des Instituts. Die Grafik zeigt die installierten RoofNet-Knoten sowie die Qualität der zwischen ihnen aufgebauten Links. Die Grafik wird ständig aktualisiert und kann (über ein Browser-Plugin) über die URL <http://sarwiki.informatik.hu-berlin.de/BerlinRoofNet> abgerufen werden.

Teilprojekt: Outdoor-Testbed für selbst-organisierende drahtlose Maschennetzwerke

Ansprechpartner: Cornelius Keller

Zusammenarbeit: Freifunk.net, OLSR-Experiment Projekte in Berlin

Es gibt mehrere Initiativen, die den Aufbau eines drahtlosen Community-Netzwerkes in Berlin zum Ziel haben. Ursprüngliche Motivation war die Versorgung von Stadtgebieten mit Internet, die durch die öffentlichen Anbieter (Telekom, ...) nicht bedient werden können. Diese Aktivitäten, die bereits beachtliche Ausmaße angenommen haben, bieten ein ideales Umfeld um praktische Erfahrungen beim Aufbau und Betreiben von 802.11 Maschennetzwerken in realen (innerstädtischen) Umgebungen zu sammeln. Im Jahr 2006 soll das Indoor-Testbed am Institut, welches jetzt bereits mehrere Häuser umspannt, auf weitere Einrichtungen in Adlershof ausgedehnt werden, so dass dem Lehrstuhl (und seinen Partnern) damit ein eigenes Outdoor-Testbed zur Verfügung stehen wird.



Abbildung 3: Installieren eines Netz-Knotens auf dem Berliner „Haus des Lehrers“

Quelle:

http://www.olsrexperiment.de/index.php?option=com_zoom&Itemid=55&page=view&catid=1&PageNo=1&key=5

Teilprojekt: Opportunistisches Mehrkanal Routing in 802.11-basierten Maschennetzen

Ansprechpartner: Dipl.-Inf. Anatolij Zubow

Mitarbeiter: Dipl.-Inf. Mathias Kurth, Jens Müller

Wireless Multi-Hop Mesh-Netzwerke spielen eine immer bedeutendere Rolle, z.B. in der Vernetzung von urbanen Gebieten. Eines ihrer größten Probleme ist jedoch, dass sie nur unzureichend mit der Anzahl der Knoten und Nutzer skalieren. Der wohl wichtigste Grund hierfür ist in der Funktionsweise eines Multi-Hop Netzwerkes zu finden: Ein Knoten ist

nicht nur für die Übertragung seiner eigenen Daten verantwortlich, sondern auch für das Weiterleiten von Paketen anderer Knoten. Nicht unbedeutender ist die Tatsache, dass sich alle Knoten in einem drahtlosen Netz das gleiche Medium für die Datenübertragung teilen müssen.

Einen aussichtsreichen Ansatz zur Verbesserung des Durchsatzes in einem drahtlosen Multi-Hop Netzwerk bietet das ExOR-Protokoll. Es versucht die Besonderheiten des Mediums explizit zu nutzen, anstatt diese zu ignorieren. Betrachtet man drahtlose Netzwerke, so fallen einem eine Reihe von Besonderheiten gegenüber ordinären drahtgebundenen Netzen auf. Zum einen besteht zwischen jedem Knotenpaar im Netzwerk eine Verbindung, wenn auch bisweilen mit sehr hoher Fehlerrate. Des Weiteren erfolgen alle Paketübertragungen mit Hilfe von Broadcasts, wodurch die Möglichkeit besteht, dass auch entfernt liegende Knoten ein ausgesendetes Paket mit einer gewissen Wahrscheinlichkeit erhalten. Letztere Eigenschaft nutzt das ExOR-Protokoll aus, weswegen es auch die Grundlage für unser Protokoll bildet.

Der IEEE 802.11-Standard, der heutzutage eine weite Verbreitung findet, ermöglicht die Verwendung von mehreren, sich nicht überlappenden Kanälen. Werden mehrere Kanäle verwendet, so können auch mehrere Übertragungen innerhalb einer Region (collision domain) gleichzeitig und ohne Interferenz stattfinden, was sich in einer Erhöhung des Netzwerkdurchsatzes positiv niederschlägt. In der Literatur existieren viele Routing-Algorithmen, welche mehrere Kanäle verwenden. Diese sind jedoch nur bedingt einsetzbar, da sie Knoten mit mehr als einem Transceiver voraussetzen. Die meisten heutzutage erhältlichen 802.11-Geräte besitzen jedoch lediglich einen Transceiver. Dies führt dazu, dass ein Knoten nicht gleichzeitig senden und empfangen kann. Möglich wird die Verwendung von mehreren Kanälen bei Vorhandensein eines Transceiver durch den schnellen Wechsel eines Kanals (80 μ s). Knoten, welche auf unterschiedlichen Kanälen betrieben werden, können nicht miteinander kommunizieren und sich damit auch nicht gegenseitig durch Interferenz beeinflussen.

In diesem Teilprojekt entwickeln wir MCEXOR - eine Synthese aus Extremely Opportunistic Routing (ExOR) und Verfahren zur Verwendung mehrerer Übertragungskanäle in multi-hop wireless Netzwerken. In den besagten Netzen bewirkt ExOR eine Reduzierung der Anzahl von Übertragungen, während die Verwendung mehrerer Übertragungskanäle zu einer Reduzierung von Interferenz führt. Im Gegensatz zu anderen Ansätzen benötigt MCEXOR nur einen Transceiver je Gerät. Es werden Algorithmen zur Kanalwahl und zur Ermittlung eines aussichtsreichen Candidate Set entwickelt und evaluiert.

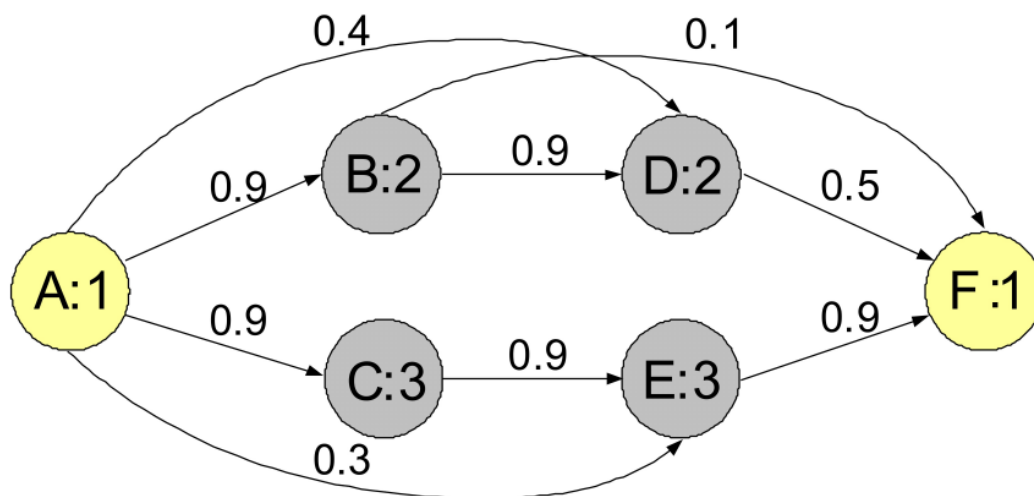


Abbildung 4: Maschennetzwerk mit verschiedenen Link-Qualitäten (ausgedrückt als Wahrscheinlichkeit dass eine Übertragung erfolgreich sein wird) und verschiedenen Übertragungsfrequenzen. Das Label eines Knotens gibt seine ID und den verwendeten Kanal an. Beispielsweise bedeutet (A:1), dass der Knoten A den Kanal 1 verwendet.

Teilprojekt: Network Coding in drahtlosen Maschennetzen

Ansprechpartner: Dipl.-Inf. Anatolij Zubow

Mitarbeiter: Stephan Allner

Der in einem drahtlosen Maschennetz erreichbare Datendurchsatz wird durch eine Reihe von Faktoren negativ beeinflusst. Herkömmliche Routingprotokolle stoßen hierbei recht schnell an ihre Grenzen. Einen interessanten Ansatz zur Erhöhung des Datendurchsatzes in einem drahtlosen Maschennetz bildet das Network Coding. Nach diesem Ansatz werden Pakete nicht lediglich als unveränderliche Einheiten betrachtet, sondern werden geeignet miteinander kombiniert um Übertragungen einzusparen. Im Falle von drahtlosen Netzen macht man sich zudem die Broadcast-Eigenschaft des Mediums zu Nutze. (siehe *Abbildung 5*)

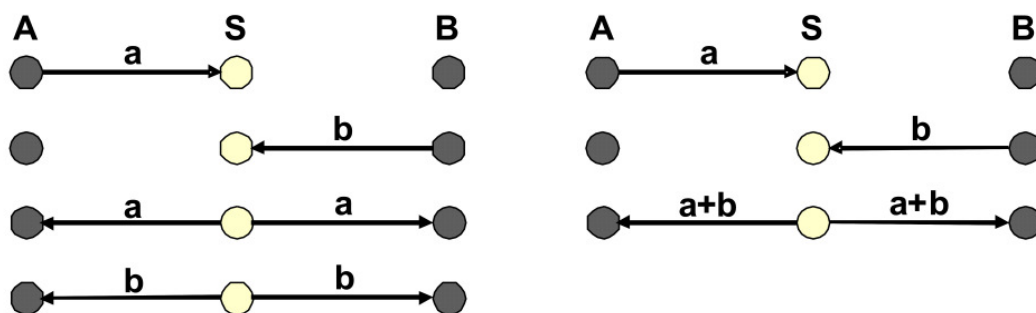


Abbildung 5: Knoten A und B möchten jeweils ein Paket austauschen. Hierzu sind sie auf die Hilfe des Knotens S angewiesen. Nach dem herkömmlichen Routing werden 4 Übertragungen benötigt (links). Bei Einsatz von Network Coding kann man sich die Broadcast-Eigenschaft des drahtlosen Mediums zu Nutze machen. Hierbei werden die Pakete (a und b) kombiniert und mit einer einzigen Datenübertragung übertragen. Es werden daher nur 3 Übertragungen benötigt (rechts).

Teilprojekt: Softwareentwicklung für die Netgear wgt634u Plattform

Ansprechpartner: Dipl.-Inf. Anatolij Zubow

Mitarbeiter: Dipl.-Inf. Mathias Kurth, Mathias Jeschke, Jens Müller

Zusammenarbeit: MIT, Computer Science and AI Laboratory: John Bicket, Sanjit Biswas

Da wir an einer weiten Verbreitung unserer Technologien interessiert sind, werden wir unsere Software für eine realistische Plattform anbieten, die für die teilnehmenden Studenten finanziell erschwinglich ist, stabil läuft und deren Radio-Eigenschaften durch die zuständigen Stellen als unbedenklich zertifiziert wurden.

Die Wahl fiel auf den Wireless LAN Router der Firma Netgear (wgt634u), der am Massenmarkt preiswert erhältlich ist, und für den uns die internen Spezifikationen weitgehend bekannt sind.

In einem studentischen Projekt wird eine Software-Cross-Entwicklungsumgebung erstellt, die es gestattet, auf einem traditionellen Intel-basierten Arbeitsplatz (mit Linux), Software für die Netgear-Plattform zu entwickeln, einschließlich eines Linux-basierten Betriebssystems, sowie einer Vielzahl zusätzlicher Programme.

Die Studenten können auf diese Weise ihr ingenieurtechnisches Können auf dem Gebiet der Softwareentwicklung für ‚Embedded Systems‘ vertiefen und praktisch erproben.



Abbildung 6: Der Netgear Wireless LAN Router bildet die Hardware-Plattform für das Berlin Roof Net Projekt. (Quelle: <http://www.netgear.com>)

Teilprojekt: Software-Verteilungs-Plattform für das Berlin Roof Net

Ansprechpartner: Bernhard Wiedemann

Drahtlose, 802.11-basierte multi-hop Netzwerke sind eine neue Forschungsrichtung, die in schneller Abfolge neue Innovationen hervorbringt. Diese machen oft den Einsatz neuer Protokolle oder neuer Systemstrukturen wünschenswert. Da es bei den betrachteten selbstorganisierenden Netzwerken jedoch keine zentrale Autorität, wie z.B. einen Operator, gibt, ist die Migration auf neue Protokolle bei einer existierenden Installation nicht trivial. Niemand hat physische oder sonstige administrative Kontrolle über die Knoten, die sich in den Wohnräumen der am Projekt teilnehmenden Studenten befinden. Wird der Austausch der Software nicht geschickt vorbereitet und koordiniert, dann kann das Netzwerk zerfallen und sich in eine nutzlose Ansammlung inkompatibler Knoten verwandeln. Eine Reparatur des Schadens über das Netzwerk ist dann natürlich nicht mehr möglich.

Es wurde daher eine Software-Verteilungs-Plattform entwickelt, speziell für die Netgear wgt634u Hardware, die die Verteilung neuer Softwareversionen unterstützt, sowie das Umschalten der Knoten auf die neue Software koordiniert. Insbesondere wird sichergestellt, dass alle Knoten gemeinsam (das ist wichtig!) zu einer früheren Softwareversion zurückkehren, sollte sich die neue Software als nicht arbeitsfähig erweisen.

In der Literatur findet man bereits Ansätze für das entfernte Aktualisieren von Software beim Endkunden. In unserem Fall entsteht jedoch eine neue Qualität, da die Software selbst die Grundlage für das Netzwerk bildet über welches die Software-Updates verteilt werden. Außerdem gehen wir davon aus, dass es in einem selbst-organisierenden System keine zentralen Instanzen, wie z.B. FTP-Server gibt. Die Knoten können daher die Software nicht von einem zentralen Server herunterladen. Die Software muss sich stattdessen verteilt in der Gemeinschaft der Knoten befinden, von denen jeder einzelne jederzeit und ohne Vorankündigung ausfallen oder vom Netzwerk separiert werden kann.

Projekt: Anwendung von Technologien des Berlin Roof Nets für Frühwarnsysteme

Ansprechpartner: Prof. Redlich, Dipl.-Inf. Mathias Kurth

Mitarbeiter: Dr. Gang Li

Zusammenarbeit: GFZ Potsdam, Institut für Geographie

Vorbemerkung: Dieses Projekt zielt auf eine langfristige Zusammenarbeit mit dem GFZ Potsdam, dem Institut für Geographie, sowie anderen Geo-Wissenschaftlern ab. In diesem Kontext wurde bei der DFG ein Graduiertenkolleg „Modellbasierte Entwicklung von Technologien für selbstorganisierende Informationssysteme - zur Anwendung im Katastrophenmanagement“ beantragt, das inzwischen bewilligt wurde.

Bei bestimmten Naturkatastrophen, wie zum Beispiel Erdbeben, vergehen zwischen den ersten Anzeichen und dem Eintreten des Schadens nur wenige Sekunden. Hier kommt es auf eine schnelle Informationserfassung und -weiterleitung an, z.B. von den Sensoren zu einem Entscheidungssystem zum Abschalten von Elektrizität/Wasser/ Gas, um Sekundärschäden zu vermeiden. Integrierbarkeit der Informationsquellen (Sensoren) mit einer Vielzahl von (zum Teil schon vorhandenen) Informationsverarbeitungswerkzeugen ist ebenfalls wichtig. Als universelles Kommunikationsmedium haben sich hierfür IP-Netze bewährt. Öffentliche Anbieter (ISPs) erscheinen ungeeignet, da deren meist zentralistisch aufgebaute und verwaltete Netze anfällig für die bei Naturkatastrophen üblichen Schäden sind (zentrale Server können ausfallen, Backbone-Leitungen unterbrochen werden); außerdem ist ihr Verhalten bei ungewöhnlicher Lastsituation, provoziert durch ungewöhnliches Nutzerverhalten, schwer kontrollierbar. Die in monatlichen Raten anfallenden Kosten sind oft erheblich. Der Aufbau eines eigenen Netzes, parallel zu den öffentlichen Anbietern, ist ökonomisch nicht vertretbar und wahrscheinlich auch nicht nötig, wie erste Erfahrungen mit drahtlosen ad-hoc Community-Netzen nahelegen. Diese Netze werden durch Geräte von der Art eines WLAN Access-Punkts gebildet, welche ungeplant (spontan) installiert werden, sich nach ihrer Installation über drahtlose Kommunikationswege automatisch finden, miteinander verbinden und schließlich selbständig zu komplexen leistungsfähigen Netzstrukturen organisieren.

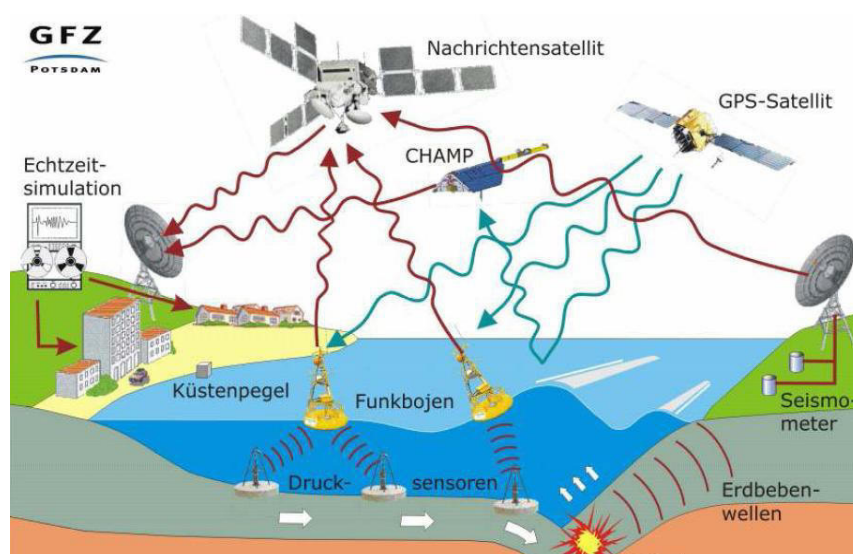


Abbildung 7: Vision eines integrierten Frühwarnsystems für Erdbeben und Tsunamis
Quelle: GFZ-Potsdam.

Ziel dieses Projekts ist die Bereitstellung eines Portfolios von Technologien für die Frühwarnung bei Naturkatastrophen. Das Portfolio soll in Berlin und Umgebung in realistischen Testbeds erprobt werden und es gestatten, einem künftigen Kunden in kurzer Zeit und basierend auf praktischen Erfahrungen eine hochwertige aber dennoch kostengünstige Lösung für seinen konkreten Anwendungsfall anbieten zu können (Bausteinsystem). In dieses Portfolio werden neuartige und angepasste Methoden und Technologien aus der Geographie und der Informatik eingehen. Der Beitrag der Informatik besteht in der Bereitstellung von Technologien zur Informationsübertragung und –Verarbeitung, die nur durch Einbeziehung neuester Erkenntnisse bzw. noch zu erforschender Erweiterungen in der Lage sind, die Anforderungen der Geographie mit der geforderten Zuverlässigkeit, Komplexität, Zeitverhalten und einem beschränkten finanziellen Budget zu erfüllen.

Projekt: Secure Personal Area Networks (SPAN)

Ansprechpartner: Prof. Dr. Jens-Peter Redlich, Dr. Wolf Müller

Mitarbeiter: Henryk Plötz, Martin Stigge

Zusammenarbeit: NEC Europe Ltd. Network Laboratories Heidelberg, NEC Japan

Forschungsförderung: NEC Europe

Die Zahl der Geräte im persönlichen Umfeld eines Menschen wächst heutzutage rasch an: PDA, Handy, Kameras, Speichergeräte, MP3-Player, Diese Geräte sind in der Lage, über drahtlose Kommunikationsschnittstellen spontan Ad-Hoc-Netze zu bilden und in Kombination neue Dienste zu erbringen, woraus ein ständig wachsender Konfigurations- und Administrationsaufwand resultiert. Das SPAN-Projekt zeigt Lösungen für dieses Problem auf. Wir schlagen vor, das Handy zum Auffinden und zur Konfiguration anderer Dienste und Geräte, sowie als zentralen Authentifikator gegenüber diesen zu verwenden. Wichtige Bausteine für *secure PANs* sind:

- **Identitätsmanagement:** Die beteiligten Parteien (Geräte/Dienste) müssen wissen, dass sie mit den richtigen Partnern kommunizieren. Dazu müssen gegebenenfalls Nachweise über Identitäten die hinter den Geräten stehen (Name) oder deren Eigenschaften („über 18 Jahre alt“) erbracht werden können. Oft wollen die Nutzer darüber hinaus nicht ihre komplette Identität preisgeben, sondern anonym bleiben.
- **Zugriffskontrollmanagement:** Nicht jedem mag es gestattet sein auf alle Geräte/Dienste zuzugreifen. Diese Rechte sollten für den Nutzer transparent verwaltet werden.
- **Reputationsmanagement:** Abhängig davon, wie der Nutzer sich in der Vergangenheit verhalten hat, kann sein Ansehen steigen oder sinken. Hat er sich unkooperativ oder destruktiv verhalten, so sollte sich das auf seine Reputation auswirken und der Zugriff auf Geräte und Dienste könnte beschränkt oder auch ganz verweigert werden.

Exemplarisch untersuchten wir exemplarisch folgendes Szenario: Unsere Anwendung besteht aus einem Datenprojektor, den eine Person verwenden will. Dies soll ohne vorherige Einstellungen oder Bekanntmachen der zwei Partner (Projektor, Benutzer) geschehen. Zugriff soll gewährt werden, wenn der Nutzer seine Identität gegenüber dem Zugriffskontrollmanager, der für den Projektor zuständig ist, zweifelsfrei nachgewiesen hat. Dazu wird ein Identitätsanbieter benutzt, dem der Zugriffskontrollmanager vertraut. Der Projektor alleine könnte diese Überprüfung nicht in dieser Art vollbringen, da er keine Verbindung

zum Core-Netzwerk besitzt. Mit Hilfe des Mobiltelefons gelingt jedoch der Nachweis von Zugriffsrechten des Benutzers in sicherer Weise. Weiterhin kann der Projektor nun automatisch an die Präferenzen des Benutzers angepasst werden (z.B. Sprache der Benutzeroberfläche) und es dient zur Steuerung des Projektors.

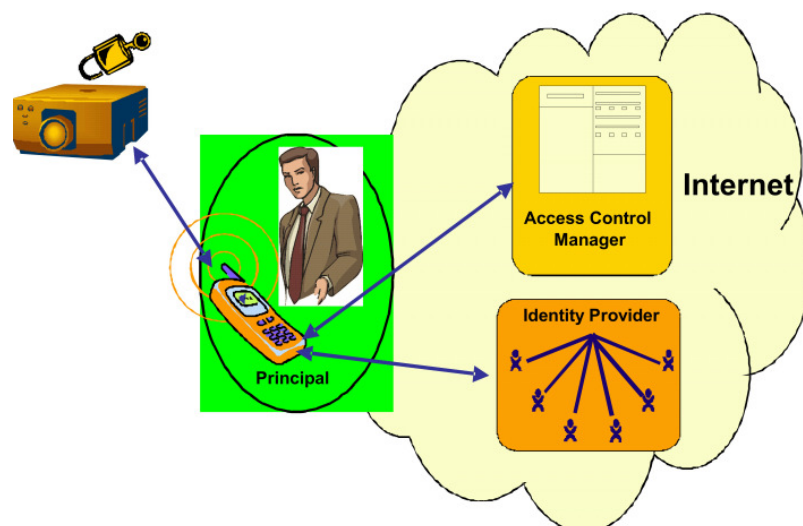


Abbildung 8: Zugangskontrolle für einen Datenprojektor mit Hilfe eines Mobiltelefons und zentraler Zugriffsrechtsverwaltung

Projekt: DV-Infrastruktur aus virtuellen Maschinen – für Lehre und Forschung

Ansprechpartner: Prof. Redlich, Dr. Bell, Herr Gandre, Herr Jeschke

Der Lehrstuhl benötigte eine leistungsfähige DV-Infrastruktur um seine Aufgaben in Forschung und Lehre ausführen zu können. Dazu gehören neben den üblichen Servern für die Web-Präsenz, FTP, Fileserver und Email auch ein Wiki sowie zahlreiche projektspezifische Server.

Es wurde eine Lösung gewählt, die alle Rechner des Lehrstuhls, mit Ausnahme des Firewalls und des Fileservers, als virtuelle Maschinen realisiert. Diese werden auf einer geringen Anzahl leistungsfähiger physikalischer Rechner betrieben, wobei die virtuellen Maschinen so konfiguriert wurden, dass sie je nach aktuellem Lastprofil dynamisch auf die physikalischen Rechner umverteilt werden können. Dieser Ansatz erspart die Beschaffung von Hardwareressourcen und ermöglicht es zudem, sehr schnell neue Maschinen für neue Projekte oder Experimente bereitzustellen. Nicht mehr benötigte Rechner können schnell stillgelegt werden, aber ebenso einfach zu einem späteren Zeitpunkt wieder reaktiviert werden. Dieser Ansatz hat sich auch bei der Durchführung von vorlesungsbegleitenden Praktika bewährt, wo den Studenten je nach Bedarf mehrere, speziell an ihre Bedürfnisse angepasste (virtuelle) Maschinen bereitgestellt werden konnten, auf denen sie Administrator-Rechte erlangen und Verwaltungsaufgaben eigenständig durchführen konnten.

Wir hoffen, dass die von uns realisierte DV-Infrastruktur als Vorbild für andere Lehrstühle dienen wird.

Rechnerbetriebsgruppe

Seit Bestehen des Instituts gehört die Betreuung des hausinternen Rechnernetzes zum Verantwortungsbereich des Lehrstuhls. Die Rechnerbetriebsgruppe sichert den laufenden Betrieb des Netzes, der zentralen Server sowie der Arbeitsplatzrechner in den Studenten-Pools. Erfahrungen des Rechnerbetriebs werden in Vorlesungen und Seminaren an die Studierenden herangetragen.

Promotionen, Diplomarbeiten, Studienarbeiten:

Der Lehrstuhl hat seine Arbeit im Sommer 2004 aufgenommen. Abschlüsse von Promotionsverfahren stehen zur Zeit noch nicht an. Eine Diplomarbeit wurde abgeschlossen: Bernhard Wiedemann: „Entwicklung einer Softwareverteilungsplattform für das Berlin Roof Net“. Mehrere Studienarbeiten wurden abgeschlossen und über die Web-Seite des Lehrstuhls publiziert.

Sonstige Aktivitäten:

Prof. Redlich:

Mitglied mehrerer Promotions- und Habilitationskommissionen
Tribunal-Mitglied an der UPC, Dept. Enginyeria Telemàtica, Barcelona
Programmkomitees: ICC 2005, ICC 2006, Mobicom 2005
Leiter der DV-Kommission des Instituts
Mehrere Arbeitstreffen mit NEC Europe, GFZ Potsdam.

- Miniworkshop mit einer Abordnung des Sino-German Joint Software Institutes der Chinesischen Akademie der Wissenschaften mit dem Ziel der Initiierung künftiger gemeinsamer Projekte. Dr. Gang Li, der sich derzeit als Gastwissenschaftler an unserem Lehrstuhl aufhält, wird von der Chinesischen Akademie der Wissenschaften und der Alexander-von-Humboldt-Stiftung finanziert
- Arbeitsbesuch einer Abordnung der Firma NEC Japan (mit Vertretern von NEC Europe) mit Miniworkshop im Juli 2005.
- Einladung zur Verleihung der Ehrendoktorwürde an Prof. Gerhard Krüger durch die Universität Rostock im April 2005.
- Teilnahme am „Treffpunkt Internet – Forum Digitale Integration“ des Bundesministerium für Wirtschaft und Technologie, November 2005.

Dr. Bell:

- Mitglied der Haushaltskommission
- DV-Verantwortlicher des Instituts

Wolf Müller:

- Mehrere Arbeitstreffen mit NEC Europe zur Einwerbung eines Drittmittelprojekts (siehe Projekt-Bericht „SPAN“)

Anatolij Zubow

- Kolloquium „Berlin Roof Net“, ausgerichtet von NEC Europe in Heidelberg, November 2005.
- Gastvortrag „Berlin Roof Net“ an der Universität Rostock im Oktober 2005.

Veröffentlichungen

Forschungsarbeiten werden im Wiki des Lehrstuhls, <http://sarwiki.informatik.hu-berlin.de>, publiziert und ständig aktualisiert.

Öffentliche Berichte (Reports) sind im Web unter <http://sar.informatik.hu-berlin.de/research/publications/index.htm> zu finden. Im Jahr 2005 wurden 8 Berichte, teils in Kooperation mit anderen Einrichtungen, veröffentlicht:

R. SPERLING: *Linux-Hardwaretreiber für die HHI CineCard-Familie*, SAR-PR-2005-01.

J.-P. REDLICH, W. MÜLLER: *State-of-the-Art in Self-Organizing Platforms and Corresponding Security Considerations*, SAR-PR-2005-02, NLE-PR-2005-59.

J.-P. REDLICH, A. ZUBOW, W. MÜLLER, M. JESCHKE, J. MÜLLER: *Hacking the Netgear wgt634u*, SAR-PR-2005-03.

T. DÄNICKE: *Sicherheit in selbstorganisierenden drahtlosen Netzen. Ein Überblick über typische Fragestellungen und Lösungsansätze*, SAR-PR-2005-04.

JENS-P. REDLICH, A. ZUBOW, J. MÜLLER: *Multi Channel Opportunistic Routing in Multi-Hop Wireless Networks using a Single Transceiver*, SAR-PR-2005-05.

J.-P. REDLICH, W. MÜLLER: *Access Control for off-line Beamer – An Example for Secure PAN and FMC*, SAR-PR-2005-06, NLE-PR-2005-81.

J.-P. REDLICH, B. WIEDEMANN: *Software Distribution Platform for Ad-Hoc Wireless Mesh Networks*, SAR-PR-2005-07.

J.-P. REDLICH, W. MÜLLER, H. PLÖTZ, M. STIGGE: *Access Control for off-line Beamer - Demo Description*, SAR-PR-2005-08, NLE-PR-2005-106.

Bewertung der Web-Präsens

Über einen externen Beobachter, den Google-Dienst „Analytics“, kann die Benutzung unserer Web-Seiten und insbesondere des Wiki objektiv beurteilt werden. An einem durchschnittlichen Arbeitstag werden ca. 1200 Seiten von ca. 450 verschiedenen Besuchern abgerufen (→ ca. 3 Seiten / Besucher / Tag). Bei den regulären Web-Seiten sind 25% der Besucher nicht aus Deutschland; beim Wiki sogar deutlich mehr als die Hälfte. Die meisten Besucher wurden als „wiederkehrend“ klassifiziert. Von unseren internationalen Lesern sind ca. 20% aus Europa, 50% aus den USA und 20% aus Asien. Besonders erfreulich ist, dass die große Mehrzahl der Leser unsere Seiten über die Suchmaschine von Google gefunden haben.

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

PROF. DR. BODO HOHBERG (BIS JUNI 05)

DR. AXEL MARTENS

PRIV.- DOZ. DR. KARSTEN SCHMIDT

DIPL.-INF. PETER MASSUTHE

DIPL.-INF. CHRISTIAN STAHL

DIPL.-INF. ANDREAS GLAUSCH (SEIT JULI 05)

DIPL.-INF. NIELS LOHMANN (SEIT OKTOBER 05)

Technikerin

DIPL.-CHEM. BIRGIT SCHIEFNER

Promotionsstipendiaten

BIXIA WU

BAVER ACU

JARUNGJIT PARNJAI

Tutorinnen und Tutoren

JAN BRETSCHNEIDER

DIRK FAHLAND (BIS SEPTEMBER 05)

CHRISTIAN GIERDS (SEIT OKTOBER 05)

ROBERT GÖTTSCHE (SEIT DEZEMBER 05)

ALEXANDRA JULIUS

NIELS LOHMANN (BIS SEPTEMBER 05)

DENNIS REINERT (SEIT OKTOBER 05)

Gast

DR. ADRIANNA ALEXANDER

In zunehmendem Umfang wird Software nicht mehr von Programmierern aus ihrem intuitiven Verständnis einer Aufgabenstellung heraus geschrieben, sondern aus Spezifikationen und Modellen generiert. Die Forschungsarbeiten des Lehrstuhls konzentrieren sich deshalb auf den modellbasierten Software- und Systementwurf, insbesondere auf Konzepte und Methoden zur Modellierung und Verifikation verteilter und reaktiver Systeme und verteilter Algorithmen für unterschiedliche Anwendungsbereiche. Abstrakte Modelle sind eine entscheidende Hilfe zur Steigerung der Effizienz 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 am Lehrstuhl 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.

Wie schon in den vergangenen Jahren, bildeten im Berichtsjahr einige Themen zur Modellierung und Analyse von Geschäftsprozessen einen Schwerpunkt der Arbeiten des Lehrstuhls. Die intensive Kooperation mit dem Böblinger Labor der IBM bildete die Grundlage für den mit 40.000 \$ dotierten Faculty Award, den der Lehrstuhl im Oktober 2005 von der IBM erhalten hat. Mit einem Drittmittelprojekt werden diese Arbeiten seit Oktober 2005 verstärkt weitergeführt. Neben Petrinetzen werden am Lehrstuhl auch andere Modellierungstechniken untersucht, insbesondere Temporale Logiken und Abstract State Machines. Seit Juli 2005 verstärkt ein Drittmittelprojekt die Arbeiten zu ASM.

Lehre

In der Lehre konzentriert sich die LFE auf zentrale Aspekte des modellbasierten Software-Entwurfs mit den regelmäßig angebotenen Vorlesungen über Verteilte Algorithmen sowie über Methoden und Modelle des Systementwurfs. Desweiteren wird die Anwendung und Analyse dieser zentralen Aspekte in den Vorlesungen Computergestützte Verifikation und Analyse von Petrinetz-Modellen gelehrt. Daneben stehen vertiefende Seminare zu diesen Themen.

Der Lehrstuhl war wesentlich beteiligt an der Ringvorlesung "Schwerpunkte der Informatik" im Sommersemester 2005, die dem Thema "Informatik im historischen Kontext" gewidmet war und als Buch erscheinen wird.

Seit dem Wintersemester 2004/2005 hat die LFE folgende Veranstaltungen angeboten:

Veranstaltungen im Grundstudium

- Praktikum zur Praktischen Informatik 1/2 (B. HOHBERG/ K. AHRENS, WiSe 2004/2005)
- Übung zur "Praktischen Informatik 1" (P. MASSUTHE, WiSe 2004/2005, WiSe 2005/2006)
- Übung zur "Praktischen Informatik 2" (K. SCHMIDT, SoSe 2005)
- Proseminar "Beauty is our Business" (W. REISIG, SoSe 2005, WiSe 2005/2006)

Kernveranstaltungen

- Vorlesung "Computergestützte Verifikation" (K. SCHMIDT, WiSe 2005/2006)
- Vorlesung "Methoden und Modelle des Systementwurfs" (W. REISIG, SoSe 2005)
- Vorlesung "Analyse von Petrinetz-Modellen" (K. SCHMIDT, WiSe 2004/2005)
- Vorlesung "Verteilte Algorithmen" (W. REISIG, WiSe 2005/2006)

- Übung "Verteilte Algorithmen" (W. REISIG, WiSe 2005/2006)
- Übung "Computergestützte Verifikation" (P. MASSUTHE, WiSe 2005/2006)

Seminare

- Seminar "Systementwurf" (W. REISIG, SoSe 2005, WiSe 2005/2006)
- Seminar/ Forschungsseminar "Angewandte Verteilte Algorithmen" (W. REISIG, SoSe 2005, WiSe 2005/2006)
- Seminar „Spezifikation, Verifikation verteilter Systeme“ (P. MASSUTHE, WiSe 2004/05)
- Seminar "Statische Programmanalyse" (P. Massuthe, WiSe 2005/2006)

Forschung

Die Forschungsarbeiten des Lehrstuhls werden im Weiteren in Teil I unter thematischen und in Teil II unter organisatorischen Aspekten dargestellt. Thematisch geht es um Methoden und Modellierungstechniken für den Entwurf verteilter, reaktiver Systeme. Organisatorisch werden diese Arbeiten teilweise aus Drittmittelprojekten finanziert, die im zweiten Teil beschrieben werden.

I. Forschungsthemen

Die Forschungsthemen umfassen Methoden, Modellierungs- und Analysetechniken für den Systementwurf. Im ersten Thema geht es um die computergestützte Verifikation. Ein am Lehrstuhl entwickeltes Software-Werkzeug zur Analyse von Petrinetz-Modellen wird ausgebaut. Zwei weitere Themen behandeln spezielle Modellierungstechniken, die gemeinsam die wesentlichen Aspekte operationeller Spezifikationstechniken abdecken. Insbesondere betonen sie logik-basierte Kompositions- und Verfeinerungskonzepte (Temporal Logic of Distributed Actions) und einen universellen, modelltheoretisch begründeten Algorithmenbegriff (Abstract State Machines). Eine Reihe von Themen behandelt die Modellierung und Analyse von Geschäftsprozessen, insbesondere Beschreibungssprachen wie BPEL und Begriffe zur Analyse von Eigenschaften wie die Bedienbarkeit und die Konstruktion von Bedienungsanleitungen.

Projekt: Computergestützte Verifikation - LoLA

Ansprechpartner: PD DR. K. SCHMIDT

Zusammenarbeit: Université Paris VI: FABRICE KORDON, Universität Stuttgart: PROF. JAVIER ESPARZA

Das Werkzeug *LoLA* unterstützt die Analyse verteilter Systeme, die als Petrinetz modelliert sind. Basierend auf Suchalgorithmen über 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 2005 wurde LoLA um eine Komponente zur gegenbeispielgesteuerten Abstraktionsverfeinerung erweitert. Das Werkzeug wurde in mehreren Fallstudien am Lehrstuhl eingesetzt. Dabei konnte LoLA wesentlich stabilisiert werden. Teile von LoLA bilden den Kern für Werkzeuge zur Analyse offener Workflownetze.

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

Projekt: Temporal Logic of Distributed Actions (TLDA)

Ansprechpartner: DIPL. INF. PETER MASSUTHE, DIPL. INF. NIELS LOHMANN

Mit einer speziellen Interpretation von Variablen zur Synchronisation von Aktionen wird am Lehrstuhl eine Variante von TLA, die *Temporal Logic of Distributed Actions* (kurz: TLDA), entwickelt. Im Gegensatz zu TLA basiert sie auf *verteilten Abläufen*, in denen Ereignisse kausal geordnet sind. Dies erlaubt eine feinere Unterscheidung der Phänomene, die in verteilten System auftreten, insbesondere Nebenläufigkeit und Nichtdeterminismus.

Mit TLDA lassen sich besonders gut Systeme modellieren, die aus Komponenten aufgebaut sind, deren Aktionen synchronisierend miteinander interagieren.

Die Spezifikation eines großen Systems entsteht nicht in einem Schritt. Es wird zunächst eine abstrakte Spezifikation entwickelt, die nur überschaubar viele Aspekte berücksichtigt. In mehreren Verfeinerungsschritten wird diese dann in eine konkrete Spezifikation verfeinert. In diesem Projekt wurden die grundlegenden formalen Begriffe zur Verfeinerung in TLDA definiert. Die Diplomarbeit [Massuthe 04] arbeitete verschiedene Methoden der Verfeinerung in TLDA heraus und demonstrierte diese an Fallbeispielen. Sie bietet noch viele Ansatzpunkte für tiefer gehende Betrachtungen. So soll die Transitionsverfeinerung weiterentwickelt werden und die Bewahrung von Eigenschaften bei der Verfeinerung näher untersucht werden. Schwerpunkt ist dabei das Erstellen einer Sammlung von Fallbeispielen und die Ableitung von Entwurfsmustern, die den verfeinernden Beweis von Eigenschaften wesentlich vereinfachen.

Die Möglichkeit der automatischen Verifikation von Eigenschaften gegebener TLDA-Spezifikationen ist wesentlich für die weitere Verbreitung der Sprache TLDA. In diesem Projekt sollen deshalb Modelchecking-Algorithmen entstehen, die speziell für TLDA-Abläufe optimiert sind. In einem ersten Schritt entsteht zunächst ein Zustands- und Aktionsbegriff für TLDA. Dieser erlaubt dann die Übersetzung von (syntaktisch eingeschränkten) TLDA-Spezifikationen in Transitionssysteme, die eine Voraussetzung für Modelchecking sind. Anschließend sollen bestehende Algorithmen zum expliziten Modelchecking an die Verifikation von TLDA-Spezifikationen angepasst und prototypisch implementiert werden. Grundlage hierzu bieten die Arbeiten [Lohmann 05a, Lohmann 05b], in denen eine schrittweise Interleavingsemantik entwickelt wurde.

Projekt: Abstract State Machines (ASM)

Ansprechpartner: PROF. DR. WOLFGANG REISIG, DIPL.-INF. ANDREAS GLAUSCH

Mitarbeiter: DIRK FAHLAND

Gurevichs "Abstract State Machines" (ASMs) sind ein Modellierungsformalismus mit dem Anspruch, den Zusammenhang eines realen oder geplanten Systems mit einem formalen Modell besonders natürlich, eng und nahtlos gestalten und formulieren zu können. Dies gelingt mit der Idee, Zustände als mathematische Strukturen aufzufassen, deren syntaktische Darstellung unerheblich ist. Wie in der Mathematik üblich, argumentiert man über

einzelne Aspekte einer solchen Struktur mit Hilfe von Termen über einer festen Signatur. Unter Verwendung dieser Terme werden dann ASM-Programme (im elementaren Fall konventionelle parallele, bedingte Wertzuweisungen) formuliert, die den aktuellen Zustand in den Nachfolgestand transformieren. Auf diese Weise erhalten die mathematischen Strukturen, ihrer Natur nach eigentlich statisch, eine Dynamik.

Die Forschung des Lehrstuhls konzentriert sich auf Fragen der Ausdruckskraft der verschiedenen Varianten von ASMs und betrachtet dabei vorrangig verteilte ASMs. Dies geschieht dabei immer in Bezug zu anderen Systemmodellen, um die Einordnung von ASMs in andere Formalismen zu ermöglichen. Weiterhin untersuchen wir die Verwendung von ASMs zur Modellierung von Web Services mit besonderem Blick auf die Sprache BPEL.

Projekt: Formale Semantik für Geschäftsprozess-Beschreibungssprachen

Ansprechpartner: DIPL.-INF. NIELS LOHMANN, PROF. DR. WOLFGANG REISIG, PD DR. KARSTEN SCHMIDT, DIPL.-INF. CHRISTIAN STAHL

Zusammenarbeit: DR. AXEL MARTENS

Mitarbeiter: DIRK FAHLAND, SEBASTIAN HINZ, PETER LAUFER

Beschreibungssprachen für Geschäftsprozesse liegen zumeist in Form einer syntaktischen XML-Spezifikation vor und sind nur umgangssprachlich formuliert, d.h. sie haben keine formale Semantik. Aufgrund der fehlenden Semantik enthalten diese Sprachen einerseits Unschärfen und andererseits ist eine Analyse von Prozessen, die in einer dieser Sprachen spezifiziert sind, mit Hilfe formaler Methoden nur schwer möglich. Stellvertretend für alle diese Sprachen untersuchen wir in diesem Projekt die *Business Process Execution Language for Web Services* (BPEL) und entwickeln für sie eine formale Semantik. Die erarbeiteten Ansätze zur Formalisierung haben einen ganz allgemeinen Charakter und lassen sich auf alle anderen Beschreibungssprachen für Geschäftsprozesse übertragen.

Als Modellierungsmethoden verwenden wir *Petrinetze* und *Abstract State Machines* (ASMs).

Petrinetze sind ein etablierter Formalismus zur Beschreibung von Geschäftsprozessen. Sie eignen sich hervorragend, um die Kausalität der Aktivitäten und damit das dynamische Verhalten eines Geschäftsprozesses zu modellieren. Mit ASMs lassen sich vor allem Strukturen und Datenaspekte adäquat abbilden. Neben der Formalisierung der Sprachen ist der *Vergleich* der beiden Ansätze ein weiterer Aspekt des Projektes.

Ein weiterer interessanter Aspekt ist die Frage der *Korrektheit*, der von uns entwickelten formalen Semantiken. Es ist nicht möglich, formal zu beweisen, dass die Semantik die zentralen semantischen Eigenschaften der Spezifikation bewahrt, da BPEL in einer informellen Spezifikation vorliegt. Aus diesem Grund können wir nur die Plausibilität anhand der informellen Spezifikation prüfen. Hier unterscheiden sich die beiden Formalismen: Da mit ASMs Systeme auf jedem beliebigen Abstraktionsgrad formal beschrieben werden können, ist ein Vergleich des ASM-Modells mit der informellen Spezifikation direkt möglich. Das Petrinetz-Modell eines konkreten Prozesses muss hingegen simuliert und temporallogische Eigenschaften der Abläufe mit dem Prozess verglichen werden.

Für die Sprache BPEL haben wir einen Compiler, BPEL2PN, implementiert, der einen BPEL-Prozess in ein Petrinetz entsprechend der am Lehrstuhl entwickelten Petrinetz-Semantik für BPEL transformiert. Aufbauend auf diesem Werkzeug analysieren wir im Thema "Analyse von Geschäftsprozess-Beschreibungssprachen" BPEL-Prozesse.

Projekt: Analyse von Geschäftsprozess-Beschreibungssprachen

Ansprechpartner: DIPL.-INF. NIELS LOHMANN, PROF. DR. WOLFGANG REISIG, PD DR. KARSTEN SCHMIDT, DIPL.-INF. CHRISTIAN STAHL

Zusammenarbeit: DR. AXEL MARTENS

Mitarbeiter: THOMAS HEIDINGER, JENS KLEINE, GERRIT MÜLLER

Wir wollen BPEL-Prozesse mit *statischer Analyse* und *Modelchecking* computergestützt verifizieren. Mit Hilfe statischer Analyse kann ein Prozess schon zur Compilezeit analysiert werden. Modelchecking ist eine Methode, die den gesamten Zustandsraum eines Modells auf eine (temporallogische) Eigenschaft untersucht. Das Hauptproblem ist dabei, die Größe des Zustandsraums in den Griff zu bekommen.

Wir entwickeln *Algorithmen* und *Methoden*, um BPEL-Prozesse statisch zu analysieren. Diese Methoden sollen es einem Anwender, der einen BPEL-Prozess implementiert, ermöglichen, schon zur Compilezeit Informationen über den Kontrollfluss oder beispielsweise uninitialisierte Variablen zu erhalten. Weiterhin ermöglichen Methoden der statischen Analyse auch die Verifikation datenabhängiger Eigenschaften.

Desweiteren arbeiten wir *Eigenschaften* für BPEL-Prozesse heraus. Zum Nachweis jeder dieser Eigenschaften benötigt man ein geeignetes Modell. Dazu werden die *Bausteine* der am Lehrstuhl entwickelten Petrinetz-Semantik für BPEL auf die jeweilige Eigenschaft angepasst. Ziel ist es, ein möglichst kleines Modell zu erhalten, um die Größe des Zustandsraumes zu minimieren.

Ein weiterer Aspekt des Projektes ist die *Verzahnung* der beiden Ansätze. Ein BPEL-Prozess wird zuerst statisch analysiert. Die Ergebnisse der statischen Analyse fließen anschließend in die Generierung eines Modells ein.

Eine andere Technik, den Zustandsraum eines Modells M wirkungsvoll zu reduzieren, ist die Konstruktion eines (bzgl. M) abstrakten Modells M' , das die in M zu untersuchende Eigenschaft bewahrt aber einen kleineren Zustandsraum als M hat. Wir entwickeln Abstraktionsalgorithmen die sowohl existierende strukturelle Reduktionstechniken von Petrinetzen erweitern bzw. an die Modelle anpassen als auch die Abläufe eines Petrinetzes berücksichtigen.

Die entwickelten Algorithmen und Methoden sind aber durchaus auf alle anderen Ansätze von Geschäftsprozess-Beschreibungssprachen übertragbar.

Projekt: Bedienbarkeit

Ansprechpartner: DIPL.-INF. PETER MASSUTHE, PD DR. KARSTEN SCHMIDT

Zusammenarbeit: DR. AXEL MARTENS

Mitarbeiter: CARSTEN FRENKLER

Wenn ein Geschäftsprozess mit anderen Prozessen interagiert, hängt es nicht mehr vom Prozess allein ab, ob er erfolgreich abgearbeitet werden kann. Dennoch kann man einem solchen Prozess die prinzipielle Eignung zu erfolgreicher Abarbeitung in einer Umgebung ansehen. Dies ist formalisiert in der Eigenschaft *Bedienbarkeit*.

Wir untersuchen Bedienbarkeit unter verschiedenen Annahmen an die Umgebung, so z.B. deren Verteiltheit oder deren Einbindung in transaktionale Anforderungen. Für die so entstehenden verschiedenen Problemstellungen entwickeln wir geeignete Begriffe, Entscheidungsverfahren und Techniken zur Effizienzsteigerung dieser Verfahren.

Bedienbarkeit wird entschieden durch die Konstruktion einer bedienenden Umgebung. Für azyklische Geschäftsprozesse kann dabei die Existenz einer eindeutig bestimmten liberalsten Umgebung gezeigt werden. Die konstruierte Umgebung kann weiterverwendet werden als "Bedienungsanleitung" für den analysierten Prozess. Durch zusätzliche Annotationen an die liberalste Umgebung kann ein Bediener *jede* bedienende Umgebung ableiten.

Die Erkenntnisse aus diesem Projekt werden erprobt in den Arbeiten zur Geschäftsprozessbeschreibungssprache BPEL.

Projekt: Bedienungsanleitungen

Ansprechpartner: DIPL.-INF. PETER MASSUTHE, PD DR. KARSTEN SCHMIDT

Mitarbeiter: KATHRIN KASCHNER

Partnerübergreifende Geschäftsprozesse werden aus *Services* zusammengesetzt und bilden ihrerseits einen partnerübergreifenden Service. Dies geschieht zunehmend im Rahmen der *Service-orientierten Architektur* (SOA): Besitzer von Services werden unterteilt in Service Provider, Service Requester und Service Broker. Der Service Provider veröffentlicht Informationen über seinen Service bei einem Service Broker. Dieser verwaltet diese Informationen vieler Provider und ermöglicht es einem Service Requester, einen zu seinem Service passenden Service eines Providers zu finden.

In diesem Forschungsthema betrachten wir Bedienungsanleitungen. Eine Bedienungsanleitung ist eine kompakte Beschreibung aller passenden Services möglicher Requester für einen gegebenen Service eines Providers. Bedienungsanleitungen dienen deshalb als die an einen Broker zu veröffentliche Information.

Weiterhin wird das Matching eines Services auf eine Bedienungsanleitung untersucht, also die Frage, ob ein Service eines Requesters eine Bedienungsanleitung respektiert. Kann diese Frage bejaht werden, so ist garantiert, dass Requester und Provider zueinander passen und vernünftig miteinander interagieren können. Matcht der Service nicht mit der Bedienungsanleitung, so können in der Interaktion beider Probleme auftreten, wie unerwartete Nachrichten oder Deadlocks.

Als Modellierungsmethoden verwenden wir Petrinetze und Automaten. Speziell annotierte Automaten dienen als kompakte Repräsentation einer Menge (nicht annotierter) Automaten. Diese annotierten Automaten werden ihrerseits als so genannte shared BDD repräsentiert.

II. Drittmittelprojekte

In Teil I entwickelte Methoden und Techniken werden in den Drittmittelprojekten des Lehrstuhls in mehreren Bereichen eingesetzt und weiterentwickelt. Den wichtigsten Einsatzbereich bilden derzeit verteilte Geschäftsprozesse.

Einen ganz anderen Anwendungsbereich bilden die global asynchronen, lokal synchronen Hardware-Schaltungen (GALS) mit verschiedenen, in Teilen neuen Varianten von Petrinetzen.

Im dritten Bereich werden Abstract State Machines weiterentwickelt.

Projekt: Tools4BPEL

Ansprechpartner: PROF. DR. WOLFGANG REISIG, PD DR. KARSTEN SCHMIDT, DIPL.-INF. NIELS LOHMANN

Zusammenarbeit: DIPL.-INF. MICHAEL HÖFT (GEDILAN TECHNOLOGIES), PROF. DR. FRANK LEYMAN (UNIVERSITÄT STUTTGART),

Mitarbeiter: DIPL.-INF. PETER MASSUTHE, DIPL.-INF. CHRISTIAN STAHL, DIPL.-INF. DANIELA WEINBERG, CHRISTIAN GIERDS, DENNIS REINERT

Forschungsförderung: Bundesministerium für Bildung und Forschung

BPEL ist eine junge Sprache zur Beschreibung von Services. Im Projekt Tools4BPEL wird die Entwicklung der Sprache mit formalen Methoden begleitet. Wir bearbeiten dabei die Schwerpunkte *Eigenschaften von Services* und *Korrektheit*.

Im Schwerpunkt *Eigenschaften von Services* betrachten wir wesentliche für Services relevante Eigenschaften, u.a. *Austauschbarkeit* (Ersetzung eines Service durch einen neuen, z.B. verbesserten), *Kompatibilität* (korrektes Zusammenarbeiten von Services) und *Bedienbarkeit* (Existenz von kompatiblen Services). Wir formalisieren diese Eigenschaften mit Hilfe *offener Workflownetze*, einer Klasse von Petrinetzen. Wir entwickeln effiziente Techniken zur Analyse dieser Eigenschaften. Weiterhin studieren wir die Generierung von *abstrakten Sichten* und *Bedienungsanleitungen* für Services. Diese Informationen dienen der Integration von Services in *service-orientierten Architekturen*.

Im Schwerpunkt *Korrektheit* übersetzen wir in BPEL beschriebene Services auf der Grundlage einer am Lehrstuhl entwickelten formalen Semantik in offene Workflownetze. Ziel der Übersetzung ist die Anbindung an Werkzeuge der computergestützten Verifikation, insbesondere LoLA, sowie an die im ersten Schwerpunkt entwickelten Werkzeuge. Zur Verringerung der Komplexität der entstehenden Petrinetzmodelle führen wir statische Analyse auf BPEL durch und nutzen die gewonnene Information zur Vereinfachung der konstruierten Petrinetze.

Die entwickelten Techniken und Werkzeuge werden anhand realistischer Fallstudien validiert und in ein kommerzielles Werkzeug zur Modellierung von Geschäftsprozessen integriert.

Projekt: GALS

Ansprechpartner: DIPL.-INF. CHRISTIAN STAHL, PROF. DR. WOLFGANG REISIG

Mitarbeiter: JAN BRETSCHNEIDER, ALEXANDRA JULIUS

Zusammenarbeit: DR. ECKHARD GRASS (IHP FRANKFURT/ODER), MILOŠ KRSTIĆ (IHP FRANKFURT/ODER), PROF. DR. BEATE MEFFERT (HU BERLIN), PROF. DR. HOLGER SCHLINGLOFF (HU BERLIN), DR. FRANK WINKLER (HU BERLIN)

Forschungsförderung: Deutsche Forschungsgemeinschaft (DFG)

Wir modellieren *global asynchrone*, aber *lokal synchrone* Systeme (GALS), eine spezielle Art von Hardware-Schaltungen (*GALS-Schaltungen*). Eine GALS-Schaltung besteht aus mehreren *Modulen*. Jedes Modul arbeitet in sich synchron und ist zugleich in einen *Wrapper* eingebettet, der eine asynchrone Schnittstelle zu anderen Modulen anbietet.

Eine GALS-Schaltung vereinigt die Vorteile synchroner und asynchroner Schaltungen: Die Taktverteilung in einer GALS-Schaltung ist einfacher umzusetzen als in einer synchronen Schaltung. Weiterhin sind der Stromverbrauch sowie die elektromagnetischen Interferenzen

zen geringer. Vor allem die letzten beiden Aspekte sind Gründe dafür, dass GALS-Schaltungen in Systemen zur mobilen Kommunikation verwendet werden.

Wir entwickeln ein Modell, das Module mit ihrem Wrapper und insbesondere den Übergang zwischen Modul und Wrapper beschreibt. Als Modellierungsmethode erscheinen Petrinetze besonders geeignet, da synchrone Systeme durch sequentielle Abläufe und asynchrone Systeme durch verteilte Abläufe beschrieben werden können.

Designer von GALS-Systemen sind an *Eigenschaften* ihrer Schaltung interessiert. Es wird zwischen funktionalen Eigenschaften (Lebendigkeits- und Sicherheitseigenschaften) und nichtfunktionalen Eigenschaften (Stromverbrauch, Datendurchsatz) unterschieden. Für jede dieser Eigenschaften entwickeln wir ein geeignetes Modell anhand dessen die entsprechende Eigenschaft analysiert werden kann. Weiterhin erarbeiten wir zum Nachweis dieser Eigenschaften *Verifikationstechniken*. Beispielsweise konnte die Frage, ob in einem Wrapper ein *Hazard* auftreten kann (ein Signal des Wrappers nimmt einen undefinierten Wert an), auf die Erreichbarkeit einer Teilmarkierung in unserem Petrinetzmodell reduziert werden. Die Frage nach der Erreichbarkeit ist ein Modelchecking-Problem. Funktionale Eigenschaften lassen sich alle auf ein Modelchecking-Problem reduzieren. Der Nachweis nichtfunktionaler Eigenschaften ist wahrscheinlich die schwierigere Aufgabe, da die bisherigen Modelle und Verifikationstechniken ungeeignet erscheinen.

Projekt: Ausdrucksstärke von Abstract State Machines

Ansprechpartner: PROF. DR. WOLFGANG REISIG, DIPL.-INF. ANDREAS GLAUSCH

Mitarbeiter: DIRK FAHLAND

Zusammenarbeit: PROF. DR. EGON BÖRGER (Universität Pisa), PROF. DR. YURI GUREVICH (Microsoft Research)

Forschungsförderung: Deutsche Forschungsgemeinschaft (DFG)

ASM schlagen einen modelltheoretisch begründeten Algorithmenbegriff vor, in dem frei wählbare semantische Objekte und Operationen die elementaren Bestandteile sind. In diesem Projekt werden verschiedene Klassen von ASM untersucht und deren Ausdrucksstärke charakterisiert, in Anlehnung an ein charakterisierendes Theorem für elementare "small step" ASM von Gurevich.

Der Schwerpunkt dieses Projektes liegt dabei auf der Klasse der *verteilten ASMs*. Eine verteilte ASM besteht aus einer Menge von elementaren ASMs, die durch das nebenläufige Ausführen von *Aktionen* den aktuellen Zustand verändern. Die kausalen Abhängigkeiten der Aktionen werden dabei durch *verteilte Abläufe* explizit modelliert. Wir betrachten und erarbeiten in diesem Projekt mehrere Varianten verteilter ASMs, die sich in Darstellung und Ausdruckskraft unterscheiden. Das letztendliche Ziel des Projektes ist die Erarbeitung einer Spezifikationstechnik, durch die sich ein verteiltes System mit Hilfsmitteln der algebraischen Spezifikation beschreiben und untersuchen lässt.

Veröffentlichungen

D. FAHLAND: Complete Abstract Operational Semantics for the Web Service Business Process Execution Language. Informatik-Berichte 190, Humboldt-Universität zu Berlin, September 2005.

D. FAHLAND, W. REISIG: *ASM-based semantics for BPEL: The negative Control Flow*. In Danièle Beauquier, Egon Börger, and Anatol Slissenko, editors, Proceedings of the 12th

International Workshop on Abstract State Machines (ASM'05), S. 131-151, Paris XII. März 2005.

C. FRENKLER, K. SCHMIDT: *Modellierung und Analyse transaktionaler Geschäftsprozesse*. In Karsten Schmidt und Christian Stahl, Hrsg., 12. Workshop Algorithmen und Werkzeuge für Petrinetze (AWPN 2005), Proceedings, S. 72-77, Humboldt-Universität zu Berlin. September 2005.

A. GLAUSCH: *Eine Charakterisierung einfacher Petrinetz-Schemata*. In Karsten Schmidt und Christian Stahl, Hrsg., 12. Workshop Algorithmen und Werkzeuge für Petrinetze (AWPN 2005), Proceedings, S. 1-6, Humboldt-Universität zu Berlin. September 2005.

E. GRASS, F. WINKLER, M. KRSTIC, A. JULIUS, CH. STAHL, M. PIZ: *Enhanced GALS Techniques for Datapath Applications*. In Vassilis Paliouras, Johan Vounckx, und Diederik Verkest, editors, Integrated Circuit and System Design: 15th International Workshop, PATMOS 2005, Leuven, Belgium, September 20-23, 2005, volume 3728 of Lecture Notes in Computer Science, S. 581-590, Springer-Verlag. August 2005.

S. HINZ, K. SCHMIDT, CH. STAHL: *Transforming BPEL to Petri Nets*. In W.M.P. van der Aalst, B. Benatallah, F. Casati, and F. Curbera, editors, Proceedings of the Third International Conference on Business Process Management (BPM 2005), volume 3649 of Lecture Notes in Computer Science, Nancy, France, S. 220-235, Springer-Verlag. September 2005.

M. KRSTIC, E. GRASS, CH. STAHL: *Request-Driven GALS Technique for Wireless Communication System*. In Proceedings of the 11th International Symposium on Advanced Research in Asynchronous Circuits and Systems (ASYNC 2005), New York, NY, USA, S. 76-85, IEEE Computer Society. März 2005.

A. MARTENS: *Consistency between Executable and Abstract Processes*. In Proceedings of Intl. IEEE Conference on e-Technology, e-Commerce, and e-Services (EEE'05), IEEE Computer Society Press. März 2005.

A. MARTENS: *Analyzing Web Service based Business Processes*. In Maura Cerioli, editor, Proceedings of Intl. Conference on Fundamental Approaches to Software Engineering (FASE'05), Part of the 2005 European Joint Conferences on Theory and Practice of Software (ETAPS'05), volume 3442 of Lecture Notes in Computer Science, Springer-Verlag. April 2005.

A. MARTENS: *Simulation and Equivalence between BPEL Process Models*. In Proceedings of the Design, Analysis, and Simulation of Distributed Systems Symposium (DASD'05), Part of the 2005 Spring Simulation Multiconference (SpringSim'05), April 2005.

A. MARTENS: *Process Oriented Discovery of Business Partners*. In Proceedings of 7th Intl. Conference on Enterprise Information Systems (ICEIS'05), Vol 3, INSTICC. Mai 2005.

P. MASSUTHE, K. SCHMIDT: *Operating Guidelines - an Alternative to Public View*. Informatik-Berichte 189, Humboldt-Universität zu Berlin, 2005.

P. MASSUTHE, K. SCHMIDT: *Matching Nondeterministic Services with Operating Guidelines*. Informatik-Berichte 193, Humboldt-Universität zu Berlin, Juni 2005.

P. MASSUTHE, W. REISIG, K. SCHMIDT: *An Operating Guideline Approach to the SOA*. 2nd South-East European Workshop on Formal Methods 2005 (SEEFM05), Ohrid, Republic of Macedonia, 2005. Erschienen auch in Informatik-Berichte 191, Humboldt-Universität zu Berlin, 2005.

P. MASSUTHE, K. SCHMIDT: *Operating Guidelines - an Automata-Theoretic Foundation for the Service-Oriented Architecture*. In Kai-Yuan Cai, Atsushi Ohnishi, and M.F. Lau, editors, Proceedings of the Fifth International Conference on Quality Software (QSIC 2005), Melbourne, Australia, S. 452-457, IEEE Computer Society. September 2005.

P. MASSUTHE, K. SCHMIDT: *Operating Guidelines for Services*. In Karsten Schmidt und Christian Stahl, Hrsg., 12. Workshop Algorithmen und Werkzeuge für Petrinetze (AWPN 2005), Proceedings, S. 78-83, Humboldt-Universität zu Berlin. September 2005.

W. REISIG: *Modeling- and Analysis Techniques for Web Services and Business Processes*. In Martin Steffen and Gianluigi Zavattaro, editors, Formal Methods for Open Object-Based Distributed Systems: 7th IFIP WG 6.1 International Conference, FMOODS 2005, Athens, Greece, June 15-17, 2005. Proceedings, volume 3535 of Lecture Notes in Computer Science, S. 243-258, Springer Verlag. Mai 2005. Erschienen auch in Informatik-Berichte 183, Humboldt-Universität zu Berlin, 2005.

W. REISIG: *On the Expressive Power of Petri Net Schemata*. In Gianfranco Ciardo and Philippe Darondeau, editors, Applications and Theory of Petri Nets 2005: 26th International Conference, ICATPN 2005, Miami, USA, June 20-25, 2005. Proceedings, volume 3536 of Lecture Notes in Computer Science, S. 349-364, Springer Verlag. Mai 2005.

W. REISIG, K. SCHMIDT, CH. STAHL: *Kommunizierende Workflow-Services modellieren und analysieren*. Informatik - Forschung und Entwicklung, S. 90-101, Oktober 2005. Erschienen auch in Informatik-Berichte 182, Humboldt-Universität zu Berlin, Februar 2005.

B.-H. SCHLINGLOFF, A. MARTENS, K. SCHMIDT: *Modeling and Model Checking Web Services*. Electronic Notes in Theoretical Computer Science: Issue on Logic and Communication in Multi-Agent Systems, 126: 3-26, März 2005.

K. SCHMIDT: *Controllability of Open Workflow Nets*. In Proc. Workshop Enterprise Modeling and Information System Architectures Klagenfurt, Lecture Notes in Informatics P-75, S. 236-249, 2005. Erschienen auch in Informatik-Berichte 180, Humboldt-Universität zu Berlin, 2005.

K. SCHMIDT, CH. STAHL (HRSG): 12. Workshop Algorithmen und Werkzeuge für Petrinetze (AWPN 2005), Proceedings. Informatik-Berichte 192, Humboldt-Universität zu Berlin, September 2005.

CH. STAHL, W. REISIG, M. KRSTIC: *Hazard Detection in a GALS Wrapper: a Case Study*. In Jörg Desel and Y. Watanabe, editors, Proceedings of the Fifth International Conference on Application of Concurrency to System Design (ACSD'05), St. Malo, France, S. 234-243, IEEE Computer Society. Juni 2005. Erschienen auch in Informatik-Berichte 184, Humboldt-Universität zu Berlin, Februar 2005

CH. STAHL: *A Petri Net Semantics for BPEL*. Informatik-Berichte 188, Humboldt-Universität zu Berlin, Juli 2005.

D. WEINBERG, K. SCHMIDT: *Reduction Rules for Interaction Graphs*. In Karsten Schmidt und Christian Stahl, Hrsg., 12. Workshop Algorithmen und Werkzeuge für Petrinetze (AWPN 2005), Proceedings, S. 60-65, Humboldt-Universität zu Berlin. September 2005.

Vorträge

J. BRETSCHNEIDER: *Tool Demo: LoLA*. Miniworkshop GALS; HU Berlin, Januar 2005.

C. FRENKLER: *Modellierung und Analyse transaktionaler Geschäftsprozesse*. 12. Workshop Algorithmen und Werkzeuge für Petrinetze (AWPN 2005), Humboldt-Universität zu Berlin. September 2005.

A. GLAUSCH: *Eine Charakterisierung einfacher Petrinetz-Schemata*. 12. Workshop Algorithmen und Werkzeuge für Petrinetze (AWPN 2005), Humboldt-Universität zu Berlin. September 2005.

A. JULIUS: *Extracting the clock of a GALS wrapper*. Miniworkshop GALS; HU Berlin, Januar 2005.

P. MASSUTHE: *Bedienungsanleitungen für Services*. Workshop AWPN, Berlin, September 2005.

P. MASSUTHE: *Operating Guidelines for Services*. Vortrag im Forschungsseminar, TU Eindhoven, November, 2005.

W. REISIG: *Hazard Detection in a GALS Wrapper: a Case Study*. Miniworkshop GALS; HU Berlin, Januar 2005.

W. REISIG: *Distributed Small Step ASM Capture Distributed Small Step Algorithms*. Redmond, USA, Februar 2005.

W. REISIG: *50 Jahre modellbasierter Entwurf: Vom Modellieren mit Programm zum Programmieren mit Modellen*. Ringvorlesung; HU Berlin, Mai 2005.

W. REISIG: *Modeling- and Analysis Techniques for Web Services and Business Processes*. MOODS 2005, Athen, Juni 2005.

W. REISIG: *Informal Introduction to Petri Nets*. Tutorial ATPN 2005; Miami, Florida, Juni 2005.

W. REISIG: *On the expressive power of Petri Net schemata*. ATPN 2005; Miami, Florida, Juni 2005.

W. REISIG: *Operating Guidelines for Workflow Services*. eingeladener Vortrag, Workshop ATPN 2005 to Coordination Workflow and Business Process Management; Miami, Florida, Juni 2005.

W. REISIG: *Bedienungsanleitungen für Workflow Services*. Start-up Workshop des BMBF-Projektes TOOLS4BPEL, HU Berlin, Oktober 2005.

W. REISIG: *Operating Guidelines for Service Oriented Architectures*. 2 South-East European Workshop on Formal Methods, Ohrid, Republik Mazedonien, November 2005.

K. SCHMIDT: *Operating Guidelines - an Automata-Theoretic Foundation for the Service-Oriented Architecture*. QSIC 05, Melbourne, Australien, September 2005.

K. SCHMIDT: *Reduction Rules for Interaction Graphs*. 12. Workshop Algorithmen und Werkzeuge für Petrinetze (AWPN 2005), Humboldt-Universität zu Berlin. September 2005.

K. SCHMIDT: KARSTEN SCHMIDT: *Controllability of Open Workflow Nets*. EMISA, Klagenfurt, Oktober 2005

CH. STAHL: Desynchronization - Transforming synchronous circuits into asynchronous ones. Miniworkshop GALS; HU Berlin, Januar 2005.

CH. STAHL: *Hazard Detection in a GALS Wrapper: a Case Study*. Fifth International Conference on Application of Concurrency to System Design (ACSD'05), St. Malo, France, June 2005.

CH. STAHL: *GALS Project at Humboldt-Universitaet zu Berlin*. Vortrag im Forschungsseminar, University of Newcastle upon Tyne, 21.7.2005.

CH. STAHL: *Transforming BPEL to Petri Nets*. Third International Conference on Business Process Management (BPM 2005), Nancy, France, September 2005.

Sonstige Aktivitäten

Prof. Dr. Wolfgang Reisig

- Mitglied im Steering-Committee "Conference on Application and Theory of Petri Nets"
- Mitglied der EG-Arbeitsgruppe "Formal Methods Europe"
- Gutachter für Graduiertenkollegs und Schwerpunktprogramme der DFG
- Gutachter für verschiedene Zeitschriften, Fachtagungen, DFG und NSF
- Editor des "Journal on Software and Systems Modeling", (SoSym) Springer-Verlag
- Vertreter der HU Berlin im Fakultätentag Informatik
- Mitglied der Aufnahmekommission des Fakultätentages Informatik
- Mitglied des Beirates des BMFT-Projektes MMiSS (Multimedia-Instruktionen in Sicheren Systemen)
- Vorsitzender des Beirates des John von Neumann Minerva Center for "Verification of Reactive Systems", Weizmann-Institut, Rehovot (Israel)

Baver Acu

- Teilnahme an AWPN 2005
- Gutachter für ICATPN 2006

Andreas Glausch

- Gutachter für ICATPN 2006
- Teilnahme an AWPN 2005

Prof. Dr. Bodo Hohberg

- Studienfachberatung Informatik
- Mitglied des Prüfungsausschuss des Instituts für Informatik

Niels Lohmann

- Teilnahme an AWPN 2005
- Gutachter für ICATPN 2006

Dr. Axel Martens

- seit April 2004 beurlaubt, Wissenschaftlicher Mitarbeiter bei IBM TJ Watson Research Center, Hawthorne, NY (USA), Component Systems Group

- Seit Juni 2004 Mitglied im OASIS Technischen Komitee zur Standardisierung von BPEL4WS

Peter Massuthe

- Gutachten für AWPN 2005, ICATPN 2006, WS-FM 2005
- Teilnahme an TAPSoft 2005, AWPN 2005, ICSOC 2005
- Teilnahme am Dagstuhl-Seminar "Service-Oriented Computing"
- Forschungsaufenthalt an der TU Eindhoven

PD Dr. Karsten Schmidt

- Editor für das Springer Journal of Universal Computer Science (J.UCS)
- Mitglied im örtlichen Wahlvorstand des Instituts für Informatik
- Mitglied der Kommission Lehre und Studium des Instituts für Informatik
- Stellv. Leiter der FG "Petrietze und verwandte Systemmodelle" der GI
- Mitglied im Programmkomitee der Int. Conf. Application and Theory of Petri Nets 2006
- Organisator des Workshops Algorithmen und Werkzeuge für Petrietze 2005
- Mitglied im Programmkomitee für den Workshop Coloured Petri Nets 05
- Gutachten, u. a. für The Computer Journal, Int Conf. Computer Aided Verification 2005, Int. Conf. Application and Theory of Petri Nets 2005, ACM Transactions on Software Engineering and Methodology, Int Journal on Software Tools for Technology Transfer, International Journal of Foundations of Computer Science, International Journal of Systems Science.

Christian Stahl

- Organisator des Workshops "Algorithmen und Werkzeuge für Petrietze" (AWPN 2005)
- Gutachten für "The International Journal of Time-Critical Computing Systems", "Workshop and Tutorial on Practical Use of Coloured Petri Nets and the CPN Tools" (CPN 2005), Workshop "Algorithmen und Werkzeuge für Petrietze" (AWPN 2005), "IEEE International Symposium on Asynchronous Circuits and Systems" (ASYNC 2006), "International Conference on Application and Theory of Petri Nets and Other Models of Concurrency" (ICATPN 2006)
- Forschungsaufenthalt University of Newcastle upon Tyne, 18.7-22.7.05
- Teilnahme an "Distinguished Lecture Series Theory and Practice of Software Development" (TAPSoft 2005), "International Conference on Application of Concurrency to System Design" (ACSD 2005), "International Conference on Business Process Management" (BPM 2005), "Algorithmen und Werkzeuge für Petrietze" (AWPN 2005)
- Teilnahme an der "ACiD-WG Winter School: Timing for Deep Submicron Chips - Mixing synchronous and asynchronous techniques to build robust systems on chip", 3.1.05 – 7.1.05, Cambridge, UK

Diplomarbeiten

CARSTEN FRENKLER: *Modellierung und Analyse transaktionaler Geschäftsprozesse*. Diplomarbeit, Humboldt-Universität zu Berlin, Juli 2005.

ANDREAS GLAUSCH: *Varianten des ASM-Theorems*. Diplomarbeit, Humboldt-Universität zu Berlin, Juni 2005.

SEBASTIAN HINZ: *Implementierung einer Petrinetz-Semantik für BPEL*. Diplomarbeit, Humboldt-Universität zu Berlin, März 2005.

NIELS LOHMANN: *Formale Fundierung und effizientere Algorithmen für die schrittweise TLDA-Interleavingsemantik*. Diplomarbeit, Humboldt-Universität zu Berlin, September 2005.

ELKE SALECKER, NORA TOUSSAINT. *Gegenbeispielgesteuerte Abstraktionsverfeinerung in High-Level-Petrinetzen*. Diplomarbeit, Humboldt-Universität zu Berlin, September 2005.

ALEXANDER BRADE: *Übersetzung graphischer Verhaltensbeschreibungen von Services in Abstract State Machines*. Diplomarbeit, Humboldt-Universität zu Berlin, November 2005.

Dissertationen

ADRIANNA ALEXANDER: *„Komposition Temporallogischer Spezifikationen – Spezifikation und Verifikation von Systemen mit Temporal Logic of Distributed Actions.“* Dissertation, Humboldt-Universität zu Berlin, September 2005.

Lehr- und Forschungseinheit

Softwaretechnik

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

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

Projektmitarbeiter

DIPL.-INF. SAM JOACHIM

Promotionsstudent

DIPL.-INF. FRANK LAMMERMANN

Tutorinnen und Tutoren

YVONNE GABRIEL

ANDREAS HIRTH

VOLKER JANETSCHKE

SAM JOACHIM

SKANDER MORGENTHALER

LUBOMIRA STANTCHEVA

Die gegenwärtigen Arbeitsschwerpunkte der Gruppe liegen auf dem Gebiet der Softwaretechnik, insbesondere im Reverse Engineering, beim Software-Test sowie dem Projektmanagement. Dabei steht auf dem Gebiet des Software Reverse Engineering ein gemeinsames Projekt mit dem Institut für Physik der Humboldt-Universität zu Berlin im Mittelpunkt. In Ergänzung zu diesem Projekt entstand eine multimediale Aufbereitung des Anwendungsbereichs, die durch beide Projektpartner genutzt wird.

Ein weiteres Projekt befasst sich in Zusammenarbeit mit einem industriellen Partner mit den besonderen Bedingungen für das Requirements Engineering und die Software-Entwicklung in klein- und mittelständischen Unternehmen mit dem Schwerpunkt auf eingebetteten Systemen.

Die Aktivitäten der Gruppe bei der Kooperation mit Universitäten in Südosteuropa wurden aufgrund des DAAD-Sonderprogramms "Akademischer Neuaufbau Südosteuropa" weiter ausgebaut. Als komplementäres Vorhaben zu diesen Kooperationsaktivitäten wurde ein EU-gefördertes TEMPUS-Projekt im ersten Projektjahr umgesetzt, in dem unter der Federführung der Gruppe ein gemeinsames Curriculum für einen Master-Studiengang zum Software Engineering an den südosteuropäischen Partner-Universitäten entsteht.

Lehre

Veranstaltungen im Grundstudium

- Vorlesungen „Praktische Informatik 1“ (K. BOTHE, WiSe 2004/2005, WiSe 2005/2006)
- Übungen „Praktische Informatik 1“ (K. SCHÜTZLER, WiSe 2005/2006)

Veranstaltungen im Hauptstudium

- Projektveranstaltungen „Software-Sanierung“ (K. BOTHE, WiSe 2004/2005, SoSe 2005, WiSe 2005/06)
- Projektveranstaltung „Requirements Engineering für Entwicklungs- und Produktionsprozesse“ (K. SCHÜTZLER, WiSe 2004/2005, SoSe 2005)
- Halbkurs/Kernmodul „Software Engineering 1“ (K. BOTHE, SoSe 2005)
- Übungen zu „Software Engineering 1“ (K. SCHÜTZLER, SoSe 2005)

Forschung

Projekt: Reverse Engineering für Steuerungssoftware in der Röntgenbeugung

Ansprechpartner: 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 2005 wurden folgende Ergebnisse erzielt:

- Arbeiten an verschiedenen Subsystemen zur Abgrenzung und zur Analyse sowie Erweiterung ihrer Funktionalität
- Ausbau der Testumgebung ATOS (automatisches System für den Regressionstest) durch Einbau eines Capture-and-Replay-Features
- Arbeiten zur Einbeziehung alternativer moderner Hardware

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 Rijeka (Kro), Universität Dubrovnik (Kro), Universität Timisoara (Rum), Universität Tirana (Alb), Universität Sarajevo (Bos-Herz), Universität 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
- Entwicklung von Kursmaterial für weitere Kurse: Projektmanagement, Objektorientierte Programmierung mit Java, Compilerbau
- Einbeziehung von eLearning-Material
- Entwicklung von Strategien und Tools für den Aufbau von mehrsprachigem Kursmaterial

Zu den Projektaktivitäten zählen u. a.

- Gastaufenthalte südosteuropäischer Wissenschaftler in Berlin
- Verbesserung der materiellen Bedingungen an den Partneruniversitäten (Hardware, Software, Literatur)
- Koordinationstreffen in Berlin, Mai 2005: Vorbereitung des Workshops
- Workshop: Software Engineering Education and Reverse Engineering, Baile Herculane, Rumänien, 28. August bis 4. September 2005

Auch in diesem Jahr stellte der Workshop unter Einbeziehung von Teilnehmern aus mittlerweile dreizehn Universitäten aus sieben Ländern den Höhepunkt der Projektaktivitäten dar, wobei es schwerpunktmäßig um die Weiterentwicklung des gemeinsamen Kurses im Bereich Software Engineering und die Übertragung der gewonnenen Erfahrungen auf neue Kooperationsfelder ging. Dem Workshop ging auch diesmal aufgrund der wachsenden Projektgröße und Teilnehmerzahl ein Koordinationstreffen der Kernmitglieder des Projekts in Berlin voraus.

Eine Reihe wissenschaftlicher Publikationen auf nationalen und internationalen Tagungen dokumentiert die Projektarbeit der vergangenen fünf Jahre.

Projekt: EU-TEMPUS-Projekt: "Joint M.Sc. Curriculum in Software Engineering"

Ansprechpartner: PROF. KLAUS BOTHE

Zusammenarbeit: Universität Novi Sad, De Montfort University Leicester, Universidad de Deusto Bilbao, Universität Skopje, Universität Belgrad, Universität Nis, Universität Timisoara, Universität Plovdiv

Forschungsförderung: Europäische Kommission

Ziel des Projektes ist die Schaffung eines gemeinsamen Curriculums für einen Masterstudiengang zum Software Engineering, der an den südosteuropäischen Partner-Universitäten angeboten werden soll.

Im Laufe des ersten Projektjahres wurden folgende Ergebnisse erzielt:

- Verabschiedung eines Curriculums: Zielstellung, Struktur, einzubeziehende Module
- Modulbeschreibungen für diese Module (Module Templates)
- Entwicklung eines Qualitätsmanagement-Verfahrens für den Studiengang
- Entwicklung erster Lehrmaterialien für ausgewählte Module

Zur Umsetzung der Projektziele wurde eine Reihe von Veranstaltungen organisiert:

- Workshop zur Verabschiedung des Curriculums, Novi Sad, Okt. 2005
- Koordinierungstreffen in Bilbao (April 2005) und Leicester (Juli 2005)

- Arbeitsaufenthalte von Wissenschaftlern aus Novi Sad, Belgrad und Skopje in Bilbao und Leicester

Projekt: Requirements Engineering für Entwicklungs- und Produktionsprozesse

Ansprechpartner: DIPL.-INF. KAY SCHÜTZLER

Zusammenarbeit: AEG netcontrol GmbH, Eichwalde bei Berlin

Beim industriellen Partner dieses Projekts werden spezielle Softwaresysteme im und um den Embedded-Software-Bereich entwickelt. Das Projekt verfolgt mehrere Ziele:

- Untersuchung geeigneter Methoden zur Anforderungsermittlung im Bereich eingebetteter Systeme,
- Verbesserung der Software-Entwicklungsprozesse im Rahmen klein- und mittelständischer Unternehmen,
- Heranführung der beteiligten Studenten an Projektmanagement.

Im Jahr 2005 standen Techniken der Anforderungsermittlung im Fokus des Projekts. Es entstand eine Reihe unterschiedlicher Anforderungsdokumente, die der Spezifikation einer individuellen Projektmanagement-Software beim Projektpartner dienen werden.

Veröffentlichungen

K. BOTHE, K. SCHÜTZLER, Z. BUDIMAC, K. ZDRAVKOVA: *Collaborative Development of a Joint Web-Based Software Engineering Course across Countries*. Proceedings of the 35th ASEE/IEEE Frontiers In Education Conference, FIE 2005, Indianapolis, USA, 2005.

K. SCHÜTZLER, S. HUSS: *Erfahrungen mit Refactorings in LabVIEW*. 7. Workshop Software Reengineering, Bad Honnef, Mai 2005, Softwaretechnik-Trends, Band 25, Heft 2.

K. BOTHE, S. JOACHIM: *Interactive Tool-based Production of Multilingual Teaching and Learning Materials*. Proceedings of the 5th IEEE International Conference on Advanced Learning Techniques, ICALT 2005, Kaohsiung, Taiwan, 2005.

S. JOACHIM, K. BOTHE: *Interaktive werkzeuggestützte Entwicklung von mehrsprachigen Lehr- und Lernmaterialien*. 3. Deutsche e-Learning Fachtagung Informatik, Rostock, September 2005, GI-Edition Lecture Notes in Informatics, P-66.

D. BOJIC, K. BOTHE: *REFMAP: Restructuring by Feature Mapping*. International IEEE Region 8 Conference on 'Computer as a tool' EUROCON, Belgrade, November 2005.

K. ZDRAVKOVA, K. BOTHE, S. JOACHIM: *Development of Multilingual Dictionaries in Software Engineering and their Implementation*. 2nd Balkan Conference in Informatics, Ohrid, November 2005.

Vorträge

K. BOTHE: *Generalized issues derived from our joint courses*. Baile Herculane, September 2005.

K. BOTHE: *Prerequisites of 1st semester students at HU*. Baile Herculane, September 2005.

K. BOTHE, Z. BUDIMAC: *An overview of activities of the Tempus project*. Baile Herculane, September 2005.

K. SCHÜTZLER: *An industrial software project as a SE module at HU*. Baile Herculane, September 2005.

K. SCHÜTZLER, K. BOTHE: *Reviewer's comments concerning our FIE 2005 paper and conclusions*. Baile Herculane, September 2005.

K. SCHÜTZLER, Z. PUTNIK: *Notes and solutions for team assignments*. Baile Herculane, September 2005.

U. SACKLOWSKI: *A style checker for Java and its application at HU*. Baile Herculane, September 2005.

S. JOACHIM: *Criteria for good e-learning materials*. Baile Herculane, September 2005.

S. JOACHIM: *E-Learning object repositories*. Baile Herculane, September 2005.

S. JOACHIM: *Adaptability of learning objects*. Baile Herculane, September 2005.

S. JOACHIM: *S-Bahn Tool: Experience and Planned Features*, Baile Herculane, September 2005.

Wissenschaftliche Kooperationen

- Institut für Physik der Humboldt-Universität zu Berlin,
- AEG netcontrol GmbH, Entwicklerzentrum
- Universität Novi Sad (Serbien-Montenegro)
- Universität Skopje (Mazedonien)
- Universität Plovdiv (Bulgarien)
- De Montfort University Leicester (Großbritannien)
- Universidad de Deusto Bilbao (Spanien)
- Universität Kragujevac (Serbien-Montenegro)
- Universität Nis (Serbien-Montenegro)
- Universität Belgrad (Serbien-Montenegro)
- Universität Zagreb (Kroatien)
- Universität Rijeka (Kroatien)
- Universität Timisoara (Rumänien)
- Universität Tirana (Albanien)
- Universität Sarajevo (Bosnien-Herzegowina)
- Universität Banja Luka (Bosnien-Herzegowina)
- DaimlerChrysler AG, Forschung und Technologie

Gastwissenschaftler

Prof. Mirjana Ivanovic, Novi Sad, Januar 2005
 Dipl.-Inf. Boro Jakimovski, Skopje, Januar 2005
 Prof. Zoran Budimac, Ph.D., Novi Sad, Februar 2005
 Dipl.-Inf. Zoran Putnik, Novi Sad, Februar 2005
 Dipl.-Inf. Natasa Ibrajter, Novi Sad, Februar 2005
 Prof. Katerina Zdravkova, Ph.D., Skopje, Mai 2005
 Prof. Ioan Jurca, Timisoara, Mai 2005
 Prof. Damir Kalpic, Zagreb, Mai 2005
 Prof. Stanimir Stoyanov, Plovdiv, Mai 2005
 Dragan Bojic, Ph.D., Belgrad, Mai 2005, Juli 2005
 Dipl.-Inf. Georgi Cholakov, Plovdiv, Juli 2005

Sonstige Aktivitäten

Prof. Dr. Klaus Bothe

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

Dissertation

FRANK LAMMERMANN: *Dimensionierung quantitativer Abbruchkriterien für den evolutionären Strukturtest mit Hilfe von Software-Maßen*. Dezember 2005.

Diplomarbeiten

SEBASTIAN HUSS: *Refactoring in LabVIEW*. Januar 2005.

NICOLAS ROCCA: *Entwicklung eines Stilbewertungsprogramms und dessen Einsatz in einem Java-Programmierpraktikum*. Juni 2005.

SAM JOACHIM: *Interaktive toolbasierte Entwicklung von mehrsprachigem Lehrmaterial*. Juni 2005.

ANDREAS HIRTH: *Automatische Generierung von Testskriptkommandos durch Capturing von Nutzereingaben in GUI-Programmen*. September 2005.

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: freytag@dbis.informatik.hu-berlin.de

Sekretariat

ULRIKE SCHOLZ

Tel.: (030) 2093 3008

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

Wissenschaftliche Mitarbeiterinnen und Mitarbeiter

DIPL.-INF. CHOKRI BEN NECIB

DIPL.-INF. OLIVER BERTHOLD

DR. RAINER ECKSTEIN

DIPL.-INF. RALF HEESE

DIPL.-INF. SVEN HERSCHEL

DIPL.-INF. FRANK HUBER

DR. STEPHAN HEYMANN

DIPL.-INF. HEIKO MÜLLER

PETER RIEGER

DIPL.-INF. (FH) KATJA THAM

Techniker

DR. THOMAS MORGENSTERN

DIPL.-ING. HEINZ WERNER

Tutorinnen und Tutoren

JAN ENGELSBERG

TIMO MIKA GLÄBER

ENNO GRÖPER

OLAF HARTIG

MIRKO HERBIG

MARTIN KOST

MAREN LENK

CLAAS REIM

BASTIAN VOIGT

HENNING ZAHN

Das Jahr 2005 war für die Lehr- und Forschungseinheit (LFE) Datenbanken und Informationssysteme (DBIS) ein Jahr des Umbruchs: da mehrere Projekte ausliefen, mussten die Arbeiten in einigen Forschungsbereichen reduziert werden, andere Projekte sind noch in der Startphase und werden erst im kommenden Jahr in die produktive wissenschaftliche

Phase eintreten. Dennoch konnte sich die LFE weiterhin im nationalen und internationalen Rahmen durch unterschiedliche Aktivitäten weiterhin einbringen und dadurch ihre Sichtbarkeit stärken.

In der Lehre und Forschung hat DBIS seine Arbeiten in den Bereichen XML & Semantic Web, Kerndatenbanken und „privacy“ konsequent ausgebaut. Hinzu kommen der neue Bereich „Daten in P2P-Netzen“ und „scientific workflows“. Industrielle Kooperationen spielen weiterhin eine wichtige Rolle für die LFE. Insbesondere die Kooperation mit der Firma Siemens hat sich mit der Entwicklung eines Anfrageoptimierers für das LDAP-Produkt DirX bestens entwickelt. Das Interesse für industrielle Verwertung der Ideen wird auch durch die Einreichung dreier Patente belegt.

Darüber hinaus geht dieser Bericht auf weitere Aktivitäten von Mitgliedern der LFE im Jahre 2005 ein. Zum einen sei auf die Ausrichtung eines Workshops und die aktive Mitarbeit bei der Ausrichtung einer Konferenz verwiesen, die die Sichtbarkeit der Gruppe gestärkt haben. Das Graduiertenkolleg „Verteilte Informationssysteme“ (GRK 316), gefördert durch die DFG, lief im Jahr 2005 aus; es hat sich neben dem BMBF-geförderten Projekt „InterVal“, an dem die LFE in zwei Teilprojekten beteiligt ist, immer noch als ein wichtiger Fokus der Zusammenarbeit der verschiedenen Datenbanklehrstühle in Berlin erwiesen. Dagegen haben die Arbeiten im Bereich der Genom-Datenbanken im Jahr 2005 trotz intensiver Bemühungen um neue Förderung deutlich abgenommen.

Lehre

Veranstaltungen im Grundstudium

- Einführung in die Programmiersprache C (PROF. J.C. FREYTAG, WS 2005/06)
- Praktikum zur Einführung in die Programmiersprache C (PROF. JOHANN CHRISTOPH FREYTAG, PETER RIEGER)
- Praktische Informatik III (Compilerbau) (PROF. J.C. FREYTAG, WS 2005/06)
- Praktikum zur Praktischen Informatik III (DR. R. ECKSTEIN, R. HEESE, WS 2005/06)

Kernveranstaltungen (Halbkurse) im Hauptstudium

- Grundlagen von Datenbanken (DBS I) (PROF. J.C. FREYTAG, WS 2004/05)
- Praktikum zu DBSI (C. BEN NECIB, H. MÜLLER WS 2004/05)
- Praktikum zu DBS II (FRANK HUBER, HEIKO MÜLLER SS 2005)
- Grundlagen von Datenbanken (DBS I) (PROF. J.C. FREYTAG, WS 2005/06)
- Praktikum zu DBS I (HEIKO MÜLLER, OLIVER BERTHOLD, FRANK HUBER, WS 2005/06)
- XML und Semantic Web (DR. RAINER ECKSTEIN, WS 2004/05)
- Praktikum zu XML und Semantic Web (DR. RAINER ECKSTEIN, WS 2004/05)

Spezialveranstaltungen im Hauptstudium

Seminare im Hauptstudium

- Syntaxanalyse von Genomdaten (Dr. S. HEYMANN, SS 2005)

Forschung

Projekt: Internetökonomie InterVa

Teilprojekt: Wissensnetze

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

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

Forschungsförderung: Berliner Forschungskonsortium "InterVal - Internet and Value Chains", im Rahmenprogramm „Internetökonomie“ des Bundesministeriums für Bildung und Forschung (BMBF)

Das Projekt untersucht die absehbare Entwicklung des zukünftigen Internets hin zu einem Netz des Wissens. Anhand zweier konkreter Szenarien werden das Potential und die wirtschaftlichen Auswirkungen eines globalen Semantic Webs sowie wertschöpfungsketteninterner Wissensnetze analysiert.

Das erste betrachtete Szenario beschreibt den Einsatz von Semantic-Web-Technologien bei der Arbeitsplatzvermittlung. Durch Annotation von Jobangeboten und -gesuchen mittels Ontologien wird eine bessere Vergleichbarkeit als bei einer schlagwortbasierten Suche erreicht. Aus diesem Projekt heraus ist in Zusammenarbeit mit der Freien Universität Berlin ein Prototyp entstanden, der die Zweckmäßigkeit des Einsatzes von Semantic-Web-Technologien demonstriert.

Das nächste Anwendungsszenario untersucht die Anwendbarkeit von Semantic-Web-Technologien in Produktinformationssystemen. Produktinformationssysteme integrieren auf semantischer Ebene Daten und Informationen über Produkte aus verschiedenen Quellen wie Hersteller oder Testorganisationen. Damit wird der Zugriff auf weiterführende Informationen zu Produkten vereinfacht, beispielsweise interessieren sich Allergiker für die Inhaltsstoffe von Lebensmitteln.

Zur Evaluierung des Szenarios wird derzeit gemeinsam mit der Freien Universität Berlin an einem Prototyp einer Arbeitsplatzvermittlung gearbeitet und es werden Gespräche mit Praxispartnern geführt. Gleichzeitig wird ein weiteres Szenario spezifiziert.

Zur Unterstützung der Entwicklung von Semantic-Web-Anwendungen wird ein Semantic-Web-Framework entwickelt. Man kann damit anhand eines Vorgehensmodells basierend auf UML-Anwendungsfall- und Aktivitätsdiagrammen von modellierten Aufgaben, die die Anwendung erfüllen sollen, die geeigneten Werkzeuge bestimmen, aus denen die Anwendung aufgebaut werden soll. In Zukunft soll es ermöglicht werden, die Werkzeuge zu komponieren. Dafür wird es notwendig sein, die Werkzeuge als Web-Services zu kapseln und weitere Informationen zu den Werkzeugen zu sammeln.

Im Rahmen des Projektes wurde in Zusammenarbeit mit der Freien Universität Berlin die Podiumsdiskussion „Standardisierung – Positiver Impuls oder Hemmnis im Markt“ auf den „Berliner XML Tagen 2005“ durchgeführt.

Teilprojekt: Datenschutzgerechte Gestaltung von Radio Frequency Identification (RFID) und Ubiquitous Computing (UC) Technologien

Ansprechpartner: DIPL.-INF. OLIVER BERTHOLD, PROF. JOHANN-CHRISTOPH FREYTAG, PH.D.

Zusammenarbeit: Humboldt-Universität zu Berlin - Institut für Wirtschaftsinformatik; Metro AG; Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein

Unter dem 1988 von Mark Weiser geprägten Begriff „Ubiquitous Computing“ wird die Allgegenwärtigkeit von Informationstechnik und Computerleistung verstanden, die in beliebige Alltagsgegenstände eindringt. Vorstellbar sind zahllose kleinste, miteinander über Funk kommunizierende Mikroprozessoren, die unsichtbar in Dinge eingebaut werden können. Mit Sensoren ausgestattet, können diese mikroskopisch kleinen Computer die Umwelt des Gegenstandes, in den sie implantiert sind, erfassen und diesen mit Informationsverarbeitungs- und Kommunikationsfähigkeiten ausstatten.

RFID (Radio Frequency Identification) stellt quasi den ersten Schritt in Richtung Ubiquitous Computing dar und bezeichnet Technologien, in denen Funksignale zur Objektidentifikation eingesetzt werden. Zu diesem Zweck werden die entsprechenden Objekte mit RFID-Chips (oder Tags) ausgestattet, welche per Funk, und somit auch durch optische Barrieren hindurch, ausgelesen werden können. Aufgrund fallender Chippreise wird die RFID-Technologie zunehmend für Anwendungen mit großen Stückzahlen nutzbar. Bedeutende Handelsunternehmen wie Metro und Wal-Mart haben bereits Zeitpläne für die Einführung von RFID zur Unterstützung ihrer logistischen Prozesse festgelegt. Vorhaben, Konsumgüter mit RFID-Tags zu versehen, werden jedoch von Datenschützern zum Teil heftig kritisiert, da jedes Objekt eine weltweit eindeutige Identifikationsnummer, den EPC (Electronic Product Code) erhalten soll.

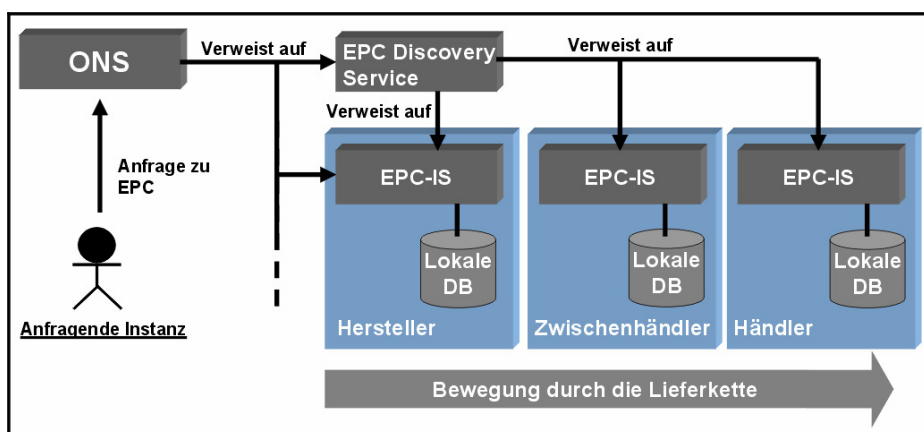


Abbildung 1: Datenverwaltung im EPC-Netzwerk

In diesem Projekt wird die UC-Technologie aus verschiedenen Blickwinkeln untersucht, wobei die Aspekte Privacy und Security einen besonderen Stellenwert einnehmen. Bisherige Untersuchungen und Ergebnisse beziehen sich auf die notwendige Gestaltung der Chip-Logik der RFID-Tags, um zu vermeiden, dass in Zukunft jeder Mensch durch mitgeführte Minisender „gläsern“ wird, d.h. überall unbemerkt identifiziert, wiedererkannt und bezüglich der mitgeführten Produkte wie z.B. der Kleidung kategorisiert werden kann. Des Weiteren untersuchen wir, wie sich eine datenschutzgerechte Struktur solcher ubiquitären

Systeme realisiert, wo der Schutz der Privatsphäre trotz allgegenwärtiger Sensoren gewahrt bleiben kann. Ein Beispiel ist die Verwaltung der EPC-Objektdaten (*ONS – Object Name Service*) von RFID-Tags, welche über das Internet erfolgen soll und so einen Datenschatten jedes einzelnen Objektes im Netz, das „Internet der Dinge“ erzeugt (*Abbildung 1*). Weitere Fragestellungen sind die Möglichkeiten der sicheren Kommunikation in sich spontan bildenden Netzen aus autonomen Sensoren.

Teilprojekt: Privater Datenzugriff

Ansprechpartner: DIPL.-INF. OLIVER BERTHOLD, PROF. JOHANN-CHRISTOPH FREYTAG, PH.D.

Zusammenarbeit: IBM Entwicklung GmbH Deutschland, Freie Universität Berlin

Mit zunehmender Nutzung des Webs 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.

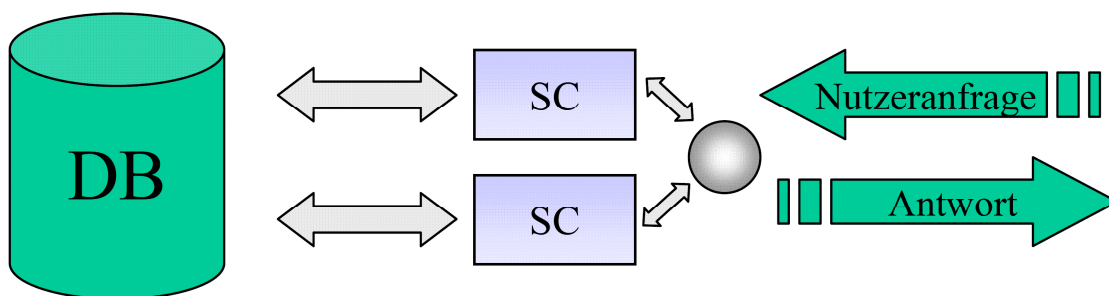


Abbildung 2: Aufbau eines PDA-Systems

In diesem Projekt werden ausgehend von existierenden Verfahren neue Ansätze des „Private Database Access“ (PDA) entwickelt, die sowohl Anforderungen aus dem Datenbankbereich berücksichtigen als auch die bisherigen Algorithmen vereinfachen. Einen relativ neuen Ansatz stellt die Nutzung von physisch geschützten (sicheren) Coprozessoren (SC) dar, welche über begrenzte interne Speicher- und Rechenkapazität verfügen (*Abbildung 2*). Da die online zu publizierenden Daten außerhalb dieser Einheiten gespeichert werden müssen, besteht die Herausforderung darin, Abhängigkeiten zwischen beobachtbaren Zugriffen zum externen Speicher und der eigentlichen Anfrage zu verringern bzw. soweit zu verschleiern, dass einen Beobachter nahezu keinen Informationsgewinn erzielt. Gleichzeitig ist der notwendige Aufwand zu minimieren. Da perfekte Unbeobachtbarkeit zwangsläufig zu einem extrem hohen Aufwand zur Bearbeitung einer Anfrage führt, größenordnungsmäßig O (*Datenbankgröße*), suchen wir nach einer vertretbaren Balance zwischen (probabilistischer) Unbeobachtbarkeit und Effizienz. *Abbildung 3* zeigt diese Abhängigkeiten für zwei von uns entwickelte Algorithmen. Die obere waagerechte Linie stellt ein beliebiges System dar, welches perfekte Unbeobachtbarkeit garantiert, allerdings zu dem oben skizzierten Aufwand.

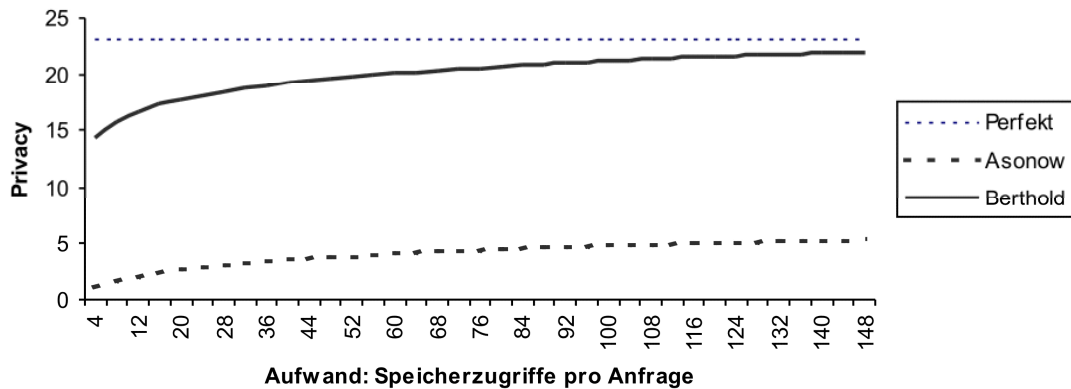


Abbildung 3: Privacy vs. Aufwand verschiedener Ansätze

Darüber hinaus werden Fragen des privaten Zugriffs im Kontext von E-Commerce untersucht. Beispielhafte Anwendungen der Technik wären unbeobachtbare, pseudonyme Community-, Publishing- oder auch MicroPayment-Systeme. Um unsere Hypothesen mit experimentellen Daten untermauern zu können, wurden Simulationstools und ein Prototyp entwickelt. Der Prototyp ist auf Basis einer von IBM entwickelten Spezialhardware und Spezialsoftware implementiert und demonstriert einige der oben beschriebenen Anwendungen.

Projekt: DirXQuE3 - DirX-Query Evaluation and Execution Engine

Ansprechpartner: DIPL.-INF. FRANK HUBER, TIMO MIKA GLÄSSER, PROF. JOHANN-CHRISTOPH FREYTAG, PH.D.

Zusammenarbeit: Siemens AG

Forschungsförderung: Siemens AG

DirX ist ein Produkt der Siemens AG, welches große Mengen von hierarchischen Daten mit Hilfe von Datenbanktechnologien verwalten kann. Im kommerziellen Einsatz waren komplexe Anfragentypen immer wieder mit einem schlechten Leistungsverhalten aufgefallen. Der Grund war, dass die benutzten Heuristiken in der Anfrageausführung in diesen Fällen falsch waren. Die Herausforderung des Projektes bestand somit darin, die bisherige auf wenigen Heuristiken basierende Anfrageausführung durch eine kostenbasierte Anfrageoptimierung und Anfrageausführung zu ersetzen.

Die Anfrageausführung kann wie in Relationalen Datenbankmanagementsystemen (RDBMSen) auf unterschiedlicher Weise erfolgen. Im Unterschied zu RDBMSen beruht die Anfragesprache in LDAP-Systemen auf der Booleschen Algebra. Es sind somit nur Ausdrücke erlaubt, die Mithilfe von Und, Oder und Nicht verknüpft werden. Ein weiterer Unterschied zu RDBMSen ist, dass DirX nur mit einer speziellen Art von Indizes arbeitet, so genannte Bit-Indizes. Diese Bit-Indizes können einfach mit Hilfe von Bit-Operatoren (AND,OR,NOT) bearbeitet und kombiniert werden. Der Zugriff auf ein Datenobjekt (Entry), eine sehr teure Operation, wird somit auf den spät möglichen Zeitpunkt verschoben. Da für die Bit-Operatoren nur je ein Ausführungsoperator existiert, spielen in diesem Fall speziell die Anordnung bzw. Reihenfolge der Operationen in der Ausführung eine große Rolle und sind maßgeblich für Antwortzeit einer Anfrage verantwortlich. Die Grundidee war, die Techniken aus der Relationalen Anfrageoptimierung für unsere Anfrageoptimierung zu benutzen und entsprechend anzupassen. Dabei musste aber auf die spe-

ziellen Bedürfnisse des hierarchischen Systems geachtet werden. Ein Problem bei der Adaption von Techniken aus der relationalen Welt waren die Statistiken. Zum Berechnen der Kosten einer möglichen Anfrageausführung und damit die Bewertung und Ordnung der einzelnen Pläne benötigt man Abschätzungen über viele Parameter, zum Beispiel die Größe des Resultates. Dazu werden Statistiken genutzt. Diese Statistiken müssen im Gegensatz zu RDBMSen nicht nur für die Gesamtheit aller Daten vorhanden sein, sondern auch für Teilmengen der Daten. Ein anderes Problem ist, die Einhaltung von Einschränkungen des Ergebnisses bezüglich Zeit und Größe. Ein Benutzer kann bestimmen, dass er nur die ersten N Entries haben möchte. Daraus resultiert, dass die Ausführung auch wirklich nach N Entries mit der weiteren Berechnung aufhören muss, um nicht unnötig Betriebsmittel und Zeit zu verschwenden.

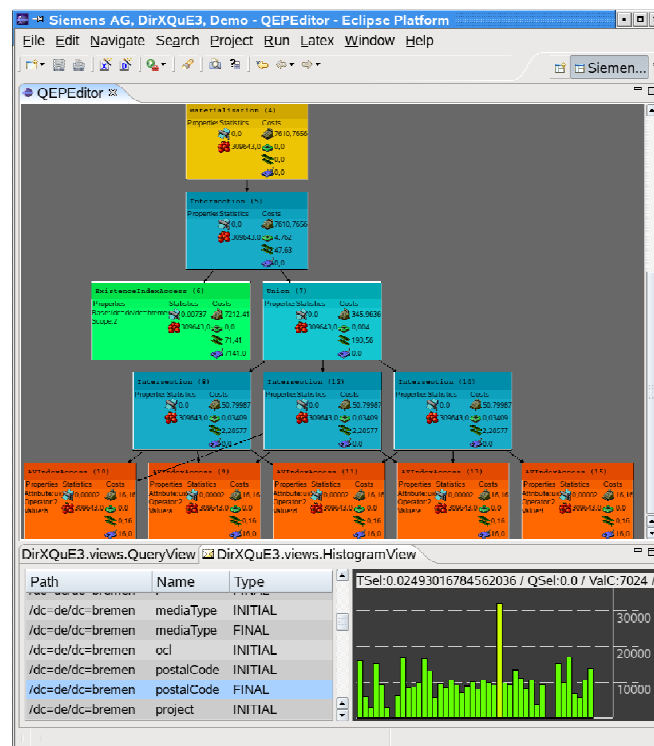


Abbildung 4 : Visualisierung eines Ausführungsplans durch die Visualisierungskomponente

Im Rahmen des Projektes entstanden so verschiedene Module, wie z.B. ein Modul zur Anfrageumschreibung, Kostenberechnung, Anfrageausführung und Visualisierung (siehe Abbildung 4). Dabei wurde speziell auf die Modularisierung und somit auf die Erweiterbarkeit beziehungsweise auf den Austausch von Komponenten geachtet, um die Weiterentwicklung und Integration in das Produkt DirX zu ermöglichen.

Projekt: Datenbanktechnologien für das „Semantic Web“ und XML

Ansprechpartner: DR. RAINER ECKSTEIN

Das Web ist eine riesige vernetzte Sammlung von Seiten, deren Inhalte von menschlichen Nutzern interpretiert werden müssen. Das Semantic Web reichert das bestehende WWW 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 und für XML erforscht.

Teilprojekt: Performante RDF-Managementsysteme

Ansprechpartner: DIPL.-INF. RALF HEESE

Die Grundlage für den Datenaustausch im Semantic Web bilden Modellierungssprachen, die auf dem Datenmodell des Resource Description Framework (RDF) aufsetzen. Die Anfragesprache SPARQL wird derzeit als Anfragesprache für den Zugriff auf RDF-Daten durch das World Wide Web Consortium (W3C) standardisiert. Ausgehend von der Idee, RDF-Datentypen in relationalen Datenbanksystemen zur Verfügung zu stellen, entwickeln wir zur Zeit die Grundlagen für die Einbettung von SPARQL in die Anfragebearbeitung von relationalen Datenbanksystemen: eine Algebra und ein Anfragegraphmodell (Query Graph Model) für SPARQL. Das Ergebnis ist ein Datenbanksystem mit einem hybriden Anfrageoptimierer (Abbildung 5, rechts), der SQL/SPARQL-Anfrage ganzheitlich betrachtet und optimiert.

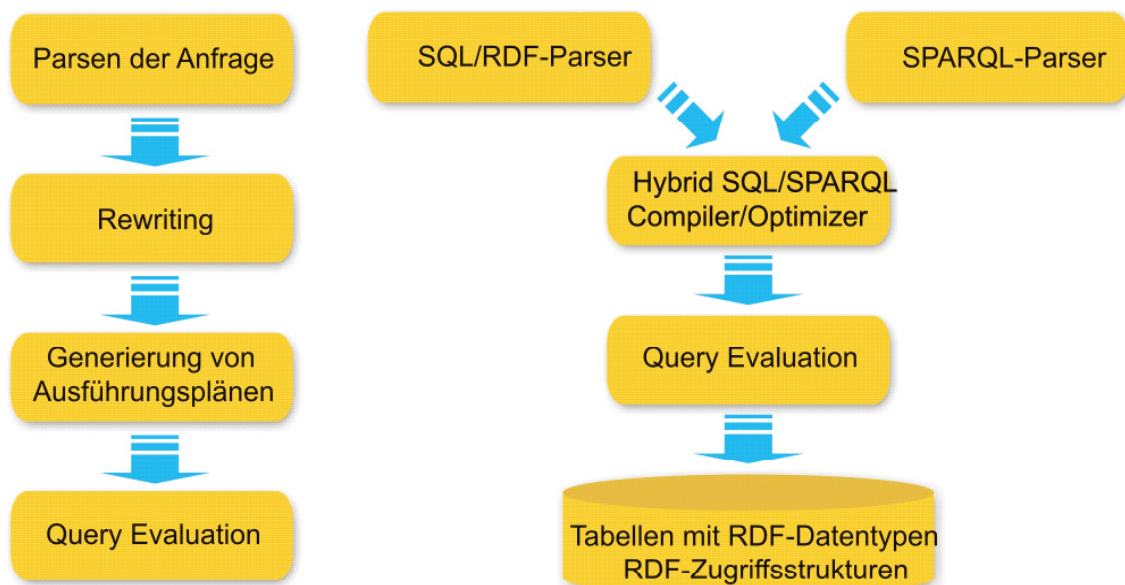


Abbildung 5: Hybride Anfragebearbeitung von SQL und SPARQL

Die Anfragebearbeitung ist in vier Phasen gegliedert: Parsen der Anfrage, Rewriting, Generierung von Anfrageplänen und die Ausführung des „besten“ Anfrageplans (*Abbildung 5, links*). Während des Parsens wird eine interne Darstellung der SPARQL-Anfrage in Form eines Query Graph Models generiert. In der nächsten Phase der Anfragebearbeitung werden semantisch äquivalente Anfragepläne gesucht, die eine bessere Performanz bei der Ausführung versprechen. Das Umschreiben der Anfragen erfolgt zum einen mit Hilfe der in der Algebra definierten Äquivalenzregeln und der Transformation des Query Graph Models.

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 wurden.

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. Dieses Werkzeug wurde um die Erzeugung von XML-Schema erweitert. Darüber hinaus wird an einem Reverse Engineering Ansatz gearbeitet, so dass es möglich sein wird, aus XML-Schema bzw. DTD-Spezifikationen UML-Klassendiagramme abzuleiten.

Projekt: Aufdecken von Regelmäßigkeiten in widersprüchlichen Daten

Ansprechpartner: DIPL.-INF. HEIKO MÜLLER

Die Auflösung von Widersprüchen in überlappenden Datenquellen ist wichtiger Bestandteil des Datenintegrationsprozesses bzw. des Teilschrittes der Datenbereinigung. Grundlage einer gezielten und zuverlässigen Konfliktauflösung ist die Kenntnis möglicher Konfliktursachen. Oftmals liegen den Konflikten Regelmäßigkeiten zugrunde, z.B., die Verwendung unterschiedlicher Maßeinheiten oder Bezeichnungen. Im Rahmen des Forschungsprojektes wurde zunächst ein Algorithmus entworfen, der solche Regelmäßigkeiten in widersprüchlichen Daten aufzeigen kann. Der Algorithmus basiert auf existierenden Ansätze des *Association Rule Mining* und sucht nach Regeln der Form „WENN *Bedingung* DANN *Konflikt*“. Diese Regeln können von einem Domänenexperten zur Aufdeckung möglicher Konfliktursachen herangezogen werden.

Ein wesentlicher Nachteil des verwendeten Ansatzes ist die Vielzahl an gefundenen Regeln mit teilweise sehr begrenzter Aussagekraft. Aus diesem Grund wurde eine Klassifikation eingeführt, welche eine Beschränkung des Algorithmus auf die Suche nach Regeln einer bestimmten Klasse erlaubt. Diese Klassifikation beruht im Wesentlichen auf den konkreten Werten, welche miteinander in Konflikt stehen und auf einer Einschränkung zulässiger Bedingungen, z.B. Bedingungen, die nur eine der Datenbanken betrachten. Eine spezielle Klasse von Regeln bilden solche, die sich in UPDATE-Operationen der Form „UPDATE *Datenbank* SET *Attribut* = *Wert* WHERE *Bedingung*“ zur Beschreibung der Unterschiede zwi-

schen zwei Datenbanken übersetzen lassen. So lässt sich beispielsweise die unterschiedliche Berufsbezeichnung weiblicher Informatikerinnen in den Tabellen aus *Abbildung 6* durch eine der drei UPDATE-Operationen erklären:

UPDATE Datenbank 1 SET Beruf = ‚Informatiker‘ *oder*

UPDATE Datenbank 1 SET Beruf = ‚Informatiker‘ WHERE Geschlecht = ‚Weiblich‘ *oder*

UPDATE Datenbank 2 SET Beruf = ‚Informatikerin‘ WHERE Geschlecht = ‚Weiblich‘

DATENBANK 1

Pers.-Nr.	Name	Geschlecht	Beruf
100	Claudia	Weiblich	Informatikerin
200	Dirk	Männlich	Informatiker
300	Sylvia	Weiblich	Informatikerin
400	Peter	Männlich	Informatiker

DATENBANK 2

Pers.-Nr.	Name	Geschlecht	Beruf
100	Claudia	Weiblich	Informatiker
200	Dirk	Männlich	Informatiker
300	Sylvia	Weiblich	Informatiker
400	Peter	Männlich	Informatiker

Abbildung 6: Konflikte in überlappenden Datenbanken

In unseren aktuellen Arbeiten suchen wir nach minimalen Sequenzen solcher UPDATE-Operationen, die eine gegebene Datenbank in eine andere gegebene Datenbank transformieren. Minimalität bezieht sich dabei auf die Anzahl der Operationen in diesen Sequenzen. Unter der Annahme, dass die Entstehung der Konflikte einer Systematik unterliegt, wird jede minimale Sequenz als die einfachste und wahrscheinlichste Erklärung der Unterschiede zwischen den Datenbanken angesehen. Die Herausforderung bei dieser Aufgabe besteht in der Größe des Suchraums und des Versagens einfacher Heuristiken zum Finden einer exakten Lösung. Aus diesem Grund arbeiten einfache Verfahren momentan nur auf sehr kleinen Datenbanken. Die von uns entwickelten Heuristiken zum Auffinden suboptimaler Lösungen sind auch für größere Datenbanken anwendbar und zeigen bereits eine hohe Genauigkeiten in ihren Ergebnissen. Die Entwicklung effizienter Algorithmen zum Finden exakter Lösungen bildet eine der zukünftigen Forschungsaufgaben.

Projekt: Ontologiebasierte Anfragenverarbeitung

Ansprechpartner: DIPL.-INF. CHOKRI BEN NECIB

Die leistungsfähige Anfragebearbeitung 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.

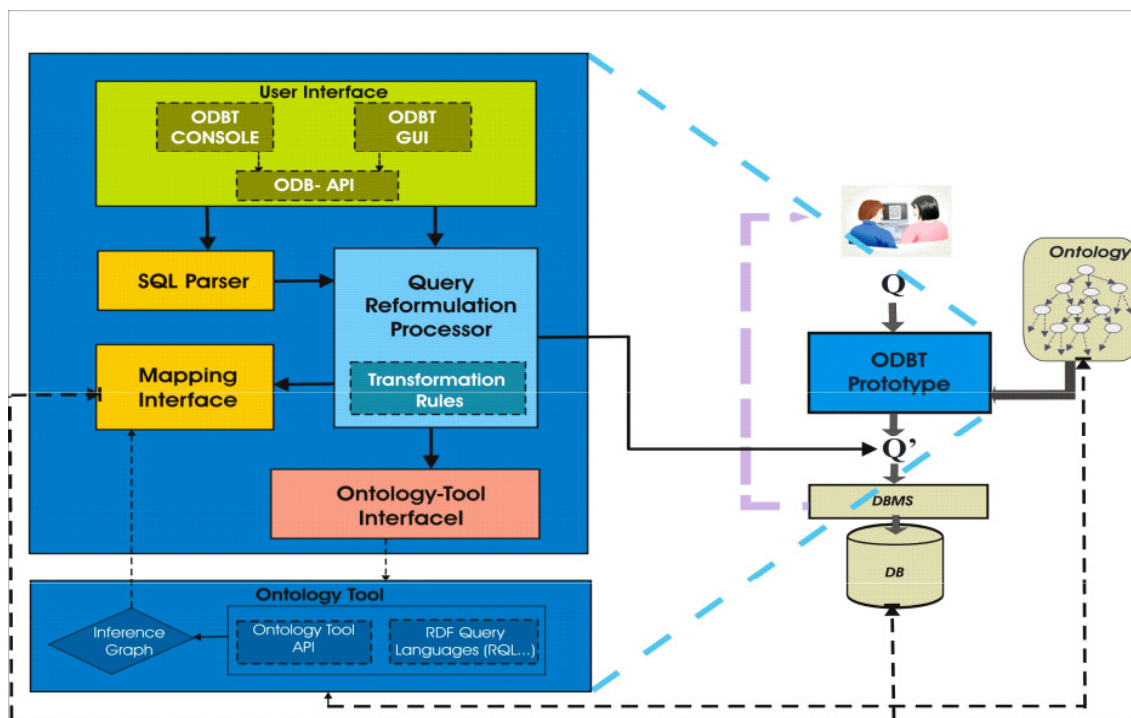


Abbildung 7: Nutzung von Ontologien für die Anfragebearbeitung

Unser Ziel ist es, einen ontologie-basierten Ansatz für die semantische Anfragebearbeitung zu entwickeln (Abbildung 7). 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. Des Weiteren sollen die Abbildungen und die Einschränkungen zwischen der Ontologie (Konzepte, Beziehungen) und der Datenbank (Schema, Instanzen) spezifiziert werden. Aufbauend auf diesem Ansatz wurde ein Prototyp entwickelt, in dem die vorgeschlagene Verfahren und Konzepte größtenteils umgesetzt sind. Für die Ontologie Darstellung werden die standardmäßige Sprachen RDF and OWL untersucht und verwendet. Für die Speicherung, den Zugriff und die Inferenzbildung von Ontologien werden in dem Prototyp Werkzeuge eingebunden, die einen Großteil der geforderten Funktionalität für die Verwendung von Ontologien umsetzen. Der Prototyp integriert zwei verschiedene Ontologie-Werkzeuge: *Sesame* und *Jena*.

Um den Prototyp nicht nur auf den speziell für Testzwecke angelegten Datenbanken und Ontologien zu testen, wurden jeweils eine real existierende Ontologie und Datenbank ausgewählt. Es wurde die *WordNet*-Ontologie verwendet, welche anhand der *GNIS*-Datenbank getestet wird. *WordNet* ist eine linguistische Ontologie, die die Bedeutung englischer Wörter beschreibt. Die *GNIS*-Datenbank beinhaltet Daten über die geographischen und kulturellen Merkmale in den Vereinigten Staaten und deren Territorien. Erstellt wurde sie vom „U.S. Geological Survey“ und „U.S. Board on Geographic Names“. Diese Datenbank wurde gewählt, da die in ihr enthaltenen Namen der geographischen Merkmale eine relativ große Übereinstimmung mit den in *WordNet* enthaltenen Konzepten besitzen.

Projekt: Benchmark Storage Area Network (SAN) vs. Lokaler Speicher

Ansprechpartner: PETER RIEGER, THOMAS MORGENSTERN

Im Zuge einer Investitionsentscheidung zum Ausbau der verfügbaren Sekundärspeicherkapazitäten am Lehrstuhl trat die Frage auf, inwieweit moderne Datenbankmanagementsysteme effizient mit zentral bereitgestelltem Speicher umgehen können.

Das physische Design einer Datenbankanwendung, d.h. die direkte Zuordnung von Primär- und Sekundärspeicher zu Datenbankobjekten wie Tabellen und Indexen ist entscheidend für die Effizienz der Anfragebearbeitung. Die Verwendung von zentral über ein SAN bereitgestellten Sekundärspeicher bietet zwar den Vorteil, die Verwaltung des Massenspeichers vom Design der einzelnen Datenbankanwendungen zu entkoppeln und damit dem Datenbankdesigner die Möglichkeit, sich auf den logischen Entwurf der Datenbank zu konzentrieren. Die Frage ist jedoch, inwieweit eine zentralisierte technische Lösung der Speicherbereitstellung die Effizienz von Datenbankanwendungen beeinflusst.

Um diese Frage zu beantworten, wurde ein standardisierter Benchmark des „Transaction Processing Performance Councils“ (TPC) durchgeführt. Das TPC stellt Benchmarks bereit, die jeweils Anforderungsprofile unterschiedlicher Aufgabenklassen von Datenbankanwendungen widerspiegeln. Da die Anwendungsschwerpunkte des Lehrstuhls, wie Bioinformatik und das Semantic Web, durch die Aufgabenklasse „Entscheidungsunterstützung und Ad-hoc Anfragen“ am besten charakterisiert werden, wurde der Standard-benchmark „TPC-H“ verwendet.

Es zeigte sich, dass insbesondere bei Ausführung komplexer Anfragen, die Verwendung von lokalem Speicher deutliche Effizienzvorteile bietet (ca. 1 Größenordnung). Da die zentrale Bereitstellung von Speicher jedoch wirtschaftlich und administrativ große Vorteile verspricht, soll der Benchmark im kommenden Jahr noch einmal gemeinsam mit den Kollegen vom CMS wiederholt werden, um zu untersuchen, ob es Möglichkeiten gibt, das SAN so zu konfigurieren, dass die Performanceunterschiede zumindest deutlich verringert werden.

Projekt: Verbindung von DBMS und Suchmaschinen

Ansprechpartner: PETER RIEGER

Eine Vielzahl der Informationen, die heute über das Internet bereitstehen, wird in Form von Texten und semistrukturierten Dokumenten bereitgestellt. Suchmaschinen wie „Google“ oder „Yahoo“ bilden zunehmend den Zugang zu fast allen über das Internet verfügbare Daten. Dabei werden von den Anwendern die einfache Anfragesprache und die empfundene Relevanz der Ergebnisse bei kurzen Antwortzeiten als Wesentliche Vorteile gesehen.

Die semantischen Klassifikation und Aufbereitung der Ergebnismengen, die angesichts massiver Bemühungen von „Spammern“, zweifelhafte Inhalte in die oberen Ränge der Ergebnismengen zu platzieren, gewinnt zunehmend auch wirtschaftlich an Bedeutung. Darüber hinaus weisen Suchmaschinen aber einen wesentlichen Nachteil gegenüber klassischen Datenbankanfragen auf: strukturelle und inhaltliche Beziehungen zwischen Dokumenten können nicht gezielt zur Qualifikation der Ergebnismenge verwendet werden.

Zu diesem Zweck soll die am Lehrstuhl entwickelte Integrationsplattform „Gene-EYE“ um eine Komponente zur semantisch gestützten Volltextsuche erweitert werden. Dieses Projekt befindet sich derzeit in der Definitionsphase.

Ziel ist es,

- ein Beschreibungssystem zu entwickeln, das es erlaubt, inhaltliche und strukturelle Bezüge zwischen Klassen von Dokumenten flexibel zu beschreiben,
- eine Anfrageumgebung zu entwickeln, die Methoden der Datenbankabfrage mit Methoden der Volltextsuche effizient kombiniert und
- eine Anfragesprache zu definieren, die dem Anwender ermöglicht, mit möglichst geringem Mehraufwand seine Anfragen zu formulieren.

Als Anwendungsgebiet soll weiterhin die Molekularbiologie dienen.

Projekt: Unvollständige Workflow-Konzepte und Methoden zur Weiterführung

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

Forschungsförderung: Promotionsstipendium gemäß Nachwuchsförderungsgesetz des Landes Berlin (NaFöG)

Traditionelle Workflow Management Systeme, ursprünglich für klar strukturierte und vorab definierbare Geschäftsprozesse konzipiert, öffnen sich zunehmend neuen Anwendungsgebieten. Stellvertretend sei das Gebiet der Genomforschung genannt, welches bestrebt ist, Gene und Genprodukte lebender Organismen zu identifizieren und zu analysieren, um so ihre Funktionen im Netzwerk der Lebensprozesse aufzuklären.

Die stete Entwicklung immer neuer Methoden zur Aufdeckung von Zusammenhängen erlaubt es, immer komplexere wissenschaftliche Forschungsfragen zu untersuchen und bedingt gleichzeitig die regelmäßige Aktualisierung vorhandener Prozessabläufe. Infolge der unvollständigen, verschwommenen Kenntnis über die Gesetzmäßigkeiten und Abläufe innerhalb lebender Organismen ist Entwicklung (Spezifikation) und Wartung (Modifikation) dieser Prozessabläufe oftmals explorativ. Erst die wiederholten Ausführung und Evaluierung verschiedener Analysemethoden anhand vorangegangener Zwischenergebnisse führt zu qualitativ hochwertigen Resultaten.

Flexibilität hinsichtlich zunächst unvollständiger bzw. erweiterbarer Workflow-Modelle ist aus diesem Grund zu einer der wichtigsten Voraussetzungen für den Einsatz von Workflow-Managementsystemen in Bereichen wie der Genomforschung geworden.

Ziel dieses Projektes ist es, das traditionelle, starre Modell einer getrennten Workflow-Spezifikation und Ausführung aufzubrechen und so eine schrittweise semiautomatische Prozessausführung zu erlauben. Dies bedeutet, dass ein Workflow-Modell explorativ, d.h. parallel zu seiner initial ausgeführten Instanz erst entsteht.

Zusätzlich wird die bisher vorab definierte Workflow-Spezifikation durch eine deklarative Beschreibung ergänzt, dies ermöglicht es einem späteren Anwender, unabhängig von den zugrundeliegenden Methoden und Algorithmen zu beschreiben, was durch den Prozess erreicht werden soll und nicht, wie dies realisiert wird. So wird es möglich, eine deklarative Prozessbeschreibung je nach Verfügbarkeit bestehender Methoden oder die Entwicklung neuer Methoden auf mehrere ausführbare Workflow-Modelle abzubilden.

Projekt: The Query Web

Ansprechpartner: DIPL.-INF. SVEN HERSHEL, DIPL.-INF. RALF HEESE

Peer-To-Peer-Netzwerke (P2P) haben nach den beachtlichen kommerziellen Erfolgen der Tauschnetzwerke auch die Aufmerksamkeit der akademischen Forschung auf sich gezogen. So werden von Forschergruppen weltweit vorhandene Algorithmen evaluiert und

neue, effizientere Zugriffsstrukturen entwickelt. Dabei stellen sich insbesondere drei Forschungsfragen: (1) effiziente Lokalisierung von Ressourcen (**Search**), (2) Optimierung von Anfragen unter Unsicherheit (**Query**) und (3) Umgang mit hoher Dynamik in der Netzwerkstruktur (**Node Joins/Node Leaves**). Herkömmliche P2P-Infrastrukturen können entweder sehr gut mit hoher Dynamik umgehen (so genannte unstrukturierte Netze) oder eine effiziente Lokalisierung von Ressourcen sicherstellen (so genannte strukturierte Netze).

Daraus abgeleitet verfolgt das Query-Web-Projekt die folgenden Ziele:

- Entwicklung und Bewertung eines hybriden Ansatzes, der die Vorteile von strukturierten und unstrukturierten Netzen vereint
- zuverlässige, effiziente Lokalisierung von Ressourcen bei hoher Dynamik des Netzes
- Anwendung dieser Architektur im Kontext des Semantic Web
- Rollenzuweisung in P2P-Netzen: Ausnutzen der speziellen Fähigkeiten (Rechenleistung, Speicher, Bandbreite) einzelner Peers zur Steigerung der Gesamtperformance des Netzes

Projekt: Indexierung unstrukturierter P2P-Netzwerke

Ansprechpartner: DIPL.-INF. SVEN HERSCHEL

Peer-To-Peer-Netzwerke haben nicht zuletzt aufgrund der erheblichen Erfolge von populären kommerziellen P2P-Netzwerken die Aufmerksamkeit der akademischen Forschung geweckt. In einem P2P-Netzwerk sind alle Knoten gleichberechtigt, nehmen also im Gegensatz zum klassischen „Client-Server“-Modell gleichzeitig Serveraufgaben (Datenauslieferung) wie Clientaufgaben (Datenanfrage) wahr.

Vorteile von unstrukturierten P2P-Netzwerken liegen in Ihrer Stabilität und Zuverlässigkeit trotz sehr hoher Dynamik, da es keinen einzelnen zentralen Angriffspunkt gibt, der das ganze Netzwerk stören würde. Diese positive Eigenschaft zieht jedoch einige negative Eigenschaften nach sich; insbesondere ist es durch das Fehlen einer zentralen Kontrollinstanz nur sehr schwer möglich, Garantien über kürzeste Wege oder Antwortzeiten, ja nicht einmal über die Existenz von Objekten innerhalb des Netzwerkes zu geben.

Das Projekt „Indexing unstructured P2P Networks“ versucht, Indexstrukturen innerhalb von unstrukturierten P2P-Netzwerken zu untersuchen. Dabei wird untersucht, welche Funktionen ein Index in diesem Rahmen überhaupt übernehmen kann, inwieweit er sich von Indizes in Datenbanksystemen unterscheidet und welche Garantien ein solcher Index liefern kann. Dies erweiternd versucht das Projekt herauszufinden, inwieweit es möglich ist, Semantic-Web-Daten in einem P2P-Netzwerk effizient zu indizieren. Das Projekt läuft voraussichtlich bis Oktober 2006.

Projekt: Alternative Spleißformen - Von der Simulation zum Nachweis

Ansprechpartner: DR. STEPHAN HEYMANN, TIMO MIKA GLÄBER

Zusammenarbeit: Forschungszentrum Rossendorf, Biomedizinischer Forschungscampus Berlin-Buch, Institut für Chemie - Humboldt-Universität zu Berlin

Forschungsförderung: Deutsche Forschungsgemeinschaft

Zahlreiche menschliche Gene instruieren unsere Körperzellen zur Bildung von mehr als nur einem Protein. Das geschieht während der Reifung der mRNA in einem Prozess, den die Molekularbiologen *Spleißen* nennen. Dabei werden distante Abschnitte aus den Kopien

der Ableseeinheit Gen linear aneinandergereiht. Der genspezifische Satz solcher Abschnitte, die als Exons bezeichnet werden, muss sich nicht immer komplett in allen mRNA-Molekülen wiederfinden. Auslassungen ganzer Exons oder Verschiebung ihrer Grenzen führen daher zu *Alternativen Spleißformen*. Die entstehenden Spleißformen bilden die 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 neue funktionelle Motive entstehen, d.h. Abschnitte des Proteins, denen eine biochemische Funktion zugewiesen werden kann. In einer vorangegangenen Projektetappe wurden mit Hilfe einer Modifikation des Viterbi-Algorithmus (dynamische Programmierung) die funktionell bedeutsamen Alternativen Spleißformen für 3583 Gene auf 17 menschlichen Chromosomen vorberechnet [M. HILLER, R. BACKOFEN, S. HEYMAN, A. BUSCH, T. M. GLÄBER, J.-C. FREYTAG: *Efficient Prediction of Alternative Spliceforms Using Protein Domain Homology, In silico Biology 2004/04/0017*].

Ziel der Projektetappe 2005 war der *Nachweis der Existenz* der zuvor simulierten Alternativen Spleißformen.

Illustration an einem Einzelbeispiel: Die Expression unterschiedlicher Spleißformen ein und desselben Rezeptors in verschiedenen Geweben hat unter anderem Konsequenzen für die Erkennung und Bindung des dazugehörigen Liganden oder Wirkstoffs. Außerhalb der Zelle angeordnete Rezeptorabschnitte beeinflussen sich gegenseitig. Die Kenntnis von Verlust oder Veränderung des Ligandenbindungsverhaltens ist in der Medizin von Bedeutung für Diagnostik und Therapie.

Abbildung 8 illustriert, wie sensitiv ein Rezeptor auf eine vergleichsweise geringe Veränderung infolge Alternativen Spleißens (Entfernung von vier der ursprünglich 418 Aminosäuren) reagieren kann: Die für die Funktion eines distanten Fragments (fett umrandete Aminosäuren) wichtige Disulfidbrücke zwischen zwei Cysteinen-SS- kann nach Wegfall der betreffenden vier Aminosäuren nicht mehr ausgebildet werden. In der Konsequenz verliert der Rezeptor seine Bindungsfähigkeit für den natürlichen Liganden.

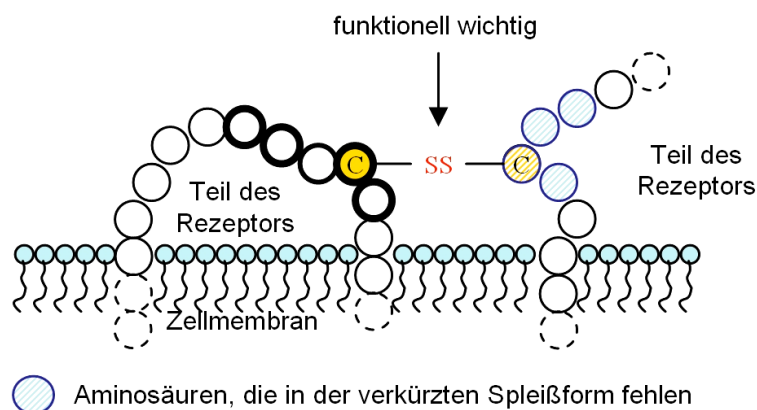


Abbildung 8: Wie die Bindungsspezifität eines Rezeptors durch Alternatives Spleißen variieren kann

Der experimentelle Nachweis von sequenziell-funktionellen Unterschieden zwischen Spleißformen erfolgte bisher im molekularbiologischen Labor *punktuell*. Nach jedem Paar Normal-Spleißform – Alternative Spleißform musste mit adäquaten Methoden extra gefahndet werden. Deshalb lag der gegenwärtigen Projektetappe die Aufgabe zugrunde, mengenwertige Betrachtungen über eine Vielzahl von Alternativen Spleißformen anzustellen, um ihre Existenz sowohl nassbiochemisch als auch computergestützt zu verifizieren oder zu falsifizieren. Entsprechend gliederte sich das Projekt in einen (i) bioinformatisch-maschinenexperimentellen Teil und in einen (ii) laborexperimentellen Teil auf. *Abbildung 9* stellt dar, wie sich die Berechnungen aufeinander aufbauend in die integrierte relationale Datenhaltung von Genom- und Proteomdaten bei DBIS einfügen.

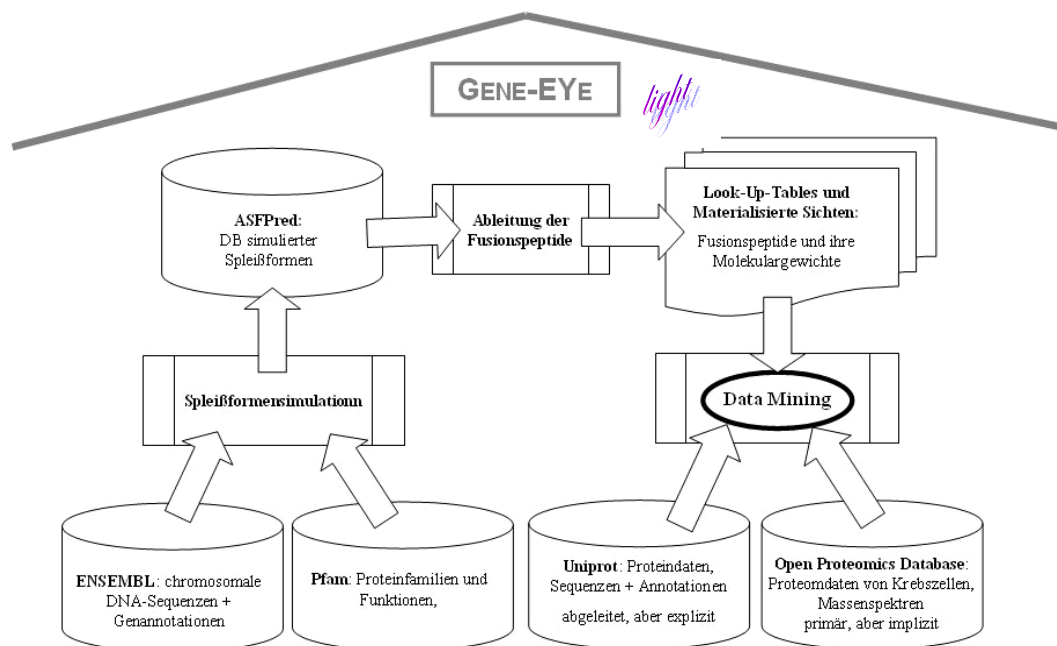


Abbildung 9: Integrierte Data Warehouse-Lösung für die Simulation von Alternativen Spleißformen anhand von Genomdaten und deren Verifikation durch Abgleich mit massenspektrometrisch gewonnenen Proteomdaten

Einschub für potenzielle Nachnutzer aus der molekularen Medizin: Aus Teil (ii) ist mit dem „Stoppⁿ-Primer-Verfahren“ eine generische, rechtsmängelfreie Methode entstanden, die Modifikationen in 2'-Position des Zuckerrests von Nukleinsäurebausteinen nutzt, um Alternative Spleißformen auch dann noch selektiv aus einer biologischen Probe zu isolieren, wenn sie im extremen Unterschuss, etwa in einem molaren Verhältnis kleiner als 1:1000 zur Normalform, darin enthalten ist [S. HEYMANN, H. JOOS, P. BENDZKO, B. KRAFFERT, R. BERGMANN, R. M. LEISER, *Verfahren zur Steuerung der abschnittweisen enzymatischen Nukleinsäurevervielfältigung über inkomplette Komplementärstränge*, Patent eingereicht am 07. 10. 2005].

Der Informatik-Teil (i) ist Gegenstand der folgenden Darlegungen. Er ist in einer Serie von Studienarbeiten am Lehrstuhl DBIS ausführlich dokumentiert und der Öffentlichkeit zugänglich. Die bioinformatische Verifikationsstrategie basiert auf der Überlegung, dass sich Alternative Spleißformen am auffälligsten an den Fusionsstellen distanter Exons unterscheiden. Daher liegt es nahe, nach den Sequenzen von Peptiden zu suchen, die an den Übergangsstellen zusammengespleißter Exons leserastergerecht entstehen. Schneidet man

diese Peptide nach den basischen Aminosäuren Arginin oder Lysin heraus, die sie flankieren, und berechnet ihre Molekülmasse, so vollzieht man am Rechner einen Prozess nach, der *in vitro* an realen Proteinen mit Hilfe des proteinspaltenden Enzyms Trypsin vorgenommen wird, bevor man Proteine massenspektrometrisch vermessen und ihre Aminosäuresequenz aus den Spektren der so genannten MS/MS-Tandemtechnik rekonstruieren kann. Das geschieht in der Absicht, Experimentaldaten aus der Massenspektrometrie ganzer Proteome (bspw. der Gesamtheit aller Proteine eines Zelltyps) nach den maschinengenerierten Spaltprodukten unserer simulierten Spleißformen zu durchforsten. Dabei sind die experimentellen Messfehler vom Gerätetyp abhängige Messgenauigkeit zu beachten. Das wird durch parametrisierte Intervallabfragen in den SQL-Statements realisiert. Chemische Modifikationen an den Aminosäuren, wie sie in lebenden Zellen sehr häufig anzutreffen sind, konnten in der vorliegenden Fassung des Datensatzes noch keine Berücksichtigung finden, sie sind aber in der Rechnerplattform bereits als Look-Up-Tables realisiert bzw. können abfrageintern durch Aufruf entsprechender nutzerdefinierter Funktionen als additive Terme auf die Fragmentmassen aufgeschlagen werden. *Abbildung 10* zeigt am Beispiel des Peptids ETASDFCSVALSK an der Fusionsstelle zweier distanter Exons aus einem Strukturgen, das Zell-Zell-Kontakte vermittelt, wie sich über die bloße Massengleichheit hinaus aus den Kollisions-fragmentserien a, b, c und x, y, z auch die gesuchte Aminosäuresequenz der fraglichen Alternativen Spleißform verifizieren lässt.

Fragment	Masse	m/z	Fehler (abs.) in u	Fehler (rel) in ppm
TASDFCSVALSK - z	1212.5710	1212.5560	0.0154	11.3627
ASDFCSVALSK - x	1154.5292	1154.5300	-0.0005	-0.4202
SDFCSVALSK - x	1083.4919	1083.4940	-0.0017	-1.2658
SDFCSVALSK - y	1055.4970	1055.4560	0.0412	30.3771
FCSVALSK - x	881.4326	881.4518	-0.0191	-14.1153
SVALSK - z	588.3487	588.3474	0.0013	0.9662
VALSK - z	501.3165	501.3203	-0.0037	-2.7889
ETASDFCSV - c	956.3920	956.3480	0.0440	32.4574
ETASDFCSVA - b	1012.4182	1012.4140	0.0044	3.2806
ETASDFCSVAL - a	1097.5075	1097.4660	0.0412	30.4257
ETASDFCSVALS - b	1212.5346	1212.5560	-0.0209	-15.4582
ETASDFCSVALS - c	1227.5457	1227.5080	0.0377	27.7912

Abbildung 10: Abfrageergebnis, das die Existenz einer vorhergesagten Alternativen Spleißform im massenspektrometrisch ermittelten Proteomdatensatz einer Krebszelllinie (Schuppenzellkarzinom) verifiziert. Chemisch modifizierte Aminosäuren fanden hierbei keine Berücksichtigung. Die Messtoleranz wurde auf 50 ppm adjustiert.

Projekt: Goya3

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 zusammengefasst 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.

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. Nachdem Goya im WS 04/05 für den Echtbetrieb freigegeben wurde, steht die Stabilisierung sowie die weitere Verbreitung des Systems nun im Mittelpunkt der Bestrebungen.

Veröffentlichungen

J.-C. FREYTAG, R. RAMAKRISHNAN, R. AGRAWAL: *Data Mining: The Next Generation (Data Mining: Die nächste Generation)*. it - Information Technology, Volume 47, Ausgabe 5, Seiten 308-312, May 2005.

C. BIZER, R. HEESE, M. MOCHOL, R. OLDAKOWSKI, R. TOLKSDORF, R. ECKSTEIN: *The Impact of Semantic Web Technologies on Job Recruitment Processes*. Tagungsband der 7. Internationale Tagung Wirtschaftsinformatik Deutschland, Bamberg, 2005.

R. ECKSTEIN, S. HEYMANN, A. GRÜNEWALD: *Integrating XML based Life Sciences Data Using UML*. International Workshop Integrative Bioinformatics, Bielefeld, 2005.

R. ECKSTEIN, R. TOLKSDORF (HRSG.): *Berliner XML-Tage 2005*. Proceedings, Berlin, Germany. ISBN 3-9810105-2-3, 12.-14. September 2005.

R. HEESE, S. HERSCHEL, F. NAUMANN, A. ROTH: *Self-Extending Peer Data Management*. Tagungsband der Konferenz Datenbanksysteme in Business, Technologie und Web (BTW), Karlsruhe, Germany, 2005.

S. HERSCHEL, R. HEESE: *Discoverer: A semantic P2P index for PDMS*. Proceedings of the International Workshop Data Integration and the Semantic Web, Porto, Portugal, 2005.

R. TOLKSDORF, M. MOCHOL, R. HEESE, R. ECKSTEIN, R. OLDAKOWSKI, C. BIZER: *Semantic-Web-Technologien im Arbeitsvermittlungsprozess*. Zeitschrift Wirtschaftsinformatik, Special Issue "Internetökonomie", Heft 1/2006.

O. BERTHOLD, O. GÜNTHER, S. SPIEKERMANN: *RFID Verbraucherängste und Verbraucherschutz*. Zeitschrift Wirtschaftsinformatik, Heft 6/2005, Vieweg Verlag, Wiesbaden, Dez. 2005.

O. BERTHOLD: *Datenschutzgerechte RFID-Technologie*. In: Hannes Federrath (ed): Sicherheit 2005, Sicherheit - Schutz und Zuverlässigkeit: Beiträge der 2. Jahrestagung des Fachbereichs Sicherheit der Gesellschaft für Informatik e.V. (GI), Lecture Notes on Informatics (LNI), S.315 - 327; Regensburg, 2005.

O. BERTHOLD: *Effizienter unbeobachtbarer Datenbankzugriff*. In: Ferstl, O. K.; Sinz, E. J.; Eckert, S.; Isselhorst, T. (Hrsg.), Wirtschaftsinformatik 2005: eEconomy, eGovernment, eSociety. Heidelberg, Physica-V, 2005.

C. BEN NECIB, J.-C. FREYTAG: *Query Processing Using Ontologies*. Proceedings of the 17th Conference on Advanced Information Systems Engineering (CAISE'05), Porto, Portugal, 2005.

C. BEN NECIB, J.-C. FREYTAG: *Semantic Query Transformation Using Ontologies*. Proceedings of the 9th International Database Applications and Engineering Symposium (IDEAS'05). Montreal, Canada, 2005.

K. ROTHER, S. TRIßL, H. MÜLLER, T. STEINKE, I. KOCH, R. PREISSNER, C. FRÖMMEL, U. LESER: *Columba: An Integrated Database of Proteins, Structures, and Annotations*. BMC Bioinformatics März 2005.

M. MIELKE, H. MÜLLER, F. NAUMANN: *Ein Data-Quality-Wettbewerb*. Datenbank-Spektrum, Heft 14, August 2005.

H. MÜLLER, M. WEIS, J. BLEIHOLDER, U. LESER: *Erkennen und Bereinigen von Datenfehlern in naturwissenschaftlichen Daten*. Datenbank-Spektrum, Heft 15, November 2005.

S. HEYMANN, P. RIEGER, A. SCHAALE: *Primary Linguistic Units in Genomes*, Jahrestagung der Deutschen Physikalischen Gesellschaft, Berlin, März 2005 (Poster).

R. ECKSTEIN, S. HEYMANN, A. GRÜNEWALD: *Integrating XML based Life Sciences Data Using UML*, International Workshop Integrative Bioinformatics, Bielefeld, Juli 2005.

L. RASCHID, S. HEYMANN, F. NAUMANN, P. RIEGER, G. MIHAILA, M. E. VIDAL, A. LEE, Y. WU, A. LASH: *Enhancing the Semantics of Links and Paths in Life Science Sources*. Workshop on Database Issues in Biological Databases (DBiBD), Edinburgh, 2005.

S. HEYMANN, P. RIEGER UND A. SCHAALE: *Primary Linguistic Units in Genomes, Symposium Biological and Social Network.*, Berlin, 2005 (Poster).

Patenteinreichungen

OLIVER BERTHOLD, JOHANN-CHRISTOPH FREYTAG, SARAH SPIEKERMANN: *Datenschutzgerechtes Radio Frequency Identifications(RFID)-System durch besitzerkontrollierte RFID-Tag Funktionalität*. Eingereicht beim Deutschen Patent- und Markenamt, München, April 2005.

FRANK HUBER, JOHANN-CHRISTOPH FREYTAG, TIMO MIKA GLÄSSER, GIOVANNI RABAIOLI: *Erzeugung von Histogrammen für Directory Services*. Eingereicht durch die Firma Siemens beim Deutschen Patent- und Markenamt, München, Oktober 2005.

STEPHAN HEYMANN, HANS JOOS, PETER BENDZKO, BEATE KRAFFERT, RALF BERGMANN, ROBERT MATTHIAS LEISER: *Verfahren zur Steuerung der abschnittweisen enzymatischen Nukleinsäurevervielfältigung über inkomplette Komplementärstränge*. Eingereicht als deutsches Patent beim Deutschen Patent- und Markenamt, München, Prioritätsdatum Oktober 2005.

Vorträge

O. BERTHOLD: *Datenschutzgerechte RFID-Technologie*. GI Jahrestagung Sicherheit 2005, April 2005, Regensburg, April 2005.

O. BERTHOLD: *Effizienter unbeobachtbarer Datenbankzugriff*. Tagung Wirtschaftsinformatik 2005, Bamberg, Februar 2005.

J.-C. FREYTAG: *XML Query Execution Engine (XEE) - Storing, querying & updating XML documents using database techniques*. ISST, Fraunhofer Gesellschaft, Berlin, Januar 2005.

J.-C. FREYTAG: *Zugriff auf Daten in Datenbanksystemen - eine Privatsache??* Dresdener Datenbankstammtisch, Dresden, Februar 2005.

J.-C. FREYTAG: *Processing Genome Data using Scalable Database Technology*. UC Davis, CA, USA, September 2005.

J.-C. FREYTAG: *Querying Databases Privately*. UC Berkeley, CA, USA, September 2005.

J.-C. FREYTAG: *Research on query processing at DBIS*, Oracle Corp., Redwood, CA, USA, September 2005.

R. HEESE: *Self-Extending Peer Data Management*. Konferenz Datenbanksysteme in Business, Technologie und Web (BTW), Karlsruhe, 2005.

R. ECKSTEIN: *Integrating XML based Life Sciences Data Using UML*. International Workshop Integrative Bioinformatics, Bielefeld, 2005.

C. BEN NECIB, J.-C. FREYTAG: *Query Processing Using Ontologies*. Proceedings of the 17th Conference on Advanced Information Systems Engineering (CAISE'05), Porto, Portugal, 2005.

C. BEN NECIB, J.-C. FREYTAG: *Semantic Query Transformation Using Ontologies*. Proceedings of the 9th International Database Applications and Engineering Symposium (IDEAS'05), Montreal, Canada, 2005.

H. MÜLLER: *Vergleich überlappender Datenquellen*. Oberseminar des Lehrstuhl für Technologie der Informationssysteme, Christian-Albrechts-Universität zu Kiel, Februar 2005.

S. HEYMANN: *Krebs-Spezifische Spleißformen des NTR1-Rezeptors*. Einladungsvortrag am Forschungszentrum Rossendorf, Dresden, 06. Februar 2005.

S. HEYMANN, H. JOOS: *Detection of Minor Components in Complex Mixtures of Nucleic Acids*. Einladungsvortrag am Biotechnologiezentrum Bergen/Norwegen, Dezember 2005.

Sonstige Aktivitäten

Prof. Johann-Christoph Freytag, Ph.D.

- Mitglied des Programmkommittees der BTW 2005 (Deutsche Datenbankkonferenz)

- Mitglied des Programmkommittees der ACM Sigmod Konferenz, Baltimore, Maryland, Juni 2005
- Mitglied des Programmkommittees der VLDB 2005, Trondheim, Norwegen
- Mitglied des Programmkommittees der Berliner XML-Tage, Berlin 2005
- Mitglied des Programmkommittees SOFEM 2006, Tschechien
- Mitglied des Programmkommittees EDBT 2006, München
- Mitglied des Programmkommittees WWW 2006, Edinburgh, Schottland
- Mitglied des Programmkommittees des Workshops on Query Languages and Query Processing 2006, München
- Mitglied des Programmkommittees des First Data and Knowledge Quality Workshops (DKQ'05)
- Gastwissenschaftler der IBM Deutschland Entwicklung GmbH (seit 2004)
- Mitorganisator des Dagstuhl-Workshops „Managing and Mining Genome Information: Frontiers in Bioinformatics“ zusammen mit Martin Vingron (Max-Planck-Institut für molekulare Genetik, Berlin) und Jacek Blazewicz (Politechnika Poznanska, Polen), November 2005
- Gutachter für das europäische ITEA Projekt Sirena, Berlin, August 2005
- Visiting Scientist, Microsoft Research, Februar - Mai 2005
- Institutsdirektor des Institutes für Informatik (bis Februar 2005)
- Stellvertretender Institutsdirektor des Institutes für Informatik (seit Februar 2005)
- Mitglied des Konzils der Humboldt-Universität zu Berlin (seit 2004)
- Mitglied des VLDB-Endowments, USA (seit 2002)
- Fachgruppensprecher der GI-Fachgruppe Datenbanken, Deutschland
- Mitglied des Herausbergremiums der Zeitschrift DB-Spektrum des dpunkt-Verlages
- Herausgeber zweier Hefte des DB-Spektrums des dpunkt-Verlages für 2005
- Mitglied des Herausbergremiums der Buchserie „Data-centric Systems and Applications“ des Springer Verlages

Dipl.-Inf. Chokri Ben Necib

- Teilnahme an der Konferenz „Advanced Information Systems Engineering (CAISE'05)“, Porto, Portugal, 2005
- Teilnahme an der Konferenz „Applications and Engineering Symposium (IDEAS'05)“, Montreal, Kanada, 2005

Dipl.-Inf. Oliver Berthold

- Teilnahme an der GI Jahrestagung Sicherheit 2005, 5.-8. April 2005, Regensburg.
- Teilnahme an der Tagung Wirtschaftsinformatik 2005, Bamberg, 23.-25. Februar 2005

Dipl.-Inf. Ralf Heese

- Organisation der Podiumsdiskussion „Standardisierung – Positiver Impuls oder Hemmnis im Markt“ im Rahmen der Konferenz „Berliner XML Tage“, Berlin, Oktober 2005
- Organisation der Konferenz „Berliner XML Tage“, Berlin, Oktober 2005

Dr. Rainer Eckstein

- Organisation der Konferenz „Berliner XML Tage“, Berlin, Oktober 2005

- Mitglied des Programmkommittees ISWC Workshop Semantic Web Case Studies and Best Practices for eBusiness - SWCASE05. 7. Nov., Galway, Ireland
- Mitglied des Programmkommittees ODBASE'05 International Conference on Ontologies, Databases and Applications of Semantics. 31. Oktober – 4. November, 2005, Zypern
- Mitglied des Programmkommittees IADIS International Conference WWW/Internet 2005, 19.-22. Oktober 2005
- Mitglied des Programmkommittees I-KNOW'05, Special Track on Knowledge Discovery and Semantic Technologies 2005 (KDaST 2005). 29. Juni. 2005, Graz, Österreich
- Gutachter: Journal IEEE Transactions on Knowledge and Data Engineering (TKDE), 2005
- Zusätzlicher Gutachter: 24th ACM SIGMOD International Conference on Management of Data 2005

Dipl.-Inf. Sven Herschel

- Einwerbung von Mitteln der Medienkommission der Humboldt-Universität für das Projekt Goya3-Entwicklung

Dr. Stephan Heymann

- Wahl zum Mitglied der Technologischen Akademie Armeniens (DKQ'05)
- Externer Gutachter im Promotionsverfahren von ROBERT GEVORGYAN an der Staatsuniversität von Armenien, Juni 2005
- Biomedizinrelevante Gemeinschaftsarbeiten mit dem Forschungszentrum Rossendorf (Kooperationsvertrag seit 2004)

Dipl.-Inf. Frank Huber

- Vorträge bei den Projekttreffen zu DirXQuE3 bei der Siemens AG München (Februar, Juni, Oktober 2005)

Dipl.-Inf. Heiko Müller

- Teilnahme am Workshop des GKs „Verteilte Informationssysteme“, Berlin, November 2005

Dipl.-Inf. (FH) Katja Tham

- Einwerbung eines Promotionsstipendium: gemäß Nachwuchsförderungsgesetz des Landes Berlin (NaFöG).
- Mitorganisation eines "Quiz-Marathons" im Rahmen des Girl's Day (Mädchen-Zukunftstag) zur Förderung der naturwissenschaftlich, technischen Interessen von Schülerinnen.
- Teilnahme am Workshop des GKs "Verteilte Informationssysteme", Mai 2005, Döllnsee.
- Teilnahme am Workshop des GKs "Verteilte Informationssysteme", November 2005, Berlin-Adlershof.

Diplomarbeiten

GERD ANDERS: *Untersuchung von Sequenzduplikaten auf funktionelle Eigenschaften im Referenzgenom von Homo Sapiens unter Verwendung eines relationalen Datenbanksystems*. März 2005.

JAN GROBE: *Eine RDF Algebra*. Mai 2005.

YVONNE GABRIEL: *Das Instituts-Gedächtnis – Ein System zur Verwaltung der Dokumente des Instituts für Informatik*. September 2005.

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 Mitarbeiterinnen und Mitarbeiter

JANA BAUCKMANN

DR. LUKAS FAULSTICH

JÖRG HAKENBERG

SILKE TRIBL

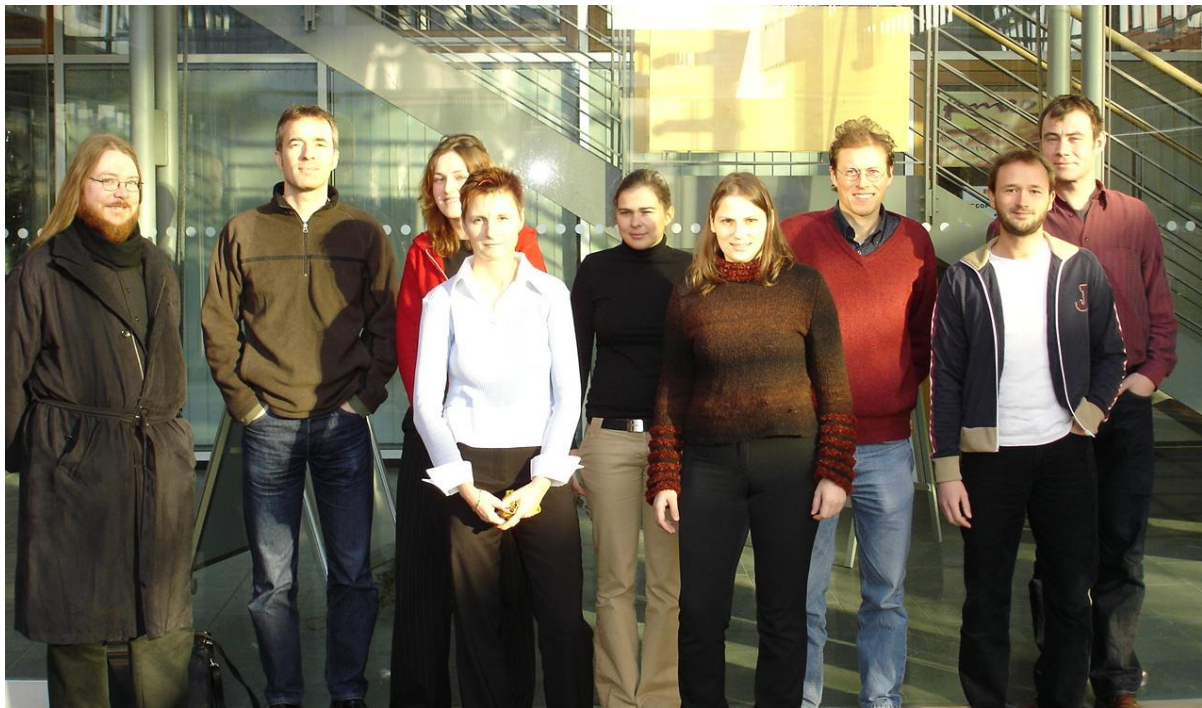
Studentische Hilfskräfte

PHILIPP HUSSELS

CHRISTIANE KRÜGER

Techniker

ROBERT MIELKE



Der Lehrstuhl beschäftigt sich mit Wissens- und Datenmanagement in der molekularbiologischen und biomedizinischen Forschung. Dazu gehören alle Aspekte der Speicherung und Analyse strukturierter und unstrukturierter Daten. In der Forschung konzentriert sich die Gruppe auf Fragestellungen in der Integration biomedizinischer Daten, der performanzorientierten Entwicklung molekularbiologischer Datenbanken sowie auf Methoden zur automatischen Analyse von biomedizinischen Veröffentlichungen (Text Mining). In der Lehre werden vor allem Themen aus dem Bereich Datenbanken, Datenanalyse und der Bioinformatik behandelt. Der Lehrstuhl ist vielfältig interdisziplinär vernetzt und Mitglied in drei interdisziplinären Zentren der Humboldt Universität: Dem Zentrum für Biophysik und Bioinformatik, dem Zentrum für Sprachliche Bedeutung sowie dem Zentrum für Ubiquitäre Informationssysteme.

Die Gruppe hat im Laufe des Jahres 2004 sowohl existierende Forschungsarbeiten erfolgreich fortgesetzt als auch neue Themen bearbeitet. Dabei sind vielfältige Arbeiten zum Text Mining, zur Analyse der Qualität von Genomdaten, zur Integration von Proteindatenbanken und zum Management von graphartigen Daten, wie sie zum Beispiel in der Systembiologie vorkommen, entstanden. Gerade das letzte Thema hat sich als neuer Forschungsschwerpunkt etabliert, der unter anderem in einer Reihe von studentischen Arbeiten vorangetrieben wird. Die Kooperation mit dem Institut für deutsche Sprache konnte erfolgreich fortgesetzt werden, was sich in der Verlängerung des Projektes durch den Berliner Senat reflektiert.

Für die bereits in 2004 durchgeführte Umfrage unter den Studenten des Instituts konnte, in Kooperation mit dem Institut für Psychologie und mit Unterstützung des Präsidiums der Humboldt-Universität, 2005 eine detaillierte Auswertung vorgenommen werden. Der dabei entstandene umfassende Bericht enthält eine Reihe wichtiger Ergebnisse, die auch schon teilweise zu Maßnahmen des Instituts geführt haben, wie zum Beispiel die Durchführung von individuellen Gesprächen mit (statistisch ermittelten) abbruchgefährdeten Studenten. Die Auswertung einer daran anschließend vorgenommenen Umfrage unter Studenten, die das Institut tatsächlich ohne Abschluss verlassen haben, ist noch nicht abgeschlossen.

Lehre im Grundstudium

Proseminar

- „Klassische Algorithmen in der Bioinformatik“, 2 SWS, Prof. Leser, SoSe 2005

Lehre im Hauptstudium

Halbkurse und Übungen

- Vorlesung „Algorithmische Bioinformatik“, 4 SWS, Prof. Leser, WiSe 2004/2005
- Übung zum HK „Algorithmische Bioinformatik“, 2 SWS, Prof. Leser, Jörg Hakenberg, Silke Trißl, WiSe 2004/2005
- Vorlesung „Implementierung von Datenbanksystemen“ (4 SWS), Prof. Leser, SoSe 2005
- Praktikum zum Halbkurs „Implementierung von Datenbanksystemen“ (2 SWS), Heiko Müller (LFE DBIS), Frank Huber (LFE DBIS), Jana Bauckmann, SoSe 2005
- Vorlesung „Algorithmische Bioinformatik“ (4 SWS), Prof. Ulf Leser, WiSe 2005/06
- Übung zur Vorlesung „Algorithmische Bioinformatik“, Prof. Ulf Leser, Jörg Hakenberg, Jana Bauckmann, WiSe 2005/06

Spezialveranstaltungen

- Vorlesung „Data Warehousing“ (2 SWS), Prof. Leser, SoSe 2005

Seminare

- Seminar „Verwandtschaft und Abstammung in Zeichenketten“ (2 SWS), Prof. Leser, WiSe 2004/2005 (interdisziplinäres Seminar zusammen mit Prof. Lüdeling und Prof. Donhauser, Institut für deutsche Sprache der Humboldt-Universität)
- Seminar „Fortgeschrittene algorithmische Bioinformatik“ (2 SWS), Prof. Leser, SoSe 2005
- Seminar „Graphmanagement in Datenbanken“ (2 SWS), Prof. Leser, Silke Trißl, WiSe 2005/06
- Interdisziplinäres Seminar „Rekonstruktion von Sprachstambäumen“ (2 SWS), Prof. Leser, WiSe 2005/2006 (interdisziplinäres Seminar mit Prof. Lüdeling, Institut für deutsche Sprache der Humboldt-Universität)
- Regelmäßiges Forschungsseminar für Diplomanden und Doktoranden, Prof. Leser, zusammen mit Prof. Naumann

Lehre im Zentrum für Biophysik und Bioinformatik

- Vorlesung „Bioinformatik für Biophysiker“, mit Softwarepraktikum, (2 SWS) Prof. Leser, WiSe 2004/2005 (Zusammen mit Prof. Frömmel, Dr. Preissner, Dr. Köpke, Prof. Herzel (Institut für Biologie und Charité)).
- Vorlesung „Bioinformatik für Biophysiker“, mit Softwarepraktikum, (2 SWS), Prof. Leser, WiSe 2005/2006 (Vorlesung zusammen mit Prof. Herzel, Dr. Köpke, Dr. Preissner (Institut für Biologie und Charité))

Forschung

Projekt „Berlin Center for Genome-Based Bioinformatics“

Ansprechpartner: Prof. 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 - Eine Datenbank über Proteinstrukturen und deren Annotationen

Ansprechpartner: Prof. Ulf Leser, Silke Trißl

Beteiligte Mitarbeiter: Philipp Hussels, Markus Pankalla

Zusammenarbeit: Universitätsklinikum Charité, Prof. Frömmel, Humboldt-Universität, Informatik, Prof. Freytag, Konrad Zuse Zentrum für Informationstechnik, Dr. Steinke, Technische Fachhochschule Berlin, Prof. Koch

Forschungsförderung: Im Rahmen des BCB

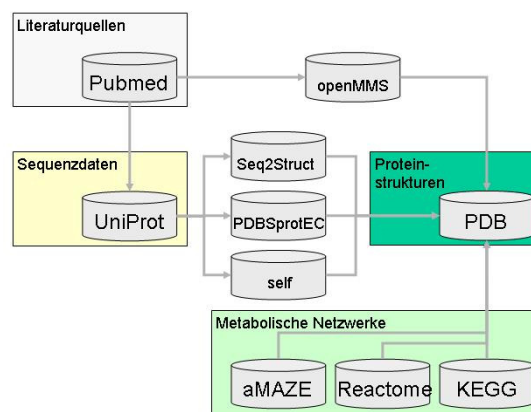
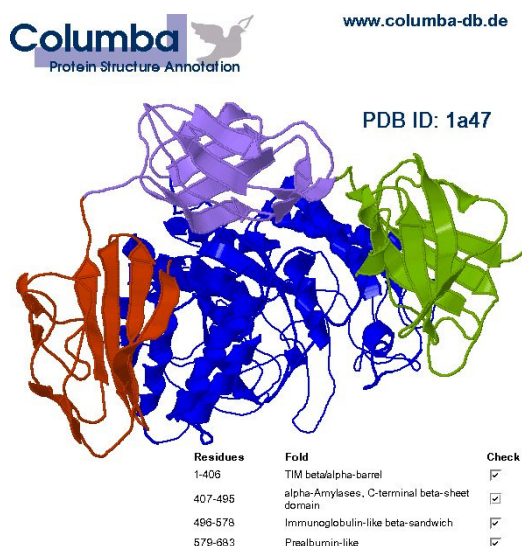
In der Biologie sind Forscher häufig nicht nur an einem Aspekt eines Proteins interessiert – wie der dreidimensionalen Struktur – sondern auch an weiteren Informationen. Diese umfassen die Funktion, die biologische Bedeutung, verwandte Proteine oder die Aufgaben und Interaktionen von Proteinen im Stoffwechsel. Die Informationen werden in verschiedenen, öffentlich zugänglichen Datensammlungen verwaltet und zur Verfügung gestellt. Ziel des Projektes Columba ist es, diese verteilten Daten, die Annotationen zu Proteinstrukturen der PDB liefern, in einer zentralen Datenbank zu sammeln und semantisch zu integrieren. Die Integration der Daten gibt dem Nutzer die Möglichkeit, die Daten anzufragen und komplexe Fragestellungen einfach zu beantworten.

Das Datenbankschema von Columba ist um Strukturen der Protein Data Bank (PDB) zentriert und stellt Annotationen aus 15 weiteren Datenquellen zur Verfügung. Diese Datenquellen enthalten Daten über den Faltungstyp eines Proteins (SCOP und CATH), über die Sekundärstruktur (DSSP), eine Beschreibung von Funktionen und Herkunftsorganismus (SwissProt, GeneOntology und NCBI Taxonomy), die Beteiligung eines Proteins an Stoffwechselfvorgängen in der Zelle (ENZYME, Boehringer, KEGG, aMAZE und Reactome), die Einteilung von Proteinen aufgrund Sequenzhomologie (PISCES) und Literaturreferenzen (PubMed).

Ein großer Teil der genannten Datenquellen ist über das Webinterface abfragbar. Die Anfrage liefert als Ergebnis eine Liste von PDB-Einträgen, wobei für jeden einzelnen Eintrag sowohl die gesamten Daten in Columba als auch die Struktur selbst angezeigt werden kann.

Die Columba-Datenbank wurde im Jahr 2005 um verschiedene Datenquellen sowie Funktionalitäten ergänzt. Beispielsweise wurde der 3D-Viewer erweitert, so dass die verschiedenen Protein-Domänen aus SCOP angezeigt werden können. Im nächsten Schritt soll das auch für CATH möglich sein. Alle Daten sind nun auch in XML verfügbar, was eine erhebliche Erleichterung für die Benutzung von Daten aus Columba darstellt. PubMed als neue Datenquelle stellt die Verknüpfung der Text Mining Arbeiten der Gruppe mit dem Columba Projekt dar. Weitere neue Datenquellen sind aMAZE und Reactome, die beide Informationen über physische Interaktionen von Proteinen liefern.

Diese verschiedenen Datenquellen differieren in der Methode der Datengewinnung oder -qualität, überschneiden sich aber inhaltlich und strukturell. Die gleiche Anfrage an verschiedene Quellen kann daher zu sehr unterschiedlichen Ergebnissen führen. In Columba wollen wir Forschern die Möglichkeit geben mehrere Datenquellen, die gleiche Informationen enthalten, parallel anzufragen. Alle Ergebnisse werden dann - sortiert nach ihrer Güte, das bedeutet in wie vielen Datenquellen das Ergebnis gefunden wurde - zurückzugeben. Im Jahr 2006 soll die Idee durch Daten zu Protein-Protein-Interaktionen noch erweitert und dem Nutzer über das Webinterface zur Verfügung gestellt werden.



Die Abbildung zeigt die dreidimensionale Struktur des PDB-Eintrags 1a47 – einer Glykosyltransferase. Die verschiedenen SCOP-Faltungstypen sind farblich markiert. In blau ist der TIM barrel-fold dargestellt, neben dem Immunoglobulin-like beta-sandwich-fold (in lila) einer der am häufigsten vorkommenden Faltungstypen in Proteinen der PDB. Die Glykosylhydrolase Domäne – verantwortlich für die eigentliche Funktion – ist rot gekennzeichnet.

Proteinstrukturen aus der PDB können über 3 verschiedene Datenquellen mit Sequenzdaten aus UniProt verknüpft werden. Je nach verwendeter Datenquelle kann dies zu unterschiedlichen Ergebnissen führen. In Columba wollen wir dem Nutzer die Möglichkeit geben, alle verfügbaren Quellen parallel anzufragen. Jedes Ergebnis wird zusammen mit der Information über wie viele und welche Quellen das Ergebnis gefunden wurde zurückgegeben.

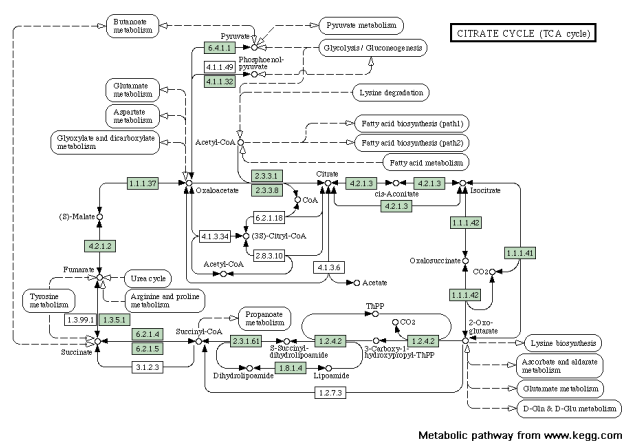
Projekt: PQL –Anfragesprachen für biologische Netzwerke

Ansprechpartner: PROF. ULF LESER

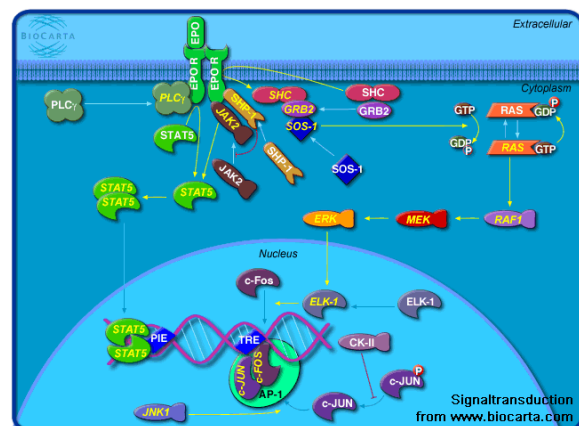
Beteiligte Mitarbeiter: SILKE TRIBL

Forschungsförderung: Im Rahmen des BCB

Jede Zelle, egal ob Bakterium oder Leberzelle, besitzt eine Vielzahl von Proteinen, die durch ihr Zusammenspiel die Funktion der Zelle aufrechterhalten. Eine Gruppe von Proteinen, die Enzyme, sind dafür verantwortlich, dass chemische Substanzen in der Zelle auf- oder abgebaut werden. Beispielsweise wird Glukose, die von der Zelle aufgenommen wird, unter Energiegewinnung zu Pyruvat abgebaut, das dann in den Zitronensäurezyklus ein- geht. Solche Stoffwechselforgänge werden in metabolischen Netzwerken dargestellt. Aber Zellen können auch auf äußere Einflüsse reagieren, wie chemische Substanzen, Veränderung der Temperatur, oder geänderte Lichteinstrahlung. Sie tun dies durch ein Zusammen- spiel von verschiedenen Proteinen. Dieses äußerst komplexe und in weiten Teilen noch unverstandene Zusammenspiel ist in den letzten Jahren unter dem Begriff „Systembiologie“ zunehmend in den Fokus der biomedizinischen Forschung gekommen.



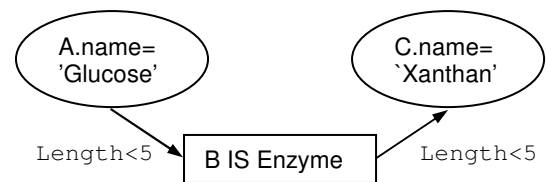
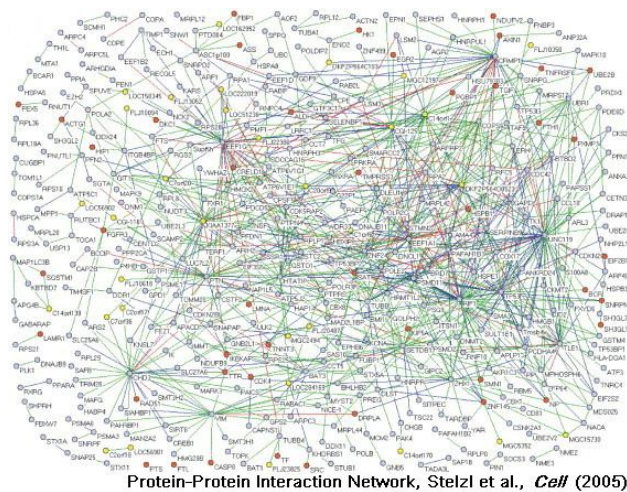
Darstellung des Zitronensäurezyklus – einem der Haupt- stoffwechselwege in jedem Organismus. Die Rechtecke stellen Enzyme dar, die chemische Substanzen (darge- stellt als Kreise) umwandeln. Die abgerundeten Boxen stehen für weitere Stoffwechselwege, in die Zwischen- produkte aus dem Zitronensäurezyklus eingehen. Beim Menschen sind bisher die farblich markierten Enzyme gefunden worden.



Die Signaltransduktion ist für Zellen ein wichtiges Instrument, um auf Signale von außen – wie chemische Substanzen oder Hitze reagieren zu können. Die Abbildung zeigt die Reaktion einer Zelle auf Erythropoetin (EPO), das als Wachstumsfaktor für rote Blutkörperchen dient. Nach der Bindung von EPO am Rezeptor an der Zelloberfläche interagiert dieser mit einer Reihe weiterer Enzyme, was dazu führt, dass die Gene für die Reifung von Erythrozyten vermehrt abgelesen werden.

Analysiert werden dabei Netzwerke, die von wenigen Proteinen oder Genen bis zu zehntausende Objekte enthalten können. Es ist essentiell, schnelle und einfach zu benutzende Möglichkeiten zur Verfügung zu haben, um diese Netzwerke zu durchsuchen und strukturell analysieren zu können. Zu diesem Zweck wurde am Lehrstuhl die Anfragesprache PQL – Pathway Query Language – für biologische Netzwerke entworfen und prototypisch implementiert. Neben dem Sprachdesign forschen wir vor allem an einer Verbesserung der Effizienz von Anfragen an sehr große Graphen durch die Verwendung einer speziellen Indexstruktur. Diese Indexstruktur ist angelehnt an Methoden für XML Dokumente, erweitert diese aber auf Graphen. In 2005 haben wir als ersten Schritt eine Anpassung auf gerichtete, azyklische Graphen entwickelt und diese dann weiter verbessert, um auch zykli-

sche Graphen indexieren zu können. Im Jahr 2006 werden wir die erweiterte Indexstruktur mit der Pathway Query Language kombinieren und damit die Möglichkeit schaffen, große Protein-Protein-Interaktionsnetzwerke anzufragen. Daneben sind bereits eine Reihe studentischer Arbeiten zur Visualisierung, Sprachverbesserung und vollständigen Implementierung angelaufen.



```
SELECT A[-s]B, B[-s]C
FROM A, B, C
WHERE A.name = 'Glucose'
      AND B IS Enzyme
      AND C.name = 'Xanthan'
      AND A[-<5]B
      AND B[-<5]C;
```

Ein Netzwerk aus Proteininteraktionen von Hefe-Zellen, bestehend aus 401 Proteinen, die durch 911 Interaktionen miteinander verbunden sind.

Diese PQL-Anfrage liefert uns den oder die kürzesten Wege in einem metabolischen Netzwerk, um von Glukose (Zucker) über ein Enzym in insgesamt höchstens zehn Schritten zu Xanthan zu kommen. Xanthan ist ein Dickungsmittel, mit dem z.B. Ketchup dickflüssiger gemacht werden kann.

Projekt „Text Mining für die Systembiologie“

Ansprechpartner: PROF. ULF LESER, JÖRG HAKENBERG

Zusammenarbeit: DR. EDDA KLIPP, DR. AXEL KOWALD, SEBASTIAN SCHMEIER, Max-Planck-Institut für Molekulare Genetik

Forschungsförderung: Im Rahmen des BCB

Das Projekt beschäftigt sich mit der Erkennung von Informationen in Fachartikeln, die für die kinetische Modellierung biochemischer Systeme relevant sind. Dazu gehören wichtige Daten wie Enzymkonzentrationen und Reaktionskonstanten und -gleichungen, die experimentell im Labor beobachtete Stoffwechselfvorgänge 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 Informationen unmöglich.

In diesem Projekt haben wir bereits in 2003 und 2004 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 erzielten Re-

sultate übertreffen die manuelle und mühsame Suche nach relevanten Publikationen bei weitem.

Darauf aufbauend haben wir in 2005 begonnen, die in den Texten diskutierten enzymatischen Reaktionen mit den zugehörigen experimentell gemessenen Parametern (Reaktionsgeschwindigkeit, Umgebungstemperatur, etc.) automatisch zu extrahieren. Derartige Informationen sind für die kinetische Modellierung komplexer Systeme unabdingbar und müssen bisher von Hand gesucht werden. Diesen Schritt haben wir soweit automatisiert, so dass ein vorgegebenes Formular per Software gefüllt und angezeigt werden kann. Vor dem Eintrag in eine Datenbank muss nun ein Experte die extrahierten Daten nur noch überprüfen. Die automatische Extraktion erzielt sehr gute Ergebnisse

(URL: <http://sysbio.molgen.mpg.de/KMedDB>)

PMID: 3036196
 Pyrimidine nucleoside monophosphate kinase hyperactivity in hereditary erythrocyte pyrimidine 5'-nucleotidase deficiency

The pyrimidine nucleoside triphosphates (CTP, UTP) increase in the pyrimidine 5'-nucleotidase (P5N) deficient red blood cell (RBC) to a greater degree than do the pyrimidine nucleoside monophosphates (CMP, UMP). Pyrimidine nucleoside monophosphate (PNMP) kinase phosphorylates CMP and UMP to their respective phosphodiester. We tested the hypothesis that increased PNMP kinase activity contributes to the disproportionate increase in CTP and UTP in the P5N deficient RBC. CMP and UMP kinase activities were increased in high reticulocyte (4.4 +/- 2.1 and 8.5 +/- 3.3 mumol/ml RBC per minute) compared to normal RBC (2.8 +/- 1.0 and 6.0 +/- 2.5 mumol/ml RBC per minute). P5N deficient RBC (n = 2) had significantly increased CMP and UMP kinase activities (14.0 and 28.5 mumol/ml RBC per minute). UMP and CDP-ethanolamine were able to increase the activity of CMP kinase in crude haemolysate and the activity of partially purified enzyme. Since the Km for CMP of CMP kinase was 33 mumol/l in P5N deficient RBC and since the CMP concentration is 25-90 mumol/l in the P5N deficient RBC, the enzyme should be nearly saturated with CMP in the P5N deficient RBC. Thus, PNMP kinase hyperactivity appears to contribute to the disproportionate increase in CTP and UTP in the P5N deficient RBC.

Relation 1: Sentence 6

Kinetic		
Constant:	Value:	Unit:
KM	33	mumol/l

Substances	
Enzyme::1 Sentence::6	Compound::1 Sentence::6
NAME:: CMP kinase	NAME:: CMP
EC::2.7.4.14 Kinetikon::2012	KEGG::C00055 Kinetikon::3901

Reaction

- REACTION::R00512 | Kinetikon::464 | EC::2.7.4.14 | COMPOUND::C00055

Suche über KMedDB: Text mit automatisch markierten Entitäten (Enzyme, chemische Moleküle) und Evidenz für eine erkannte Reaktion zweier Stoffe (gelb).

Projekt „Informationsextraktion für die Genetik: Adipositas-assoziierte Gene“

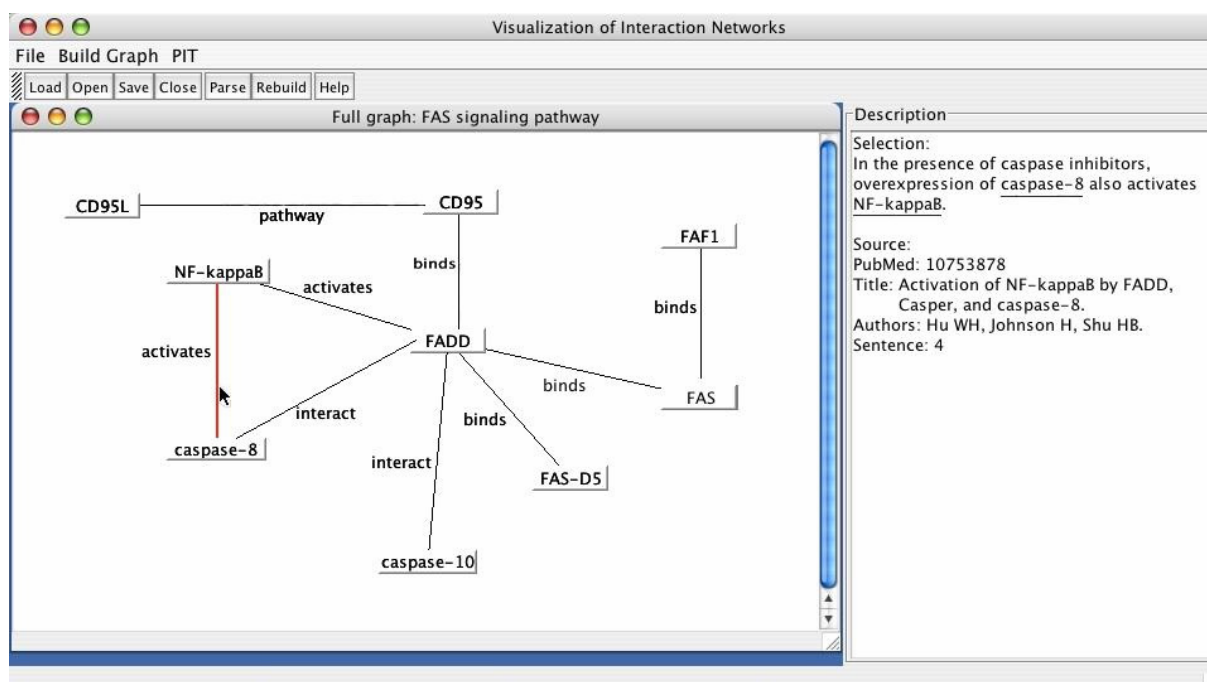
Ansprechpartner: PROF. ULF LESER, JÖRG HAKENBERG

Beteiligte Mitarbeiter: CHRISTIANE KRÜGER, CONRAD PLAKE

Zusammenarbeit: PROF. GUDRUN BROCKMANN, DR. ARMIN SCHMITT, Abteilung Züchtungsbiologie und molekulare Genetik, Institut für Nutztierwissenschaften, Humboldt-Universität

In der Zusammenarbeit mit der Gruppe von Prof. Brockmann beschäftigen wir uns mit der Extraktion von Informationen zu Genen, die in Zusammenhang mit Fettleibigkeit (Adipositas) in Fachartikeln diskutiert werden. Dies umfasst verschiedene Schritte des Text Minings, wie die verbesserte Literatursuche durch Textklassifizierung, das automatische Erkennen von Gen- und Proteinnamen im Text und die Extraktion von Interaktionspartnern von Adipositas-assoziierten Genen und Genprodukten. Die Suche nach Interaktionen basiert auf der Erkennung von bestimmten sprachlichen Mustern, die Proteine in Texten miteinander in Beziehung stellen. Diese Muster werden aus manuell annotierten Beispielen gelernt und auf unbekannte Dokumente angewendet.

Der manuelle Aufwand ist hierbei verhältnismäßig gering, die automatisch auf bekannte Interaktionen untersuchte Textmenge hingegen umfasst ca. 16 Millionen Fachartikel. Die fruchtbare Zusammenarbeit mit Domänenexperten aus der Biologie ermöglicht es, konkrete und exakte Fallbeispiele zu finden sowie die Methoden zu evaluieren und geeignet zu verbessern.



Aus einer großen Menge von Texten gefiltertes Interaktionsnetzwerk zwischen verschiedenen Proteinen (Rechtecke). Zu jeder entdeckten Interaktion wird dem Benutzer Art und auf Wunsch der Ursprung (z. B. ein Fachartikel, rechte Seite) angezeigt.

Projekt „Interdisziplinärer Forschungsverbund Linguistik - Bioinformatik“

Ansprechpartner: PROF. LESER

Beteiligte Mitarbeiter: DR. LUKAS FAULSTICH, MIRKO HOCHMUTH, THORSTEN VITT

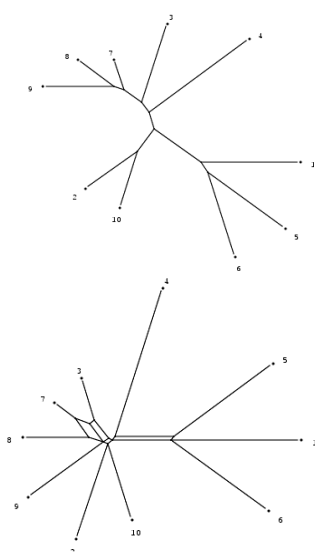
Zusammenarbeit: PROF. LÜDELING, PROF. DONHAUSER, Institut für Deutsche Sprache, Humboldt-Universität

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. Den ersten Bereich, das Management komplexer Annotationen, verfolgen wir im Rahmen der Vorbereitung des Projekts DDD (siehe dortige Beschreibung). Im zweiten Bereich erforschen wir die Übertragung phylogenetischer Verfahren von der Bioinformatik auf die Sprachforschung.

Verwandtschaft und Abstammungsbeziehungen werden in der modernen Biologie zunehmend durch den computerbasierten Vergleich von DNA Sequenzen ermittelt. Diese Methoden werden von uns nun auf Sprachen und Texten erprobt. Nachdem im Jahr 2004 Vorarbeiten zur Verwendung dieser Algorithmen abgeschlossen werden konnten, widmeten wir uns 2005 verstärkt der Entwicklung eigener Modelle und Methoden. Dabei konzentrieren wir uns auf die Abschätzung der Anwendbarkeit der in der Biologie etablierten Verfahren. Diese ist nicht unmittelbar gegeben, da Sprachen ständig miteinander in Kontakt stehen, was zum Austausch von Wörtern, Redewendungen und Phrasen führt. Diese verletzen die „klassische“ Struktur von Sprachstammbäumen viel stärker als das biologische Mechanismen des Gentransfers bei den meisten Organismen tun.

Wir entwickeln daher ein Modell, mit dem man Sprachkontaktphänomene simulieren und deren Auswirkungen auf die Präzision von phylogenetischen Verfahren beurteilen kann. Das Modell simuliert die Entstehung, Veränderung und Beeinflussung von Sprachen auf einem einfachen geographischen Netzmodell. Eine „Sprache“ wird dabei durch eine 200 Einträge umfassende Liste repräsentiert. Verschiebungen auf Lautebene werden dann über einen großen Zeitraum zufällig und in zufällig gewählten Sprachen erzeugt. Da der komplette Prozess protokolliert wird, ist am Ende einer Simulation genau bekannt, in welchen Abstammungsverhältnissen die dabei erzeugten Sprachen stehen. Diese wahren Beziehungen können dann mit den Beziehungen verglichen werden, die phylogenetische Algorithmen nur aufgrund der Sprachen im Endzustand berechnen. Es zeigt sich aber bereits jetzt, dass Verfahren, die über die Rekonstruktion von Bäumen (im Unterschied zu phylogenetischen Netzwerken) hinausgehen, noch kaum existieren.



Vergleich zweier Verfahren zur Berechnung von Sprachstammbäumen auf identischem Datensatz. Das linke Verfahren („Neighbor-Joining“) kann ausschließlich Bäume berechnen; das rechte Verfahren („Splits-Tree“) entdeckt auch Verhältnisse, die einer Baumstruktur widersprechen und markiert sie durch Rauten. Ein Vergleich beider Methoden muss von der Erscheinung der Graphen abstrahieren und nur deren Topologie berücksichtigen.

Viele phylogenetische Verfahren basieren auf Abstandsmaßen zwischen Sprachen. Abstandsmaße müssen dabei die Natur des untersuchten Objektes berücksichtigen. Für Sprachen muss die Aussprache und nicht die Schreibweise verglichen werden. Dazu müssen Ähnlichkeitsmatrizen von Lauten berechnet werden.

Projekt „Deutsch. Digital. Diachron.“

Ansprechpartner: PROF. LESER

Beteiligte Mitarbeiter: DR. LUKAS FAULSTICH

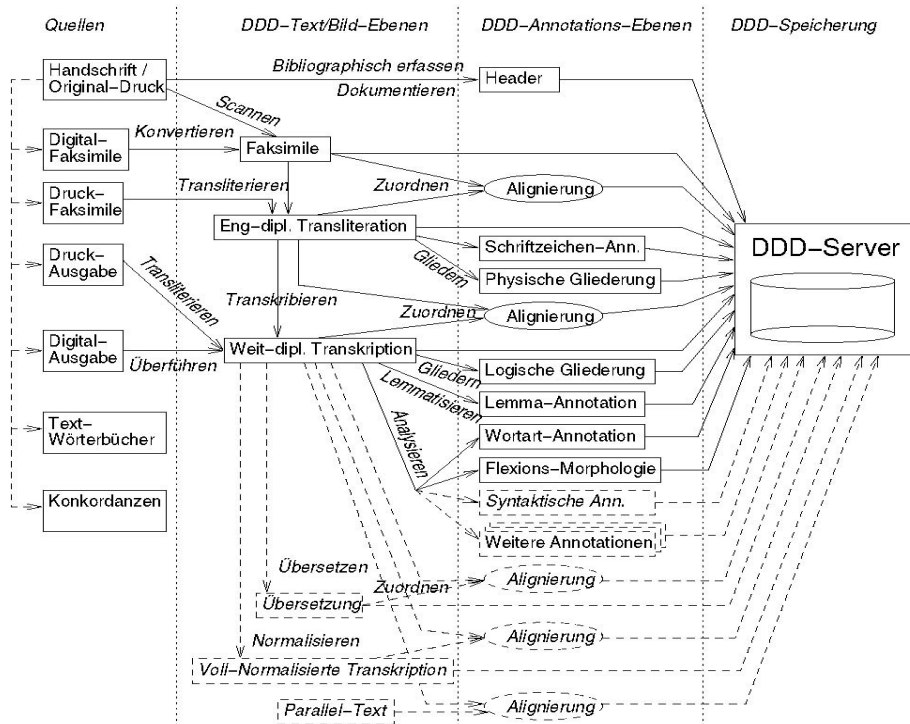
Zusammenarbeit: PROF. LÜDELING, PROF. DONHAUSER, Institut für Deutsche Sprache, Humboldt-Universität Berlin, Bundesweit weitere 15 Gruppen im Bereich Linguistik, Germanistik und Philologie

Forschungsförderung: Beantragt

Texte in deutscher Sprache sind seit dem 9. Jahrhundert in Handschriften und Drucken überliefert. Sie stellen einen wesentlichen Aspekt der europäischen (Kultur-)Geschichte dar 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 verschiedene Sprachepochen bearbeiten und die philologisch geprüften und mit linguistischen An-

notationen versehenen elektronischen Texte an eine zentrale Datenbank in Berlin leiten. Diese gliedert die Texte in einer für alle Sprachstufen übereinstimmenden Form in die digitale Bibliothek ein und macht sie über verschiedene Schnittstellen einer weltweiten Recherche zugänglich.



Geplanter Arbeitsablauf zur verteilten Digitalisierung und Annotierung von Texten in DDD. Dabei gilt es über ein Dutzend bundesweit verteilte Gruppen zu organisieren. Gesammelt werden eine Vielzahl unterschiedlicher Texte und Informationen über deren Herkunft, Aufbau, Struktur, etc. Die Annotation geht über Sätze, Wörter mit Wortarten und Wortstämmen bis auf einzelne Zeichen hinunter – was bei mittelalterlichen Texten oft Erweiterungen des „normalen“ Zeichensatzes erfordert.

Im vergangenen Jahr konnte der 2004 eingereichte DDD Vorantrag überarbeitet und in endgültiger Fassung zur Begutachtung eingereicht werden. Dabei wurde ein detaillierter Katalog der zu digitalisierenden Texte erarbeitet. Dieser führt über 1300 Texte und Textsammlungen (z.B. Zeitungsjahrgänge) mit zusammen mehr als 40 Millionen Wortformen auf. Der Antrag sieht einen Projektverbund mit dem von der Berlin-Brandenburgischen Akademie der Wissenschaften geplanten Korpusprojekt „Deutsches Text-Archiv“ vor. Bereits im Vorfeld konnten wir außerdem die Firma Oracle als Sponsor von Datenbank-Lizenzen für DDD gewinnen.

Gleichzeitig haben wir uns weiter mit Vorstudien zur Repräsentation, Speicherung und Anfrage von annotierten Korpora befasst. Ziel dieser Untersuchungen ist eine beschleunigte Implementierung der digitalen Bibliothek in der Anfangsphase von DDD, um so die einzelnen Sprachstufen-Arbeitsgruppen möglichst bald wirkungsvoll technisch unterstützen zu können. So wurde ein Satz grundlegender Anfrageoperationen identifiziert und in Form von logischen Prädikaten spezifiziert, ein Anfrageübersetzer für Prädikaten-logische

Formeln nach SQL entwickelt sowie im Rahmen einer Diplomarbeit die darauf aufbauende, an XPath orientierte, linguistische Anfragesprache DDDquery entwickelt. Für weitere Informationen und Publikationen siehe <http://www.deutschdiachrondigital.de>.

Projekt: ALADIN – (Nahezu) Automatische Integration von Datenbanken in den Life Sciences

Ansprechpartner: PROF. ULF LESER

Beteiligte Mitarbeiter: JANA BAUCKMANN, VÉRONIQUE TIETZ

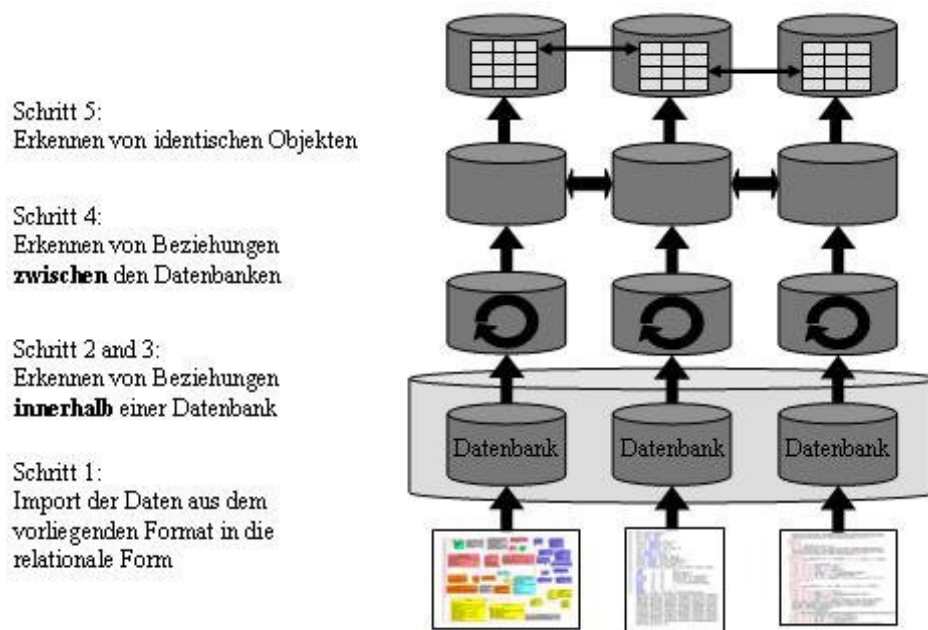
Zusammenarbeit: PROF. NAUMANN, Humboldt-Universität, Informatik

Forschungsförderung: Im Rahmen des BCB sowie DFG

Aladin hat, genau wie das Projekt Columba, das Ziel, Datenbanken in den Life Sciences semantisch zu integrieren. Im Gegensatz zu Columba erfolgt die Integration der einzelnen Datenbanken jedoch nicht manuell, sondern soll automatisch erfolgen.

Das Problem bei der Integration von Datenbanken ist, Informationen zu identifizieren, die in beiden Datenbanken – gegebenenfalls unterschiedlich – repräsentiert sind. Betrachtet man zum Beispiel zwei Datenbanken, die Informationen zu Personen enthalten, so kann man sich folgendes Szenario vorstellen: Datenbank A beinhaltet eine Tabelle „Person“ mit den Spalten „Vorname“ und „Nachname“; Datenbank B besteht ebenfalls aus einer Tabelle namens „Mitarbeiter“ mit den Spalten „Name“ und „Geburtsdatum“. Offensichtlich enthalten beide Datenbanken Informationen über Personen, die Repräsentation des Namens ist jedoch unterschiedlich und der Detailgrad ebenso. Die meisten Projekte zur (semi-) automatischen Datenbankintegration verwenden zum Erkennen solcher Informationen die Struktur der Datenbank: Für relationale Datenbanken, die aus (evtl. verketteten) Tabellen mit Spalten bestehen, sind dies die Tabellen- und Spaltennamen sowie die Struktur zwischen diesen. Wie bereits in obigem Beispiel ersichtlich wird, ist dies jedoch problematisch.

Besonders im Bereich der Life Sciences werden gleiche Begriffe von unterschiedlichen Datenbanken unterschiedlich verwendet. Darüber hinaus sind in den relationalen Repräsentationen der Datenbanken die Strukturen zwischen den einzelnen Tabellen nur selten definiert – obwohl sie implizit in den Daten erkennbar sind. Aus diesen Gründen arbeitet Aladin nicht auf den Strukturinformationen, sondern auf den Daten der einzelnen Datenbanken. Darüber hinaus verwendet Aladin domänenspezifische Informationen, d.h. Informationen über die Struktur der (meisten) Life Science Datenbanken. Die Integration soll dabei in den in der Abbildung dargestellten 5 Schritten erfolgen.



Zum derzeitigen Stand des Projektes wurden 4 Datenbanken – PDB, UniProt, CATH, SCOP – importiert (Schritt 1). Weiterhin wurden Verfahren zum Erkennen von Beziehungen innerhalb einer Datenbank (Schritt 2+3) realisiert und gegeneinander getestet. In Zukunft sollen diese Verfahren optimiert werden und auf das Erkennen von Beziehungen zwischen unterschiedlichen Datenbanken (Schritt 4) angewandt werden.

Vorträge

J. HAKENBERG: *Text Mining for Life Sciences*. AG Textmining. Fraunhofer Institute for Algorithms and Scientific Computing (SCAI), St. Augustin, 14. Oktober 2005

J. HAKENBERG: *From Sequence Alignment to Sentence Alignment*. BCB-Symposium, Ph.d. student workshop. Berlin, 22. Mai 2005.

J. HAKENBERG: *Tuning Text Classification for Hereditary Diseases with Section Weighting*. Symposium on Semantic Mining in Biomedicine (SMBM). Hinxton, UK, 10.-13. April 2005.

J. HAKENBERG: *Optimizing Syntax Patterns for Discovering Protein-Protein Interactions*. ACM Symposium for Applied Computing. Santa Fe, USA, 13.-17. März 2005.

S. TRIBL: *Querying ontologies in relational database systems*. BCB Symposium, Berlin, Juni 2005.

S. TRIBL: *Querying Ontologies in Relational Database Systems*. 2nd Workshop on Data Integration in the Life Science, San Diego, USA, Juli 2005.

S. TRIBL: *Anfragen an Ontologien in relationalen Datenbanken*. 17. Workshop über Grundlagen von Datenbanken, Wörlitz, Mai 2005.

U. LESER: *(Almost) Hands-Off Information Integration for the Life Sciences*. Conference on innovative Data Research, Asilomar (CA), Januar 2005

U. LESER: (*eingeladener Vortrag*), *Columba: Multidimensionale Integration von Proteinstrukturannotationen*. Workshop zu Datenbanken im Nationalen Genomforschungsnetzwerk (NGFN), März 2005.

U. LESER: *A Query Language for Biological Networks*. European Conference on Computational Biology, Madrid, Spain, September 2005.

U. LESER: *Querying Biological Networks*. Dagstuhl Seminar on Managing and Mining Bioinformatics Data, Oktober 2005.

Veröffentlichungen

In Zeitschriften

K. ROTHER, M. DUNKEL, E. MICHALSKY, S. TRIBL, A. GOEDE, U. LESER, R. PREISSNER: *A structural keystone for drug design*. Journal of Integrative Bioinformatics (to appear).

U. LESER, J. HAKENBERG: *What Makes a Gene Name?* Named Entity Recognition in the Biomedical Literature. Briefings in Bioinformatics, 6(4):357-369, 2005.

H. MÜLLER, M. WEIS, J. BLEIHOLDER, U. LESER: *Erkennen und Bereinigen von Datenfehlern in naturwissenschaftlichen Daten*. Datenbank Spektrum, Vol. 15, pp. 26-35, 2005.

J. HAKENBERG, ST. BICKEL, C. PLAKE, U. BREFELD, H. ZAHN, L. FAULSTICH, U. LESER, T. SCHEFFER: *Systematic Feature Evaluation for Gene Name Recognition*. BMC Bioinformatics, 6(Suppl 1):S9, 2005.

U. LESER, J. HAKENBERG: *Text Mining: Automatische Analyse biomedizinischer Texte*. GenomXPress, 4.05, pp. 20-22, 2005.

U. LESER: *A Query Language for Biological Networks*. 3rd European Conference on Computational Biology, Madrid, Spain. Also in Bioinformatics, 21(suppl_2):ii33-ii39, 2005.

D. SCHOBER, U. LESER, J. REICH: *GandrKB: Ontological Microarray Annotation and Visualization*. Bioinformatics, 1;21(11):2785-2786, 2005.

K. ROTHER, S. TRIBL, H. MÜLLER, T. STEINKE, I. KOCH, R. PREISSNER, C. FRÖMMEL, U. LESER: *Columba: An Integrated Database of Proteins, Structures, and Annotations*. BMC Bioinformatics, 6:81, 2005.

K. ROTHER, E. MICHALSKY, U. LESER: *How well are protein structures annotated in secondary databases?* PROTEINS, 60(4):571-76., 2005.

A. LÜDELING, TH. POSCHENRIEDER, L. FAULSTICH: *DeutschDiachronDigital - Ein diachrones Korpus des Deutschen*. Jahrbuch für Computerphilologie (to appear).

Auf Konferenzen und Workshops

J. BAUCKMANN, U. LESER, F. NAUMANN: *Efficiently Computing Inclusion Dependencies for Schema Discovery*. InterDB (with ICDE), Atlanta (to appear).

L. FAULSTICH, U. LESER, T. VITT: *Implementing a Linguistic Query Language for Historic Texts*. International Workshop on Query Languages and Query Processing, München, (to appear).

J. HAKENBERG, J. RUTSCH, U. LESER: *Tuning Text Classification for Hereditary Diseases with Section Weighting*. Proc. of First International Symposium on Semantic Mining in Biomedicine (SMBM), pp.34-39. Hinxton, UK, 2005.

C. PLAKE, J. HAKENBERG, U. LESER: *Optimizing Syntax Patterns for Discovering Protein-Protein Interactions*. Proc. of ACM Symposium on Applied Computing, Bioinformatics track, 1:195-201. Santa Fe, USA, 2005.

J. HAKENBERG, C. PLAKE, U. LESER, H. KIRSCH, D. REBHOLZ-SCHUHMAN: *LLL'05 Challenge: Genic Interaction Extraction with Alignments and Finite State Automata*. Proc Learning Language in Logic Workshop (LLL'05) at ICML 2005, pp. 38-45. Bonn, 2005.

C. PLAKE, J. HAKENBERG, U. LESER: *Learning Patterns for Information Extraction from Free Text*. Workshop des Arbeitskreises Knowledge Discovery, Karlsruhe, 2005.

S. SCHMEIER, J. HAKENBERG, A. KOWALD, E. KLIPP, U. LESER: *Textmining zur Unterstützung von Datenbankkuratoren*. Workshop Datenbanken im NGFN, Köln, 2005.

S. TRIßL, U. LESER: *Querying ontologies in relational database systems*. 2nd Workshop on Data Integration in the Life Sciences (DILS05), San Diego, USA (2005). In Lecture Notes in Computer Science, Volume 3615, Pages 63 – 79, 2005

F. MOUSSOUNI, E. GUREIN, G. MARQUET, A. BURGUN, O. LOREAL. L. BERTI-EQUILLE, U. LESER: *Integrating and Warehousing Liver Gene Expression Data and Related Biomedical Resources in GEDAW*. 2nd Workshop on Data Integration in the Life Sciences (DILS05), San Diego, USA. Also as LNCS, Volume 3615:158 – 174, 2005.

U. LESER, F. NAUMANN: *(Almost) Hands-Off Information Integration for the Life Sciences*. Conference on Innovative Database Research (CIDR), Asilomar, CA, 2005.

Weitere Veröffentlichungen

R. RAMAKRISHNAN, R. AGRAWAL, J.-C. FREYTAG, T. BOLLINGER, C. W. CLIFTON, S. DZEROSKI, J. HIPPE, D. KEIM, S. KRAMER, H.-P. KRIEGEL, U. LESER, et al.: *Data Mining: The Next Generation*. Report from a Dagstuhl Perspectives Workshop. Seminar Proceedings, Internationales Begegnungs- und Forschungszentrum (IBFI), Schloss Dagstuhl, Germany, 2005.

U. LESER: *A Query Language for Biological Networks*. Institut für Informatik der Humboldt-Universität zu Berlin, Informatik-Bericht Nr.187, 2005.

S. TRIßL: *Anfragen in Ontologien in relationalen Datenbanken (German)*. 17. GI-Workshop über 'Grundlagen von Datenbanken', Wörlitz, Germany, 2005.

S. SCHMEIER, J. HAKENBERG, A. KOWALD, U. LESER, E. KLIPP: *Recognition and Extraction of Entities related to Enzyme Kinetics*. BioText - a biomedical text mining workshop, London, UK, 2005.

S. SCHMEIER, A. KOWALD, J. HAKENBERG, U. LESER, E. KLIPP: *Finding Kinetic Data Using Text Mining*. 5th Workshop on Bioinformatics and Systems Biology, Berlin, 2005.

ST. RIECHE, U. LESER: *Versionierung in relationalen Datenbanken*. Studierendenprogramm der 11. GI-Fachtagung für Datenbanksysteme in Business, Technologie und Web, Karlsruhe, 2005.

L. FAULSTICH, U. LESER, A. LÜDELING: *Storing and querying historical texts in a relational database*. Institut für Informatik der Humboldt-Universität zu Berlin, Informatik-Bericht Nr.176, 2005.

L. FAULSTICH, U. LESER: *Implementing Linguistic Query Languages Using LoToS*. Institut für Informatik der Humboldt-Universität zu Berlin, Informatik-Bericht Nr.195, 2005.

Sonstige Aktivitäten

Ulf Leser

- Mitglied des Institutsrates und der Kommission für Lehre und Studium
- Mitglied der Kommission für Lehre und Studium des Instituts
- Mitglied des Leitungsgremiums des Berlin Centre for Genome-Based Bioinformatics
- Mitglied des Leitungsgremiums des interdisziplinären Zentrums für Biophysik und Bioinformatik der Humboldt-Universität Berlin
- Mitglied des Leitungsgremiums des interdisziplinären Zentrums für Sprachliche Bedeutung der Humboldt-Universität Berlin
- Mitglied des interdisziplinären Zentrums für Ubiquitäre Informationssysteme der Humboldt-Universität Berlin
- Mitglied des Leitungsgremiums des Verein "Berliner Centrum für genombasierte Bioinformatik e.V." (in Gründung)
- Mitglied des wissenschaftlichen Beirats der SMP Bioinformatics, NGFN-II
- Gutachter für Bioinformatics, BMC Bioinformatics, Briefings in Bioinformatics, Nucleic Acids Research, SIGMOD Record, VLDB Journal, Web Semantics Journal
- Mitglied im Programmkomitee von SAC2005, DILS2005, DIQ2005, SMDB2005, CompLife2005, ICDE2006, EDBT PhD Workshop 2006, WWW2006, ICDE2007
- General Chair des 3rd Int. Workshop on Data Integration for the Life Science (DILS) 2006
- Gutachter für die Deutsche Forschungsgemeinschaft und die Deutsch-Israelische Forschungskooperation

Jörg Hakenberg

- Organisation des 6. Treffens der bundesweiten AG Textmining. Berlin, 12. Dezember 2005.
- Gutachter für ICDE 2006, GIW 2005, SMBM 2005, Critical Reviews in Biotechnology
- Forschungsaufenthalt am European Bioinformatics Institute (EMBL-EBI, Hinxton/Cambridge, Großbritannien) in der Gruppe von Dr. Rebholz-Schuhmann, finanziert durch ein Doktorandenstipendium des Deutschen Akademischen Austauschdienstes (DAAD; Referenznummer D/05/26768).

Jana Bauckmann

- Mitorganisation des "Girls' Day" des Instituts in Zusammenarbeit mit Katja Tham.

Diplomarbeiten

SEBASTIAN SCHMEIER (Masterarbeit Bioinformatik): “Automated Recognition and Extraction of Entities Related to Enzyme Kinetics from Text”

THORSTEN VITT (Diplomarbeit): „Anfragen an komplexe Korpora“

PETER ADOLPHS (Diplomarbeit): „Maschinelle morphologische Analyse für das Deutsche. Ein Überblick.“

CONRAD PLAKE (Diplomarbeit): „Bestimmung von Text-Pattern für die Informationsextraktion“

Lehr- und Forschungseinheit

Künstliche Intelligenz

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

Leiter

PROF. DR. HANS-DIETER BURKHARD

Tel.: (030) 2093 3165

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 Mitarbeiterinnen und Mitarbeiter

DIPL.-INF. DANIEL GÖHRING

DIPL.-MATH. MANFRED HILD

DIPL.-PHYS. JAN HOFFMANN

DIPL.-INF. MATTHIAS JÜNGEL

DR. GABRIELA LINDEMANN-VON TRZEBIATOWSKI

DIPL.-INF. MIRJAM MINOR

DR. DAGMAR MONETT DÍAZ

DIPL.-INF. KAY SCHRÖTER

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

Technische Mitarbeiterin

DIPL.-MATH. OLGA SCHIEMANGK

Promotionsstudenten

DIPL.-INF. MOHAMMED MAYYAS

DIPL.-INF. HELMUT MYRITZ

M. SC. JESÚS EMETERIO NAVARRO BARRIENTOS

M. SC. VICTOR EMANUEL UC CETINA

Tutorinnen und Tutoren

BENJAMIN ALTMEYER

FERRY BACHMANN

RALF BERGER

MARTIN BERTHEAU

CHRISTIAN BRANDT

MONIKA DOMANSKA

MATTHIAS GAUDIG

VIVIANA GOETZKE

PETER HARTIG

CLAUDIA HÄRTEL

DANIEL HEIN
DANIELA HÖHNE
JULIA JABLINSKI
THOMAS KRAUSE
JULIA KÖHLER
ROBIN MALITZ
HEINRICH MELLMANN
JAKOB VON RECKLINGHAUSEN
DANIEL RENZ
TANJA RICHTER
ANSELM RINGLEBEN
MICHAEL SPRANGER
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 umfas-

sende 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.

Unsere Projekte im Bereich RoboCup haben sich in diesem Jahr erneut sehr erfolgreich an internationalen und nationalen Wettbewerben beteiligt. Das GermanTeam, das wir in Kooperation mit den Universitäten Bremen, Darmstadt und Dortmund entwickeln, wurde beim RoboCup 2005 in Japan Weltmeister in der „Sony four-legged Robot League“ und gewann beim „Technical Challenge“ den 3. Platz. Das Aibo Team Humboldt – unser eigenes Team in der Liga der vierbeinigen Roboter – wurde Vizemeister bei den RoboCup GermanOpen in Paderborn.

Im Jahr 2006 werden die 10. Weltmeisterschaften des RoboCup vom 14.-20. Juni in Bremen stattfinden. Wir koordinieren die vom BMBF geförderte „Bundesweite Initiative RoboCupJunior“ zur Vorbereitung von Jugendlichen auf die Wettbewerbe.

Die Lehr- und Forschungseinheit beteiligt sich an den beiden DFG-Schwerpunktprogrammen „Sozionik: Erforschung und Modellierung Künstlicher Sozialität“ und „Kooperierende Teams mobiler Roboter in dynamischen Umgebungen (RoboCup)“. Es bestehen langjährige Kooperationen mit anderen Forschungseinrichtungen und Partnern in der Wirtschaft. Insbesondere engagieren wir uns für die Zusammenarbeit am Standort Adlershof.

Lehre

Veranstaltungen im Grundstudium

- *Vorlesung* „Praktische Informatik II“ (PROF. H.-D. BURKHARD, SoSe 2005)
- *Übungen* „Praktische Informatik II“ (DIPL.-INF. DANIEL GÖHRING SoSe 2005)
- *Übungen* „Praktische Informatik II“ (DR. G. LINDEMANN- V. TRZEBIATOWSKI SoSe 2005)
- *Übungen* „Praktische Informatik II“ (DIPL.-INF. MIRIAM MINOR SoSe 2005)
- *Übungen* „Praktische Informatik I“ (DR. G. LINDEMANN- V. TRZEBIATOWSKI, WiSe 2005/2006)

Kernveranstaltungen (Halbkurse)

- *Vorlesung* „Einführung in die Künstliche Intelligenz“ (PROF. H.-D. BURKHARD, WiSe 2005/2006)
- *Übungen* „Einführung in die Künstliche Intelligenz“ (DIPL.-INF. D. GÖHRING, WiSe 2005/2006)
- *Vorlesung* „Kognitive Robotik“ (PROF. H.-D. BURKHARD WiSe 2005/2006)
- *Übungen* „Kognitive Robotik“ (PROF. H.-D. BURKHARD WiSe 2005/2006)

Seminare

- „Modelle für Wahrnehmung“ (PROF. H.-D. BURKHARD, SoSe 2005)
- „Neuronale Netze und autonome Roboter“ (M. HILD, WiSe 2005/2006)

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, in der Geoinformatik und im eCommerce. Neben entscheidungsunterstützenden Systemen haben sich in den letzten Jahren auch Assistenzsysteme für Wissensmanagement und Erfahrungsmanagement 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.

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.

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).

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, das Verhalten in dynamischen Umwelten, und für autonome intelligente Roboter.

Projekte

Projekt: Integration kooperationsfähiger Agenten in komplexen Organisationen – INKA

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

Beteiligte Mitarbeiter: DR. DAGMAR MONETT DÍAZ, DIPL.-INF. KAY SCHRÖTER, DIPL.-KFM. DIPL.-INF. DIEMO URBIG, CHRISTIAN BRANDT, NORA HANS, CLAUDIA HÄRTEL, ROBIN MALITZ, ANSELM RINGLEBEN, DANIEL RENZ, CHRISTIAN WIECH

Zusammenarbeit: PROF. DR. WERNER RAMMERT, Technische Universität Berlin, Institut für Soziologie, Fachgebiet Techniksoziologie

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 agentenbasierte 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.

Nachdem im Jahr 2004 ein erster Prototyp unseres Multi-Agenten Systems (MAS) unter Nutzung der Java-basierten Plattform JADE fertig gestellt und im Experiment angewendet werden konnte, wurde der Prototyp im Jahr 2005 stark erweitert. Die Entscheidungsfindung der Agenten ist gemäß des C-IPS Ansatzes in Komponenten zur Wahl von Verhand-

lungsthemen (engl.: Issue), Verhandlungspartnern (engl.: Partner) und Verhandlungsschritten (engl.: Step) gegliedert. Das C im Kürzel C-IPS steht für die explizite Modellierung der Beschränkungen (engl.: Constraints) dieser Wahlprozesse. Die vorhandenen automatischen Komponenten wurden um Module ergänzt, die es den Agenten mittels Methoden des Fallbasierten Schließens (FBS) ermöglichen, aus früheren Verhandlungserfahrungen zu lernen, erfolgreiche Verhandlungsmuster zu wiederholen, schließlich Verhaltensroutinen auszubilden und diese ggf. auch wieder zu verwerfen. Im Rahmen der interdisziplinären Konzeption wurden soziologische Konzepte zur Routinebildung adäquat mit FBS-Techniken zur Anwendung in einem MAS verknüpft. Manuelle Komponenten dienen wie bisher menschlichen Verhandlungsteilnehmern als Schnittstelle zum System.

In einem zweiten Interaktivitätsexperiment wurde das Zusammenspiel dieser zu Routinen fähigen Agenten mit Menschen untersucht. Das Experiment erstreckte sich über zwei Tage, in denen mehrere Szenarien durchgespielt wurden. Wie schon im letzten Jahr waren 12 menschliche Probanden an dem Experiment beteiligt. Die Ergebnisse sollen insbesondere auch mit denen des vorigen Interaktivitätsexperimentes verglichen werden, um herauszufinden welche Einfluss die Berücksichtigung von Verhandlungserfahrungen und Routinen auf das sozionische Gesamtsystem haben.

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

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

Kooperation: DIPL.-MATH. JAN LORENZ, Universität Bremen

Einstellungen bilden im Marketing eine Grundlage zur Prognose von Verhalten. Bei der Modellierung solcher Dynamiken kommen sowohl sehr komplexe multidimensionale Ansätze als auch einfache eindimensionale Ansätze zum Einsatz. In diesem Jahr stand ein multidimensionaler Ansatz im Vordergrund. Basierend auf der Annahme hierarchischer Verarbeitungsprozesse konnte im Rahmen einer Diplomarbeit eine neue Dynamik identifiziert werden. Diese Dynamik führt zu einer Balancierung von Einstellungen und zur Ausbildung von tendenziell extremeren Meinungen in Gruppen. Die Effekte wurden sowohl in agentenbasierten Modellen als auch in auf Markovketten basierenden Modellen gefunden und untersucht.

Projekt: Architekturprinzipien und Lernen auf der Basis mentaler Modelle

Ansprechpartner: PROF. DR. HANS-DIETER BURKHARD, DIPL.-INF. MATTHIAS JÜNGEL

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.-INF. MATTHIAS JÜNGEL

Beteiligte Mitarbeiter: DIPL.-INF. DANIEL GÖHRING, DIPL.-PHYS. JAN HOFFMANN, BENJAMIN ALTMAYER, THOMAS KRAUSE, MICHAEL SPRANGER

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 (AIBO von SONY) gegeneinander an. Seit 2001 arbeiten wir mit den oben genannten Universitäten in der deutschen Nationalmannschaft „German Team“ zusammen.

Ein Forschungsschwerpunkt ist der Wahrnehmungsprozess des Roboters, also die Datenaufnahme der Sensoren (Kamera usw.), 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. 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 den diesjährigen GermanOpen in Paderborn belegte unsere Mannschaft (Aibo Team Humboldt) den 2. Platz. Bei den Weltmeisterschaften in Osaka, Japan, wurde das GermanTeam WELTMEISTER.

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 in der Vorlesung „Kognitive Robotik“ Anwendung.



Das German Team gewinnt den RoboCup 2005 in Osaka.

2 AT Humboldt (RoboCup Simulationsliga)

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

Beteiligte Mitarbeiter: RALF BERGER, DANIEL HEIN

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

Informationen: <http://www.robocup.de/AT-Humboldt/index.shtml>

In der Simulationsliga treten pro Team 11 autonome Programme (Softwareagenten), sowie ein Coach-Programm, auf einem virtuellen Spielfeld gegeneinander an. Die Raum- /Zeitsimulation des Spielgeschehens erfolgt durch den „Soccerserver“. Er nimmt die Kommandos der einzelnen Agenten entgegen und sendet ihnen Informationen über Wahrnehmung. Gegenstand unserer Arbeiten in diesem Jahr war vor allem die weitere Entwicklung und Analyse der von uns entwickelten „Doppelpass-Architektur“ sowie deren Migration in die neue Agentenplattform „ATH-3D“ für den neuen 3D-Soccerserver mit ODE-Simulation. Auch in diesem Jahr konnte die Größe der darin implementierten Verhaltensbäume wieder gesteigert werden, wobei sich die letztjährigen Verbesserungen im Laufzeitverhalten der Architektur sowie Erweiterungen an der Beschreibungssprache sehr positiv auswirkten. Ein weiterer Schwerpunkt war der Einsatz von Verfahren des Reinforcement-Learning für low- und midlevel Controller im Simulationsagenten. Im Rahmen einer Diplomarbeit wurde eine umfangreiche generische Softwarebibliothek entwickelt und deren effizienter Einsatz an Beispielen dargestellt. Speziell für die RoboCup Weltmeisterschaften wurde das Lernsystem um Fähigkeiten der automatischen, unüberwachten Durchführung ganzer Reihen von Lernexperimenten mit teilweise selbstadaptiven Lernparametern erweitert.

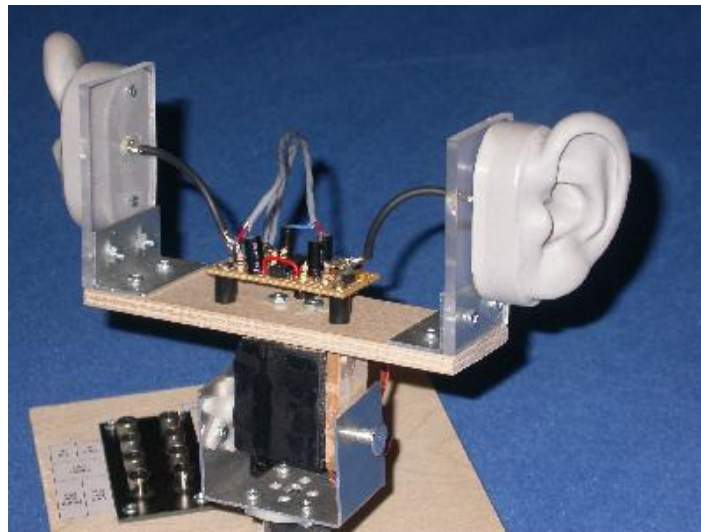
Bei den German Open in Paderborn konnte unser Team im klassischen 2D-Wettbewerb den zweiten Platz erringen. Bei den diesjährigen Weltmeisterschaften in Osaka (Japan) kam unser 3D-Team auf den achten Platz und ist damit für die Weltmeisterschaft 2006 in Bremen wieder qualifiziert.

Projekt: Evolution und Robotik

Ansprechpartner: DIPL.-MATH. MANFRED HILD

Zusammenarbeit: PROF. FRANK PASEMANN, Fraunhofer Institut Autonome Intelligente Systeme (AIS), Sankt Augustin.

Die Evolutionsmechanismen der Natur (natürliche Auslese, Kreuzung und Mutation) lassen sich auch in der Robotik verwenden und können mobile autonome Roboter mit interessanten Verhaltensweisen hervorbringen. Im Rahmen der dreiteiligen Seminarreihe „Evolution und Robotik“ und der beiden Projekthalbkurse „Kognitive Robotik“ wurden verschiedene Experimente zum Artificial Life Ansatz durchgeführt.



*Durchführung von Evolutionsexperimenten:
Phonotaxis.*

Bei der künstlichen Evolution rekurrenter, neuronaler Netze werden zunächst die Merkmale intelligenter Verhaltensweisen identifiziert und ohne weitere Annahmen über zugrundeliegende Funktionsmodule evolviert. Die Analyse der hervorgebrachten, neurodynamischen Steuerungsstrukturen und der Vergleich mit Erkenntnissen aus den modernen Kognitionswissenschaften bietet die Möglichkeit, universelle Neuromodule intelligenten Verhaltens zu identifizieren.

Neben den inhaltlichen Aspekten, zeigt die Durchführung der Lehrveranstaltungen und Betreuung der wissenschaftlichen Arbeiten außerdem, dass die eingesetzte Methodik didaktisch besonders geeignet ist, Studenten die Thematik „Kognitive Robotik“ zu vermitteln.

Projekt: Maßnahmen des RoboCup zum Wissenschaftsjahr

Ansprechpartner: PROF. DR. HANS-DIETER BURKHARD, DIPL.-MATH. MANFRED HILD

Beteiligte Mitarbeiter: BENJAMIN ALTMAYER, FERRY BACHMANN, MARTIN BERTHEAU, MONIKA DOMANSKA, MATTHIAS GAUDIG, PETER HARTIG, DANIELA HÖHNE, JULIA JABLINSKI, TANJA RICHTER

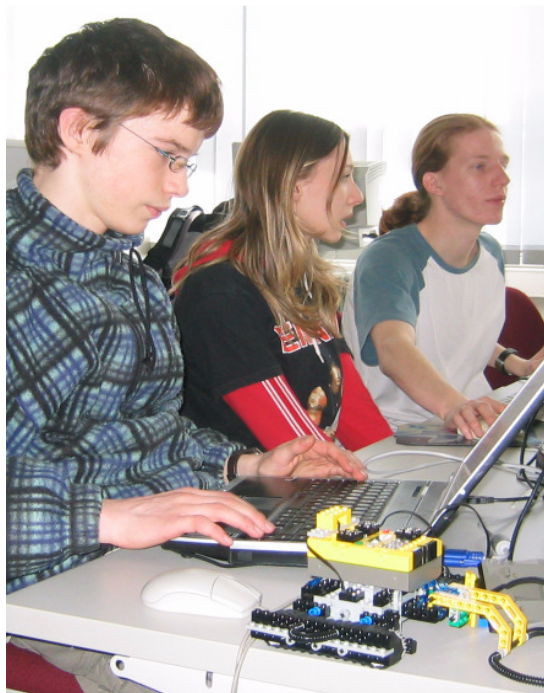
Zusammenarbeit: PD DR. UBBO VISSER, Universität Bremen, Technologie-Zentrum Informatik (TZI)

Forschungsförderung: BMBF, im Rahmen des Wissenschaftsjahres Informatikjahr 2006

Informationen: <http://www.robocupjunior.de>

Die 10. Weltmeisterschaften des RoboCup vom 14.-20. Juni 2006 in Bremen werden ein Höhepunkt im Informatikjahr sein. Bereits im Jahre 2005 haben wir im Rahmen der BMBF-geförderten „Bundesweiten Initiative RoboCupJunior“ damit begonnen Jugendliche zur IT-Forschung zu animieren („Bau dir deinen Roboter“).

Die Maßnahmen zielen darauf ab, die Anzahl teilnehmender Teams aus Deutschland massiv zu erhöhen und die Vorbereitung der Teams auf die Weltmeisterschaft durch inhaltlichdidaktische Unterstützung zu fördern. Mit Hilfe von Informationsveranstaltungen, TV-Auftritten, Print-Reportagen und speziellen Events soll zugleich eine breite Wahrnehmung dieser Aktivitäten in der Bevölkerung erreicht werden.



Schüler im RoboCup Junior

Alle Aktionen werden durch unsere zentrale Projektgruppe koordiniert. Sie organisiert Workshops, erarbeitet und verteilt Informationsmaterial, und sie stellt die Vernetzung der Projekte her. Dabei arbeitet sie eng mit den erfahrenen Gruppen zusammen (TCSI, RoCCI, Roberta usw.) und vermittelt Beratung, Anleitung und Erfahrungsaustausch. Als zentrales Informationsportal dient hierbei die spezielle Projektwebseite www.robocupjunior.de. Während der Projektlaufzeit im Jahr 2005 wurden in folgenden Städten Informationsver-

anstaltungen und Workshops durchgeführt: Berlin, Bernau, Braunschweig, Bremen, Claus-
thal, Dresden, Göttingen, Ilmenau, Jena, Kiel, Köthen, Lüneburg, Magdeburg, Merseburg,
Münster, Naumburg, Potsdam, Schwerin und Stuhr-Brinkum.

Projekt: Theorie und Techniken des fallbasierten Schließens

Ansprechpartner: PROF. DR. HANS-DIETER BURKHARD

Beteiligte Mitarbeiter: DIPL.-INF. MIRJAM MINOR, DIPL.-INF. KAY SCHRÖTER

Zusammenarbeit: Fell & Kernbach, Dipl.-Inf. SANDRO KÖPPEN, UNIVERSITÄT
HILDESHEIM, DIPL.-INF. ALEXANDRE HANFT

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, Vernetzung und Wartung von Erfahrungswissen und zur Integration von Ontologien bilden einen weiteren Schwerpunkt der aktuellen Forschungsarbeiten. Die Ergebnisse werden in den anwendungsorientierten Projekten intensiv genutzt.

Geoinformatik: 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 wird ein fallbasiertes System entwickelt, das aus Daten von geologischen Aufschlüssen (z.B. Bohrprofilen, geologischen Karten) ein 3D-Modell über den Aufbau eines Untergrundes erstellt. Dazu werden bekannte Strukturmodelle aus der Fallbasis hinzugezogen und deren Struktur wieder verwendet. In Erweiterung eines üblichen Retrieval-Verfahrens wird eine grobe Vorauswahl an Fällen ermittelt, für die eine detaillierte Ähnlichkeitsberechnung überhaupt sinnvoll ist. Erst im zweiten Schritt der Berechnung werden die ähnlichsten Fälle sortiert.

Collaborative Maintenance

Dieses Teilprojekt beschäftigt sich mit der Entwicklung einer kollaborativen Strategie zur Wartung von fallbasierten Assistenzsystemen. Die Benutzer selbst werden in einem automatisch ausgelösten Verfahren gebeten, ältere Fälle zu begutachten und gegebenenfalls zu verbessern oder zu löschen. So wird die Wartung und Pflege der Falldaten zwar von Menschen ausgeführt, aber automatisch gesteuert.

Vergleich von Musiksammlungen anhand von Tonträgermetadaten

In diesem Teilprojekt werden Methoden des fallbasierten Schließens eingesetzt, um Sammlungen von Tonträgerdaten wie CDs, DVDs usw. zu modellieren und ähnliche Sammlungen zu bestimmen. Semantische Netze mit objektiven und subjektiven Beziehungen zwischen Begriffen bilden die Grundlage für die Ähnlichkeitsberechnung, um zum Beispiel Kandidaten für den Austausch von Tonträgern zu finden.

Vages Matchen in Datenbanken

Ziel dieses Teilprojektes ist es, die Funktionalität eines Case-Retrieval-Netzes in einem Datenbanksystem nachzubilden. Der Einsatz eines Datenbanksystems erlaubt eine effizien-

tere Pflege der Falldaten. Zunächst wurde eine Datenstruktur entwickelt, die es erlaubt die Falldaten und die für den Aufbau eines Case-Retrieval-Netzes erforderlichen Informationen für beliebige Fallformate, die sowohl textuelle Abschnitte, als auch Attribut-Wert-Tupel enthalten können. Darauf aufbauend wurden zwei unterschiedliche Ansätze verfolgt: Zum einen wurde das Retrieval mit Stored Procedures und Funktionsbibliotheken umgesetzt; zum anderen wurde es ausschließlich mit SQL-Anfragen realisiert. Ersteres bietet den Vorteil höherer Geschwindigkeit und ist damit für Produktionssysteme geeignet, während letzteres auf einfache, effiziente Weise Modifikationen an Fallstruktur und Retrieval-mechanismus ermöglicht und somit für Entwicklungssysteme geeignet ist. Beide Ansätze wurden prototypisch umgesetzt und in Form von zwei Diplomarbeiten dokumentiert.

Projekt: Intelligente Techniken im Knowledge Management

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

Beteiligte Mitarbeiter: DR. DAGMAR MONETT DÍAZ, FABIAN MÜLLER

Zusammenarbeit: DR. MARIO LENZ, DIPL. INF. WOLF-DIETRICH MATERNA, empolis, Zentrum für Biomaterialien der Universität Havanna,

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. Eine ganzheitliche Herangehensweise, die auch psychosoziale Aspekte berücksichtigt, ist dabei unumgänglich. 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.

- ***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 das Langzeitverhalten eines solchen Systems im praktischen Einsatz.

- ***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. Weblog-Analysen untersuchen das Benutzerverhalten im Hinblick auf organisatorische Massnahmen wie Gespräche, Vorträge, schriftliche Gruppenarbeit oder Email-Kommunikation sowie soziotechnische Massnahmen wie Einbettung des Systems in den normalen Arbeitskontext der Studierenden oder zusätzliche Kommentarfelder. Das System hat nach einer aktuellen Überarbeitung des Contents stabile Benutzerzahlen erreicht, so dass zunächst keine weiteren organisatorischen oder soziotechnischen Massnahmen mehr erforderlich sind.

- ***Agentenbasierte Konfiguration von metaheuristischen Algorithmen***

Dieses Thema betrachtet einen agentenbasierten Ansatz zur Konfiguration von Algorithmen. Der Hauptfokus des Ansatzes, die +CARPS-Architektur (+CARPS: Multi-Agent System for Configuring Algorithms in Real Problem Solving), liegt darin, agentenbasierte Verfahren für die Anpassung algorithmischer Parameter anzuwenden. Vor allem wird die Parameterkonfiguration (d.h. die Anpassung der unbekannt Parameter) mittels verteilter, autonomer, kooperierender und spezialisierter Agenten erreicht. Die Agenten führen ein Optimierungsverfahren aus. Es handelt sich hierbei um ein Verfahren, das Parameterkonfigurationen optimiert. Die Hauptidee hinter dem vorgeschlagenen Ansatz liegt in der Vorstellung, Agenten mit der Fähigkeit zur "Anpassung metaheuristischer Parameter" in der Art auszustatten, wie dies ein Experte tun würde. Die agentenbasierte Konfiguration von Algorithmen ist ein passender, robuster und wirksamer Ansatz, der selbstständig das Verfahren der Konfiguration durchführt, wobei die Agenten die Algorithmen automatisch ausführen und konfigurieren.

- ***Grundlagen und Anwendungen des Erfahrungsmanagement***

Erfahrungsmanagement ist eine Spezialform von Wissensmanagement, die sich mit Erfahrungswissen beschäftigt. Das Anwenden vorhandener Erfahrung trägt dazu bei, Probleme erfolgreich zu lösen. Ziel von Erfahrungsmanagement 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.

Projekt: Unterstützung der Durchführung klinischer Studien durch ein Multi-Agenten-System

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) wird ein Multi-Agenten-System (MAS) installiert, welches die Durchführung klinischer Studien überwacht und bei deren Ablauf helfend eingreift. Ziel hierbei ist es die Ausfallquote von Studienteilnehmern einer Studie zu minimieren und somit die Aussagekraft der Studie zu stärken. Allen beteiligten Personen (Patienten, Ärzte, Studienpersonal) wird es ermöglicht mit den Agenten via email zu kommunizieren, um einen schnellen Informationsfluss zu garantieren. Die Qualität des MAS soll sowohl durch den Vergleich der Ausfallquoten zu früheren klinischen Studien als auch direkt durch das medizinische Personal beurteilt werden. Neben der Mensch-Maschine-Interaktion sind Ontologien und Robustheit von Agentensystemen weitere Schwerpunkte des Projekts.

Gastwissenschaftler

Dipl.-Inf. Mohammed Mayyas, externer Promotionsstudent

Herr Mayyas beschäftigt sich mit „Experience Management mit CBR (Case-Based Reasoning) Methoden“. Er arbeitet auf dem Gebiet Textuelles Fallbasiertes Schließen, Informationssysteme und Semantic Lexicon Learning. Ziel seiner Arbeit ist es, den Inhalt eines Lexikons, mit Hilfe der ontologiebasierten Wissensextraktion, automatisch zu erweitern bzw. zu verbessern. Die Daten werden aus einer bestimmten Domäne durch verschiedenen Methoden extrahiert und für das CBR vorbereitet.

M. Sc. Jesús Emeterio Navarro Barrientos, Centro de Investigacion en Matematicas, Guanajuato, Guanajuato, Mexiko

Forschungsförderung: DAAD

Herr Navarro forscht im Bereich der Multi-Agenten-Systeme und der Agent-Based Computational Economics (ACE). Im Speziellen entwickelte er Modelle und Simulationsprogramme, die es ermöglichen, Investitionsentscheidungen sowie Netzwerke von koalierenden Agenten unterschiedlicher "Intelligenz" zu untersuchen. Ziel seiner Forschung ist es, solche Investitionstrategien und Interaktion-Netzwerke zu verstehen und globale Regelmäßigkeiten zu demonstrieren, die sich "bottom-up" aus der Performanz der Agenten und den Netzwerken ergeben.

M. Sc. Victor Emanuel Uc Cetina, Universität Yukatan, Merida, Mexiko

Forschungsförderung: PROMEP

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 hochdynamischen 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

H.-D. BURKHARD: *Experiences with the design and development of autonomous agents and robots for RoboCup* (Abstrakt) Proc. of CS&P 2005, 69-70.

H.-D. BURKHARD: *Warum sind wir schlauer als Roboter*, In: Humboldt Kinderuni, Lohmann-Verlag 2005, 14-23.

H.-D. BURKHARD, U. VISSER, M. JÜNGEL, A. BREDENFELD, T. CHRISTALLER: *Herausforderung für KI und Robotik*, erscheint in KI Heft2/2006.

U. DÜFFERT, J. HOFFMANN: *Reliable and Precise Gait Modelling for a Quadruped Robot*. In: 9th International Workshop on RoboCup 2005 (Robot World Cup Soccer Games and Conferences), Osaka, Japan, Lecture Notes in Artificial Intelligence. Springer (to appear).

D. GÖHRING: *Object Modeling in Dynamic Environments by Mobile Agents*. In: Proceedings of Concurrency, Specification and Programming (CS&P) 2005.

A. HANFT, M. MINOR: *A Low-Effort, Collaborative Maintenance Model for Textual CBR*. In Steffi Brüninghaus, editor, *ICCBR 2005 Workshop Proceedings*, pages 138 -- 149, Chicago, Illinois, August 2005, DePaul University, Chicago, USA.

J. HOFFMANN, M. SPRANGER, D. GÖHRING, M. JÜNGEL: *Making Use Of What You Don't See: Negative Information In Markov Localization*. In: Proceedings of IEEE/RSJ International Conference of Intelligent Robots and Systems (IROS) 2005. (to appear).

J. HOFFMANN, M. SPRANGER, D. GÖHRING, M. JÜNGEL: *Exploiting the Unexpected: Negative Evidence Modeling and Proprioceptive Motion Modeling for Improved Markov Localization*. In: 9th International Workshop on RoboCup 2005 (Robot World Cup Soccer Games and Conferences), Lecture Notes in Artificial Intelligence. Springer (to appear).

J. HOFFMANN, M. SPRANGER, D. GÖHRING, M. JÜNGEL: *Negative Information and Proprioception in Monte Carlo Self-Localization for a 4-Legged Robots*. In Nineteenth International Joint Conference on Artificial Intelligence (IJCAI) 2005, Workshop on Agents in Real-Time and Dynamic Environments.

LINDEMANN, G., OSSOWSKI, S. PADGET, J., SALCEDA, J.V. (eds.): Proceedings Workshop "ANI@REM -Agents, Norms and Institutions for Regulated Multiagent Systems", AAMAS 2005, Utrecht, The Netherlands, 2005

M. MINOR: *Introduction Strategy and Feedback from an Experience Management Project*. In Klaus-Dieter Althoff, Andreas Dengel, Ralph Bergmann, Markus Nick, and Thomas Roth-Berghofer, editors, *Professional Knowledge Management. WM 2005 post-conference proceedings*, LNAI 3782, pages 284 -- 292, Springer-Verlag, 2005.

M. MINOR: *Assistant Agents With Personal Ontologies*. In Juriaan van Diggelen, Virginia Dignum, Ludger van Elst, and Andreas Abecker, editors, *Proceedings of AAMAS 2005 Workshop AMKM-2005*, pages 41 -- 51, Utrecht, The Netherlands, Universiteit Utrecht, The Netherlands, July 2005.

M. MINOR, CH. BIERMANN: *Case-Acquisition and Semantic Cross-Linking for Case-Based Experience Management Systems*. In Du Zhang, Taghi M. Khoshgoftaar, and Mei-Ling Shyu, editors, Proceedings of the 2005 IEEE International Conference on Information Reuse and Integration (IRI-2005), pages 433 -- 438, IEEE Systems, Man and Cybernetics Society (SMC), 2005.

M. MINOR, CH. BIERMANN: *Learning and Linking Textual Cases*. In Steffi Brüninghaus, editor, *ICCBR 2005 Workshop Proceedings*, pages 128 -- 137, Chicago, Illinois, August 2005, DePaul University, Chicago, USA.

M. MINOR, S. KÖPPEN: *Design of Geologic Structure Models with Case Based Reasoning*. In Ulrich Furbach, editor, *KI2005: Advances in Artificial Intelligence. Proceedings*, LNAI 3698, pages 79 -- 91, Springer-Verlag, 2005.

M. MEISTER, D. URBIG, K. SCHRÖTER, R. GERSTL: Agents Enacting Social Roles. Balancing Formal Structure and Practical Rationality in MAS Design. In: K. Fischer/ M. Florian/ T. Malsch (Eds.): *Socionics: Scalability of Complex Social Systems*. Lecture Notes in Computer Science, 3413, Berlin, Heidelberg, Springer, 2005, 104-131.

H. MYRITZ, G. LINDEMANN: *Monitoring Patient Behaviour in Clinical Studies by Multi-Agent Techniques*. In: Proceedings of Concurrency, Specification and Programming (CS&P) 2005 Ruciane, Polen.

K. SCHRÖTER, D. URBIG, N. HANS: Social Formation of Negotiation Space and Groups for Non-Isolated Multilateral Negotiations. *Fundamenta Informaticae*, vol. 67, no. 1-3, IOS Press, 2005, pp. 187-201.

D. URBIG: Weight-based negotiation mechanisms: Balancing personal utilities, *Fundamenta Informaticae*, vol. 67, no.1-3, IOS Press, 2005, pp. 271-285.

D. URBIG, R. MALITZ: Dynamics of structured attitudes and opinions. Troitzsch, K.G. (ed.): *Representing Social Reality. Pre-Proceedings of the Third Conference of the European Social Simulation Association (ESSA' 05)*, September 5-9, Koblenz, Germany, 2005, pp. 206-212.

D. URBIG, K. SCHRÖTER: From Full Autonomy to Manual Control. Specifying Negotiating Agents. *Proceedings of the Workshop Concurrency, Specification & Programming (CS&P 2005)*, Ruciae-Nida, Poland, September, 2005, pp. 520-531.

D. URBIG, K. SCHRÖTER: Negotiating Agents: From Full Autonomy to Degrees of Delegation. *Proceedings of the 4th International Joint Conference on Autonomous Agents and Multi Agent Systems (AAMAS 2005)*, Utrecht, July 25-29, ACM, pp. 1167-1168.

Vorträge

H.-D. BURKHARD: *RoboCup Junior Förderung im Wissenschaftsjahr 2006*, Informations-Workshop Roboter in Schulen - Initiativen und Erfahrungen Paderborn, 09.04.2005.

H.-D. BURKHARD: *Software development experience of the RoboCup project at HU*, 5th Workshop Software Engineering Education and Reverse Engineering, 29.8. - 03.09.2005 Baile Herculane (Romania) from 28 August to 4.

H.-D. BURKHARD, MANFRED HILD, JOHANNES KLOTZ, SEBASTIAN POLLYy: *RoboCupJunior zum Wissenschaftsjahr 2006*, 10.-11-6.2005, EduNetwork 05, Naumburg.

H.-D. BURKHARD: *Programming Bounded Rationality*, 25.-28.9.2005 University Days, Beida, Peking.

H.-D. BURKHARD: *50 Jahre Künstliche Intelligenz*, 08.12.2005 OTA Berlin.

DANIEL GÖHRING: *Object Modeling in Dynamic Environments by Mobile Agents*. In: *Proceedings of Concurrency, Specification and Programming (CS&P) 2005*, Ruciane, Polen, 28.09.2005.

DANIEL GÖHRING: *Negative Information and Proprioception in Monte Carlo Self-Localization for a 4-Legged Robots*. Nineteenth International Joint Conference on Artificial Intelligence (IJCAI) 2005, Workshop on Agents in Real-Time and Dynamic Environments, Edinburgh, Großbritannien, 30. Juli 2005.

J. HOFFMANN: *Reliable and Precise Gait Modelling for a Quadruped Robot*. In: 9th International Workshop on RoboCup 2005 (Robot World Cup Soccer Games and Conferences), Osaka, Japan, Lecture Notes in Artificial Intelligence. Springer (to appear).

J. HOFFMANN: *Making Use Of What You Don't See: Negative Information In Markov Localization*. International Conference of Intelligent Robots and Systems (IROS) 2005, Montreal, Kanada.

J. HOFFMANN: *Exploiting the Unexpected: Negative Evidence Modeling and Proprioceptive Motion Modeling for Improved Markov Localization*. In: 9th International Workshop on RoboCup 2005 (Robot World Cup Soccer Games and Conferences), Osaka, Japan.

M. MINOR: *Introduction Strategy and Feedback from an Experience Management Project*. Conference on Professional Knowledge Management, Kaiserslautern, 11.-13. April 2005.

M. MINOR: *Erfahrungsmanagement mit fallbasierten Assistenzsystemen*. Forschungsseminar Wirtschaftsinformatik, Universität Trier, 14. April 2005, Forschungsseminar Informatik, Universität Kassel, 15. April 2005.

M. MINOR: *Assistant Agents With Personal Ontologies*. AAMAS Workshop on Agent-Mediated Knowledge Management (AMKM 2005), Utrecht, Niederlande, 25. Juli 2005.

M. MINOR: *Case-Acquisition and Semantic Cross-Linking for Case-Based Experience Management Systems*. IEEE International Conference on Information Reuse and Integration (IRI-2005), Las Vegas, USA, 15.-17. August 2005.

M. MINOR: *Learning and Linking Textual Cases*. ICCBR 2005 Workshop on Textual Case-Based Reasoning, Chicago, USA, 23.-26. August 2005.

M. MINOR: *Design of Geologic Structure Models with Case Based Reasoning*. 28th Annual German Conference on Artificial Intelligence (KI 2005), Koblenz, 11.-14. September, 2005.

H. MYRITZ, G. LINDEMANN: *Monitoring Patient Behaviour in Clinical Studies by Multi-Agent Techniques*. In: Proceedings of Concurrency, Specification and Programming (CS&P) 2005 Ruciane, Polen, 28.09.2005.

J.E. NAVARRO-BARRIENTOS: *Risk propensity Strategies in Random Environments*. Poster auf der 69. Jahrestagung der Deutschen Physikalischen Gesellschaft, AKSOE, Berlin, März 2005.

K. SCHRÖTER, M. MEISTER: *Routinehandeln für Soziale Agenten*. Forschungskolloquium Technik- und Wissenschaftsforschung, TU Berlin, Institut für Soziologie, 12. Januar 2005.

K. SCHRÖTER, M. MEISTER: *Ergebnisse⁺ des INKA-Projektes*. Jahrestreffen des SPP Sozionik, Kloster Seeon, 24. – 27. Juni 2005.

D. URBIG, K. SCHRÖTER: *Negotiating Agents: From Full Autonomy to Degrees of Delegation*. Poster presentation at 4th International Joint Conference on Autonomous Agents and Multi Agent Systems (AAMAS 2005), Utrecht, Niederlande, 25.-29. Juli 2005.

D. URBIG: *Dynamics of structured attitudes and opinions*. Third Conference of the European Social Simulation Association (ESSA 2005), Koblenz, 5.-9. September 2005.

C. Wiech: *From Full Autonomy to Manual Control. Specifying Negotiating Agents*. Workshop on Concurrency, Specification and Programming 2005 (CS&P 2005), Ruciae-Nida, Poland, September 2005.

Präsentationen zum RoboCup

Demonstrationen des Aibo Team Humboldt gab es bei folgenden Veranstaltungen:

- *Girlsday*, Berlin - Adlershof, 28. April 2005
- *Tag der Mathematik*, Berlin - Adlershof, 04. Juni 2005

- *Lange Nacht der Wissenschaften*, Berlin - Adlershof, 11. Juni 2005
- *Adlershof Roadshow*, 08. September 2005
- *Tage der Forschung in Adlershof*, 22. - 23. September 2005

Im Rahmen der vom BMBF geförderten bundesweiten Initiative RoboCupJunior zum Wissenschaftsjahr *Informatikjahr 2006* wurden im Jahr 2005 bereits Informationsveranstaltungen und Workshops in folgenden Städten durchgeführt (in chronologischer Reihenfolge):

Bremen, Naumburg, Köthen, Potsdam, Berlin, Münster, Merseburg, Stuhr-Brinkum, Schwerin, Ilmenau, Jena, Dresden, Göttingen, Lüneburg, Braunschweig, Clausthal, Bernau, Kiel.

Pressespiegel (Auswahl)

- Interview zu Sozionik und INKA-Projekt. „Ersatzmensen oder Menschenersatz?“, WDR, Q21 - Wissen für morgen, 7. Juni 2005.
- „Robocup entwickelt sich zum Renner“, Handelsblatt, 10. April 2005.
- „RoboCup Paderborn 2005“, WDR – Servicezeit: Technik, 14. April 2005.
- „RoboCup-WM, Deutschland ist (mehrfacher) Weltmeister!“, Heise-Online, 17. Juli 2005.
- „WM im Roboterfußball – Deutschland ist Weltmeister“, Spiegel-Online, 18. Juli 2005.
- „Berlins Robo-Kicker bei der WM nicht zu schlagen“, Berliner Morgenpost, 19. Juli 2005.
- „Wir wollen der Welt zeigen, was wir können“, spektrumdirekt, 18. August 2005.
- „Tage der Forschung 2005 in Adlershof“, BerliNews, 22. September 2005.
- „Dribbeln im Dienste der Wissenschaft“, Kölner Stadtanzeiger, 08. November 2005.
- FAZ – Net. „Das große Fußballspiel der Moderne“, Natur und Wissenschaft, 28. Dezember 2005.

Kooperationen

Innerhalb der Humboldt-Universität

- Charité – Universitätsmedizin Berlin, Klinik für Kardiologie, Angiologie und Pulmologie
- Charité – Universitätsmedizin Berlin, Klinik für Mund-, Kiefer- und Gesichtschirurgie – Klinische Navigation und Robotik
- Institut für Entrepreneurship / Innovationsmanagement
- Institut für Psychologie

Mit anderen Universitäten

- Technische Universität München, Fakultät für Informatik, Arbeitsgruppe Intelligente Autonome Systeme
- Universität Bremen, Fachbereich Mathematik und Informatik
- 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“
- Universität Hildesheim, Institut für Mathematik und Angewandte Informatik, Arbeitsgruppe „Intelligente Informationssysteme“
- University Novi Sad, Serbien-Montenegro
- Universität Plovdiv, Bulgarien
- University of Technology and Information Management, Rzeszow, Polen
- Universität Warschau, Institute für Informatik und Mathematik, Polen
- Zentrum für Biomaterialien der Universität Havanna, Kuba
- Institut für Nationale Planung, Kairo, Ägypten

Mit Außeruniversitären Forschungseinrichtungen und Partnern in der Industrie

- Empolis GmbH, Kaiserslautern
- Fell & Kernbach GmbH, Berlin
- Fraunhofer Institut für Autonome Intelligente Systeme, Sankt Augustin
- Fraunhofer Institut für Rechnerarchitektur und Softwaretechnik, Berlin
- Gesellschaft zur Förderung angewandter Informatik e.V., Berlin
- Outermedia GmbH, Berlin
- Siemens ElectroCom Postautomation, Berlin
- Siemens Medical Solutions, Erlangen
- SONY Corporation, Japan
- ThinkCell, Berlin

Weitere Aktivitäten

Prof. Dr. Hans-Dieter Burkhard

- Vizepräsident der RoboCup Federation
- Sprecher des Deutschen Nationalkomitee "RoboCup"
- Mitglied Koordinationsgruppe DFG-SPP 1125 "Kooperierende Teams mobiler Roboter in dynamischen Umgebungen"
- ECCAI Fellow
- EDITORIAL COUNCIL of ComSIS (Computer Science and Information Systems)
- Gasteditor für KI, Heft 2/2006
- Mitarbeit im BMBF Strategiekreis "Wissenschaftsjahr 2006"
- Mentor im Programm "ProFiL. Professionalisierung für Frauen in Forschung und Lehre: Mentoring, Training, Networking" 2004/05
- Mitglied im Pädagogischen Beirat "Zukunftscamp" der Deutschen Telekom
- Co-Chair 10. International RoboCup Competitions and Conference 2006 (Bremen)
- OC-Mitglied: German Open, RoboCup 2005, CS&P2005
- Mitglied in PCs: AAMAS05, AAMAS06, ATOP05, BIS2005, CS&P2005, Exact2005, GI WS "Roboter in der Informatikausbildung" IAT05, ICCBR05, IDA2005, IJCAI-WS Agents in Real-Time and Dynamic Environments, IROBOT05, ISMIS05, KI05, MATES05, MRC2005, Robocup2005, RSFDGrC2005, SCI05

Dipl.-Inf. Matthias Jünger

- League-Chair der Sony Four Legged League bei den RoboCup German Open 2005

Dr. Gabriela Lindemann-v. Trzebiatowski

- Frauenbeauftragte der Mathematisch-Naturwissenschaftlichen Fakultät II
- Mitglied des Konzils der Humboldt-Universität (seit Juni 2004)
- Stellvertretendes Mitglied im Akademischen Senat der Humboldt-Universität (seit Juni 2004)
- Mitglied der Kommission für Frauenförderung (KFF) des Akademischen Senats der Humboldt-Universität (seit Juni 2004)
- stellvertretendes Mitglied der Entwicklungsplanungskommission (EPK) des Akademischen Senats der Humboldt-Universität (seit Juni 2004)
- Mitglied der Fachgruppenleitung „Verteilte Künstliche Intelligenz“ der Gesellschaft für Informatik
- Co-Chair vom Workshop ANI@REM auf der AAMAS'05
- Co-Chair von der CS&P 2005
- Mitglied im Programmkomitee ISMIS 2005, Mates 2005, AAMAS Workshop AMKM, AAMAS Hauptkonferenz, RSFDGrC 2005, IDA 2005, IDPT 2005, IAT 2005, IEEE ISM 2005
- Gutachterin für das Journal Fundamenta Informaticae
- Gutachterin für das Journal JASSS

Dipl.-Inf. Mirjam Minor

- Stellvertretende Sprecherin der Fachgruppe „Wissensmanagement“ der Gesellschaft für Informatik
- Mitglied im Programmkomitee 3rd Conference Professional Knowledge Management, Kaiserslautern, 2005.
- stv. Mitglied der Kommission Lehre und Studium des Instituts für Informatik

Dr. Dagmar Monett Díaz

- Lehrbeauftragte für die Vorlesung: „Aktuelle Themen der Informatik, Evolutionäre Algorithmen“ an der Berufsakademie Berlin, FHW.
- Lehrbeauftragte für die Vorlesung: „Wissensbasierte Systeme“ an der Berufsakademie Berlin, FHW.

Dipl. Kfm. Dipl.-Inf. Diemo Urbig

- Mitglied im Programmkomitee European Conference on Artificial Life (ECAL) 2005
- Mitglied in Gesellschaft für Informatik (GI), Academy of Management (AoM), European Social Simulation Association (ESSA)
- Tätigkeit als Reviewer für Fachzeitschriften: Journal of Artificial Societies and Social Simulation (JASSS)

Dissertation

M. SC. DAGMAR MONETT DIAZ: *Agent Based Configuration of (Metaheuristic) Algorithms.*

Diplomarbeiten

MICHAEL GOLLIN: *Implementation einer Bibliothek für Reinforcement Learning und Anwendung in der RoboCup Simulationsliga.*

MARC HOCHSTRATE: *Veränderung der Ähnlichkeitsstruktur durch kontextuelle Merkmale und Entwicklung eines Klassifikationsalgorithmus zur Simulation dieser Veränderungen.*

FABIAN MÜLLER: *Integration von TFBS-Systemen und Ontologien.*

SEBASTIAN MÜLLER: *Automatic Chart Labeling.*

DAVID SALZ: *3DView2: Eine dreidimensionale Visualisierungs- und Steuerungskomponente für die MicroPsi-Multiagentenplattform.*

MARTIN STIEL: *CDWeb – eine Plattform zur Bestimmung der Ähnlichkeit von Tonträgermetadaten.*

Lehr- und Forschungseinheit

Wissensmanagement

http://informatik.hu-berlin.de/Forschung_Lehre/wm/

Leiter

PROF. DR. TOBIAS SCHEFFER

Tel.: (030) 2093 3107

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

Sekretariat

SABINE DZIWISZ

Tel.: (030) 2093 3044

Fax: (030) 2093 3045

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

Wissenschaftliche Mitarbeiterinnen und Mitarbeiter

STEFFEN BICKEL

ULF BREFELD

MICHAEL BRÜCKNER

ISABEL DROST

SASCHA SCHULZ

Tutoren

CHRISTOPH BÜSCHER

PETER HAIDER

ROLF SCHIMPFKY

HAGEN ZAHN

Unser Forschungsinteresse gilt Problemen der Entdeckung von Wissen in großen Datenbanken und Textsammlungen und der Nutzung dieses Wissens für die Lösung schwieriger, wissensintensiver Probleme. Wir beschäftigen uns sowohl mit der Entwicklung und Untersuchung von Verfahren, als auch mit Anwendungen aus den Bereichen Information Retrieval und Dokumentenmanagement.

Forschungsschwerpunkte

Lernen aus Texten

Aus der Analyse von Texten lassen sich Modelle gewinnen, die sowohl die Struktur vernetzter Textsammlungen als auch Eigenschaften des Aufbaus von Sprache beschreiben. Anwendungen hierfür schließen Wortarten- und Eigennamenerkennung, Parsieren, Textklassifikation und Clusteranalyse ein. Eine große Herausforderung für das Lernen aus natürlicher Sprache ist gegenwärtig die Entwicklung diskriminativer Verfahren, die strukturierte Ausgaben wie Parsbäume oder gemeinsame Strukturen aus Wortart- und Eigennamen-Tags in einem gemeinsamen Optimierungsproblem erzeugen können. Heute werden

verschiedene Sprachverarbeitungsschritte sequentiell aneinandergereiht, Fehlerraten summieren sich dabei auf. Diskriminative Zielkriterien maximierende Verfahren sind darauf hin optimiert, etwa die korrekte Wortart eines Wortes von allen inkorrekten zu unterscheiden, während die heute vorwiegend verwendeten generativen Modelle ein Kriterium maximieren, das weniger eng mit der Erkennungsrate korreliert.

Halbüberwachte Lernverfahren nutzen zusätzlich zu (wenigen) klassifizierten Trainingsbeispielen auch unklassifizierte Beispiele aus. Besonders beim Lernen aus Texten – Klassifikation, Eigennamenerkennung, Informationsextraktion – bieten unklassifizierte Texte (etwa das Web) zusätzliche Informationen, wie sie effektiv für diskriminative Lernaufgaben ausgenutzt werden können ist eine spannende Forschungsfrage.

Information Retrieval, Adversarial Learning

Zu den wichtigsten offenen Problemen des Information Retrieval gehört die Bewertung der Relevanz von Dokumenten in Umgebungen, in denen Anbieter von Informationen Relevanzindikatoren manipulieren um Aufmerksamkeit zu gewinnen. Versender von Spam-E-mails passen ihre Nachrichten den verwendeten Spam-Filtern an, Anbieter von Webseiten erzeugen Link-Farmen um den PageRank und damit ihre Platzierung bei Google zu verbessern. Die Bewertung von Relevanz lässt sich als Spiel ansehen, bei dem ein Spieler nach sicheren Relevanzindikatoren sucht während der zweite Spieler Relevanz vortäuschen möchte. Ob und wie Relevanzindikatoren konstruiert werden können, die robust gegen Manipulation sind, ist eine offene Frage. Eine weitere, damit zusammenhängende offene Frage ist, wie das Informationsbedürfnis eines Benutzers mit Hilfe von Kontextparametern (z.B. Ort, Zeit), Clickstream-Daten und persönlichen Informationen (etwa Informationen über soziale Netze, die sich aus Blogs gewinnen lassen) besser charakterisieren lässt. In diesem Themenkomplex liegen viele Probleme, zu deren Lösung das maschinelle Lernen beitragen kann.

Assistenzsysteme, Wissensmanagement

Im Mittelpunkt dieses Themengebietes stehen Textverarbeitungsprozesse wie das Beantworten häufig gestellter Fragen in Emails und das häufige Schreiben ähnlicher Texte. Wir entwickeln Verfahren, die Wissen aus in der Vergangenheit geschriebenen Texten extrahieren und dieses nutzen um einen Benutzer bei Dokumentenverarbeitungsprozessen zu unterstützen. Wir untersuchen, wie sich Informationen aus Werkstattberichten extrahieren und nutzen lassen, um Service-Mitarbeiter mit den für eine Problembeschreibung relevanten Informationen zu versorgen.

Text Mining in der Bioinformatik

Um ein besseres Verständnis für die Zusammenhänge zwischen genetischer Information und zellulären Prozessen zu gewinnen, müssen sehr große Mengen einzelner wissenschaftlicher Ergebnisse, die chemische und funktionale Eigenschaften einzelner Gene und ihrer Proteine betreffen, berücksichtigt und in Modellen zusammen geführt werden. Wir untersuchen Verfahren, die Wissen über Eigenschaften von Genen aus Veröffentlichungen extrahieren und daraus Modelle bestimmter funktionaler Zusammenhänge bilden.

Wissensentdeckung in großen Datenbanken

Viele Datenbanken sind zu groß als dass ein Data-Mining-Algorithmus alle Transaktionen bearbeiten könnte. Wir verfolgen die Frage, wie Lernalgorithmen konstruiert werden können, die dennoch – mathematisch beweisbar – die interessantesten Zusammenhänge mit hoher Wahrscheinlichkeit finden. Wir verfolgen zudem die Frage, wie Hintergrundwissen

über ein Gebiet so in den Data-Mining-Prozess integriert werden kann, dass nur unbekannte, unerwartete Zusammenhänge entdeckt werden.

Lehre

Veranstaltungen im Hauptstudium

- Vorlesung „Maschinelles Sprachverarbeitung“ Halbkurs (T. SCHEFFER, SoSe 2005).
- Übung „Maschinelle Sprachverarbeitung“ (U. BREFELD, SoSe 2005).
- Seminar „Text Mining und Anwendungen“ (T. SCHEFFER, S. BICKEL, U. BREFELD, M. BRÜCKNER, WiSe 05/06).
- Forschungsseminar „Machine Learning Journal Club“ (T. SCHEFFER, SoSe / WiSe).

Veranstaltungen im Grundstudium

- Vorlesung „Praktische Informatik 1 für Bachelor-Studenten“ (T. SCHEFFER, WiSe 2005/2006).

Forschung

Projekt: „Data and Text Mining in Quality and Service“

Ansprechpartner: PROF. DR. TOBIAS SCHEFFER.

Forschungsförderung: DAIMLERCHRYSLER AG

Projektlaufzeit: 08. 2005 bis 07. 2008

Projektmitarbeiter: SASCHA SCHULZ

Wir untersuchen das Problem, neue Entwicklungen und Trends in Produktions- und Garantiendatenbanken und Werkstattberichten automatisch zu entdecken und deren versteckte Ursachen zu ermitteln. Aus Anwendungssicht ist das Ziel, durch automatische Datenanalyse Fehler in Konstruktion und Produktion früher zu erkennen sowie Service-Prozesse zu verbessern.

Projekt: „Serverseitige Identifikation von Spam-Emails“

Ansprechpartner: PROF. DR. TOBIAS SCHEFFER.

Forschungsförderung: STRATO RECHENZENTRUM AG

Projektlaufzeit: 07. 2005 bis 06. 2007

Projektmitarbeiter: MICHAEL BRÜCKNER

Die STRATO Rechenzentrum AG betreibt die Infrastruktur des Webproviders STRATO. Wir analysieren das Problem der „Adversarial Classification“, der Klassifikation von Emails als Spiel zwischen zwei Gegnern (Spam-Filter und Spam-Versender), die jeweils auf die „Züge“ des anderen reagieren um einander zu überlisten; Spam-Versender versehen ihre Emails etwa mit zusätzlichen Texten, welche die charakteristischen Merkmale von Spam-Emails verwischen. Wir suchen eine Gewinnstrategie, mit der wir auch zukünftige Spam-Mails erkennen können. Während bei vielen anderen Klassifikationsaufgaben klassifizierte Beispieltex-te zum Trainieren verwendet werden können, ist die Ausgangslage hier komplizierter. Zwar steht uns ein Strom von Emails zur Verfügung, doch ist es unpraktisch und zugleich wegen der

Vertraulichkeit der Daten nicht zulässig, Beispieltex te manuell in Spam und Nicht-Spam zu unterteilen. Ein im Rahmen unserer Kooperation entwickelter Spam-Filter wird schon heute eingesetzt und filtert etwa 15 Millionen Emails von Strato-Kunden täglich.

Projekt: „Text Mining: Wissensentdeckung in Textsammlungen und Effizienz von Dokumentenverarbeitungsprozessen“

Ansprechpartner: PROF. DR. TOBIAS SCHEFFER

Forschungsförderung: DEUTSCHE FORSCHUNGSGEMEINSCHAFT

Projektlaufzeit: 06. 2003 bis ca. 12. 2008

Beteiligte Mitarbeiter: STEFFEN BICKEL, ULF BREFELD, ISABEL DROST

Das Volumen an Dokumenten, die in Archiven und im Web verfügbar sind, unterliegt einem explosionsartigen Wachstum. Daraus erwächst ein Bedarf an Verfahren, die große Textmengen automatisch analysieren, klassifizieren und filtern können, und die das in ihnen enthaltene Wissen extrahieren, darstellen und es nutzbar machen können.

Ein großer Teil unserer Arbeitsprozesse besteht 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 die dieses gewonnene Wissen nutzen, um künftige Dokumentenverarbeitungsprozesse zu unterstützen und effizienter zu gestalten.

Weitere Kooperationen

Wir sind am „Interdisziplinären Zentrum Sprachliche Bedeutung“ der Humboldt-Universität zu Berlin beteiligt. An diesem Zentrum beleuchten Linguisten, Psychologen, Informatiker, Philosophen und Neurologen Aspekte der Semantik von Sprache.

Wir sind am „Interdisziplinären Zentrums Ubiquitäre Information“ (Sprecher: Prof. Oliver Günther) beteiligt. Wirtschaftswissenschaftler, Informatiker, Mediziner und Rechtswissenschaftler untersuchen hier die Grundlagen und Auswirkungen der Informatisierung von Wirtschaft und Alltag.

Gemeinsam mit der ID-Berlin GmbH untersuchen wir, wie sich das Verfassen von Arztberichten mit Assistenzsystemen unterstützen lässt. Wir ermitteln etwa, ob und mit welcher Genauigkeit sich der ICD-10-Diagnosecode basierend auf der Freitexteingabe des Arztes bestimmen lässt.

Mit der GSD – Gesellschaft für Systemforschung und Dienstleistungen im Gesundheitswesen mbH – führen wir ein Pilotprojekt für die Assistenz bei der Beantwortung technischer Anfragen durch.

Gemeinsam mit AI Insight, Inc., untersuchen wir skalierbare Verfahren zum Lernen von Regressionsbäumen für die Bewertung der Risiken von Krankenversicherungen.

Veröffentlichungen

- D. VOGEL, ST. BICKEL, P. HAIDER, R. SCHIMPFKY, P. SIEMEN, ST. BRIDGES, T. SCHEFFER: *Classifying search engine queries using the web as background knowledge*. SIGKDD Explorations, in print.
- J. HAKENBERG, ST. BICKEL, U. BREFELD, H. ZAHN, L. FAULSTICH, U. LESER, T. SCHEFFER: *Systematic feature evaluation for gene name recognition*. BMC Bioinformatics 6(1), 2005.
- T. SCHEFFER: *Finding association rules that trade support optimally against confidence*. Intelligent Data Analysis 9(3), 2005.
- T. SCHEFFER: *Semi-Supervised Learning*. In Wang, editor, Encyclopedia of Data Warehousing and Data Mining. 2005.
- ST. BICKEL, T. SCHEFFER: *Mining Email Data*. In Wang, editor, Encyclopedia of Data Warehousing and Data Mining. 2005.
- M. KOCKELKORN, T. SCHEFFER: *The Responsio email management system*. In Zanasi, editor, Text Mining and its Applications. WIT Press, 2005.
- A. HOFFMANN, H. MOTODA, T. SCHEFFER: *Discovery Science*. Proceedings of the 8th International Conference. Springer LNCS 3734, 2005.
- ST. RÜPING, T. SCHEFFER: editors. Proceedings of the ICML Workshop on Learning with Multiple Views, 2005.
- ECML Best Paper Award*: U. BREFELD, CH. BÜSCHER, T. SCHEFFER: *Multi-view discriminative sequential learning*. Proceedings of the European Conference on Machine Learning. 2005.
- S. JAROSZEWICZ, T. SCHEFFER: *Fast discovery of unexpected patterns in data relative to a Bayesian network*. Proceedings of the SIGKDD International Conference on Knowledge Discovery and Data Mining. 2005.
- ST. BICKEL, P. HAIDER, T. SCHEFFER: *Predicting sentences using N-gram language models*. Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2005.
- ST. BICKEL, T. SCHEFFER: *Estimation of mixture models using Co-EM*. Proceedings of the European Conference on Machine Learning. 2005.
- ST. BICKEL, P. HAIDER, T. SCHEFFER: *Learning to complete sentences*. Proceedings of the European Conference on Machine Learning. 2005.
- I. DROST, T. SCHEFFER: *Thwarting the nigritude ultramarine: learning to identify link spam*. Proceedings of the European Conference on Machine Learning. 2005.
- M. BRÜCKNER: *The p-center machine*. Proceedings of the International Joint Conference on neural Networks, 2005.
- M. BRÜCKNER, W. DILGER: *A soft Bayes perceptron*. Proceedings of the International Joint Conference on neural Networks, 2005.
- I. DROST, ST. BICKEL, T. SCHEFFER: *Discovering communities in linked data by multi-view clustering*. Proceedings of the Conference of the German Classification Society, 2005.
- U. BREFELD, T. SCHEFFER: *AUC Maximizing support vector learning*. Proceedings of the ICML Workshop on ROC Analysis in Machine Learning, 2005.

U. BREFELD, C. BÜSCHER, T. SCHEFFER: *Multi-view hidden Markov perceptrons*. Proceedings of the German Workshop on Machine Learning (FGML), 2005

ST. BICKEL, T. SCHEFFER: *Estimation of mixture models using Co-EM*. Proceedings of the ICML Workshop on Learning with Multiple Views. 2005.

Vorträge

ST. BICKEL: *Predicting sentences using N-gram language models*. International Conference on Empirical Methods in Natural Language Processing, 2005.

ST. BICKEL: *Estimation of mixture models using Co-EM*. Proceedings of the ICML Workshop on Learning with Multiple Views. 2005.

U. BREFELD: *Multi-view discriminative sequential learning*. European Conference on Machine Learning. 2005.

U. BREFELD: *AUC Maximizing support vector learning*. ICML Workshop on ROC Analysis in Machine Learning, 2005.

U. BREFELD: *Multi-view hidden Markov perceptrons*. German Workshop on Machine Learning (FGML), 2005.

M. BRÜCKNER: *The p-center machine*. International Joint Conference on neural Networks, 2005.

M. BRÜCKNER: *A soft Bayes perceptron*. Proceedings of the International Joint Conference on neural Networks, 2005.

I. DROST: Eingeladener Vortrag: *Maschinelles Lernen und Suchmaschinentechnologie*, Chaos Communication Congress, CCC, Berlin.

I. DROST: *Thwarting the nigrityde ultramarine: learning to identify link spam*. European Conference on Machine Learning. 2005.

I. DROST: *Discovering communities in linked data by multi-view clustering*. Proceedings of the Conference of the German Classification Society, 2005.

T. SCHEFFER: Eingeladener Vortrag: *Swarm intelligence, social software, and information spam*. European Commission Consultation Workshop on the Future of Search Engines.

T. SCHEFFER: Eingeladener Vortrag: *Discriminative sequential learning from labeled and unlabeled data*. DaimlerChrysler Workshop on Information and Relation Extraction for Text and Data Mining.

T. SCHEFFER: Eingeladener Vortrag: *Multi-view clustering and link spam detection*. Dagstuhl-Seminar Probabilistic, logical, and relational learning: towards a synthesis.

T. SCHEFFER: *Estimation of mixture models using Co-EM*. European Conference on Machine Learning. 2005.

T. SCHEFFER: *Learning to complete sentences*. European Conference on Machine Learning. 2005.

Tagungsorganisation

T. SCHEFFER: Program Co-Chair, International Conference on Discovery Science, Singapore, 2005.

T. SCHEFFER: Co-Chair, ICML Workshop on Learning with multiple Views.

Weitere Aktivitäten

Steffen Bickel

- Program Committee, ICML Workshop on learning with multiple views.
- Reviewer, International Conference on Machine Learning,
- Reviewer, European Conference on Machine Learning,
- Reviewer, European Conference on Principles and Practice of Knowledge Discovery in Databases,
- Reviewer, International Conference on Discovery Science.

Ulf Brefeld

- Program Committee, ICML Workshop on Learning with multiple views.
- Reviewer, International Conference on Machine Learning,
- Reviewer, European Conference on Machine Learning,
- Reviewer, European Conference on Principles and Practice of Knowledge Discovery in Databases,
- Reviewer, International Conference on Discovery Science.

Michael Brückner

- Reviewer, International Conference on Machine Learning,
- Reviewer, European Conference on Machine Learning,
- Reviewer, European Conference on Principles and Practice of Knowledge Discovery in Databases.

Isabel Drost

- Reviewer, IEEE Transactions on Knowledge and Data Engineering,
- Reviewer, International Conference on Machine Learning,
- Reviewer, European Conference on Machine Learning,
- Reviewer, European Conference on Principles and Practice of Knowledge Discovery in Databases,
- Reviewer, International Conference on Discovery Science.

Tobias Scheffer

- Editorial Board, Data Mining and Knowledge Discovery Journal.
- Steering Committee, International Conference on Discovery Science.
- Gutachter für die Dissertation von Sambu Seo, Technische Universität Berlin.
- Program Co-Chair, International Conference on Discovery Science.
- Co-Chair, ICML Workshop on Learning with multiple Views.
- Program Area Chair, International Conference on Machine Learning,
- Program Area Chair, European Conference on Machine Learning,

- Program Area Chair, European Conference on Principles and Practice of Knowledge Discovery and Data Mining.
- Program Committee, International Conference on Genome Informatics,
- Program Committee, EPAI Workshop on Computational Methods in Bioinformatics.
- Referee for the ICML student scholarship program.
- Reviewer, Machine Learning Journal,
- Reviewer, Data Mining and Knowledge Discovery Journal,
- Reviewer, Journal of Machine Learning Research,
- Reviewer, BioSystems Journal,
- Reviewer, Journal of Human Computer Studies,
- Reviewer, International Joint Conference on Artificial Intelligence,
- Reviewer, International Conference on Logic for Programming,
- Reviewer, Artificial Intelligence and Reasoning.

Abschlussarbeiten

CHRISTOPH BÜSCHER: *Halbüberwachtes lernen zur Named-Entity-Recognition mit der Hidden-Markov-Support-Vektor-Maschine*. Diplomarbeit. 2005.

STEFAN TRÄGER: *Korpora aus dem Netz: Die Erstellung eines Fachkorpus aus Webseiten und Möglichkeiten der sprachwissenschaftlichen Nutzung*. Magisterarbeit.

BIRGIT WANDREI: *Satzvervollständigung mit latent semantischer Analyse*. Diplomarbeit.

THOMAS POSNER: *„Interaktive Satzgenerierung mit N-Gramm-Sprachmodellen: Evaluierung der Praxistauglichkeit“*. Diplomarbeit.

MARTINA BRECHLER: *Analyse des Navigationsverhaltens verschiedener Nutzergruppen auf Websites mit Hilfe des Web Usage Mining und Ontology Learning*. Diplomarbeit.

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

Die Lehr- und Forschungseinheit vertritt die Gebiete „Stochastische Aspekte der Informatik“ und „Computergestützte Statistik“.

Lehre

Vorlesungen

- Einführung in die Informations- und Kodierungstheorie (E. RÖDEL, WiSe 2004/2005)
- Stochastische Methoden der Globalen Optimierung (E. RÖDEL, SoSe 2005)
- Werkzeuge der empirischen Forschung (W. KÖSSLER, SoSe 2005)

Übungen

- Einführung in die Informations- und Kodierungstheorie (E. RÖDEL, WiSe 2004/2005)
- Stochastische Methoden der Globalen Optimierung (E. RÖDEL, SoSe 2005)
- Werkzeuge der empirischen Forschung (W. KÖSSLER, SoSe 2005)

Seminare

- Proseminar Fuzzy-Logik und Fuzzy-Mengen (E. RÖDEL, WiSe 2004/2005)
- Zirkel der Mathematischen Schülergesellschaft, 12. Klasse (W. KÖSSLER, WiSe 2004/2005, SoSe 2005)

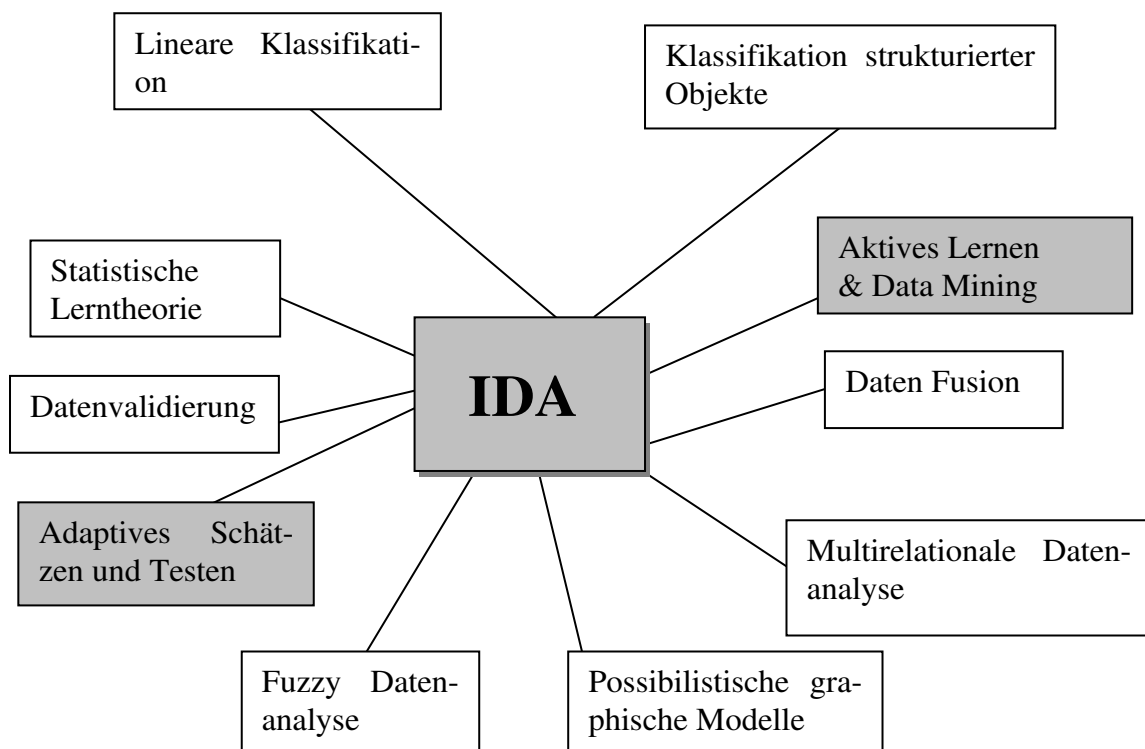
Forschung

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 vier 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.

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 wird in Zusammenarbeit mit dem Institut für Wirtschaftsinformatik der FU Berlin weitergeführt.

Projekt: Statistische Prüfung von Daten

Ansprechpartner: PROF. DR. EGMAR RÖDEL, DR. WOLFGANG KÖSSLER

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 „Statistik“ (Zusammenarbeit mit dem Universitätsrechenzentrum)

Ansprechpartner: DR. WOLFGANG KÖSSLER

Im vergangenen Jahr wurden unter anderem folgende Projekte unterstützt:

- Analyse von Konzentration und Leistung im Tagesverlauf (Institut für Psychologie)
- Soziologische Untersuchungen zur Schuleintrittszeit (Institut für Erziehungswissenschaften)
- Analyse der Zeitdauer ausgesprochener Vokale (Institut für Linguistik)
- Einkaufsverhalten und Verkehrsmittelnutzung (Institut für Geographie)
- Analyse von Kompetenz und Aggression von Grundschulern (Institut für Psychologie)

Projekt: Unabhängigkeitstests

Ansprechpartner: PROF. E. RÖDEL, DR. WOLFGANG KÖSSLER

Eine grundlegende statistische Fragestellung ist die nach der Unabhängigkeit von zwei Populationen. Wir untersuchen die wichtigsten in der statistischen Theorie und Praxis verwendeten Modelle, und berechnen asymptotische absolute und relative Effizienzen sowie die Gütefunktion verschiedener linearer Rangtests. Rangtests sind ein beliebtes Werkzeug in der Statistik, da sie nicht auf der Normalverteilungsannahme beruhen und somit recht robust sind. Weiterhin sind die Güteeigenschaften von Rangtests meistens besser als die der klassischen parametrischen Tests. Die Güteeigenschaften hängen stark von dem (angenommenen) Modell ab. Wie auch in vielen anderen Problemstellungen erweisen sich die einfachen, mittlerweile klassischen Rangtests, hier insbesondere der Spearman-Test, als eine gute Wahl.

Projekt: Lokationstests unter Verwendung von U-Statistiken

Ansprechpartner: DR. WOLFGANG KÖSSLER

U-Statistiken sind eine interessante Klasse von Statistiken, bei denen alle möglichen Teilstichproben von bestimmtem vorgegebenem Umfang aus der zugrunde liegenden Stichprobe gezogen werden. Für das Zweistichproben-Lageproblem ist der Mann-Whitney-Wilcoxon Test der klassische Vertreter dieser Klasse. Wir betrachten einige Verallgemeinerungen und untersuchen die asymptotischen Eigenschaften. Weiterhin konstruieren wir adaptive Tests, bei denen zunächst die Verteilung anhand einer Selektorstatistik bezüglich Tailstärke (und Schiefe) geschätzt, und dann ein geeigneter, auf U-Statistiken beruhender Test ausgewählt wird. Der vorgeschlagene adaptive Test hat gute asymptotische und finite Güteeigenschaften.

Projekt: SAS-Bediensystem

Ansprechpartner: DR. WOLFGANG KÖSSLER

Beteiligte Studenten und Mitarbeiter: Dirk-Johann Deters, Wolf F. Lesener

Für den Vergleich zweier unabhängiger Messreihen gibt es eine Reihe von parametrischen und nichtparametrischen Methoden (z.B. t-Test, Wilcoxontest), aber auch erst in jüngerer Zeit entwickelte adaptive Verfahren. Der Anwender steht u.U. vor der Frage, welche der vielen Methoden anzuwenden ist. Wir haben ein in der SAS-Macrosprache geschriebenes

Bediensystem entwickelt, das den Anwender durch seine Problemstellung führt, und, je nach seinem Kenntnisstand über die Daten, eine geeignete Methode auswählt. Es werden beliebige (stetige) Verteilungen zugelassen, also sowohl symmetrische als auch Verteilungen mit kurzen oder langen Tails. Für den Fall unbekannter Verteilungen sind adaptive Tests implementiert worden, bei denen zunächst die Form der Verteilung anhand einer Selektorstatistik klassifiziert und dann ein geeigneter Rangtest ausgewählt wird.

Veröffentlichungen

E. RÖDEL: *Statistical Analysis of Discrete Memoryless Channels*. Problems of Information Transmission, angenommen.

W. KÖSSLER UND E. RÖDEL: *The asymptotic efficacies and relative efficiencies of various linear rank tests for independence*. Metrika, angenommen.

W. KÖSSLER: *Some c-Sample Rank Tests for Homogeneity Against Umbrella Alternatives With Unknown Peak*, Journal of Statistical Computation and Simulation, 76 (2006), 57-74.

W. KÖSSLER: *Some c-Sample Rank Tests for Homogeneity Against Ordered Alternatives Based on U-Statistics*. Journal of Nonparametric Statistics 17 (2005), 777-795.

W. KÖSSLER, N. KUMAR: *An adaptive test for the two-sample location problem based on U-Statistics*. eingereicht.

Vorträge

W. KÖSSLER: *Ein SAS-Bediensystem für Lokationstests*. 9. Konferenz der SAS-Anwender in Forschung und Entwicklung, Humboldt-Universität zu Berlin, März 2005.

W. KÖSSLER: *Ein auf U-Statistiken beruhender adaptiver Test für das Zwei-Stichproben-Lokationsproblem (Poster)*. 9. Konferenz der SAS-Anwender in Forschung und Entwicklung, Humboldt-Universität zu Berlin, März 2005.

Sonstige Aktivitäten

Prof. Dr. Egmar Rödel, Dr. Wolfgang Kössler

- Beratung von Anwendern statistischer Methoden

Prof. Dr. Egmar Rödel

- Gutachtertätigkeit für Zeitschriften: Statistics, Biometrical Journal, Statistical Papers, Allgemeines Statistisches Archiv, Statistics and Computing
- Studienfachberater für die Lehramtsstudiengänge Informatik
- Mitglied im Institutsrat
- Mitglied der Internationalen Association for Statistical Computing (IASC)
- Mitglied der Fachgruppe Stochastik der DMV

Dr. Wolfgang Kössler

- Gutachtertätigkeit für die Zeitschriften Statistics, Statistical Papers, Allgemeines Statistisches Archiv
- Verantwortlicher für die Evaluierung der Lehre im Institut für Informatik
- Aktive Mitarbeit im Mentoring Programm

- Mitglied der Deutschen Statistischen Gesellschaft
- Mitglied der Kommission Lehre und Studium
- Mitglied des erweiterten Fakultätsrates
- Korrektor bei der Mathematikolympiade

Wissenschaftliche Kooperationen

- Institut für Mathematik und Informatik der Universität Havanna/Kuba
- Institut für Statistik der Universität Chandigarh, Panjab, Indien
- Institut für Wirtschaftsinformatik der Freien Universität Berlin
- Institut für Statistik und Ökonometrie der Freien Universität Berlin
- Institut für Informatik der TU Berlin
- Universität Magdeburg

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 Mitarbeiterinnen und Mitarbeiter

DIPL. INF. PETER BITTNER

DR. HABIL. CHRISTIAN DAHME

DR. VOLKER GRASSMUCK (DFG)

DR. JOCHEN KOUBEK

DIPL. INF. ROLAND KUBICA

DIPL. INF. JENS-MARTIN LOEBEL (DFG)

DIPL. INF. CONSTANZE KURZ

Technische Unterstützung

FRANK WOZOBULE

Tutorinnen und Tutoren

ELZBIETA JASINSKA

STEFAN KLUMPP

AGATA KROLIKOWSKI

MARCUS RICHTER

Die Arbeitsgruppe „Informatik in Bildung und Gesellschaft“ erforscht die Rolle der Informatik auf dem Weg zur Informationsgesellschaft. Die dabei untersuchten Wechselwirkungen der Informatik umfassen die unterschiedlichsten Aspekte. Neben historischen, sozialen, kulturellen Fragen betrifft dies auch ökonomische, politische, ökologische, ethische, didaktische und selbstverständlich technische Aspekte. Die entstehende global vernetzte *Informationsgesellschaft* wird als zentrale Herausforderung für den Bereich Informatik und Gesellschaft gesehen, in der die Informatik als technische Grundlagenwissenschaft eine definierende Rolle spielen kann. Dies bildet die Ausgangslage für die Forschungen der Arbeitsgruppe.

Schwerpunkte in Forschung und Lehre

Die Arbeiten der Arbeitsgruppe *Informatik in Bildung und Gesellschaft* konzentrierten sich im Jahr 2005 im wesentlichen auf folgende Themen:

Digitale Medien

Die Digitalisierung tradierter Medien verändert den Umgang mit der Erzeugung, Verteilung und Nutzung von Kulturgütern im weitesten Sinne. Digitale Medien ermöglichen neue Benutzerschnittstellen, neue Darstellungsformen z.B. in der Computergrafik und neue Nutzungsmöglichkeiten. Es stellen sich eine Reihe von Fragen z.B. nach der Zuverlässigkeit, der Abbildungstreue oder der Haltbarkeit solcher Medien. Der Einsatz Digitaler Medien in Bildung und Lehre ist ein langjähriger Schwerpunkt unserer Arbeit.

Informatische Bildung

Das klassische Bildungsideal eines selbst bestimmten und verantwortungsvollen Handelns in der Welt bezieht sich in der Informationsgesellschaft auch auf einen ebensolchen Umgang mit Informationstechnologien. Wir untersuchen die Rahmenbedingungen für eine solche Bildung in der Schule, im Studium, im Beruf sowie 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.

Informationssicherheit, Datenschutzrecht und Informationelle Selbstbestimmung

Die Abhängigkeit der Informationsgesellschaft von Informationstechniken erhöht ihre Verletzbarkeit durch Angriffe von innen und aussen auf eben diese Infrastruktur. Nutzer müssen die digitalen Spuren kennen, die ihre Handlungen in verschiedenen Netzen und Systemen hinterlassen. Angesichts schärferer politischer Diskussionen zur inneren Sicherheit werden Bürgerrechte zur Disposition gestellt. Mit dem Anspruch, die Sicherheit der Bürger zu erhöhen, wurden Einschnitte in Persönlichkeitsrechte in der EU wie in den USA ergriffen. Bei der Einführung des ePasses in Deutschland, bei der Ausweitung der Videoüberwachung, beim verstärkten Einsatz biometrischer Verfahren oder bei der beschlossenen Vorratsdatenspeicherung aller Telekommunikationsverbindungen wird die Informatik zur Hilfswissenschaft staatlicher Überwachung. Neben diesen staatlichen Überwachungsansprüchen wachsen vor allen durch das Internet oder durch neue Techniken wie RFIDs die Zugriffe auf persönliche Daten. Auch dies ist ein wichtiger Untersuchungsgegenstand.

Urheberrecht und Geistiges Eigentum

Der Computer als universelle Kopiermaschine bedroht immaterielle Besitzstände. Die Diskussion um den zweiten Korb der Urheberrechtsnovelle beschäftigte auch im Jahr 2005 Jahr die Politik und die beteiligten Interessengruppen, darunter die Wissenschaft. Mit Regelungen zum Umgang mit Kopien für Lehre und Forschung werden Weichen für eine künftige Wissensordnung gestellt, die in ihren Konsequenzen den Verantwortlichen nicht immer deutlich zu sein scheint. Zwar erlauben die Schrankenbestimmungen grundsätzlich die Verwendung urheberrechtlich geschützter Materialien für wissenschaftliche und auch private Zwecke, faktisch wird eine solche Nutzung durch „technische wirksame Maßnahmen“ wie zum Beispiel *Digital Rights Management* ausgehebelt und kriminalisiert. Um Hollywood-Filme und Pop-Musik zu schützen, werden die Möglichkeiten der Erzeugung, Verbreitung, Aneignung und Nutzung von Wissen massiv eingeschränkt, auch dort, wo es gar nicht um Hollywood-Filme oder um Pop-Musik geht, nämlich im Bereich Bildung und

Wissenschaft. *Open Source, Open Access, Creative Commons* und andere Entwicklungen werden von uns unterstützt, um die historisch gewachsene Balance zwischen den unterschiedlichen Interessen in diesem Bereich zu erhalten.

Geschichte der Informatik

Ein Blick in die Geschichte belegt, dass wir nicht in der besten aller, sondern nur in einer gewordenen möglichen Welt leben und dass Entwicklungen selten inneren Notwendigkeiten folgen, sondern häufig auch einen anderen Weg hätten einschlagen können. Vor diesem Hintergrund kann ein Bewusstsein für die Verantwortung für aktuelle technische und gesellschaftliche Entwicklungen entstehen.

Akzente in der Forschung

Die Transdisziplinarität ihrer Themen findet einen deutlichen Ausdruck in unserer Mitarbeit im „Hermann von Helmholtz-Zentrum für Kulturtechnik“, wo im Rahmen der DFG-Forschergruppe „Bild-Schrift-Zahl“ ein Projekt zur „Visuellen Argumentation“ im Kontext der Digitalen, rechnergestützten Medien bearbeitet wurde und ein neues Projekt zu „Urheberschaft und geistigen Eigentum“ verfolgt wird.

- Wir beteiligen uns seit vielen Jahren an der »HyperKult«-Tagung im Juli in Lüneburg, einer Plattform für engagierte Wissenschaftler aus so unterschiedlichen Bereichen wie der Informatik, Kunstgeschichte oder den Geisteswissenschaften, die sich mit Künstler treffen. Aktuell steht die Planung der HyperKult XV im Juli 2006 an.

Die letztjährigen Tagungen wurden kürzlich dokumentiert (Martin Warnke, Wolfgang Coy, Christoph Tholen, HyperKult II – Zur Ortsbestimmung analoger und digitaler Medien, Bielefeld: Transcript-Verlag 2004).

Wir sind bei der Fachtagung „Informatik-Bildung in Berlin und Brandenburg“ aktiv; die 6. Tagung zur Schulinformatik GI-FIBBB 2006 wird von uns organisatorisch betreut und vor Ort in Adlershof durchgeführt.

Anfang 2006 wird eine große Tagung „Shapes of Things to Come – Die Zukunft der Informationsgesellschaft“ von uns an der HU organisiert und zusammen mit dem Berliner Studienverbundkolleg der Alcatel-SEL-Stiftung durchgeführt.

- Weitere Einzelheiten: <http://waste.informatik.hu-berlin.de/Tagungen>

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 bzw. Hauptgebäude) und Berlin-Adlershof (im Informatikgebäude) durchgeführt. Die 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, ohne die Diplomstudierenden zu vernachlässigen.

Neu war die Teilung der Einführungsveranstaltung „Praktische Informatik I“, die von uns erstmals für den neuen Bachelor-Studiengang im Lehramt, für Magisterstudierende und für

Informatik im Nebenfach getrennt von der Einführungsveranstaltung für Diplomstudierende durchgeführt wurde.

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 neben Kulturwissenschaftlern, Kunstwissenschaftlern, Mathematikern, Altorientalisten und Philosophen als Gründungsmitglied beteiligt ist.

<http://www2.rz.hu-berlin.de/kulturtechnik/bsz/bsz.htm>

Dies hat zum Aufbau einer DFG-Forschergruppe „Bild-Schrift-Zahl“ geführt (Sprecher: Wolfgang Coy), deren Förderung bis zum März 2007 verlängert wurde. Dieser DFG-Forschungsverbund wurde auch im Rahmen einer umfassenden externen Beurteilung des Hermann von Helmholtz-Zentrums sehr positiv evaluiert, was zu einer weiteren fünfjährigen Verlängerung der Zentrumseinrichtung durch den Akademischen Senat führte.

<http://waste.informatik.hu-berlin.de/Forschung/>

Lehre

URL: <http://waste.informatik.hu-berlin.de/Lehre/>

Veranstaltungen im Grundstudium

- Praktische Informatik I (W. COY WiSe 2004/2005)
- Übungen zu Praktische Informatik I (R. KUBICA, WiSe 2004/2005)
- Übungen zu Praktische Informatik I (P. BITTNER, WiSe 2004/2005)
- Unterrichtspraktikum (CH. DAHME, WiSe 2004/2005)
- Einführung in die Fachdidaktik (KOUBEK, WiSe 2004/2005)
- Unterrichtspraktikum (CH. DAHME, SoSe 2005)
- Wurzeln der Informatik (C. KURZ, SoSe 2005)
- Informatik und Gesellschaft (J. KOUBEK, SoSe 2005)
- Einführung in die Fachdidaktik (W. COY, WiSe 2005/2006)
- Unterrichtspraktikum (CH. DAHME, WiSe 2005/2006)
- Mentoren-Tutoren-Modellprojekt, Mentorentätigkeit (CH. DAHME, WiSe 2005/2006)
- The Weakest Link - Sicherheit von Online-Systemen (J.-M. LOEBEL, WiSe 2005/2006)

Kernveranstaltungen im Hauptstudium

- Konzepte der frühen Phasen der Softwareentwicklung (CH. DAHME, SoSe 2005)
- Informatik und Informationsgesellschaft I: Digitale Medien (W. COY mit J. KOUBEK, WiSe 2005/2006)

Spezialveranstaltungen

- Geschichte der Informatik (W. Coy, WiSe 2004/2005)
- Kooperatives Prototyping (Ch. Dahme, WiSe 2004/2005)
- Kooperatives Prototyping (Ch. Dahme, WiSe 2005/2006)

Seminare

eHealth – Verdatung im Gesundheitswesen (P. BITTNER, WiSe 2004/2005)
 Multimedia in der Schule (J. KOUBEK, WiSe 2004/2005)
 Informationelle Selbstbestimmung (P. BITTNER, SoSe 2005)
 Biometrie im Kontext (P. BITTNER mit PROF. DR. . MEFFERT, SoSe 2005)
 Internet und Politik (P. Bittner, SoSe 2005)
 Multimedia in der Schule (J. Koubek, SoSe 2005)
 Geschichte der Kryptologie (C. Kurz und R. Kubica, SoSe 2005)
 Langzeitarchivierung (W.Coy, WiSe 2005/2006)
 Informatische Bildung (J. Koubek, WiSe 2005/2006)
 Urheberrecht zwischen digitaler Rechtekontrolle und Wissensfreiheiten (C.Kurz und V.Grassmuck, WiSe 2005/2006)
 Hauptseminar Fachdidaktik (C. Kurz, WiSe 2005/2006)
 Gesellschaftliche, technische und juristische Perspektiven von Filesharing-Netzen (C. Kurz, WiSe 2005/2006)
 Der Fehler im System (Roland Kubica, WiSe 2005/2006)
 3D-Grafik (Roland Kubica, WiSe 2005/2006)

Forschung

URL: <http://waste.informatik.hu-berlin.de/Forschung/>

Bild, Schrift, Zahl in der TuringGalaxis. Die technischen und soziokulturellen Hintergründe geistigen Eigentums unter den Bedingungen multimedialer Digitalisierung und globaler Vernetzung.

Teilprojekt im Rahmen des DFG-Forschungsverbands “Bild - Schrift - Zahl” am Helmholtz-Zentrum für Kulturtechnik der Humboldt-Universität

Ansprechpartner: Prof. Dr. Wolfgang Coy

Beteiligte Mitarbeiter: Dr. Volker Grassmuck

Zusammenarbeit: Hermann von Helmholtz-Zentrum für Kulturtechnik, DFG-Forschergruppe „Bild-Schrift-Zahl“

Forschungsförderung: DFG (bis 2007)

URL: <http://waste.informatik.hu-berlin.de/Forschung/bsz/>

Derzeit ist ein heftiger Umbruch im Umgang mit geistigem Eigentum zu vermerken. Das deutsche Urheberrecht wurde auf Grund internationaler Vorgaben an digitale Techniken angepasst. Patentrecht wird im EU-Parlament kontrovers diskutiert und steht vor einer wesentlichen Erweiterung. Nicht nur Medienunternehmen formulieren weitgehende ökonomische Ansprüche, auch Wissenschaft und Öffentlichkeit besitzen gewachsene kulturelle Interessen an Zugang und Nutzung von Wissen.

Auslöser solch heftiger Debatten sind Digitalisierung und Vernetzung und damit einhergehende technische Veränderungen bei Herstellung, Speicherung und Distribution multimedialer Artefakte. In der Folge befindet sich die gewachsene Wissensordnung in einem Strukturwandel, der kulturelle Praktiken, ökonomische Beziehungen, technologische Trajektorien, ebenso wie seinen politischen Regulierungsrahmen grundlegend verändert. Dies betrifft sogar Basisbegriffe wie Autor, Werk und Wissen. Es ist somit dringend geboten, neben den rechtlichen und ökonomischen Bedingungen des „digitalisierten“ geistigen Ei-

gentums auch dessen technische Basis und seine kulturelle Tradition zu betrachten.

Im Zentrum des Projektes steht die Frage nach Ausgleich unterschiedlicher Ansprüche: der persönlichkeits- und vermögensrechtlichen Interessen von Autoren und Erfindern, der Verwertungsinteressen von Verlegern und anderen Parteien sowie der Interessen der Allgemeinheit. An Hand konkreter Fragestellungen wird das Feld „Bild-Schrift-Zahl“ in seinen kulturtechnischen Verzahnungen beleuchtet. Ziel ist es, die aktuellen Debatten um geistiges Eigentum aus der juristisch-ökonomischen Engführung herauszulösen und eine offenere Diskussion auf dem Weg zur Turing-Galaxis anzuregen.

Veröffentlichungen

Viele der Veröffentlichungen können als PDF- oder HTML-Dateien unter URL: <http://Waste.Informatik.HU-Berlin.de/> kopiert werden.

P. BITTNER, E. HORNECKER: *A micro-ethical view on computing practice*. In: Bertelsen, Olav W.; Bouvin, Niels Olof; Krogh, Peter G.; Kyng, Morten (eds.): *Proceedings of the 4th Decennial Conference on Critical Computing* (Aarhus, 20.-24.08.2005), New York: ACM Press, 2005, S. 69-78; <http://doi.acm.org/10.1145/1094562.1094571>

W. COY, Analog/Digital, in: M. Warnke, W. Coy, Ch. Tholen (Hrsg.), *HyperKult II – Zur Ortsbestimmung analoger und digitaler Medien*, Bielefeld: transcript-Verlag 2004

W. COY, Zum Streit der Fakultäten, in Claus Pias (Hrsg.): *Cybernetics. The Macy-Conferences 1946-1953*, Zürich-Berlin: Diaphanes, 2005

W. COY: The contest of faculties, in John Impagliazzo, John A. N. Lee, *History of Computing in Education*, Boston Dordrecht London: Kluwer Academic Publishers, 2005

W. COY: Between the Disciplines, Editorial zu *Inroads - SIGCSE Bulletin*, New York: ACM, 2005

W. COY: Internetgesellschaft Version 0.9 Beta, in: P. Gendolla: *Wissensprozesse in der Netzwerkgesellschaft*, Bielefeld: transcript-Verlag, 2005

W. COY, Ideen und ihre Verwertung in globalen Gemeinschaften, in: D.Klumpp, H. Kubicek, A. Roßnagel, W. Schulz (Hrsg.) *Medien, Ordnung und Innovation*, Berlin-Heidelberg-New York: Springer, 2005

W. COY, Informatik...im Großen und Ganzen, *LOGIN* 6/2005

W. COY, Rechnen als Kulturtechnik, in: J. Brüning, H. Knobloch (Hrsg.), *Die mathematischen Wurzeln der Kultur*, München: Fink Verlag, 2005

M. WARNKE, W. COY, CHRISTOPH THOLEN (Hrsg.), *HyperKult II – Zur Ortsbestimmung analoger und digitaler Medien*, Bielefeld: transcript-Verlag 2004

V. GRASSMUCK: *Das Ende der Universalmaschine*, in: Claus Pias (Hrsg.), *Zukünfte des Computers*, Diaphanes, Zürich-Berlin 2005, S. 241-268.

V. GRASSMUCK: *Ein Plädoyer für durchsetzbare Schrankenbestimmungen für Privatkopie, Zitat und Filesharing*, in: Zeitschrift für Urheber- und Medienrecht (ZUM), Sonderheft zum Symposium *Urheberrecht in der Informationsgesellschaft – der Referentenentwurf zum Zweiten Korb* des Bundesministeriums der Justiz in Zusammenarbeit mit dem Institut für Urheber- und Medienrecht, 2. November 2004, München, Jahrg. 49, 2/2005, S. 104-108.

V. GRASSMUCK: Gutachten zu *Freie Verbreitung von Informationen im Internet – kollektives Gut oder kollektives Übel?* für Medien & Kommunikationswissenschaft, Bredow-Institut.

V. GRASSMUCK: *LiMux - Free Software for Munich*, in: Joe Karaganis & Robert Latham, Social Science Research Council (SSRC) (eds.), *The Politics of Open Source Adoption*, May 2005, S. 14-36
<http://www.ssrc.org/programs/ccit/publications/POSA1.0.pdf>, Fortschreibung unter:
<http://www.ssrc.org/wiki/POSA>

V. GRASSMUCK: *Wissenskontrolle durch DRM: von Überfluß zu Mangel*, Preprint auf <http://edoc.huberlin.de>; erscheint in: Jeanette Hofmann (Hrsg.), *Eigentum und Wissen*, Bundeszentrale für Politische Bildung, Berlin 2006.

E. HORNECKER, P. BITTNER: *Jenseits von Whistleblowing*. FIF-Kommunikation 1/2005, S. 8-11

J. KOUBEK: *Geistiges Eigentum lehren. Das Rerum-Unterrichtsmaterial-Copyrights im digitalen Zeitalter*. In Eva Matthes / Carsten Heinze (Hrsg.): *Das Schulbuch zwischen Lehrplan und Unterrichtspraxis*, Klinkhardt, 2005.

J. KOUBEK: *Normative Mathematik in der Betriebswirtschaftslehre*. In: Brüning; Knobloch (Hrsg.): *Die mathematischen Wurzeln der Kultur*. Fink, 2005.

J. KOUBEK: *Zur Kulturgeschichte des physikalischen Raums*. In Warnke; Coy; Tholen (Hrsg.): *Hyperkult II. Zur Ortsbestimmung analoger und digitaler Medien*. transcript: 2005.

JOCHEN KOUBEK: *The Review-System*. Tagungs-CD-ROM zur TESI 2005.

JOCHEN KOUBEK: *Standards für informatische Bildung. E-Mail-Kompetenzen*. In: LOG IN 135, *Standards in der informatischen Bildung*, 2005.

J. KOUBEK: *Der Informatiker und Gesellschaft*. In: LOG IN 136/137, *Gesellschaftliche Themen in der Informatischen Bildung*, S. 22-23, 2005.

J. KOUBEK: *Recht und Informatische Bildung*. In: LOG IN 136/137, *Gesellschaftliche Themen in der Informatischen Bildung*, S. 36-38, S. 40, 2005.

J. KOUBEK: *Die Allegorie von Flächenland*. In: LOG IN 136/137, *Gesellschaftliche Themen in der Informatischen Bildung*, S. 14-17, 2005.

J. KOUBEK: *Aus großer Macht folgt große Verantwortung*. In: LOG IN 136/137, *Gesellschaftliche Themen in der Informatischen Bildung*, S. 39, 2005.

C. KURZ, JAN KRISLER: *Die Ergebnisse der BioP2-Studie zur Leistungsfähigkeit biometrischer Systeme*. FIF-Kommunikation 4/2005, S. 4-6.

C. KURZ: *Sicherheit durch Biometrie?*. Die Datenschleuder, Nr. 87.

C. KURZ: *Die Auswertung der BioP-II-Studie des BSI*. Die Datenschleuder, Nr. 87.

C. KURZ: *Biometrie und ePass*. LOG IN Nr. 136/138.

Vorträge

P. BITTNER: *Videoüberwachung durchschauen*. Vortrag im Rahmen der FIFF-Vortragsreihe »Überwachung«, Villa Ichon, Bremen, 12.02.2005.

P. BITTNER: *Strukturen des gesichtslosen Blicks*. Vortrag im Rahmen des Symposiums »Datenspuren 2005 – Privatsphäre war gestern«, Kulturzentrum Scheune, Dresden, 08.05.2005.

W. COY: *Long-term preservation of Multimedia Artifacts*. PKI Lab, Dartmouth College, 17.4.2005.

W. COY: Is there Digital Property? Dartmouth College, Dept. of Computer Science, 3.5.2005

W. COY: *Roots & Inconsistencies of Intellectual Property*, Dartmouth College, Dept. of Philosophy, 18.5.2005

W. COY: *Informatik & Kino, Science & Fiction*, Keynote zur Tagung „Neue Medien in der Informationsgesellschaft - Film und Computer“ NMI 2005, 2.7.2005, TU Berlin

W. COY: *Die Geburt des globalen Netzes aus dem Geist der Zahl*, Friedrich-Nietzsche-Archiv, Weimar 1.7.2005

W. COY: *Schrift-Bild-Zahl*, Innsbruck, 7.10.2005

W. COY: *Geistiges Eigentum in einer digitalen Welt*, Fiff Jahrestagung, München, 5.11.2005

W. COY: *Visuelle Argumente – Denken mit Bildern*, Technische Universität Cottbus, Philosophie der Technik, 23.11.2005

W. COY: *Schrödingers Katze wird 70*, cms Berlin 29.11.2005

CH. DAHME: *Was lässt sich an einer Tätigkeit objektivieren?* 2. Workshop „Tätigkeitstheorie und kulturhistorische Schule“ vom 8. bis 10. Juli 2005, im Haus Ohrbeck bei Osnabrück.

CH. DAHME: *Möglichkeiten und Grenzen der Tätigkeitstheorie am Beispiel der frühen Phasen der Softwareentwicklung*. Tagung „Netzwerk Organisation: Bildungstechnologie – Tätigkeitstheorie – Regulation – Lernen und Ethik“, am 18. und 19.11.2005 im Clubhaus der Freien Universität Berlin.

V. GRASSMUCK: *iRights.info-Copyrights in the Digital World*, exhibition & lecture together with Matthias Spielkamp, ValieDjordjevic, at the Transmediale 05, Haus der Kulturen der Welt, Berlin, 5.2.2005.

V. GRASSMUCK: *Creative Commons*, Goethe Institut Hongkong, 12.03.2005.

V. GRASSMUCK: Teilnehmer Podiumsdiskussion *Privatkopie*, Heise-Forum auf der CeBIT, Hannover, 13.05.2005

V. GRASSMUCK: „Bezahlen in der Wissensgesellschaft: pauschal oder a la carte?“ auf Workshop „Informationsgesellschaft“ der Berliner Stiftungsverbundkollegiatur, 29.4.-1.5.05, Schloß Reichenow 1. Mai 2005.

V. GRASSMUCK: „LiMux - Free Software for Munich“, held at: 6th Forum Internacional Software Livre, PUCRS, Porto Alegre, Brazil, 3.6.2005;
<http://fisl.softwarelivre.org/6.0/>

- V. GRASSMUCK: „LiMux - Free Software for Munich“, held at: Computer Department of USP, Sao Carlos, Brazil, 7.6.2005.
- V. GRASSMUCK: Vortrag „DRM“ auf „Neue Medien in der Informationsgesellschaft - Film & Computer“, TU Berlin, 01.07.2005.
- V. GRASSMUCK: Vortrag „Japanische Schrift“ auf dem Workshop „Alphabete“ des Helmholtz-Zentrums, HU Berlin, 09.07.2005
- V. GRASSMUCK: Workshop „Licensing for free publication of intellectual or artistic content in the www: the BBC Creative Archive“, ars electronica, Linz, 03.09.2005.
- V. GRASSMUCK: Panel „(No) Free Access to Information in so-called Information Society?“, ars electronica, Linz, 04.09.05
- V. GRASSMUCK: WOS at WSFII, London, 30.9.05, <http://www.okfn.org/wsfii/>
- V. GRASSMUCK: Panel-Speaker auf „Copyright in the P2P Era: An Exclusive Right, a Remuneration Right or What Exactly?“, org. Prof. Alain Strowel, Facultés universitaires Saint-Louis, Centre Innovation Propriété intellectuelle (CIPI), Bruxelles, 14.10.2005.
- V. GRASSMUCK: „The Digital Commons in Development“, at Submidialogia, Campinas, Brasilien, 29.10.2005.
- V. GRASSMUCK: „LiMux“, Vortrag auf FOSS.IN, 29.11.-2.12.05 Bangalore, 02.12.2005.
- V. GRASSMUCK: „Creative Commons“, Keynote auf FOSS.IN, 29.11.-2.12.05, Bangalore, 2.12.2005.
- V. GRASSMUCK: „A Hitchhiker’s Guide to the Turing Galaxy“, at kuda.org, Novi Sad, 15.12.2005.
- V. GRASSMUCK: „A Hitchhiker’s Guide to the Turing Galaxy“, at the Faculty of Philosophy, University of Belgrade, 16.12.2005.
- V. GRASSMUCK: WOS4 Workshop, auf 22C3, Berlin, 29.12.2005.
- J. KOUBEK: *Informatik und Gesellschaft*. Workshop bei der Fachtagung der GI-Fachgruppe IBBB, Potsdam, 24.02.2005.
- J. KOUBEK: *The Review System*. TESI 2005, Maastricht, 22.03.2005.
- J. KOUBEK: *Informatik und Gesellschaft lehren*. 12. Königsteiner Gespräche zur Fachdidaktik, 14.03.2005.
- J. KOUBEK: *Informatik, Mensch, Gesellschaft – Kompetenzen*. 12. Königsteiner Gespräche zur Fachdidaktik, 15.03.2005.
- J. KOUBEK: *Fernuni for the Rest of Us*. Oberseminar »Informatik und Gesellschaft«, Lanke, 03.08.2005.
- J. KOUBEK: *Informatische Allgemeinbildung*. INFOS 2005, Dresden, 28.09.2005.
- J. Koubek: *E-Paper is Tomorrow. The Future of Disposable Information*. MIST 2005, Berlin, 19.09.2005.
- J. KOUBEK: *Informatik – Mensch – Gesellschaft*. Keynote bei der HILL 2005, Hamburg, 29.10.2005.

J. KOUBEK: *Informatische Bildung*. Oberseminar »Informatik und Gesellschaft«, Weimar, 10.12.2005.

R. KUBICA: *Post-Mortem Analyse eines Servereinbruchs*. Oberseminar »Informatik und Gesellschaft«, Weimar, 11.12.2005.

C. KURZ: *Biometrische Feldtests in Europa*. Oberseminar »Informatik und Gesellschaft«, Weimar, 10.12.2005.

C. KURZ: *Wir haben etwas zu verbergen - Die Informationelle Selbstbestimmung und der Schutz der Privatsphäre*. Projekt P - Berlin05. Festival für junge Politik des BMF und der Bundeszentrale für politische Bildung, 11.06.2005.

C. KURZ, C. YXEN: *Datenspuren im Netz - Ich weiß, was du gestern gesurft hast*. Projekt P - Berlin05. Festival für junge Politik des BMF und der Bundeszentrale für politische Bildung, 12.06.2005.

C. KURZ: *Biometrie und ePass*. Oberseminar »Informatik und Gesellschaft«, Lanke, 03.08.2005.

C. KURZ: *Die BioP-II-Studie des BSI*. 22c3 Chaos Communication Congress, Berlin, 27.12.2005.

C. KURZ, R. KUBICA: *Biometrics in Science Fiction*. 22c3 Chaos Communication Congress, Berlin, 27.12.2005.

J.-M. LOEBEL: *Securing MacOS X Tiger for everyday use*. Oberseminar »Informatik und Gesellschaft«, Lanke, 03.08.2005.

J.-M. LOEBEL: *Langzeitarchivierung multimedialer Objekte*. Oberseminar »Informatik und Gesellschaft«, Weimar, 10.12.2005.

Hörspielwerkstatt

11.06.2005 Lange Nacht der Wissenschaften: *Neue und Aktuelle Erkenntnisse aus den Forschungsbereichen «Lehrreiches Wissen» und «Nützliche Argumente»*. Vorgetragen und zu Gehör gebracht von der Hörspiel-Werkstatt der Humboldt Universität zu Berlin.

29.06.2005, What The Hack: *Proof by Erasure. The Life of John von Neumann*.

28.12.2005, 22c3 Chaos Communication Congress: *Have you hugged your Mac today? Die Geschichte von Apple*.

DVD-Produktion *The Incomputable Alan Turing*.

Sonstige Aktivitäten

Peter Bittner

- Stellv. Mitglied im Institutsrat
- Mitglied der Bibliothekskommission des Instituts
- Stellvertretender Vorsitzender des Forum InformatikerInnen für Frieden und gesellschaftliche Verantwortung (FifF) e.V.
- Delegierter bei European Digital Rights (EDRI)
- Fachexperte im Leitungsgremium des Fachbereichs „Informatik und Gesellschaft“ (FB IuG) der Gesellschaft für Informatik (GI)

- Mitglied im Programmkomitee der 21. FIF-Jahrestagung „Versteckte Computer – Unkontrollierbare Vernetzung“, München, 05.-06.11.2005
- Mitglied im Beirat des Webportals iRights.info (Verbraucherinformation zum Urheberrecht)
- Mitglied der Initiative privatkopie.net
- Mitglied der Berliner Aktion für Datenschutz und informationelle Selbstbestimmung
- Mitglied der Fachgruppe „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)
- Mitbegründer und Reihenherausgeber (gemeinsam mit Jens Woinowski, München) der Buchreihe „Kritische Informatik“ beim LIT-Verlag
- Ehrenmitglied der Fachschaft Informatik der Universität Kaiserslautern

Wolfgang Coy

- Harris German Distinguished Visiting Professor am Dartmouth College, Hanover, New Hampshire von März bis Mai 2005
- Mitglied des Konzils der HU
- Mitglied des Akademischen Senats der HU
- Vorsitzender der zentralen Medienkommission des Akademischen Senats
- Prodekan der mathematisch-naturwissenschaftlichen Fakultät II
- Mitglied des Fakultätsrates der mathematisch-naturwissenschaftlichen Fakultät II
- Mitglied des Institutsrates Informatik
- Aufsichtsrat der Multimedia Hochschulgesellschaft MHSg der Berliner Hochschulen
- Deutscher Delegierter im Technical Committee 9 „Computers and Society“ der International Federation for Information Processing (IFIP)
- Sprecher der IFIP Working Group 9.5 „Virtuality“
- Fellow der Gesellschaft für Informatik (GI)
- Fachexperte des FB8 „Informatik und Gesellschaft“ der GI
- Mitglied der Grünen Akademie der Heinrich Böll-Stiftung
- Mitherausgeber des „Informatik Spektrum“ (Springer Verlag)

Christian Dahme

- Mitglied des Konzils
- Mitglied des Konzilvorstandes
- Mitglied der Verfassungskommission des Konzils
- Mitglied des Akademischen Senats
- Mitglied der Kommission Lehre und Studium des Akademischen Senats
- Mitglied der Arbeitsgruppe Mittelbau beim Präsidium der HU
- Prodekan für Lehre und Studium – Studiendekan der mathematisch-naturwissenschaftlichen Fakultät II
- Mitglied des Fakultätsrates der mathematisch-naturwissenschaftlichen Fakultät II
- Mitglied der Arbeitsgruppe zur Reform der Lehramtsausbildung (BA, MA) in Informatik in Berlin
- Member of editorial board of (der Zeitschrift) tripleC (<http://triplec.uti.at/>)
-

Volker Grassmuck

- Steuerungsgruppe des Aktionsbündnisses „Urheberrecht für Bildung und Wissenschaft“
- Co-Chair der IP Working Group der European Digital Rights Initiative (EDRI)
- Projektleiter des Informationsportals „iRights.info“
- Projektleiter der Konferenzserie „Wizards of OS“

Jochen Koubek

- Sprecher der GI-Fachgruppe *Computer als Medium* des FB 8 der GI
- Mitglied der Fachgruppe *IBBB – Informatik-Bildung in Berlin und Brandenburg* des FB 7 der GI
- Mitglied im Gremium *Treffpunkt Internet – Forum Digitale Integration* des BMWA
- Mitglied im Expertenteam zur Erarbeitung von Bildungsstandards im FB 7 der GI
- Mitglied der Arbeitsgruppe zur Reform der Lehramtsausbildung in Informatik in Informatik in Berlin
- Leiter der Schüler-AG »Informatik - Mensch - Gesellschaft«
- Redaktionelle Verantwortung für das LOG IN Heft Nr. 136/137 »Gesellschaftliche Themen in der informatischen-Bildung«
- Betreuung verschiedener Schülerpraktika
- Leiter, Autor und Sprecher der Hörspiel-Werkstatt
- Organisation der Campus-Konzerte Adlershof

Constanze Kurz

- Mitglied in der Fachgruppe „Informatik und Verantwortung“ des FB 8 der GI
- Leiterin der Schüler-AG »Informatik - Mensch - Gesellschaft«
- Redakteurin der »Datenschleuder – Das wissenschaftliche Fachblatt für Datenreisende«
- Hessischer Rundfunk (Frankfurt am Main), Sendung Nightline: zweistündige Radiosendung zum Thema ePass, Biometrie, Kryptographie, Datenschutz und Informationelle Selbstbestimmung
- Autorin und Sprecherin der Hörspiel-Werkstatt

Betreute und abgeschlossene Diplomarbeiten 2005

ROBERT STOLL: *Separierung des Vordergrundes in Echtzeit aus Live-Videoströmen in immersiven Telekonferenzszenarien: technologischer Überblick, zwischenmenschliche Aspekte*

JAN KINDER: *Langzeitarchivierung mittels Emulatoren*

Heinz-Günter Kuper, *"Please, Sir, I Want Some More": An Adaption of Oliver Twist as a Textual Interactive Narrative*

MICHAEL PLATE: *Erweiterung des Dokumentenservers Dspace: Internationalisierung und Implementation des URN-Systems zur permanenten Dokumentenadressierung (Masterarbeit)*

SEBASTIAN THIEBES: *Synthese von Gesichtsbewegungen durch Analyse von Sprache*

KAY SCHUBERT: *Dokumentation des songpatternbasierten Stepsequenzers SOCAWO*

CONSTANZE KURZ: *Some Rights Reserved? Sichtweisen auf geistiges Eigentum*

DANIEL MATZKE: *3D-Visualisierung von Wasser – eine Betrachtung von Simulations-, Animations- und Shadingalgorithmen unter Echtzeitaspekten.*

MARCUS RICHTER: *Das wohltemperierte Massenmedium: Chancen inhaltlicher Interaktivität für die Online-Angebote klassischer Massenmedien.*

DANIEL GÖRLICH: *Geschichte und Zukunft von Eingabeschnittstellen in der Mensch-Computer-Interaktion.*

Lehr- und Forschungseinheit

Parallele und Verteilte Systeme

<http://www.zib.de/csr/>

Leiter

PROF. ALEXANDER REINEFELD

Tel.: (030) 84185-130

E-Mail: ar@zib.de

Sekretariat

PETRA FEHLHAUER

Tel.: (030) 84185-131

E-Mail: fehlhauer@zib.de

Wissenschaftliche Mitarbeiterinnen und Mitarbeiter

DR. ARTUR ANDRZEJAK

MIKAEL HÖGQVIST

SEBASTIAN HEIDL

FELIX HUPFELD

TOBIAS LANGHAMMER

PATRICK MAY

KATHRIN PETER

THOMAS RÖBLITZ

FLORIAN SCHINTKE

THORSTEN SCHÜTT

DR. THOMAS STEINKE

Studentische Mitarbeiterinnen und Mitarbeiter

GUILLERMO ARIAS DEL RIO

MATTHIAS BEICK

MEHMET CEYRAN

HANS-CHRISTIAN EHRlich

TINOSCH KOLAGHEICHI-GANJINEH

MINOR GORDON

RENE HEEK

ULF HERMANN

FALKO KRAUSE

JONAS MAASKOLA

JÖRG MELTZER

MONIKA MOSER

DANIEL MAUTER

STEFAN PLANTIKOW

CHRISTIAN VON PROLLIUS

ROLAND TUSCHL

SASCHA WILLUWEIT

JOCHEN WITTE

Die wissenschaftliche Arbeit am Lehrstuhl konzentriert sich auf die Entwicklung von Methoden und Algorithmen für parallele und verteilte Systeme, derzeit insbesondere im Bereich des Grid Computing. Der Lehrstuhlinhaber leitet zugleich den Bereich Computer Science (mit Hochleistungsrechenzentrum) am Zuse-Institut Berlin (ZIB) wodurch neueste Erkenntnisse des wissenschaftlichen Hochleistungsrechnens auf direktem Wege in den Lehr- und Forschungsbetrieb der Humboldt-Universität einfließen. Die Lehrveranstaltungen und Forschungsarbeiten finden institutsübergreifend an beiden 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 (IBM p690, Cray XD1, Cluster) und bietet seinen Anwendern fachspezifische Beratung zur effizienten Nutzung der Rechnersysteme. 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 ca. 5 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 können im Web unter www.hlrn.de nachgelesen werden.

Der Lehrstuhl Parallele und Verteilte Systeme deckt ein breites Forschungsspektrum ab, das von systemnahen Arbeiten an 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.

Ein zentraler Forschungsschwerpunkt liegt in der Entwicklung von Methoden zum Management geographisch verteilter Ressourcen wie z. B. Rechenleistung, Speicherkapazität, Netzwerke oder Visualisierungskomponenten. Unser Ziel ist es, diese Ressourcen in einem überregionalen Lastverbund verfügbar zu machen, so dass die Nutzer möglichst einfach darauf zugreifen können. Das System 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.

Derzeit konzentrieren sich unsere Forschungsarbeiten schwerpunktmäßig auf das Management sehr großer Datenmengen in verteilten Grid-Systemen. Unter der Bezeichnung „ZIB-DMS“ entwickeln wir – basierend auf Peer-to-Peer-Techniken – skalierbare Software zur sicheren und effizienten Speicherung attributierter Dateien in verteilten Systemen. Mehrere Doktoranden, Diplomanden und Studenten arbeiten an dieser komplexen Thematik mit dem Ziel, eine Software zum effizienten Management von Millionen Dateien im Grid zu entwickeln. Da die Speicherknoten im Grid jederzeit transient oder permanent ausfallen können, ist eine geeignete Datenreplikation unabdingbare Voraussetzung. Wir entwickeln Methoden für Datenplatzierung, Caching, Staging, Routing, Scheduling und Synchronisation.

Zum praktischen Test unserer Datenmanagement-Verfahren dienen komplexe Anwendungen der Bioinformatik, die gleichfalls an unserem Lehrstuhl entwickelt werden. Die von Dr. Thomas Steinke geleitete 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

Vorlesungen

- Verteilte Systeme: Konzepte und Basisalgorithmen (A. REINEFELD, A. ANDRZEJAK, SoSe 2005)
- Einführung in das Molecular Modeling (T. STEINKE, SoSe 2005)
- Data Mining - Algorithmen und Anwendungen (A. ANDRZEJAK, WiSe 2005/2006)

Seminare

- Ressourcen-Management im Cluster und Grid (A. REINEFELD, T. RÖBLITZ, SoSe 2005)
- Forschungsseminar Grid Computing (A. REINEFELD, SoSe 2005)
- Anwendungen und spezielle Themen in Data Mining, (A. ANDRZEJAK, SoSe 2005)
- Parallele Algorithmen in der Bioinformatik (A. REINEFELD, T. STEINKE, WiSe 2005/06)
- Forschungsseminar Grid Computing (A. REINEFELD, WiSe 2005/06)

Projekt

- FPGA-Programmierung für die Bioinformatik auf der Cray XD1 (A. REINEFELD, T. STEINKE, T. SCHÜTT, WiSe 2005/06)

Forschung

Projekt: CoreGrid

Ansprechpartner: Dr. ARTUR ANDRZEJAK

Beteiligte Mitarbeiter: FELIX HUPFELD, PROF. ALEXANDER REINEFELD, THOMAS RÖBLITZ, FLORIAN SCHINTKE, THORSTEN SCHÜTT, DR. THOMAS STEINKE

Forschungsförderung: Europäische Union

Das Projekt CoreGrid („European Research Network on Foundations, Software Infrastructures and Applications for Large-Scale Distributed, Grid and Peer-to-Peer Technologies“) unterstützt die wissenschaftliche Zusammenarbeit von 41 europäischen Universitäten und Forschungsinstituten. Als so genanntes *Network of Excellence* soll es die Qualität der Forschung im Bereich Grid und Peer-to-Peer-Computing steigern und die Ergebnisse einer breiteren Öffentlichkeit verfügbar machen. Dies geschieht durch gemeinsame Forschungsaktivitäten, Symposien sowie gegenseitige Kurzbesuche und Stipendien für Postdokoranden.

Wir sind an den virtuellen Instituten „System Architecture“ und „Resource Managment“ beteiligt, wobei das Hauptgewicht auf dem Institut „System Architecture“ (WP4) liegt. Dieses Institut wird von Dr. Artur Andrzejak seit der Entstehung des Netzwerkes geleitet. Die zentralen Themen dieses Institutes liegen in der Untersuchung von Skalierbarkeit, A-

daptivität und Fehlertoleranz von Grid-Architekturen. Im Jahr 2005 haben wir folgende Treffen der beteiligten Partner organisiert:

- First CoreGRID WP4 Meeting – 18. Januar 2005, Heraklion, Griechenland,
- Second CoreGRID WP4 Meeting – 14. Juli 2005, Barcelona, Spanien,
- Informal WP4 meeting – 31. August 2005, Lissabon, Portugal.

Darüber hinaus haben wir – zusammen mit dem Partner FORTH, Griechenland – den *1st CoreGRID Workshop on GRID and P2P Systems Architecture* organisiert. An diesem Workshop waren neben CoreGrid-Mitgliedern folgende eingeladene Sprecher beteiligt: Mema Roussopolous (Harvard Univ.), Nektario Koziris (NTU Athens) und Ian Taylor (Cardiff Univ.). Größere Beiträge zur Leitung von WP4 umfassten die Erstellung der *Roadmap Version 1 of the Institute on System Architecture (D.SA.01)* (zusammen mit anderen Partnern) und die Vorbereitung des *Joint Program of Activities JPA2*, die die Arbeit des virtuellen Instituts WP4 in den Projektmonaten 18 bis 30 spezifizierten.

Projekt: GridCoord

Ansprechpartner: PROF. ALEXANDER REINEFELD

Beteiligte Mitarbeiter: FLORIAN SCHINTKE, THOMAS RÖBLITZ

Zusammenarbeit: DIST (I), UP (I), HLRS (DE), QUB (IR), EPSRC (UK), INRIA (F), UNSA (F), ZIB (DE), UvA (NL), MTA SZTAKI (HU), UPM (ESP), VINNOVA (SE), PSNC (PL)

Forschungsförderung: Europäische Union

Zur Weiterentwicklung des Grid Computing sind in Europa viele nationale und europäische Forschungsprogramme etabliert worden, die es sich zum Ziel gesetzt haben, neue Methoden, Technologien und Anwendungen zu entwickeln. Eine Koordination zwischen den Initiativen ist notwendig, um sowohl die nötige kritische Masse an Förderung zu erreichen, als auch Doppelförderung und Fragmentierung zu vermeiden, damit die Herausforderungen zielstrebig gemeistert werden können.

Ein zentrales Ziel von GridCoord („ERA Pilot on a Co-ordinated Europe-Wide Initiative in Grid Research“) liegt darin, die Kooperation zwischen den nationalen Förderinstitutionen und der EU zu stärken, um die Planung zukünftiger Grid-Forschung besser koordinieren zu können. Zusätzlich sollen bestehende Kooperationen zwischen Forschern und Anwendern intensiviert werden. Auf Basis der Ergebnisse der Analyse europäischer und nationaler Forschungsförderung soll ein abgestimmtes europäisches Forschungsförderprogramm formuliert werden, das es Europa erlaubt, eine führende Rolle im Bereich der Grid-Technologien und Anwendungen einzunehmen.

Projekt: C3-Grid

Ansprechpartner: FLORIAN SCHINTKE

Beteiligte Mitarbeiter: PROF. ALEXANDER REINEFELD, TOBIAS LANGHAMMER

Forschungsförderung: BMBF

Im D-Grid-Projekt C3-Grid („Collaborative Climate Community Data and Processing Grid“) soll eine hochproduktive, grid-basierte Umgebung für die deutsche Erdsystemforschung geschaffen werden, die effizientes verteiltes Datenprocessing und interinstitutionellen Austausch von hochvolumigen Modell- und Beobachtungsdaten erlaubt. Das kollaborative Gesamtsystem soll einen einheitlichen und transparenten Zugriff auf geographisch

und institutionell verteilte Archive erlauben und die Verarbeitung großer Datenmengen erleichtern.

Bei der Lösung der Problemstellung wird aufgrund der großen Datenvolumina nicht nur, wie in den meisten bisherigen Grid-Umgebungen die Rechenkapazität als geographisch verteilte Ressource behandelt, sondern auch der persistente Speicher auf Festplatten oder Magnetbändern, sowie Netzwerkleitungen. Durch geeignete Planung können voluminöse und lang andauernde Datenübertragungen vermieden werden, indem Rechnungen bevorzugt auf Rechenressourcen in der Nähe der Daten ausgeführt werden. Zudem sollen große Datentransfers, falls dennoch nötig, nicht einfach bei Bedarf angestoßen werden, sondern automatisch zeitlich geplant werden, um Zeiträume mit Freikapazitäten, beispielsweise in der Nacht, effektiver zu nutzen. Unsere Arbeitsgruppe entwickelt hierfür einen geeigneten Grid-Informationdienst mit dazugehörigem Grid-Datenmanagement.

Projekt: DGI

Ansprechpartner: FLORIAN SCHINTKE

Beteiligte Mitarbeiter: TOBIAS LANGHAMMER, DANIEL MAUTER

Forschungsförderung: BMBF

Das DGrid-Projekt DGI („D-Grid-Integrationsprojekt“) hat das Ziel, in Deutschland eine robuste, flexible und nachhaltig nutzbare Grid-Infrastruktur zu schaffen, die für Wissenschaft und Industrie zur Basis neuer e-Science-basierter Arbeitsformen werden wird. Im Teilbereich, an dem unsere Arbeitsgruppe beteiligt ist, sollen zum Management großer Datenvolumina Stagesysteme im Grid auf Basis existierender und in einzelnen Communities bereits eingesetzter Software bereitgestellt und möglichst transparenter Zugriff ermöglicht werden. Unsere Arbeitsgruppe wird hier insbesondere die Einsatzmöglichkeiten des Storage Resource Brokers *SRB* des San Diego Supercomputer Center untersuchen, einem etablierten Datenmanagement-Softwarepaket zur Handhabung verteilter und gegebenenfalls replizierter Daten.

Projekt: AstroGrid-D

Ansprechpartner: THOMAS RÖBLITZ

Beteiligte Mitarbeiter: MIKAEL HÖGQVIST

Forschungsförderung: BMBF

Das Projekt AstroGrid-D („German Astronomy Community Grid“) ist ein Forschungsprojekt zur Unterstützung des verteilten wissenschaftlichen Arbeitens im Rahmen der deutschen D-Grid-Initiative. Ziel dieses Vorhabens ist die Schaffung einer Grid-basierten Infrastruktur für die deutsche astronomische und astrophysikalische Forschung, aufbauend auf dem Rechnerverbund der teilnehmenden Institute und Rechenzentren, um die existierenden Hardwareressourcen effektiver nutzen zu können.

Das resultierende *German Astronomy Community Grid* wird bereits bestehende, geographisch verteilte Hochleistungsrechner-Ressourcen mit großen astronomischen Datenarchiven, ferngesteuerten Radio- und Roboterteleskopen sowie bodengebundenen Gravitationswellen-Detektoren in einem kohärenten Rechen- und Daten-Grid integrieren. Standardisierte Benutzerschnittstellen ermöglichen einen einheitlichen, ortsunabhängigen Zugang zu den vorhandenen Rechenressourcen, was die Durchführung astrophysikalischer, numerischer Simulationen vereinfacht und effizienter gestaltet. Der transparente Zugriff sowohl auf lokale als auch im Grid verteilte astronomische Datenarchive soll für die in der Com-

munity entwickelten Datenanalyse-Anwendungen eine vereinheitlichte Verwaltung, Extraktion und Verarbeitung von komplexen Datensätzen ermöglichen.

Projekt: MediGRID – Modul Ressourcenfusion

Ansprechpartner: DR. THOMAS STEINKE

Beteiligte Mitarbeiter: KATHRIN PETER

Forschungsförderung: BMBF

Das MediGRID-Vorhaben („Ressourcenfusion für Medizin und Lebenswissenschaften“) realisiert für geeignete Anwendungsfälle mittels sechs miteinander verzahnter Module, die durch ein Koordinations- und ein eScience-Modul flankiert werden, eine Grid-Infrastruktur für Medizin und biomedizinische Informatik. Das Modul Ressourcenfusion beinhaltet Forschungs- und Entwicklungsaufgaben mit der langfristigen Zielsetzung, eine Grid-Infrastruktur zur breiten Integration der bioinformatischen und medizinischen Forschung zu ermöglichen, sowie die technisch-organisatorische Koordination der Ressourcenprovider für MediGRID-Anwendungen.

Projekt: Co-Reservierung von Ressourcen im Grid

Ansprechpartner: THOMAS RÖBLITZ

Beteiligte Mitarbeiter: JÖRG MELTZER, PROF. ALEXANDER REINEFELD, FLORIAN SCHINTKE

Forschungsförderung: IBM Deutschland Entwicklungs GmbH

Die Lösung wissenschaftlicher Probleme und das Bestreben der Industrie Rechenleistung flexibel anbieten zu können, stellen komplexere Anforderungen an die Verwaltung der verfügbaren Ressourcen. Insbesondere ist eine Koordination mehrerer Aktivitäten, z.B. die Arbeitsschritte in einem Workflow, erforderlich.

Die Arbeiten dieses Projektes zielen darauf ab, die Koordination mit im voraus getroffenen Qualitätsvereinbarungen zu realisieren. Reservierungen von Ressourcen sind eine Möglichkeit, solche Vereinbarungen zu gewährleisten.

Ein Reservierungssystem muss dabei gleichzeitig die Anforderungen verschiedener Beteiligter – Ressourcenprovider, normale Benutzer und Benutzer, die Reservierungen verwenden – berücksichtigen. Die Auswahl erfolgversprechender Reservierungsparameter ist insbesondere in Grid-Umgebungen, in denen Zustandsinformationen über die Ressourcen unvollständig (Datenschutzbeschränkungen) oder veraltet sind, schwierig. Im Berichtszeitraum haben wir die bereits entwickelten Algorithmen für die Bearbeitung flexibler Reservierungsanfragen weiter verbessert und umfangreiche Simulationen durchgeführt, um allen Beteiligten Empfehlungen für die Benutzung eines solchen Systems geben zu können. Die bisher entwickelten Verfahren für die Auswahl der Reservierungsparameter können nur auf wenige Anfragemuster – CPU-Ressourcen, die für das gleiche Zeitintervall reserviert werden sollen – angewendet werden. Im nächsten Schritt sollen diese Verfahren verallgemeinert werden, um unterschiedliche Ressourcen, die zu verschiedenen oder gleichen Zeiträumen benötigt werden, in Form einer einzigen Reservierung behandeln zu können. Darüber hinaus soll die Leistungsfähigkeit des Verfahrens bei konkurrierenden Anfragen ermittelt werden.

Projekt: Alignment und Threading auf massiv parallelen Rechnern**Ansprechpartner:** DR. THOMAS STEINKE**Beteiligte Mitarbeiter:** CHRISTIAN EHRLICH, RENÉ HEEK, FALCO KRAUSE, JONAS MAASKOLA, PATRICK MAY**Zusammenarbeit:** FU Berlin, HU Berlin, Max-Delbrück-Centrum für Molekulare Medizin Berlin-Buch, Max-Planck-Institut für Molekulare Genetik, TFH Berlin.**Forschungsförderung:** BMBF, BCB Forscher-Nachwuchsgruppe

In 2005 lagen schwerpunktmäßig die Forschungs- und Entwicklungsaufgaben der Nachwuchsgruppe „Alignment und Threading auf massiv parallelen Rechnern“ des „Berliner Centrums für genombasierte Bioinformatik“ (BCB) bei der Implementierung eines vollständig auf Web Services basierenden Workflows zur Proteinstrukturvorhersage. Kernkomponente des Workflows ist die Implementierung des Threading-Ansatzes zur Proteinfaltungserkennung. Zusätzlich sind im Workflow auch Web Services für das Homology Modeling integriert, um auch Sequenzen mit einer hohen Sequenzähnlichkeit zu einer bekannten Struktur modellieren zu können. Damit werden Voraussetzungen geschaffen, um in 2006 mit einem automatischen Strukturvorhersage-Server am CAFASP-Wettbewerb teilzunehmen und bei Bedarf diesen Dienst für externe Interessenten anbieten zu können.

Die Threading-Methode sucht auf der Basis experimenteller Daten aus einer gegebenen Menge von Proteinstrukturen für eine Sequenz mit unbekannter Struktur diejenige aus, die der nativen Struktur am ähnlichsten ist. Die Strukturvorhersage-Pipeline umfasst in der Phase der Vorprozessierung die Analyse der Sequenz, dem folgt das eigentliche Threading, und schließlich die 3D-Modellierung.

Neben dem in der Arbeitsgruppe entwickelten Framework zur Faltungserkennung (Theus) werden in der Strukturvorhersage-Pipeline Legacy-Codes wie PsiBlast (Sequenzanalyse), PsiPred (Sekundärstrukturvorhersage), und Modeller (Homology Modeling) verwendet. Alle im Workflow benutzten Programme wurden als Web Services gekapselt. Die rechenintensiven Schritte sind als asynchrone Web Services implementiert. Diese Web Services interagieren mit dem Ressourcenmanagementsystem (Torque) eines Compute-Clusters, um einerseits die Ressourcenanforderungen für serielle und parallele Jobs zu befriedigen und andererseits eine effektive Nutzung der vorhandenen Rechenressourcen zu garantieren.

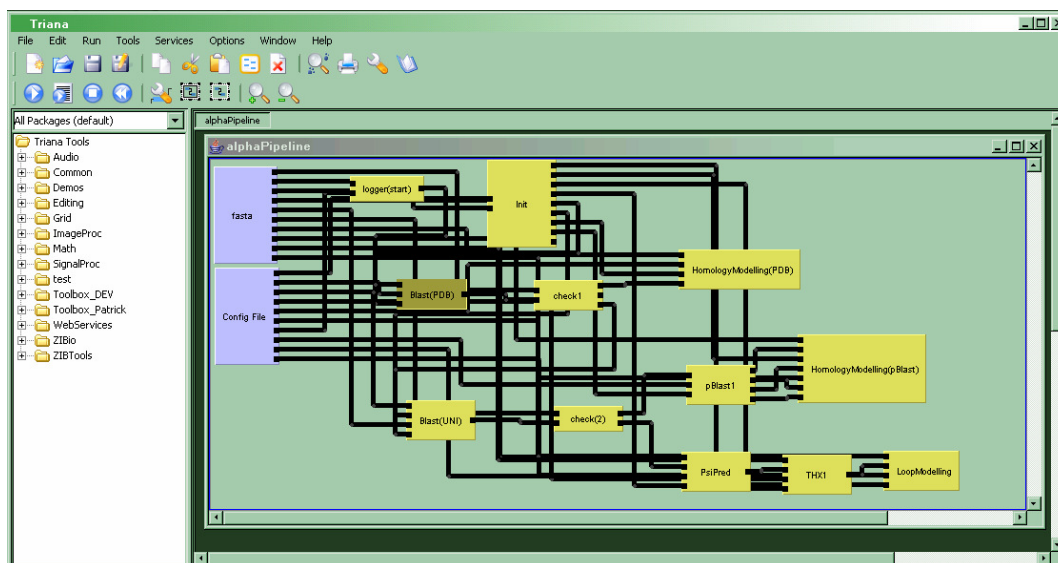


Abbildung 9: Kompakte Darstellung des Workflows zur Proteinstrukturvorhersage in der Triana-Workflow-Engine

Aufgrund der guten Unterstützung bei der Orchestrierung von Web-Service-basierten Workflows wurde als Workflow-Engine Triana gewählt.

Abbildung 9 zeigt eine kompakte Darstellung des Designs für den Strukturvorhersage-Workflow in Triana. Im Vergleich zu einer monolithischen Perl-Implementierung des Workflows ist der zusätzliche Aufwand durch die Triana/Web Service basierte Umsetzung unter 1%, so dass die Vorteile des modularen Designs praktisch überwiegen.

Projekt: StorageBox – Effiziente optimistische Replikation für mobile Datenbanken

Ansprechpartner: FELIX HUPFELD

Beteiligte Mitarbeiter: MINOR GORDON

Als persönliches Datenbanksystem ist StorageBox (www.storagebox.org) darauf ausgelegt, alle beim alltäglichen Gebrauch von PCs anfallenden Daten strukturiert zu speichern, und für den Benutzer und andere Programme komfortabel zugreif- und verwaltbar zu machen. Um den Benutzer bei der Verwaltung seiner Daten auf mehreren Geräten zu entlasten, sorgt StorageBox selbständig für Verteilung und Abgleich der Daten. Damit kein Gerät als zentrale Datenbank dienen muss, kann der Abgleich geänderter Daten auch über Dritte geschehen.

Hauptproblem eines solchen Abgleichs ist es, zu erkennen, welche Änderungen zur gleichen Zeit auf verschiedenen Kopien vorgenommen wurden und somit Konflikte enthalten können. Für die Datenbank von StorageBox wurde hierzu ein Algorithmus entwickelt, der zu einer Klasse von Ansätzen gehört, die auf aufgezeichneten Änderungsprotokollen basieren. Im Gegensatz zu anderen Ansätzen zeichnet der Algorithmus die Abgleichoperationen mit auf, und nutzt diese um gleichzeitige Änderungen zu erkennen und zu benennen. Da außer diesen mitprotokollierten Abgleichoperationen keine weiteren Informationen benötigt werden, ist der Algorithmus sehr speicher- und kommunikationseffizient.

Projekt: Zoetrop – Ein Webserver aus Verarbeitungsstufen

Ansprechpartner: FELIX HUPFELD

Beteiligte Mitarbeiter: MINOR GORDON

Server Software muss so strukturiert sein, dass sie eine hohe Anfragerate parallel verarbeiten kann. Da die Computerhardware nur eine physische Ausführungseinheit (Prozessor) besitzt, wird diese vom Betriebssystem in virtuelle Ausführungseinheiten, sogenannte Threads, unterteilt, die in kurzer Abfolge durchgeschaltet werden.

Neben dem klassischen Ansatz, bei dem pro ankommender Anfrage ein solcher Thread auf Daten vom Netzwerk wartet und alle nötigen Schritte zur Bearbeitung einer Anfrage abarbeitet, kann Serversoftware um die ankommenden Daten herum strukturiert werden. Dazu werden die nötigen Verarbeitungsschritte in einzelne Stufen zerlegt, die wie auf einem Förderband angekommene Ereignisse verarbeiten und daraus resultierende Ereignisse an die Warteschlange der jeweils nächsten Stufe weiterleiten. In Systemen mit mehreren Prozessoren legt man pro Prozessor eine solche Pipeline an, um den Kommunikationsbedarf zwischen den Prozessoren klein zu halten.

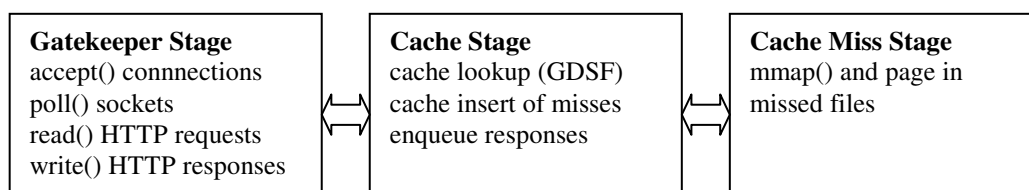


Abbildung 10: Die Verarbeitungsstufen des Zoetrop Webservers

Die einzelnen Verarbeitungsstufen werden jeweils von nur einem Thread ausgeführt. Daraus folgt insbesondere, dass die Verarbeitungslogik einer Stufe keine Operationen verwenden darf, die den Thread blockieren und somit die Verarbeitung von weiteren Ereignissen verzögern. Sind solche blockierenden Operationen notwendig, müssen sie in separate Verarbeitungsstufen ausgelagert werden. Vor diesem Hintergrund wurde die Abarbeitung von Webanfragen in Zoetrop in Stufen zur Netzwerkkommunikation (Warten auf Netzwerkeignisse), zum Versenden ge-cachter Dateien (Nachsehen im Cache), und zum Nachladen nicht ge-cachter Dateien (Warten auf Festplattenzugriff) gegliedert. Mit Hilfe des SPECweb99 Benchmarks konnte gezeigt werden, dass Zoetrop hinsichtlich Geschwindigkeit und Skalierbarkeit mit anderen Forschungswebservern mindestens vergleichbar ist.

Projekt: Grid-Systeme mit Peer-to-Peer Techniken

Ansprechpartner: THORSTEN SCHÜTT, FLORIAN SCHINTKE

Beteiligte Mitarbeiter: MONIKA MOSER, PROF. ALEXANDER REINEFELD, ROLAND TUSCHL, 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 Monitoring & Discovery Server (MDS) in Globus, durch hierarchisch vernetzte Implementationen ersetzt.

In letzter Zeit hat sich besonders die Forschung im Bereich Peer-to-Peer Computing (P2P) mit Skalierbarkeit und Zuverlässigkeit beschäftigt. Die bekanntesten Verfahren aus diesem

Bereich sind Lookup-Dienste und verteilte Hashtabellen. Die flachen, nichthierarchischen 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. Dadurch ist jeder Peer gleich wichtig bzw. unwichtig. Den Dienst erbringen alle Peers gemeinsam.

Die bekannten P2P-Systeme sind größtenteils monolithisch aufgebaut, was die Substitution einzelner Komponenten sowie deren Wiederverwendung in anderen Systemen erschwert. Grid-Systeme sind aber in der Regel sehr komplex und werden deshalb in viele kleine Dienste aufgeteilt. 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 in unserem Artikel der „Annual Reviews of Scalable Computing“ vorgestellt worden.

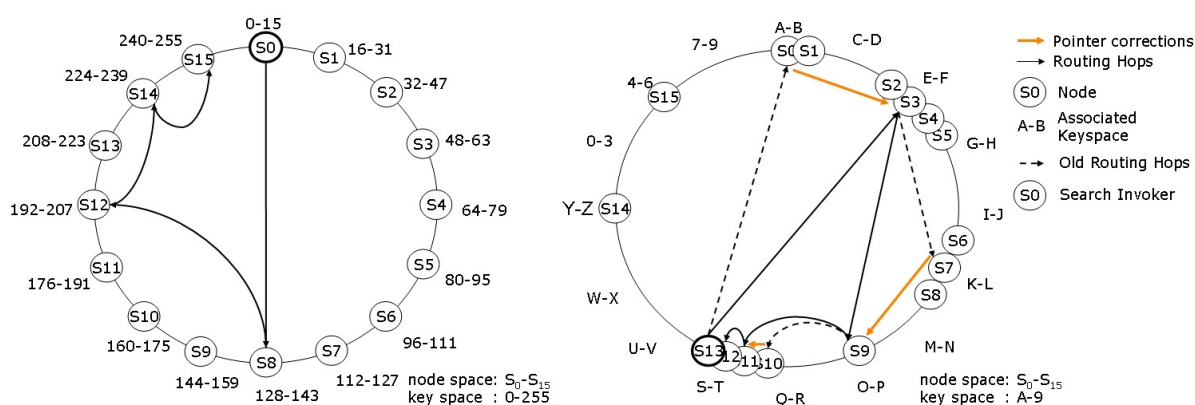


Abbildung 11: Von Chord nach Chord#

Existierende verteilte Hashtabellen wie z. B. Chord unterstützen nur Suchen nach exakten Werten. Um auch Bereichsabfragen unterstützen zu können, haben wir Chord# entwickelt, das neben flexibleren Suchmöglichkeiten, auch weniger Verwaltungsaufwand zum Aufrechterhalten des Systems benötigt. Weitere Informationen zu Chord# sind in unserem ZIB-Report nachzulesen.

Projekt: Adaptive Workflows für automatisches Management von Computersystemen

Ansprechpartner: DR. ARTUR ANDRZEJAK

Beteiligte Mitarbeiter: ULF HERMANN

Zusammenarbeit: Hewlett-Packard Labs

Forschungsförderung: ZIB

Automatisierung des Managements von Ressourcen, verteilten Computersystemen oder ganzen Rechenzentren ist in den letzten Jahren zu einem wichtigen Thema der Forschung avanciert. Die Hintergründe dafür sind die wachsende Komplexität und Anzahl der Komponenten von Systemen, die das manuelle Eingreifen zum Flaschenhals des Managements

machen, sowie der zunehmende Anteil der (traditionellen) Managementkosten an dem „Total Cost of Ownership“.

In diesem Projekt fokussieren wir die automatische Generierung von Workflows zum Zwecke der adaptiven Systemverwaltung. Derartige Workflows können komplexe Vorgänge durch Systemmanagement ausführen, wie Installation von Software in verteilten Systemen, Aufbau virtueller Rechnerfarmen in Grids und Rechenzentren, oder Bereitstellung der Ressourcen für spezielle Jobs. In unserem Ansatz verwenden wir deklarative Spezifikationen von möglichen Managementaktionen, welche mit Hilfe von Techniken des automatischen Planens zu einem Workflow zusammengesetzt werden, um einen vorgegebenen Zielzustand des Systems zu erreichen. Ein solches Ziel kann sein, die vorhandenen Ressourcen zu einer virtuellen Serverfarm zusammenzuschließen, wobei die benötigte Software als Zwischenschritt konfiguriert wird. Das Feedback über das Erreichen oder Nichterreichen des Zielzustands wird eingesetzt, um nach einer Neuberechnung des Systemzustands diesen Prozess gegebenenfalls zu wiederholen, wodurch die adaptive Anpassung des Workflows an die neuen Ausführungsbedingungen erreicht wird.

Nach der Umsetzung dieses Ansatzes in dem prototypischen System „FeedbackFlow“ fokussierten wir das Problem der einfacheren Formalisierung (Spezifikation) der Inputs, d. h. der Aktionen, Systemzustände, und Planungsziele. Die Hauptaktivität in diesem Bereich besteht in der Integration eines existierenden Editors für Petri-Netze (Renew) mit unserem Prototyp, um grafische Petri-Netz-basierte Spezifikationsarten zu ermöglichen. Neben der Fortführung dieses Vorhabens werden wir uns künftig mit dem Problem beschäftigen, wie bestehende Planer, die auf der Basis eines „Constraint Satisfaction Problem Solvers“ arbeiten, so modifizieren werden können, dass neben den Planungsnebenbedingungen auch beliebige andere Constraints mit berücksichtigt werden.

Projekt: Modellierung und Vorhersage des Bedarfs von Anwendungen in verteilten Umgebungen

Ansprechpartner: DR. ARTUR ANDRZEJAK

Beteiligte Mitarbeiter: MEHMET CEYRAN, STEFAN PLANTIKOW, PROF. ALEXANDER REINEFELD

Zusammenarbeit: Hewlett-Packard Labs (USA), University of Coimbra (Portugal)

Forschungsförderung: ZIB

Die Modellierung und Vorhersage des Bedarfs von Anwendungen (sowie komplementär, der Kapazität von Ressourcen) werden zum einen bei dem automatischen Management von Ressourcen eingesetzt, zum zweiten sind sie notwendig bei der gemeinsamen Nutzung von Ressourcen (resource sharing) in Rechenzentren.

Die Zuverlässigkeit von Systemen gewinnt mit jedem Jahr an Bedeutung, was durch die zunehmende Verbreitung von funktionskritischen Computersystemen in vielen Lebensbereichen bedingt ist. Das automatische Management von Ressourcen hat daher u. a. die Gewährleistung der Zuverlässigkeit zum Ziel. Hier hilft die Bedarfsmodellierung und -vorhersage bei der Erkennung von Systeminfektionen (intrusion detection). Sie erlaubt es, Fehlfunktionen von Servern oder Netzwerken automatisch zu erkennen und sie unterstützt Entscheidungen, ob und wann ein System neu hochgefahren werden soll, um eine bessere Leistung zu erzielen (software rejuvenation).

Die gemeinsame Nutzung von Ressourcen ist wirtschaftlich motiviert. Die Erfahrung zeigt, dass die Ressourcen in industriellen wie auch wissenschaftlichen Rechenzentren nicht immer vollständig genutzt werden. Die effiziente Ausnutzung dieser Reserven ist besonders interessant im Bereich des Utility Computing, der im Rahmen der Zusammenarbeit mit Hewlett-Packard Labs angegangen wird. Hier hilft die Modellierung des kumulativen Applikationsbedarfs (sog. workloads) bei der langfristigen Kapazitätsplanung sowie dem kurzfristigen Scheduling. Ein weiterer Anwendungsbereich dieser Techniken ist die Zuordnung von Jobs zu Ressourcen bei Desktop Grids, also Ansammlungen von Desktop Computern, die in größeren Unternehmen oder Universitäten zusätzlich zu ihrer primären Aufgaben in einen virtuellen Rechenverbund oder Grid zusammengeschlossen wurden. Durch Einsatz von Modellierung kann die Ausnutzung der freien Kapazitäten maximiert werden, ohne dabei die primären Benutzer zu stören.

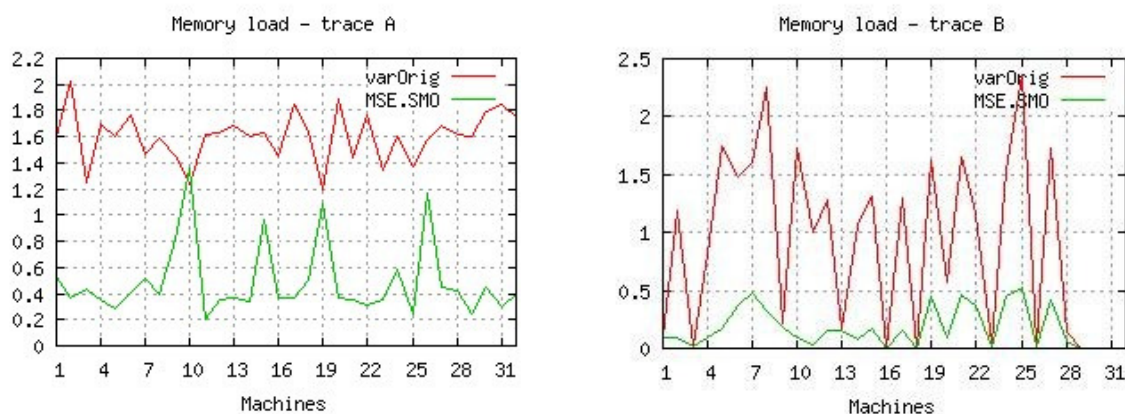


Abbildung 12: Mittlerer quadratische Fehler der Vorhersagen von Speicherauslastung (durchgezogen, 30 Min in voraus) gegenüber der Signalvarianz (gestrichelt) für zwei Datensätze über 39 bzw. 17 Tage

Im Rahmen des Projektes wurden mehrere Studien an realen Daten durchgeführt, sowie ein Softwareframework aufgebaut, das flexibles Erkunden der Daten unter Einbeziehung diverser Algorithmen ermöglicht.

Bei den Studien haben wir (in Zusammenarbeit mit der University of Coimbra) untersucht, inwieweit man Charakteristika von Desktopcomputern in universitären Umgebungen modellieren und kurzfristig vorhersagen kann. Die eingesetzten Methoden beinhalten eine ausführliche Datenvorbereitung, automatische Attributselektion, sowie Klassifikation der erwarteten Desktopkapazität mit Hilfe von Verfahren aus dem Bereich Machine Learning/Data Mining, insbesondere Klassifikatoren wie Naive Bayes oder Entscheidungsbäume. Die Ergebnisse sind ermutigend: *Abbildung 12* zeigt, dass die Vorhersagen für alle untersuchten Rechner sinnvolle Werte liefern, und für das Management von solchen Desktop Grids (z. B. Jobzuordnung) eingesetzt werden können. Eine weitere Studie galt der Vorhersage von Rechnerverfügbarkeit in universitären Desktoppools. Auch hier wurden gute Vorhersageergebnisse erzielt. Die gegenwärtige Arbeit umfasst Studien über den Einsatz dieser Techniken für kurzfristiges Scheduling in Desktop Grids.

Das im Laufe dieses Projektes aufgebaute Softwareframework „OPENSERIES“ ermöglicht einfache Spezifikation einer „Kette“ von Verarbeitungsschritten (wie Datenaufbereitung, Klassifikation usw.) sowie automatische Durchführung von Batchexperimenten für die

Untersuchung verschiedener Szenarien. Das Framework beinhaltet eine Vielzahl von Algorithmen (aus fremden Bibliotheken wie Weka 3.4) sowie eigene Implementierungen, z. B. ARIMA-basierte Zeitreihenanalyse oder Klassifikationsalgorithmen, die Techniken der Genetic Fuzzy Logic einsetzen. Die gegenwärtige Arbeit fokussiert die Erweiterung und Modifikation von „OPENSERIES“, um sowohl Echtzeit- als auch Batchbetrieb ohne wesentliche Anpassungen der Konfigurationen durchführen zu können.

Projekt: ZIB-DMS

Ansprechpartner: FLORIAN SCHINTKE

Beteiligte Mitarbeiter: THORSTEN SCHÜTT, TINOSCH KOLAGHEICHI-GANJINEH, MONIKA MOSER, CHRISTIAN VON PROLLIUS, PROF. ALEXANDER REINEFELD, ROLAND TUSCHL, JOCHEN WITTE

Forschungsförderung: ZIB

ZIB-DMS („ZIB Distributed Data Management System“) ist ein verteiltes System zur Verwaltung replizierter Daten im Grid. An diesem Forschungsprototyp untersuchen wir Aspekte, wie z. B. Skalierbarkeit, effiziente verteilte Speicherung von Metadaten, neue Organisationsmechanismen für sehr große Anzahlen von Dateien, Integrationsmöglichkeiten in existierende Systeme, Datenplatzierung, Verfügbarkeit replizierter Dateien, effizienter Zugriff in verteilten Systemen, topologieüberblickende Synchronisation und Möglichkeiten zur Selbstoptimierung und Adaption.

Der Prototyp wurde um weitere Funktionalität für flexible Replikatplatzierung und datenabhängiges Scheduling erweitert und insbesondere das Metadatenbackend wurde durch ein Redesign leistungsfähiger gemacht. Weitere Forschungsergebnisse, die im Rahmen dieses Projektes entstanden sind, konnten auf Konferenzen und in wissenschaftlichen Journalen veröffentlicht werden; das funktionierende Softwaresystem wurde auf verschiedenen wissenschaftlichen Veranstaltungen präsentiert.

Veröffentlichungen der Arbeitsgruppe

M. COSNARD, V. GETOV, D. LAFORENZA, A. REINEFELD (EDS.): *Future Generation Grids*. Dagstuhl Seminar Proceedings, Vol. 04451, <http://drops.dagstuhl.de/opus/volltexte/2005/26/>, Januar 2005.

J. WENDLER, F. SCHINTKE: *Executing and observing CFD applications on the Grid*. Future Generation Computer Systems the International Journal of Grid Computing: Theory, Methods and Applications (FGCS), 21(1), 11.-18. Januar 2005.

SLOOT, P.M.A.; HOEKSTRA, A.G.; PRIOL, T.; REINEFELD, A.; BUBAK, M. (EDS.): *Advances in Grid Computing - EGC 2005*. Revised Selected Papers, Series: Lecture Notes in Computer Science, Vol. 3470, European Grid Conference, Amsterdam, Niederlande, 14.-16. Februar 2005.

A. ANDRZEJAK, M. CEYRAN: *Characterizing and Predicting Resource Demand by Periodicity Mining*. In: Journal of Network and System Management, special issue on Self-Managing Systems and Networks, Vol. 13, No. 1, März 2005.

G. ALLEN, K. DAVIS, T. GOODALE, A. HUTANU, H. KAISER, T. KIELMANN, A. MERZKY, R. VAN NIEWPOORT, A. REINEFELD, F. SCHINTKE, T. SCHÜTT, E. SEIDEL, B. ULLMER: *The Grid Application Toolkit: Toward Generic and Easy Application Programming Interfaces for the Grid*. In: Proceedings of the IEEE, 93(3): 534-550, März 2005.

T. RÖBLITZ, A. REINEFELD: *Co-Reservation with the Concept of Virtual Resources*. In Proceedings of the 5th IEEE International Symposium on Cluster Computing and the Grid, Cardiff, Wales, UK, 09.-12. Mai 2005.

T. RÖBLITZ, F. SCHINTKE, A. REINEFELD: *Resource Reservations with Fuzzy Requests*. Concurrency and Computation: Practice and Experience, 2005.

T. SCHÜTT, F. SCHINTKE, A. REINEFELD: *Chord#: Structured Overlay Network for Non-Uniform Load Distribution*. Technical Report ZR-05-40, August 2005.

V. GETOV, D. LAFORENZA, A. REINEFELD (EDS.): *Future Generation Grids*. ISBN: 0-387-27935-0, November 2005.

P. VAN ROY, A. GHODSI, S. HARIDI, J-B. STEFANI, T. COUPAYE, A. REINEFELD, E. WINTER, R. YAP: *Self Management of Large-Scale Distributed Systems by Combining Peer-to-Peer Networks and Components*. Institute on System Architecture, CoreGRID Technical Report, TR-0018, 14. Dezember, 2005.

Vorträge

ALEXANDER REINEFELD: *Requirements for Grid Operating Systems: scalability, autonomy, mobility*. IST Call 5 Preparatory Workshop, Brüssel, 01.02.2005.

ALEXANDER REINEFELD: *Requirements for GridOSs from application and user view*. Workshop on Network Centric Operating Systems, Brüssel, 16.03.2005.

ALEXANDER REINEFELD: *Verteiltes Storage- und Datenmanagement*. IBM Academic Days, Böblingen, 18.03.2005.

ALEXANDER REINEFELD: *The German D-Grid Initiative and the Data-Management Challenge*. Keynote, Western Grid Summit, Banff, Kanada, 20.04.2005.

ARTUR ANDRZEJAK: *System Architecture in CoreGRID*. EGTD05 - Concertation Meeting-Technical Group 2, Architecture, Brüssel, 01.06.2005.

ALEXANDER REINEFELD: *Die Entwicklung der Spielprogrammierung: Von John von Neumann zur hochparallelen Schachmaschine*. Ringvorlesung, HU Berlin, 02.06.2005.

ALEXANDER REINEFELD: *Data Grids: Challenges in World-Wide Data Management*. Hermann-von-Helmholtz-Symposium, Berlin, 14.06.2005.

PATRICK MAY: *CASP6: Insights and results*. BCB-Symposium, Berlin, 22.06.2005.

ALEXANDER REINEFELD: *Scalable Data Management on Distributed Systems and Grids*. International Supercomputer Conference, Heidelberg, 21.-24.06.2005.

THOMAS STEINKE, CHRISTIAN EHRLICH: *Structure Prediction Pipeline @ ZIB*. HOBIT-Meeting, Berlin, 13.10.2005.

Sonstige Aktivitäten

Advisory Boards

Dr. Artur Andrzejak

- CoreGrid Executive Committee

Prof. Alexander Reinefeld

- Global Grid Forum Advisory Committee (GFAC)
- D-Grid Lenkungsausschuss
- DFN Technischer Ausschuss
- EU Next Generation Grid Advisory Board

Dr. Thomas Steinke

- Vizevorsitzender des Management Committees der EU COST Action D23, METACHEM - Metalaboratories for Complex Computational Applications in Chemistry

Editorial Boards und Conference Advisory Boards**Dr. Artur Andrzejak**

- Multiagent and Grid Systems - An International Journal

Prof. Alexander Reinefeld

- Future Generation Computing Systems, Elsevier
- Journal of Grid Computing, Kluwer Academic Publisher
- IEEE/ACM Symposium on Cluster Computing and the Grid (CCGrid)
- Euro-Par

Veranstaltete Tagungen und Workshops**Prof. Alexander Reinefeld**

- EGC 2005 – European Grid Conference, Amsterdam, Februar 2005 (Co-Organisator)
- CoreGRID Workshop on Network Centric Operating Systems, Brüssel, März 2005 (Co-Organisator)
- GridCoord Retreat, Berlin, Oktober 2005 (Organisator)

Mitarbeit in Programmkomitees**Dr. Artur Andrzejak**

- IEEE/ACM International Symposium on Cluster Computing and the Grid (CCGrid2005), Cardiff
- European Grid Conference (EGC2005), Amsterdam
- IEEE 7th International Workshop on Policies for Distributed Systems and Networks (POLICY 2005)
- 1st CoreGRID Workshop on GRID and P2P Systems Architecture, (Chair)

Prof. Alexander Reinefeld

- 13th Annual Mardi Gras Conference "Frontiers of Grid Applications and Technologies", New Orleans, Februar 2005
- CIC 2005 – Sixth International Conference on Communications in Computing, Las Vegas, Juni 2005
- GP2P'05 – Global and Peer-to-Peer Computing Workshop, Cardiff, USA, Mai 2005
- HPDC-14 – IEEE Symposium on High Performance Distributed Computing, Research Triangle Park, NC, Juli 2005

- e-Science 2005 – 2nd IEEE Intl. Conf. on e-Science and Grid Computing, Melbourne, Oktober 2005
- SC05 – Supercomputing 2005, Seattle, November 2005

Betreute Diplom-, Bachelor-, und Masterarbeiten

Patrick May

SANDRO ANDREOTTI: *An Integer Linear Programming Approach for the Protein Threading Problem*, BSc.-Arbeit, Oktober 2005.

Prof. Alexander Reinefeld

JOCHEN WITTE: *Implementierung einer NFS-Schnittstelle im Kontext eines Grid-Datenmanagementsystems*, Diplomarbeit, Humboldt-Universität zu Berlin, Januar 2005.

JÖRG MELTZER: *Vergleich von Co-Reservierungsalgorithmen in Grid-Umgebungen*, Diplomarbeit, Freie Universität Berlin, Februar 2005.

RENÉ HEEK: *Entwurf und Implementation von Grid-Diensten zur Verwaltung von Dateien, Aufträgen und Arbeitsabläufen in der Bioinformatik*, Diplomarbeit, Freie Universität Berlin, April 2005.

THE ANH PHAM: *Einbettung neuer Verwaltungsmethoden in die hierarchische Dateisystemsicht*, Diplomarbeit, Technische Universität Berlin, September 2005.

MINOR GORDON: *Staged Design for Highly Concurrent Web Servers*, Diplomarbeit, Technische Universität Berlin, November 2005.

TOM LICHTENSTEIN: *Implementation eines Ressourcenbrokers im Kontext eines Grid-Datenmanagement-Systems*, Diplomarbeit, Humboldt-Universität zu Berlin, Dezember 2005.

Dr. Thomas Steinke

TIMO RÖSSNER: *Fragment-Bibliothek für Threading-Methoden*, BSc.-Arbeit, Juli 2005.

FALKO KRAUSE: *Implementation of Web Services for the ZIB Protein Structure Prediction Pipeline*, BSc.-Arbeit, September 2005.

ROLAND BÖHRENZ: *Verteiltes Rechnen in der Bioinformatik*, MSc.-Arbeit Oktober 2005.

Lehr- und Forschungseinheit

Spezifikation, Verifikation and Testtheorie

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

Leitung

Prof. Dr. H. 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 Spezifikationssprachen, 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 im Verkehrswesen, Mobilfunk oder bei Chipkarten. 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“ der Abteilung EST beim FIRST. Dort werden die erzielten Ergebnisse in konkreten Anwendungsprojekten mit industriellen Partnern, etwa DaimlerChrysler, Alcatel oder Siemens 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 Automobil (modellbasierte Entwicklung von Steuergeräten), Schienenverkehr (Zuverlässigkeit eines fehlertoleranter Stellwerkscomputers), Automatisierungstechnik (Modellprüfung einer Sicherungsschicht), Zahlungsverkehr (Formalisierung einer Kartenterminal-Spezifikation) und bei Internetdiensten (Korrektheit von Webservices).



Lehre im Jahr 2005

- Vorlesung WiSe 2004/2005: Qualitätssicherung von Software (Halbkurs)
- SoSe 2005: Softwarequalität – Geschichte und Trends (Ringvorlesung am 9.6.2005)
- Vorlesung WiSe 2005/2006: Software-Engineering II: Eingebettete Systeme (Halbkurs)
- sowie diverse industrielle Schulungen und Weiterbildungsveranstaltungen

Forschung

In der Arbeitsgruppe von Prof. Schlingloff am Fraunhofer Institut FIRST werden Industrieprojekte, öffentlich geförderte Projekte und Fraunhofer-Vorlaufforschungsprojekte durchgeführt. Nähere Informationen sind auf den offiziellen Webseiten des Instituts <http://www.first.fraunhofer.de> zu finden.

Darüber hinaus gibt es einige anderweitig finanzierte Projekte, die am Institut für Informatik der Humboldt Universität durchgeführt werden.

Projektname: ep2

Beteiligter Mitarbeiter: Satish Mishra

Zusammenarbeit: Univ. Swansea, Zühlke Engineering AG

Mittelgeber: IAL System GmbH, Berlin

ep2 ist der Name eines neuen Schweizer Bezahlkartensystems für elektronischen Geldtransfer. In diesem Projekt wird die Modellierung von ep2 in der neuen algebraisch/prozessalgebraischen Spezifikationssprache CSP-CASL untersucht. Anhand der formalen Spezifikation werden Sicherheitseigenschaften des Protokolls bewiesen sowie Testfälle für die zentrale Komponente, das ep2 Terminal, generiert.

Projektname: SWIFI – Software-implementierte Fehlerinjektion**Beteiligter Mitarbeiter:** Sasa Vulinovic (TU Berlin)

In diesem Projekt wird der Einsatz von Techniken der softwareimplementierten Fehlerinjektion in einer modellbasierten Entwicklungsumgebung untersucht. Die intendierte Klasse von Zielsystemen sind dabei verteilte Steuerungsfunktionen und vernetzte Steuergeräte im Automobil. Wir analysieren die auf den verschiedenen Ebenen möglichen Fehlerursachen und Ausfälle, identifizieren Fehlerklassen und Fehlermodelle für die Injektion, und implementieren eine Software-Umgebung für die Durchführung der Simulationsläufe.

Projektname: Model Checking eines Automatisierungsprotokolls**Beteiligter Mitarbeiter:** Farooq Khan**Mittelgeber:** Siemens AG

Im Bereich der Automatisierungstechnik wird die Bus-orientierte Punkt-zu-Punkt-Kommunikation zunehmend durch Ethernet-basierte Protokolle ersetzt. Dabei spielt die Sicherheit der Übertragung und das Realzeitverhalten eine entscheidende Rolle. In diesem Projekt wurden Teile eines neuen internationalen Standards in der Automatisierungstechnik formal verifiziert. Das Ablaufverhalten der Sicherungsschicht ist dabei mit UML-StateCharts spezifiziert; im Projekt wurden die Verklemmungsfreiheit und andere Eigenschaften automatisch mittels Modellprüfung nachgewiesen.

Veröffentlichungen

H. SCHLINGLOFF (ed.): *Proceedings of the 4th Workshop “Methods for Modalities”*. Informatik-Bericht Nr. 194, Humboldt-Universität zu Berlin. 305 pp, ISSN 0863-095X, Nov. 2005.

H. SCHLINGLOFF, S. VULINOVIC: *Model based dependability evaluation for automotive control functions*. In SimSafe05, Modeling and simulation for public safety, P. Fritzson (ed.), Linköping, May 2005.

H. Schlingloff, S. MISHRA: *Loose semantics in the verification of communicating systems*. Short note at AVIS'05 - Fourth International Workshop on Automated Verification of Infinite-State Systems ; @ ETAPS 2005, Edinburgh, (Apr 2005).

A. MARTENS, H. SCHLINGLOFF, K. SCHMIDT: *Modeling and Model Checking Web Services*. In: ENTCS - Electronic Notes in Theoretical Computer Science; Volume 126, Pages 1-114, Elsevier (March 2005). Issue on *Logic and Communication in Multi-Agent Systems*, edited by W. van der Hoek; A. Lomuscio; E. de Vink; M. Wooldridge.

S. VULINOVIC, H. SCHLINGLOFF: *Zuverlässigkeitsprüfung eingebetteter Steuergeräte mit modellgetriebener Fehlerinjektion*. In: ASIM/GI-Tagung "Simulations- und Testmethoden für Software in Fahrzeugsystemen" Berlin (März 2005).

Ausgewählte Vorträge

- 16.1.2005: Modelling and Model Checking of Web Services. Feldafinger Kreis, Bad Honnef.
- 25.1.2005: Ausfall im Keim ersticken – Sichere Bauteil- und Systementwicklung. IIR, München.

- 2.3.2005: Zuverlässigkeitsprüfung eingebetteter Steuergeräte mit modellgetriebener Fehlerinjektion. ASIM, TU Berlin.
- 12.4.2005: Test Generation from Process Algebra. Workshop, Univ. Swansea, Wales.
- 22.4.2005: Temporal Logic: Past, Present, and Future. Festkolloquium anlässlich des 60. Geburtstags von Prof. Kröger, LMU München.
- 30.5.2005: Model based dependability evaluation for automotive control functions. SimSafe, Linköping, Schweden.

Sonstige Aktivitäten

- Mitglied des Organisationskomitees der Tagungen
 - M4M-4 – Methods for Modalities, Berlin, Dez. 2005
 - Dagstuhl Seminar 06411 Specification, Verification and Test of Open Systems, Okt. 2006
 - CS&P – Concurrency, Specification and Programming, Berlin, Okt. 2006
 - ISoLA 2006 - IEEE Workshop on Leveraging Applications. of Formal Methods, Verification, and Validation, Zypern, Nov. 2006
 - CAV – Computer Aided Verification, Berlin, Juni 2007
- Mitglied des Programmkomitees folgender Tagungen und Konferenzen
 - RelMiCS/AKA 2006 – Relational Methods in Computer Science
 - 2nd Int. Conf. on Intelligent Computer Communication and Processing: Static and Runtime Verification, Cluj-Napoca, Romania, Sept. 2006
 - AIML 2005 – Advances in Modal Logic 2005
 - LCMAS 2005 – 3rd International Workshop on Logic and Communication in MultiAgent Systems (satellite workshop of IJCAI 2005), Edinburgh
 - SimSafe 2005 – Modeling and simulation for public safety, Linköping
- Gutachter für verschiedene wissenschaftliche Zeitschriften und Kongresse, für die DFG, das BMBF sowie Mitglied in mehreren Promotions- und Habilitationskommissionen

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 Mitarbeiterinnen und Mitarbeiter

JENS BLEIHOLDER

MELANIE WEIS

Tutorinnen und Tutoren

CHRISTOPH BÖHM

KARSTEN DRABA

VÉRONIQUE TIETZ

ERIC WITZMANN

Lehre

Veranstaltungen im Hauptstudium

- Vorlesung „Informationsintegration“ (PROF. NAUMANN, WiSe 2004/2005)
- Praktikum „Informationsintegration“ (PROF. NAUMANN, BLEIHOLDER, WEIS, WiSe 2004/2005)
- Ringvorlesung „Große Datenmengen im Web“ (WiSe 2004/2005) (PROF. Naumann mit PROF. FREYTAG, PROF. LESER, und PROF. REINEFELD)
- Spezialvorlesung „XML & Datenbanken“
- Vorlesung „Informationsintegration“ (Prof. Naumann, WiSe 2005/2006)
- Praktikum „Informationsintegration“ (Prof. Naumann, Bleiholder, WiSe 2005/2006)

Forschung

Projekt: "MAC: Merging Autonomous Content"

Ansprechpartner: PROF. DR. FELIX NAUMANN

Beteiligte Mitarbeiter: JENS BLEIHOLDER, MELANIE WEIS

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

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: JENS BLEIHOLDER, MELANIE WEIS, CHRISTOPH BÖHM, KARSTEN DRABA

Webpage: <http://www.informatik.hu-berlin.de/mac/hummer/>

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: MELANIE WEIS

Ein Problem der Datenintegration ist das Auftreten von verschiedenen Datensätzen in der Zielrepräsentation, die das gleiche Objekt beschreiben. Um die Qualität der Daten in der Zielrepräsentation zu erhöhen, müssen solche Duplikate erkannt werden. Ziel dieses Projektes ist es, Duplikate in geschachtelten und semi-strukturierten XML Daten mittels eines geeigneten Algorithmus zu identifizieren. Anforderungen an den Algorithmus sind, dass er unabhängig von der Anwendungsdomäne ist und nur wenig Eingabe eines Nutzers benötigt, um Duplikate effizient und effektiv zu identifizieren.

Teilprojekt: "Relationale Datenfusion - Zusammenführen von Daten"

Ansprechpartner: JENS BLEIHOLDER

Dieses Projekt befasst sich mit der Fusion von Informationen über gleiche Objekte aus Datenquellen. Informationen über gleiche Objekte aus heterogenen Datenquellen können sich sowohl ergänzen als auch widersprechen. Ziel dieses Projektes ist es, dem Nutzer eines integrierenden Informationssystems eine Spezifikation der Lösung der (möglicherweise) auftretenden Datenkonflikte zu ermöglichen. Dies geschieht durch die Erweiterung der Anfragesprache SQL um ein entsprechendes Konstrukt, die „FUSE BY“-Anweisung. Für das Hum-

Mer System wurde ein Modul erstellt, das solche Anweisungen entgegennimmt und ausführt. Dieses wird kontinuierlich um weitere Konfliktlösungsfunktionen ergänzt. Darauf aufbauend wird im Weiteren die Optimierung solcher relationaler Fusionsanfragen betrachtet.

Projekt: "BioFast: Efficient and seamless access to Internet accessible life sciences data sources"

Ansprechpartner: PROF. DR. FELIX NAUMANN, JENS BLEIHOLDER

Kooperation: PROF. DR. LOUIQA RASCHID, University of Maryland, PROF. DR. ZOÉ LACROIX, Arizona State University

Webpage: <http://www.umiacs.umd.edu/labs/CLIP/BFEnt02/>

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

Kooperation: 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.

Projekt: "XQuery Generierung"

Ansprechpartner: PROF. DR. FELIX NAUMANN, CHRISTOPH BÖHM

Webpage: <http://www.informatik.hu-berlin.de/mac/xquery/>

Basierend auf dem XML Schema und dem XQuery Standard entwerfen wir ein Softwaretool, das es Nutzern erlaubt, graphisch XQuery Anfragen zu spezifizieren. Zu diesem Zweck wird ein XML Schema als Baum visualisiert. Nutzer können Knoten des Baumes auswählen und mit Prädikaten annotieren. Das Tool wandelt diese Auswahl in eine entsprechende XQuery um und zeigt gegebenenfalls das Anfrageergebnis.

Projekt: "System P"

Ansprechpartner: PROF. DR. FELIX NAUMANN

Kooperation: ARMIN ROTH, Daimler Chrysler Research

Webpage: <http://www.informatik.hu-berlin.de/mac/SystemP>

System P ist ein relationales Peer Data Management System (PDMS): Peers evaluieren konjunktive Punkt- oder Bereichsanfragen auf den eigenen Datenbeständen und schicken die Anfrage außerdem an benachbarte Peers. Zur Umformulierung von Anfragen nutzen die Peers in System P Schema-Mappings zwischen den Schemata der Peers, die ebenfalls durch konjunktive Anfragen spezifiziert sind. Die Anfragebearbeitung in PDMS ist aufgrund vieler

teilweise redundanter Mapping-Pfade im Allgemeinen sehr ineffizient und daher schlecht skalierbar. Andererseits bewirken Selektionen und Projektionen in den Schema-Mappings Informationsverluste beim Transfer der Ergebnisdaten entlang von Mapping-Pfaden. Im Zusammenhang mit System P ist Forschungsziel, durch gezielte Auswahl verlustarmer Schema-Mapping Pfade die Effizienz der Anfragebearbeitung bei akzeptabler Ergebnisqualität deutlich zu steigern und so sehr große PDMS zu ermöglichen. Eine Testumgebung ermöglicht eine automatische Generierung sehr großer PDMS-Instanzen, um die Wirkung der oben genannten Optimierungsansätze experimentell zu erproben.

Projekt: "Aladin"

Ansprechpartner: PROF. DR. FELIX NAUMANN

Kooperation: PROF. ULF LESER, JANA BAUCKMANN (Wissensmanagement in der Bioinformatik, HU)

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

Aladin hat das Ziel, Datenbanken in den Life Sciences semantisch zu integrieren. Im Gegensatz zu vielen anderen Ansätzen erfolgt die Integration der einzelnen Datenbanken jedoch nicht manuell, sondern automatisch. Besonders im Bereich der Life Sciences werden gleiche Begriffe von unterschiedlichen Datenbanken unterschiedlich verwendet. Darüber hinaus sind in den relationalen Repräsentationen der Datenbanken die Strukturen zwischen den einzelnen Tabellen nur selten definiert – obwohl sie implizit in den Daten erkennbar sind. Aus diesen Gründen arbeitet Aladin nicht auf den Strukturinformationen, sondern auf den Daten der einzelnen Datenbanken. Darüber hinaus verwendet Aladin domänenspezifische Informationen, d.h. Informationen über die Struktur der (meisten) Life Science Datenbanken. Die Integration erfolgt in 5 Schritten, die sukzessive die Rohdaten mit Struktur und Zusammenhängen anreichern.

Eine ausführlichere Darstellung des Aladin Projekts befindet sich im Bericht des Lehrstuhls „Wissensmanagement in der Bioinformatik“.

Sokrates: Studentenaustausch in Europa

Ansprechpartner: PROF. DR. FELIX NAUMANN

Webpage: http://www.informatik.hu-berlin.de/mac/socrates/socrates_06_07.html

Seit Oktober 2005 bestehen einige Sokrates Austauschabkommen mit Europäischen Universitäten. Ab dem Wintersemester 2006/2007 können sich Studenten für die subventionierten Studienplätze bewerben. Zugleich freuen wir uns auf Studenten aus dem Ausland, die an Lehrveranstaltungen des Instituts teilnehmen wollen oder ihre Abschlussarbeit hier schreiben wollen.

Veröffentlichungen

Conferences

- Declarative Data Fusion - Syntax, Semantics, and Implementation
Jens Bleiholder and Felix Naumann: *Proceedings of the International Conference on Advances in Databases and Information Systems (ADBIS) 2005*, Tallin, Estonia. In: *LNCS* 3631, pp. 58-73, ©Springer-Verlag
- DogmatiX Tracks down Duplicates in XML
Melanie Weis and Felix Naumann: *Proceedings of the International Conference on Management of Data (SIGMOD) 2005*, Baltimore, MD
- Self-Extending Peer Data Management
Ralf Heese, Sven Herschel, Felix Naumann, Armin Roth: *Proceedings of the German Conference in Databases (BTW) 2005*, Karlsruhe, Germany
- (Almost) Hands-Off Information Integration for the Life Sciences
Ulf Leser and Felix Naumann: *Proceedings of the Conference in Innovative Database Research (CIDR) 2005*, Asilomar, CA
- Schema Matching using Duplicates
Alexander Bilke and Felix Naumann: *Proceedings of the International Conference on Data Engineering (ICDE 05)* Tokyo, Japan

Workshops

- Benefit and Cost of Query Answering in PDMS
Armin Roth and Felix Naumann. *VLDB 2005 workshop on Databases, Information Systems and Peer-to-Peer Computing (DBISP2P'05)*, Trondheim, Norway
- A Relational Operator Approach to Data Fusion
Jens Bleiholder. In *VLDB 2005 PhD Workshop*, Trondheim, Norway
- Fuzzy Duplicate Detection on XML Data
Melanie Weis. In *VLDB 2005 PhD Workshop*, Trondheim, Norway
- A Data Model and Query Language to Explore Enhanced Links and Paths in Life Science Sources
George Mihaila, Felix Naumann, Louiqa Raschid, Maria-Esther Vidal. In *SIGMOD WebDB 2005*, Baltimore, MD
- Enhancing the Semantics of Links and Paths in Life Science Sources.
S. Heymann, F. Naumann, P. Rieger and L. Raschid. In *ICDT Workshop on Database Issues in Biological Databases (DBiBD 2005)*, Edinburgh, Scotland

Books & Book Chapters

- Proceedings of the International Conference on Information Quality (IQ)
Felix Naumann, Michael Gertz, and Stuart Madnick, Cambridge, MA 2005

Poster & Demos

- Automatic Data Fusion with HumMer (demo)
Alexander Bilke, Jens Bleiholder, Christoph Böhm, Karsten Draba, Felix Naumann, Melanie Weis. *Proceedings of the Int. Conf. on Very Large Databases (VLDB) 2005*, Trondheim, Norway

Vorträge

Neben den Vorträgen zu den oben genannten Konferenz- und Workshopbeiträgen wurden die folgenden Vorträge gehalten.

- PROF. DR. FELIX NAUMANN:
 - *HumMer: Data fusion and more*
Kolloquium am IBM Almaden Research Center, San Jose
 - *Schemamanagement*
DBTT auf der BTW, Karlsruhe
 - *Information quality in integrated information systems*
Kolloquium I2R, Singapur
 - *Schema Mapping im Kontext der Datenfusion*
Kolloquium, TU Dresden, Mai
 - *Schema Mappings in PDMS*
Kolloquium, Uni Rostock, Juni
 - *Datenqualität aus Datenbanksicht*
DIA Seminar zusammen mit Prof. Kai-Uwe Sattler, Berlin
 - *HumMer and the three steps of information integration*
Universität Modena, November
- JENS BLEIHOLDER
 - *Data fusion and the HumMer system*
 - Universität Modena, November
- MELANIE WEIS
 - SIGMOD IQIS Gong Show, Juni
 - *Duplicate Detection in XML*
INRIA Futurs, Frankreich, November
 - *Algebraic Operators and Optimizations in XClean*
IST Tagus Park, zusammen mit Ioana Manolescu

Weitere Aktivitäten

- PROF. DR. FELIX NAUMANN
 - Gastwissenschaftler: Prof. Louiqa Raschid, University of Maryland, Februar
 - Besuch am I2R, Singapur, Februar
 - Gastwissenschaftler: Armin Roth, Daimler-Chrysler, August & November
 - Organisation der International Conference on Information Quality (IQ) 2005 als Program Co-Chair.
 - Konferenzteilnahmen:
 - CIDR, Asilomar
 - BTW, Karlsruhe
 - Emmy-Noether Jahrestreffen, Potsdam
 - VLDB & VLDB PhD workshop, Trondheim
 - ICIQ, Boston
 - DB2 community meeting, Böblingen
 - Programmkomiteemitgliedschaften
 - DEXA
 - DILS
 - ICDE/EDBT PhD Workshop
 - Informatik 2004 DQ workshop

- Organisation der 2. Humboldt-Informatik Gong-Show (HIGS 05) am Tag der Informatik 2005
- Neugestaltung des WWW
- JENS BLEIHOLDER
 - Konferenzteilnahmen
 - VLDB & VLDB PhD Workshop, Trondheim
 - ADBIS, Tallinn
 - Besuch der Gruppe von Prof. Sonia Bergamaschi, Universität Modena, Italien
- MELANIE WEIS
 - Besuch der FUZZY! Informatik AG, Stuttgart, März
 - Auslandsaufenthalt bei INRIA Futurs, Frankreich, November 2005 bis Februar 2006
 - Besuch am IST Tagus Park, Portugal, Dezember
 - Konferenzteilnahme
 - SIGMOD & SIGMOD IQIS Workshop Baltimore
 - VLDB & VLDB PhD Workshop, Trondheim
- Jens Kruppa: *Minimum Union als DB2 table function*

Diplomarbeiten

FRANK LEGLER: *Datentransformation mittels Schema Mapping*

LENKA IVANTYSYNOVA: *Kombination von Schema Matching Verfahren zur semi-automatischen Integration von Fahrzeug-Daten*

SVEN PUHLMANN: *Duplikaterkennung in XML Daten mit der Sorted Neighborhood Methode*

JENS HARZER: *Klassifikation von Schema Mappings und deren Anwendung in einer Portal-lösung*

JANA BAUCKMANN: *Schemaintegration auf der Grundlage von Schema-Mappings*

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 Mitarbeiterinnen und Mitarbeiter

DIPL.-INF. PETER IBACH

DIPL.-INF. JAN RICHLING

BURAK SIMSEK

DR. SIEGMAR SOMMER

DR. KATINKA WOLTER

Technische Mitarbeiterinnen und Mitarbeiter

CHRISTINE HENZE

DIPL.-CHEM. BIRGIT SCHIEFNER

DIPL.-ING. STEFFEN TSCHIRPKE

Promotionsstudenten

DIPL.-INF. GÜNTHER HOFFMANN

DIPL.-ING. NIKOLA MILANOVIC

DIPL.-ING. BRATISLAV MILIC

DIPL.-ING. FELIX SALFNER

VLADIMIR STANTCHEV

DIPL.-INF. (FH) JENS STEGEMANN

Tutoren

STEFFEN BUHLE

TOBIAS HAMPEL

MARKUS MERGNER

PHILIPP REINECKE

JOHANNES ZAPOTOCZKY

Die Lehr- und Forschungseinheit Rechnerorganisation und Kommunikation arbeitet auf dem Gebiet des zuverlässigen verteilten und parallelen Rechnens mit den Schwerpunkten Fehlertoleranz, Fehlervorhersage, Echtzeitfähigkeit und Komponierbarkeit.

Zur Lehr- und Forschungseinheit gehört die Netzwerkbetriebsgruppe. Ihre Arbeit wird in **Abschnitt VI** genauer beschrieben.

Lehre

Veranstaltungen im Grundstudium

- Technische Informatik II (M. MALEK, SoSe 2005)
- Einführung in die Technische Informatik (K. WOLTER, WiSe 2005/2006)

Veranstaltungen im Hauptstudium

- Spezielle Techniken der Rechnerkommunikation (Halbkurs mit Praktikum, S. SOMMER, SoSe 2005)
- Grundlagen der Rechnerkommunikation (Halbkurs mit Praktikum, S. SOMMER, WiSe 2005/2006)
- Eigenschaften Mobiler und Eingebetteter Systeme (Halbkurs mit Projekt, J. RICHLING, WiSe 2005/2006)
- Zuverlässige Systeme (Halbkurs mit Projekt, J. RICHLING/K. WOLTER, WiSe 2005/2006)

Seminare

- *NOMADS* – Networks of Mobile Adaptive Dependable Systems (M. MALEK UND P. IBACH, SoSe 2005)
- CERO – CE Robots Community (M. MALEK, P. IBACH UND J. RICHLING, SoSe 2005)
- Spezielle Probleme von Echtzeitsystemen (J. RICHLING, SoSe 2005)

Forschung

NOMADS – Networks of Mobile Adaptive Dependable Systems

Vernetzt leben – gesellschaftliche Vernetzung in einer vernetzten Gesellschaft – so lautet unsere Vision der Zukunft: Menschen, ihre Bedürfnisse, Alltags- und Geschäftsprozesse werden in einer Welt allgegenwärtiger Informationstechnologie umfassend unterstützt. Wir nennen diese Welt „NOMADS Republic“: Milliarden von Sensoren, Aktuatoren, mobilen Geräten, Netzwerken, Rechnern aller Leistungsklassen und Nutzern interagieren unter dynamischen Bedingungen. Wir untersuchen Interoperabilität, Selbstorganisation, Mobilität und Adaptivität von NOMADS unter Dienst-Orientierung und Peer-to-Peer Umgebung und erreichen Verlässlichkeit (Zuverlässigkeit, Echtzeit und Sicherheit) über verschiedene Methoden.

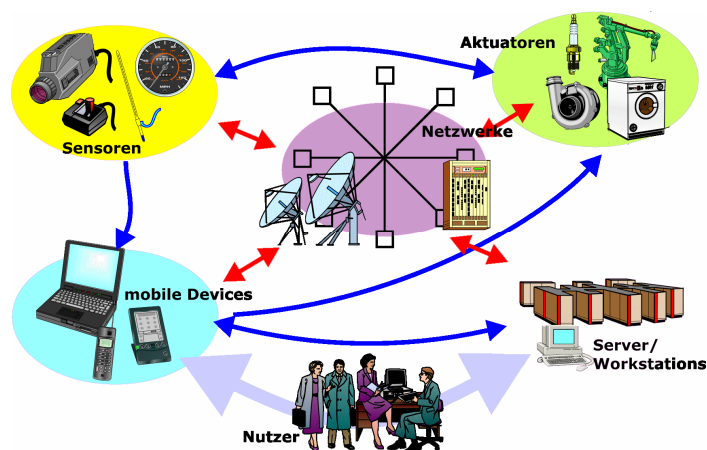


Abbildung 1: *NOMADS* in einer Welt allgegenwärtiger Informationstechnologie

Wir gehen von der in Abbildung 2 dargestellten Struktur der einzelnen Problemgebiete aus und konzentrieren unsere Untersuchungen auf:

- Komponierbarkeit
- Modelle und Architekturen
- Fehlervorhersage
- Mobilität

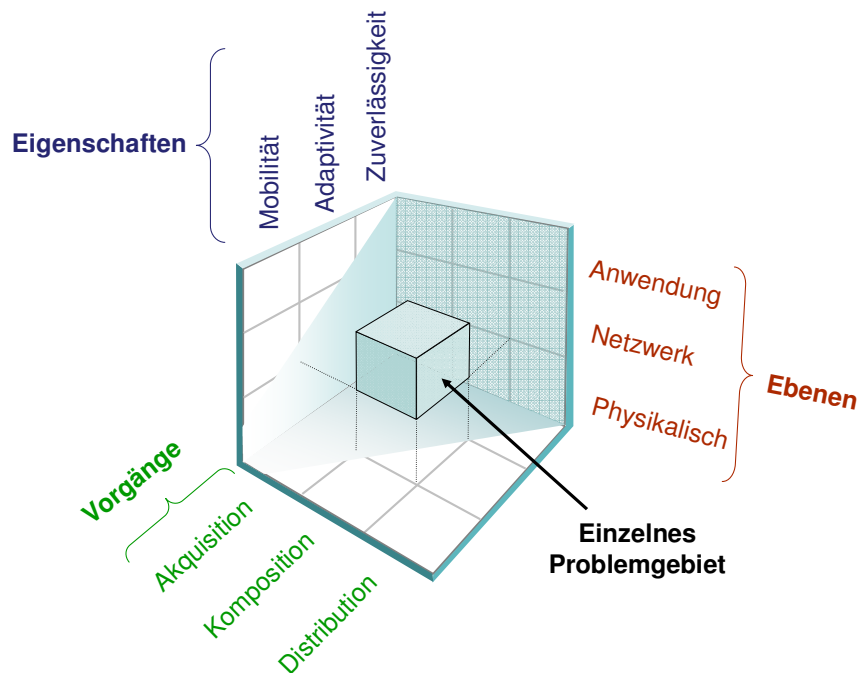


Abbildung 2: Strukturierung der Problemgebiete in Eigenschaften, Prozesse und Designebenen

Forschungsgebiet „Komponierbarkeit“

Durch den zunehmenden Einsatz sehr komplexer Software gibt es einen großen Bedarf an Verfahren zum Umgang mit dieser Komplexität. Konzepte der strukturierten Programmierung, der Objektorientierung und der komponentenbasierten Softwareentwicklung haben in diesem Zusammenhang sehr an Bedeutung 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 das fehlerfreie Verhalten des Systems zusätzlich von weiteren Eigenschaften wie beispielsweise dem zeitlichen Verhalten, dem Ressourcenbedarf, der Fehlertoleranz und anderen nichtfunktionalen Eigenschaften abhängt. Es genügt in einem solchen Szenario für das Zusammenfügen zweier Komponenten nicht, wenn ihre Schnittstellen funktional zusammen passen, es ist zusätzlich erforderlich, dass auch solche nichtfunktionalen Aspekte berücksichtigt werden.

Ziel dieses Forschungsgebietes ist es, Komponierbarkeit in Bezug auf nichtfunktionale Eigenschaften umfassend zu behandeln. Der Schwerpunkt liegt dabei auf den folgenden beiden Aspekten:

- Entwicklung von Terminologien, Taxonomien und Modellierungsverfahren für diese Art der Komponierbarkeit
- Entwicklung von Techniken zur Komponierbarkeit am Beispiel von Fallstudien (Projekte „Komponierbarkeit von Echtzeitsystemen“ und „Komposition von Web-Diensten“), die sich jeweils auf ausgewählte Eigenschaften konzentrieren und damit Erkenntnisse liefern, die zum Zwecke der Verallgemeinerung benutzt werden können.

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

Innerhalb des Forschungsgebietes „Komponierbarkeit“ werden in diesem Projekt zwei Schwerpunkte betrachtet:

- Bestimmung des Begriffes „Komponierbarkeit in Bezug auf nichtfunktionale Eigenschaften“.
Der Begriff Komponierbarkeit für die Komposition in Bezug auf nichtfunktionale Eigenschaften wurde in den vergangenen Berichtszeiträumen 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 bereits beim Entwurf einer (komponierbaren) Architektur berücksichtigt und verifiziert werden können, so dass durch Komposition erzeugte Systemen die betreffenden Eigenschaften per Konstruktion haben. 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.
- Anwendung des Konzeptes der Komponierbarkeit auf das Beispiel „Komponierbarkeit in Bezug auf das zeitliche Verhalten“ und Entwicklung, Verifikation und Implementation einer Beispielarchitektur.
- Unter dem Namen "*Message Scheduled System*" wurde eine gemischt ereignis- und zeitgesteuerte Architektur für eingebettete Systeme entworfen und schrittweise 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 drei Schwerpunkte:

- Vollendung der technischen Beschreibung von MSS einschließlich Kompositionsalgorithmen und Abschätzungen zur Komplexität
- Auf Basis der im vergangenen Berichtszeitraum erstellten komponierbaren Spezifikation wurde Untersuchung der Architektur MSS in Hinblick auf das garantierte zeitliche Verhalten fortgesetzt und der formale Nachweis der zeitlichen Garantien abgeschlossen.
- Auf Basis des im vergangenen Berichtszeitraums von Sebastian Ssmoller entwickelten statischen Prototypen von MSS wurde mit der Entwicklung eines weiteren Prototypen begonnen, der auch die dynamischen Funktionen von MSS, insbesondere vorhersagbare Kompositionen zur Laufzeit, unterstützt.

Das Projekt wird fortgesetzt, wobei der Abschluß der Entwicklung des dynamischen Prototypen im Mittelpunkt steht.

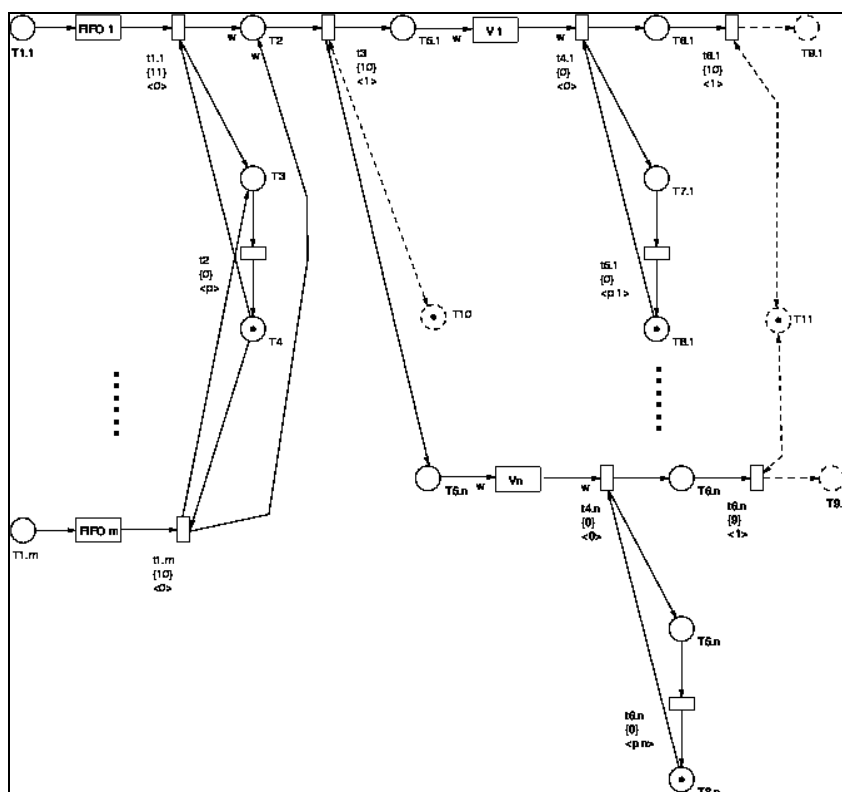


Abbildung 3: Teil der formalen Spezifikation von MSS

Projekt: Automatische Komposition von Web-Diensten mit Korrektheitsgarantien

Ansprechpartner: NIKOLA MILANOVIC

Web-Dienste dienen zunehmend als Standard-Plattform bei der Entwicklung von zukünftigen verteilten Systemen. Sie sind die Antwort auf die Herausforderungen der globalen Wirtschaft, in der an einem Geschäftsprozess viele Unternehmen beteiligt sind und bei geschäftlichen Transaktionen viele juristische und technologische Hindernisse sowie Vertrauensbarrieren zu überwinden sind.

Dienste sind selbst-beschreibende und offene Komponenten und ermöglichen eine schnelle und kostengünstige Komposition von verteilten Anwendungen. Dienste werden von Anbietern (Service Provider) angeboten - das sind Unternehmen, die Dienst-Anwendungen vermitteln und damit verbundene technische und betriebswirtschaftliche Betreuung (Unterstützung) anbieten.

Die erste Ebene der dienst-orientierten Architektur liefert sogenannte natürliche Fähigkeiten: Bekanntgabe, Feststellung, Auswahl und Verbindlichkeit. Die Dienste müssen standardisierte und maschinenlesbare Beschreibungen besitzen, die in Verzeichnissen aufgelistet werden, in denen sie auch gesucht werden können. Wenn der Kunde sich entschieden hat, welchen Dienst er nutzen will, beginnt der Prozess der Verbindlichkeit: der Kunde ruft den Dienst auf und anschließend wird die erforderliche Operation durchgeführt.

Die zweite Ebene der dienst-orientierten Architektur beschäftigt sich mit dem Problem der Dienst-Komposition. Sie umfasst Dienst-Koordination, Konformität, Überwachung und Dienstqualität. Die Koordination ermöglicht die Orchestrierung der Reihenfolge der kom-

ponierten Dienste, Konformität überprüft die Integrität des komponierten Dienstes, Überwachung beinhaltet Fehlererkennung und Korrekturmechanismen und die Dienstqualität hilft bei der Bestimmung der Zweckmäßigkeit des komponierten Dienstes und ermöglicht es, ihn mit anderen zu vergleichen.

Die oberste Ebene beschäftigt sich mit Themen des globalen Dienst-Marktes, beispielsweise wie beglaubigte Bescheinigungen ausgestellt werden können, wie ein zuverlässiges Reputationssystem für Dienste aufgebaut werden kann oder wie sich die Haftbarkeit feststellen lässt in einer komplexen Kette von komponierten Diensten vieler verschiedener Unternehmen.

Bisher ist nur die erste Ebene (Basisdienste) standardisiert worden. Es werden folgende Technologien benutzt: WSDL (Web Service Description Language) zur Beschreibung der Eigenschaften des Dienstes, UDDI (Universal Description Discovery and Integration) zur Speicherung und Suche von Dienstbeschreibungen und SOAP (Simple Object Access Protocol) zum Aufrufen eines Dienstes und zum Austausch von Mitteilungen. Die zwei oberen Ebenen (zusammengesetzte Dienste und verwaltete Dienste) bieten noch viele Möglichkeiten für Forschung und Entwicklung.

Wir beabsichtigen mit diesem Projekt auf dem Gebiet des dienst-orientierten Rechnens einen Beitrag zur Erweiterung der Dienstbeschreibungssprachen und zur Entwicklung von Mechanismen zur Prüfung der Korrektheit und der automatischen Dienstkomposition zu leisten.

Wir stellen zwei Hauptforschungsziele vor:

- Entwicklung eines Dienstkompositionsframework, das die Prüfung der Korrektheit einer beliebigen Komposition unterstützt
- Entwicklung eines Mechanismus für die automatische Dienstkomposition, bei der der Nutzer lediglich das Ziel spezifizieren muss und ein "intelligentes" Kompositionssystem dann adäquate Dienste auswählt und die Komposition durchführt

In diesem Jahr wurde die Prüfung der Korrektheit bei der Entwicklung eines formalen Dienstkompositionsframeworks bearbeitet:

- Entwicklung einer Dienstbeschreibungssprache, die sowohl funktionale als auch nicht-funktionale Eigenschaften abdeckt
- Definition des Vertragsmodells für die Dienstbeschreibung
- Auswahl von Vertragsinformationen von vorhandenen Architekturen
- Vorstellung von formalen Methoden zur Dienstbeschreibung (abstrakte Maschinen)
- Definition von elementaren Kompositionsoperatoren
- Überprüfung der Eigenschaften der komponierten Dienste

Forschungsgebiet „Modelle und Architekturen für Verteilte Systeme“

Verteilte Systeme werden immer komplexer – neben der höheren Anzahl beteiligter Knoten und ihrer höheren Komplexität, machen auch mobile ad-hoc Verbindungen das Verhalten des Gesamtsystems schwer vorhersagbar. Änderungen an einer Stelle können weitreichende Folgen für alle beteiligten Knoten haben, die nur schwer erörtert werden können. Die Entwicklung hin zur service-orientierten Architekturen (SOA) führt dazu, dass ein verteiltes System zunehmend aus verteilten Einzeldiensten besteht, anstatt aus verteilten Anwendungen.

Im Rahmen einer langfristigen Kooperation mit Partnern aus Industrie und Universitäten versuchen wir diese Komplexität abzuschätzen und zu beherrschen. Verteilte Systeme umfassen mehrere Hardware- und Softwareschichten – etwa Kommunikationsmedien und Verfahren, Ausführungsplattformen, Betriebssysteme, Middleware / Komponentenframeworks, Applikationskomponenten und Dienstzugriffspunkte. Wir entwickeln und testen Modelle dieser Schichten, die in einem Gesamtmodell des Systems integriert werden. Aus diesen Modellen erstellen wir Architekturen, die es erlauben bestimmte Eigenschaften wie Echtzeitverhalten, Zuverlässigkeit und Mobilität des Gesamtsystems zu beeinflussen, indem optimale Änderungen in unterschiedlichen Schichten des Systems eingeführt werden. Wir bezeichnen diesen Ansatz als *architectural translucency*.

Projekt: MONARCH – Modelle und Architekturen für Zuverlässige Systeme

Ansprechpartner: VLADIMIR STANTCHEV

Zusammenarbeit: Microsoft Research, Cambridge, UK.

Das Projekt MONARCH konzentriert sich auf zuverlässige, dienst-basierte Forschung. Im Rahmen des Projekts NOMADS (Network of Mobile Adaptive Dependable Systems) wird die Architektur einer service-basierten Umgebung gestaltet. Dabei wird auch die Integration vorhandener Technologien behandelt, insbesondere Component Frameworks wie .NET und EJB.

Bei der Evaluierung des Ansatzes von *architectural translucency* haben wir PN-basierte Modelle definiert und Quantifizierungsmethodiken für nichtfunktionale Eigenschaften entworfen. Weitere Fragestellungen, die wir bearbeitet haben, umfassen die Auswirkung einer prinzipiell ähnlichen Maßnahme (Replikation) wenn diese auf unterschiedliche Schichten innerhalb einer Architektur angewendet wird.

Projekt: RES.NET – Responsive .NET

Ansprechpartner: VLADIMIR STANTCHEV

Beteiligte Mitarbeiter: NIKOLA MILANOVIC, BRATISLAV MILIC

Unter dem Titel RES.NET untersuchen wir im Rahmen des Projekts MONARCH neue Ansätze zur Anbindung von responsiven virtuellen Experimenten in Enterprise Komponenten-Frameworks wie .NET und EJB für den Einsatz in verteilten Szenarien. Diese 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. Ziel ist es, nichtfunktionale Eigenschaften und vorhersagbares Antwortverhalten in einer schwach gekoppelten dienstorientierten Architektur zu unterstützen.

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 Speichermodells nicht in der Lage ohne weiteres Echtzeitgarantien einzuhalten. Ziel unserer Forschungsaktivitäten ist eine .NET basierte Softwareumgebung für responsive Dienste, deren Konzepte sich auch auf EJB anwenden lassen. Dabei wenden wir Konzepte wie Synchronisation mittels Konsensalgorithmen, Interprozesskommunikation via Remoting und Web-Services, sowie aspektorientierte Programmierung an. Mit der Hilfe eines Contract-Negotiating-Frameworks sind die einzelnen Komponenten in der Lage bestimmte funktionale und nichtfunktionale Eigenschaften untereinander auszuhandeln. Im aktuellen Berichtszeitraum

haben wir eine Architekturerweiterung erstellt, die es ermöglicht, funktionale und nicht-funktionale Eigenschaften von .NET Komponenten auszuwerten und in dem von uns spezifizierten Format (CDL – Contract Definition Language) abzuspeichern.

Projekt: Responsive Remote Control of Virtual Presence Experiments

Ansprechpartner: VLADIMIR STANTCHEV

Innerhalb des Projekts MONARCH verwenden wir die RES.NET-Umgebung für den speziellen Fall der Remote-Control-Experimente (durch Zustandsmaschinen modelliert). Diese Experimente eignen sich aufgrund ihrer Anschaulichkeit und den spezifischen Responsivitätsanforderungen hervorragend als Fallstudien zu der Problemstellung. Die RES.NET Umgebung gewährleistet die Kohärenz der Präsenz an der Gegenstelle mit der physikalischen Präsenz. Die Kontrollinfrastruktur erlaubt die Erstellung und Verarbeitung von CDL-Spezifikationen. Wir haben auch unsere Remote-Control-Experimente teilweise in einer service-basierten Umgebung integriert.

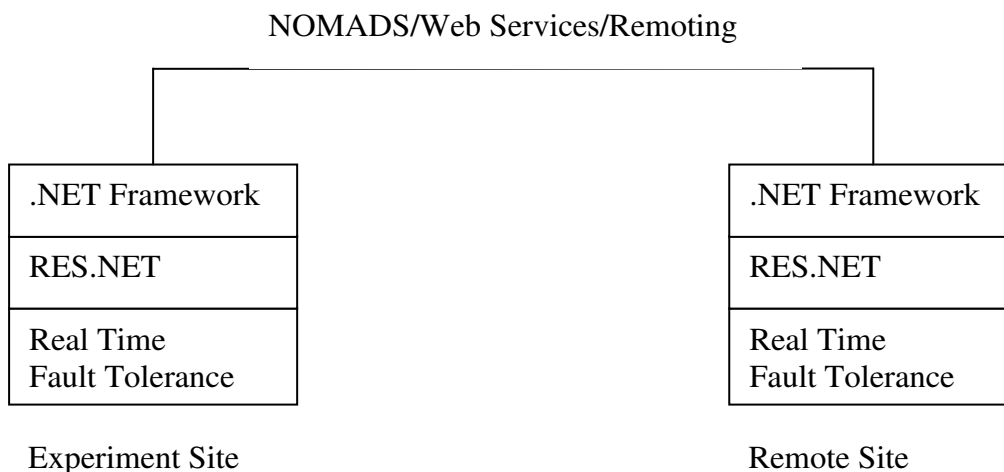


Abbildung 4: RES.NET als zusätzliche Schicht, die Responsivität gewährleistet

Projekt: CERO - CE Robots Community

Ansprechpartner: JAN RICHLING

Beteiligte Mitarbeiter: VLADIMIR STANTCHEV, NIKOLA MILANOVIC, PETER IBACH

Forschungsförderung: Microsoft Research

Roboter sind ein klassisches Beispiel für eingebettete Systeme mit nichtfunktionalen Anforderungen. Ein wesentliches Problem bei universitärer Robotik-Forschung sind die erheblichen Kosten, die bereits für eine einfache Ausrüstung erforderlich sind. Insbesondere die Forschung an kooperierenden Robotergruppen erfordert bislang einen enormen Kostenaufwand.

Idee bei CERO ist es darum, Robotik auf Basis von Standard-Hardware- und Software-Komponenten zu betreiben, um flexible Robotiklösungen zu vertretbaren Kosten möglich zu machen. Dabei können existierende Ansätze übertragen, sowie die neuen Ergebnisse einem breiten Einsatzbereich verfügbar gemacht werden.

CERO nutzt dazu das Betriebssysteme Microsoft Windows CE, welches sowohl auf PDAs als auch auf eingebetteten Boards lauffähig ist. Diese kann man mit flexibler modularer Robotik-Hardware (LEGO-Mindstorm) verbinden.

Ziel des CERO-Projektes ist es, eine generische Plattform für die Forschung an kooperativen Robotiksystemen bereitzustellen, in der modulare Basiskomponenten wie Ad-hoc-Routingverfahren oder Konsensprotokolle eine Gruppenkommunikation und Kooperation unter Robotern erlauben und weitere, darauf aufbauende anwendungsspezifische Entwicklungen erheblich vereinfachen. Besonderes Augenmerk bei der Entwicklung der Plattform und der Basiskomponenten gilt den nichtfunktionalen Eigenschaften: Echtzeitfähigkeit, Fehlertoleranz Mobilität und Adaptivität in dynamischen Situationen.

Im aktuellen Berichtszeitraum wurde das im vorangegangenen Berichtszeitraum entstandene Steuerungsprotokoll weiterentwickelt, das Steuerrechner und Roboter-Hardware verbindet. Parallel dazu wurde ein Unterprojekt „LIRO“ initiiert, bei dem sehr kostengünstige eingebettete Linux-Rechner – hier in Form von DSL/WLAN-Routern (ASUS 500G bzw. Linksys WRT54GL) – zur Steuerung der vorhandenen Robotik-Hardware benutzt werden. Damit ergibt sich eine flexible Lösung, in deren Mittelpunkt das selbstkonfigurierende Steuerungsprotokoll steht, so dass sowohl die Steuerrechner als auch die Robotik-Hardware in einem gewissen Rahmen ausgetauscht werden können.

Projekt: Mobile Solution Composer

Ansprechpartner: VLADIMIR STANTCHEV

Zusammenarbeit: MARIAN SCHERZ, Fachgebiet Systemanalyse und EDV, Institut für Wirtschaftsinformatik, TU Berlin

Im Rahmen des Projekts "Mobile Solution Composer" wird der Prototyp eines neuartigen Werkzeuges zur schnellen und anwendungsfreundlichen Implementierung von plattform- und device-unabhängigen, erweiterbaren Webanwendungen für die Zielgruppe kleiner mittelständischer Unternehmen (KMU) erstellt. Technisch basiert der Prototyp ausschließlich auf Open Source Komponenten. Über das Tool sind zum einen Geschäftsprozesse grafisch modellierbar, die Darstellungsschicht lässt sich erzeugen und es erfolgt eine automatische Dokumentation des aktuellen Systems. Projektziel ist es, KMU die Möglichkeit zu verschaffen mit bezahlbaren, aber trotzdem sicheren, robusten, verteilten und einfachen Systemen „Net Business“ zu betreiben.

Unser Beitrag im Rahmen des Projekts umfasst die nichtfunktionalen Eigenschaften des Systems, insbesondere Zuverlässigkeit, Sicherheit und *architectural translucency*. In diesem Jahr haben wir Konzepte zur Komponenten-Replikation in den Prototypen eingebaut und diese evaluiert.

Projekt: Quality-of-Service Management in Kommunikationssystemen

Ansprechpartner: DR. KATINKA WOLTER

Beteiligte Mitarbeiter: BURAK SIMSEK

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

Beteiligte Mitarbeiter: PHILIPP REINECKE

Zusammenarbeit: AAD VAN MOORSEL, Universität Newcastle, Großbritannien, Forschungsvorhaben „Quality-of-Service in rekonfigurierbaren, teilweise drahtlosen Rechen-systemen: 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 und

Antworten auf einige der oben formulierten Fragen in den Publikationen der letzten Monate veröffentlicht.

Forschungsgebiet „Mobilität“

Das Internet drängt in den mobilen Bereich. 2005 wird die Milliardengrenze der vernetzten Menschen und Geräte überschritten werden. 2007 werden es 1,5 Milliarden Nutzer sein, wobei dann bereits geschätzte 70% überwiegend drahtlosen Zugang haben werden. Die Anwendungen in einem „drahtlosen und allgegenwärtigen Netz“ werden sich dabei verändern: Typischerweise spielt die Position und die Interaktion mit der räumlich nahen Umgebung bei der mobilen Nutzung eine viel stärkere Rolle. Dies wird eine enge Verzahnung von physikalischen und informatorischen Prozessen mit sich bringen. Navigationslösungen weisen uns heute bereits den Weg durch den Verkehrsdschungel. Geoinformationsdienste und weitere Location Based Services warten auf expandierende Märkte. In wenigen Jahren werden voraussichtlich viele Milliarden mobiler und eingebetteter Systeme – einschließlich durch RFID-Chips bestückte Konsumobjekte – die Grundlage für eine durchdringende drahtlose Vernetzung und Funkortung schaffen. Wir begegnen bei den Entwicklungen, die dramatische Veränderungen unserer Geschäfts- und Alltagsprozesse bedeuten, einer Reihe von Zukunftsängsten – aber auch Chancen und technologischen Herausforderungen: ad-hoc Vernetzung, präzise Ortung, nahtlose Integration lokationsbezogener Dienste, globales Location Roaming, mobile Anwendungen und, nicht zuletzt, Schutz der Privatsphäre.

Projekt: PLAS.MA (Person Loss Avoidance System for Mobile Applications)

Ansprechpartner: NIKOLA MILANOVIC

Das PLAS.MA-Projekt entstand im Rahmen des Computer Society International Design Wettbewerbs 2004 der IEEE. Unter dem Thema "Making the world a safer place" waren weltweit Studenten aufgerufen ihre Projekte einzureichen. Die Humboldt Universität Berlin wurde durch 4 Studenten vertreten (Steffen Buhle, Dan Kreiser, Johannes Zapotoczky, Mario Hensel und Mentor Nikola Milanovic) und schaffte es mit PLAS.MA bis in die Endauswahl der 10 besten Teams von insgesamt 277 Teams zur Präsentation des Projektes nach Washington (IEEE CSIDC 2004 World Finals).

PLAS.MA - ausgesprochen "person loss avoidance system for mobile applications" - ist ein System zur Überwachung oder neutraler Beobachtung und Zusammenhaltung von Gruppen. Die Idee zu PLAS.MA gründet ursprünglich auf der Beobachtung und Absicherung von Kindergruppen. Diese zeigen auf Ausflügen oft eine besonders starke Eigendynamik und sind deshalb nur schwer zu überwachen. Um diese Gruppe vor dem Verlust einzelner Teilnehmer zu schützen, beobachtet das System vorbeugend kontinuierlich alle Teilnehmer. Entfernt sich ein Teilnehmer zu weit von der Gruppe oder ist er gar nicht mehr auffindbar, schlägt PLAS.MA Alarm und warnt einen verantwortlichen Beobachter akustisch und assistiert bei der Suche nach dem entfernten oder verlorenen Gruppenmitglied.

Die Besonderheit des Systems liegt in der Beobachtung von mobilen Gruppen. Jeder Teilnehmer der Gruppe trägt bei der Beobachtung ein intelligentes Hardware-Tag. Diese Tags tauschen während des Einsatzes des Systems untereinander kontinuierlich drahtlos Informationen über den Status der einzelnen Teilnehmer der Gruppe aus. Diese Informationen werden dann über ein spezialisiertes Hardware-Tag aufbereitet und dem verantwortlichen Betreuer der Gruppe visuell und akustisch zur Verfügung gestellt. Er ist somit jederzeit in

der Lage, die aktuelle Situation der Gruppe einzuschätzen. Durch die kontinuierliche Beobachtung und Analyse des Status der einzelnen Mitglieder der Gruppe kann PLAS.MA schnell und sicher fehlende Teilnehmer ermitteln und melden.

Die Überwachung erfolgt für die Nutzer des Systems transparent und ohne Einschränkung ihrer Bewegungsfreiheit. Dabei ist es nicht von Bedeutung, ob sich die einzelnen Mitglieder relativ zu einander bewegen oder die gesamte Gruppe ihre Position beliebig ändert. Das entscheidende Merkmal von PLAS.MA ist die dezentrale Organisation des Systems, welche es unabhängig von anderer Infrastruktur macht, wie etwa Basis- oder Relaystationen. Die nötige Infrastruktur für die Kommunikation der einzelnen Hardware-Tags wird komplett von den teilnehmenden Knoten selbst bereitgestellt. Diese Eigenschaft macht das System für die Anwendung in vielfältigen Szenarien interessant. Außerdem ist es dadurch in den unterschiedlichsten Umgebungen einsetzbar. So zum Beispiel innerhalb als auch außerhalb von Gebäuden. Die Informationen zur Beobachtung der Teilnehmer werden unter den Hardware-Tags verteilt und systematisch während der Kommunikation untereinander ergänzt. Somit ist es dem System möglich, seine Funktion unabhängig von einzelnen Tags aufrecht zu erhalten.

Projekt: MANET (Mobile Ad Hoc Network)

Ansprechpartner: BRATISLAV MILIC

Es ist vorgesehen, mobile Ad-Hoc-Netzwerke (MANET) in Situationen, in denen keine Kommunikationsinfrastruktur existiert oder eine existierende gestört ist, als Mittel anzuwenden um sich selbst organisierende Kommunikationsinfrastrukturen aufzubauen. Anwendungsszenarien reichen vom täglichen Leben bis hin zu lebensrettenden Situationen. Die verschiedenen Anwendungsfälle stellen unterschiedliche Anforderungen, wie effizientes Routing oder zuverlässiges Multicasting. Eine Eigenschaft, die für alle Anwendungsfälle und die Netzwerkfunktionalität generell bedeutend ist, ist die Konnektivität (Verbindungsfähigkeit).

Aus diesem Grund ist es wichtig, die Konnektivitätseigenschaften solcher Netzwerke zu verstehen und Partitionierungen vorherzusehen und zu verhindern. Meine Thesen liefern drei wesentliche Beiträge: ein stochastische Modell für ortsbezogene MANET, Partitionierungs-Vorhersagealgorithmen und Partitionierungs-Verhinderungsmethoden. Partitionierungs-Vorhersage- und Verhinderungsalgorithmen sollten verteilt und skalierbar sein, um effizient ein unerwünschtes Netzwerkverhalten zu verhindern. Ein stochastisches Modell ortsbezogener MANET wird bei der Abschätzung des Verhaltens und der Eigenschaften von Ad-Hoc-Netzwerken helfen, ohne aufwändig simulieren zu müssen.

Bisher haben wir ein graphen-basiertes mathematisches Modell entwickelt, das zur Implementierung eines Protokolls zur Vorhersage der Konnektivität verwendet wurde.

Der Vorhersage-Algorithmus des Modells ist verteilt und Knoten benötigen nur Wissen über die lokale Topologie um das globale Netzwerkverhalten vorherzusagen. Das Modell ist nachrichtenbasiert - um eine endgültige Entscheidung zu treffen, müssen die Knoten entlang des kritischen Pfades Nachrichten austauschen. Die Nachrichten-Sequenz muss so früh starten, dass sie vor der Partitionierung des Netzwerks beendet ist. Ebenso sollte die Nachrichten-Sequenz so spät wie möglich starten, um die durch den Vorhersage-Algorithmus erzeugte Netzwerklast und die Anzahl vorzeitige Alarme so gering wie möglich zu halten.

Der stochastische Ansatz wird zur Bestimmung der optimalen Startzeit der Vorhersage-Sequenz, anderer Protokollparameter und allgemeiner Netzwerk-Verbindungseigenschaften als Funktion der Netzwerkkonfiguration (Knotendichte, Knoten-Kommunikationsradius, Platzierung, Bewegungsmuster).

Die Konnektivitäts-Eigenschaften unseres Netzwerkmodells korrespondieren zu den Eigenschaften des Gebiets-Durchfluß-Modells (bekannt aus der Physik) - z.B. entspricht der Konnektivitäts-Wert im Netzwerkmodell der freien Energie im Gebiets-Durchfluß-Modell. Es ist geplant, die Resultate dieser Theorie bei der Modellierung der Netzwerk-Konnektivität zu verwenden, zu adaptieren und zu erweitern.

Projekt: Architekturen und Geschäftsmodelle für die Internet-Ökonomie

Ansprechpartner: PETER K. IBACH

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?

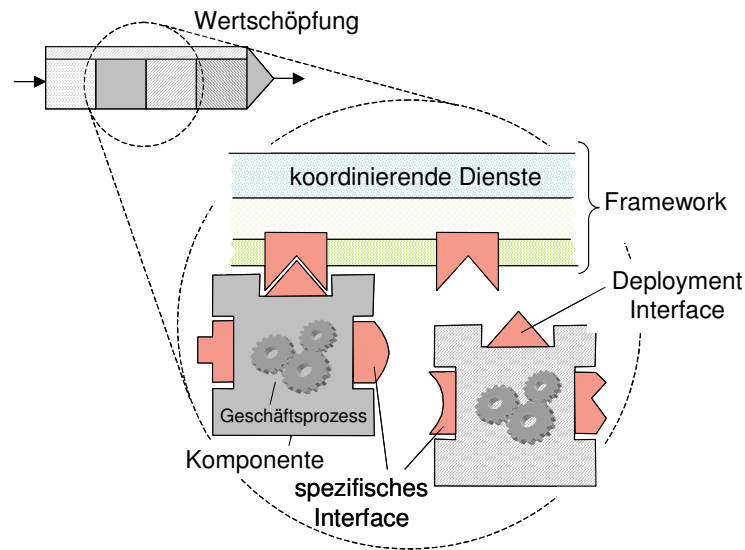


Abbildung 5: Dienstbasierte Architekturen

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 gestalten und plattform- bzw. umgebungsunabhängig wieder zu verwenden.

Projekt: Location Based Services

Ansprechpartner: PETER K. IBACH

Die Komposition von Diensten wurde im Umfeld von Telekommunikationsdienstleistungen für Location Based Services untersucht. Adaptive Komposition entsprechend der positionsabhängigen Verfügbarkeit von elementaren Diensten kann hohe Verfügbarkeit und Flexibilität auf der Ebene der zusammengesetzten Dienste erreichen (siehe Abbildung 6, Einzelheiten in [Ibach, Horbank 2004]). Standardisierte Schnittstellen, Adaptivität und Vertrauen sind Voraussetzungen für die flexible Komposition von Geschäftsprozessen zu „dynamischen Wertschöpfungsnetzen“ (vgl. [Ibach, Tamm, Horbank 2005]).

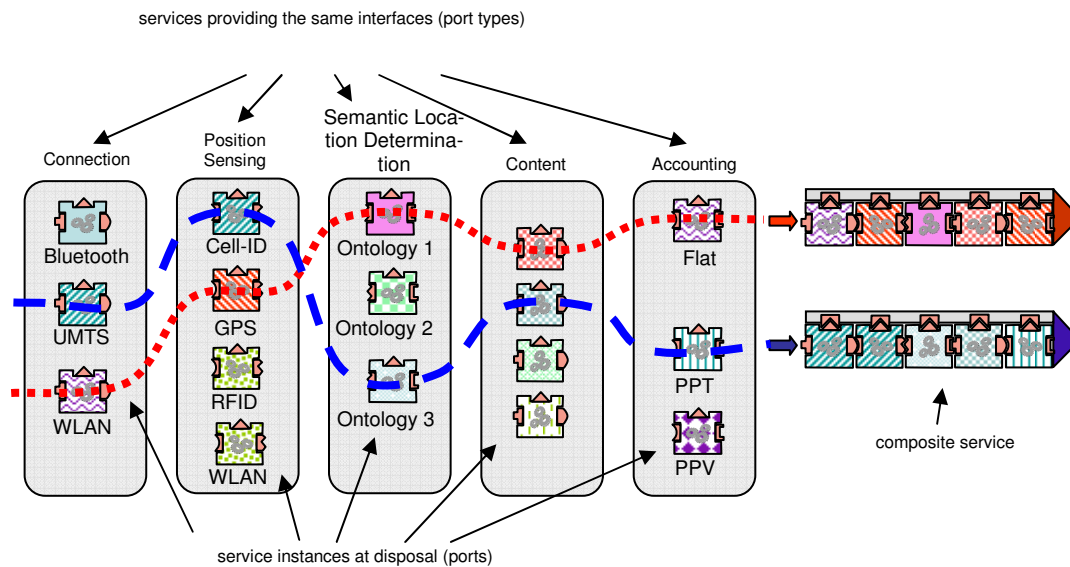


Abbildung 6: Adaptive Service Composition

Projekt: MagicMap

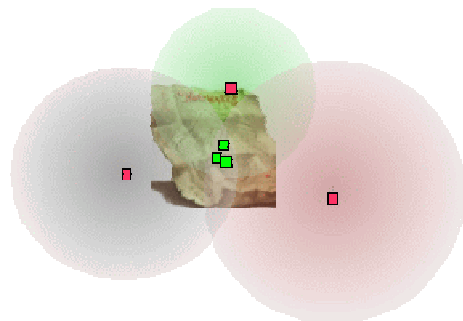
Ansprechpartner: PETER K. IBACH

Mit Hilfe der „Karte des Rumtreibers“ konnte Harry Potter die Positionen aller Personen in und um Schloss Hogwarts überblicken. Wir haben eine solche "magische Karte" zur Positionsbestimmung mobiler Systeme basierend auf WLAN-Signalstärke-Auswertung entwickelt. MagicMap ist eine Softwarelösung, die bei den mobilen Systemen außer einer konventionellen WLAN-Ausstattung keine weitere Hardware erfordert. Die WLAN Access Points können beliebig verteilt sein und es sind weder Eingriffe an der AP-Hardware noch an der Software nötig. Das hybride Verfahren ist in [Ibach, Hübner, Schweigert 2004] publiziert und kombiniert:

- Signalstärke-Triangulation zu Access Points (zukünftig auch zu Peer-Knoten)
- Radio Map (weighted averaged k-nearest neighbor)
- Probabilistisches Graph-Mapping nach Maximum-Likelihood-Methode

Unsere Versuche am Uni-Campus Berlin-Adlershof zeigen in typischen Büroumgebungen eine Positionierungsgenauigkeit im einstelligen Meter-Bereich. Die eigene Sichtbarkeit unterliegt dabei der vollen Kontrolle der Peer-Knoten, die den Zugriff auf ihre Positionsinformation individuell festlegen können. Die Software ist Open Source und steht zum freien Download auf der MagicMap Homepage zur Verfügung:

www.informatik.hu-berlin.de/rok/MagicMap



Forschungsgebiet „Fehlervorhersage“

Moderne Softwaresysteme erreichen heute oft eine Komplexität, die nur mit hohem personellen Aufwand zu verwalten ist. Ausfälle derartiger Systeme verursachen zudem enorme finanzielle Schäden. Wir entwickeln Verfahren, um Fehler in großen Softwaresystemen frühzeitig zu erkennen. So werden Wartungsaufwände deutlich reduziert. Weiterhin arbeiten wir an Verfahren zur Vorhersage von Fehlern, um Probleme zu erkennen, bevor sie Auswirkungen auf die Dienstqualität des Gesamtsystems haben. Als Datenquelle für diese Auswertungen dienen uns sehr große Mengen von Logging- und Monitoring-Daten, die wir in Zusammenarbeit mit Siemens erheben. Die Datenbanken sowie Filter- und Analyseverfahren werden an unserem Lehrstuhl entwickelt. Die derzeit verwendeten Daten sind Logdaten aus Testsystemen.

Graduiertenkolleg "Stochastische Modellierung und quantitative Analyse großer Systeme"

Sprecher: PROF. G. 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 von Teilen des Systems vorherzusagen, um präventive Maßnahmen ergreifen zu können. Im Rahmen dieses Forschungsvorhabens ist eine Methode entwickelt worden, die eine solche Vorhersage für große Softwaresysteme ermöglicht. Die verwendete stochastische Modellierung verfolgt den bislang wenig beachteten Ansatz, das zeitliche Fehlerverhalten 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 Projekt „Early Fault Detection“ durchgeführt.

Projekt: Early Fault Detection

Ansprechpartner: DIPL.-ING. STEFFEN TSCHIRPKE, DIPL.-ING. FELIX SALFNER

Zusammenarbeit: DIPL.-INF. GÜNTHER A. HOFFMANN (Content Pro AG)

Wichtige Ziele bei der Entwicklung moderner Computersysteme sind 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: (a) Eine Modellierung des Systems anhand von analytisch gewonnenem Systemwissen, (b) ein auf Markov-Ketten basierender Ansatz und (c) ein Funktionsapproximationsverfahren.

Beim analytischen Verfahren werden die Kommunikationsstrukturen des Systems dynamisch ermittelt. Aus Fehlermeldungen wird dann auf globale Systemzustände geschlossen. Das Markov-Ketten basierte Verfahren modelliert das zeitliche Auftreten von Fehlermustern. Die Muster werden algorithmisch aus Fehlermeldungen des Systems gewonnen. Das Funktionsapproximationsverfahren basiert auf radialen Basisfunktionen. Nach Auswahl aussagekräftiger Systemvariablen werden die wichtigsten 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. Integraler Bestandteil der Arbeit ist ein großer Datensatz, der von einer kommerziellen Telekommunikationsplattform gewonnen wurde. Anhand dieser Daten werden Praxistauglichkeit und Qualität der Vorhersageverfahren für industrielle Softwaresysteme ermittelt. Um das Potenzial der Verfahren in Bezug auf die Erhöhung der oben genannten Eigenschaften zu untersuchen, werden im Rahmen des Projekts zur Zeit mehrere Studien- / Diplomarbeiten geschrieben.

Projekt: Motherboard Failure Prediction

Ansprechpartner: Dipl.-Ing. Felix Salfner

Forschungsförderung: Intel Corporation

Das Projekt beschäftigt sich mit der Vorhersage von Ausfällen auf Ebene der Komponenten eines Intel-basierten Motherboards. Es werden verschiedene Vorhersagetechniken evaluiert und in Bezug auf ihre Einsetzbarkeit auf Motherboards bewertet. Der Fokus liegt dabei auf Vorhersageverfahren, die Daten von vorhandenen Motherboardsensoren sowie Fehlermeldungen einzelner Komponenten auswerten. Ferner wird untersucht, welche Strategien zur Vermeidung von Ausfällen oder zur Verkürzung von Reparaturzeiten bei Motherboards einsetzbar sind, um dem übergeordneten langfristigen Ziel der Verfügbarkeitserhöhung näher zu kommen.

Veröffentlichungen

M. MALEK: *The NOMADS Republic – A Case for Ambient Service-Oriented Computing*, Keynote Address, IEEE Proceedings of International Workshop on Service-Oriented System Engineering, SOSE 2005, Peking, Oktober 2005.

M. MALEK, M. REITENSPIESS, J. KAISER (EDS.): *Service Availability*, First International Service Availability Symposium, ISAS2004, München, Lecture Notes in Computer Science, Vol. 3335, Springer-Verlag, 2005.

M. MALEK, E. NETT, N. SURI (EDS.): *Service Availability*, Second International Service Availability Symposium, ISAS2005, Berlin, Lecture Notes in Computer Science, Vol. 3694, Springer-Verlag, 2005.

P. IBACH, G. TAMM, M. HORBANK: *Dynamic Value Webs in Mobile Environments Using Adaptive Location-based Services*, Hawaii International Conference on System Sciences (HICSS-38), Hawaii, USA, Januar 2005.

P. IBACH, M. HORBANK: *Towards a Global Spatial Data Infrastructure using Web Services*, GSDI-8, Kairo, April 2005.

P. IBACH, N. MILANOVIC, J. RICHLING, V. STANTCHEV, A. WIESNER, M. MALEK: *CERO: CE Robots Community*. IEE Proceedings Software, Special Issue on Embedded Systems, Volume 152, Nr. 5, Juni 2005.

P. IBACH, F. SCHREINER, V. STANTCHEV, H. ZIEMEK: *Ortung drahtlos kommunizierender Endgeräte mit GRIPS/MagicMap*, Workshop Mobiles Informationsmanagement und seine Anwendungen, 35. Jahrestagung der Gesellschaft für Informatik, Bonn, September 2005.

P. IBACH, V. STANTCHEV, F. LEDERER, A. WEISS, T. HERBST, T. KUNZE: *WLAN-based Asset Tracking for Warehouse Management*, Proceedings of the International Conference e-Commerce, Porto, Portugal, Dezember 2005.

N. MILANOVIC: *Contract-based Web Service Composition Framework with Correctness Guarantees*, Proceedings of the 2nd International Service Availability Symposium (ISAS2005), Berlin, April 2005.

N. MILANOVIC, M. MALEK: *Architectural Support for Automatic Service Composition*, Proceedings of the IEEE International Conference on Services Computing (SCC2005), Orlando, Florida, USA, 2005.

B. MILIC, N. MILANOVIC, M. MALEK: *Prediction of Partitioning in Location-aware Mobile Ad Hoc Networks*, Proceedings of the Hawaii International Conference on System Sciences, HICSS-38, (Minitrack on Quality of Service in Mobile and Wireless Networks), Hawaii, USA, Januar 2005.

F. SALFNER: *Predicting Failures with Hidden Markov Models*. Proceedings of the 5th European Dependable Computing Conference (EDCC 5), Student Forum, Budapest, April 2005.

F. SALFNER: *Homöopathie: Hochverfügbare Rechner heilen sich selbst*. c't Magazin für Computer Technik; 14, 156-162; Heise Verlag, Juni 2005.

F. SALFNER, G.A. HOFFMANN, M. MALEK: *Prediction-based Software Availability Enhancement*. In: *Self-Star Properties in Complex Information Systems*, O. Babaoglu, M. Jelasity, A. Montresor, C. Fetzer, S. Leonardi, A. van Moorsel and M. van Steen (Eds.), LNCS, vol. 3460, Springer-Verlag, Berlin 2005.

F. SALFNER, M. MALEK: *Proactive Fault Handling for System Availability Enhancement*. Proceedings of 10th IEEE Workshop on Dependable Parallel, Distributed and Network-Centric Systems (DPDNS '05), Denver, April 2005.

B. SIMSEK, K. WOLTER, H. COSKUN: *Analysis of the QBSS Load Element Parameters of 802.11e for a priori Estimation of Service Quality*, Proc. UK Performance Engineering Workshop (UKPEW 2005), Newcastle, UK, Juli 2005.

B. SIMSEK, K. WOLTER, H. COSKUN: *Analysis of the QBSS Load Element of IEEE 802.11e for a priori Estimation of Service Quality*, Simulation: Systems, Science & Technology, Dezember 2005.

V. STANTCHEV, M. MALEK: *Modeling Translucent Architectures*, Enterprise Application (Hrsg. S. Aier, M. Schönherr), Volume III, ISBN 3-936771-50-2, Gito Verlag, Berlin, August 2005.

V. STANTCHEV, M. MALEK: *Architectural Translucency – Vision and Approach*, International Conference Automatics and Informatics, Sofia, Bulgarien, Oktober 2005.

V. STANTCHEV, M. MALEK: *Architectural Translucency in Service-oriented Architectures*, to appear in IEE Proceedings Software.

K. WOLTER: *Self-Management of Systems through Automatic Restart*. In: *Self-Star Properties in Complex Information Systems*, O. Babaoglu, M. Jelasity, A. Montresor, C. Fetzer, S. Leonardi, A. van Moorsel and M. van Steen (Eds.), LNCS, vol. 3460, Springer-Verlag, Berlin, 2005.

A. VAN MOORSEL, K. WOLTER: *A Short Investigation into an Underexplored Model for Retries*, Proc. Seventh International Workshop on Performability of Computer and Communication Systems (PMCCS-7)}, Turin, Italien, September 2005.

Vorträge

M. MALEK: *The Future of ICT Power Systems: Emerging Security Challenges*, 2. Februar 2005, Brüssel.

M. MALEK: *Keynote Address zum 20. Jahrestag der Informatik an der Universität Rostock*, 21. März 2005, Rostock.

M. MALEK: *A Road to NOMADS (Networks of Mobile, Adaptive Dependable Systems) – A Perspective on Distributed Computing*, Ringvorlesung am Institut für Informatik, 12. Mai 2005, HU Berlin.

M. MALEK: *Dependability Modeling of Computer/Communication Systems*, Ringvorlesung, Graduiertenkolleg "Stochastische Modellierung und quantitative Analyse großer Systeme", 2. Juni 2005, TU Berlin.

M. MALEK: *The NOMADS Republic – A Case for Ambient Service-Oriented Computing*, Keynote Address, International Workshop on Service-Oriented System Engineering, SOSE 2005, 20. Oktober 2005, Peking.

M. MALEK: *Failure Prediction for Availability Enhancement*, Chinese University of Hong Kong, 27. Oktober 2005, Hongkong.

M. MALEK: *Ambient Service-Oriented Computing*, Shanghai Jiang Tong University (SJTU), 7. Oktober 2005, Shanghai.

M. MALEK: *The NOMADS Republic*, Chinese University of Hong Kong, 17. November 2005, Hongkong.

M. MALEK: *Evolution of NOMADS*, Silvercord Center, 19. November 2005, Hongkong.

Wissenschaftliche Kooperationen

- 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
- Microsoft Research, Cambridge, UK
- Consiglio Nazionale di Ricerca, Pisa, Italien
- New York University, USA
- Stanford University, USA
- Siemens AG
- Geografisches Institut, HU Berlin, Prof. Doris Dransch, Frank Schwedler
- Fraunhofer IIS, Nürnberg/Erlangen, Steffen Meyer
- Fraunhofer Fokus, Berlin, Holger Ziemek, Florian Schreiner
- Forschungs- u. Anwendungsverbund Verkehrssystemtechnik, Herr Emert, Herr Schreiter
- BSS GmbH, Eisenach
- AJE Consulting, Roes
- VAB GmbH, Berlin
- GTalk GmbH, Berlin
- TSI GmbH, Berlin

- Adisoft GmbH, Berlin
- Wincor Nixdorf, Berlin

Sonstige Aktivitäten

International Service Availability Symposium

Prof. Malek hat das „International Service Availability Symposium“ mitgegründet. Das erste Symposium ISAS2004 hat am 13. und 14. Mai 2004 in München stattgefunden und Prof. Malek leitete das Programm. Das zweite Symposium fand am 25. und 26. April 2005 in Berlin statt und Prof. Malek war General Chair.

Prof. Miroslaw Malek

- Program Committee Member: 10th IEEE Workshop on Dependable Parallel, Distributed and Network-Centric Systems, Denver, Colorado, USA, zusammen mit: 19th International Parallel and Distributed Processing Symposium, 4.-8. April 2005.
- General Chair: 2nd International Service Availability Symposium (ISAS2005) Berlin, April 2005.
- Student Forum Chair: 5th European Dependable Computing Conference (EDCC5), Budapest, Ungarn, April 2005.
- Program Committee Member: 5th International Conference on Application of Concurrency to System Design (ACSD), St. Malo, Frankreich, Juni 2005.
- Vice Program Chair: 4th International Workshop on Assurance in Distributed Systems and Networks (ADSN), Columbus, Ohio, USA, Juni 2005.
- Program Committee Member: 8th Euromicro Conference on Digital System Design (DSD), Porto, Portugal, August/September 2005.
- Program Committee Member: High Assurance Systems Engineering (HASE 05) Heidelberg, 12.-14. Oktober 2005.
- Program Committee Member: 24th Symposium on Reliable Distributed Systems (SRDS 2005), Orlando, Florida, USA, 26.-28. Oktober 2005.

Bratislav Milic

Program Committee Member: Wireless Ad hoc and Sensor Networks Symposium at IEEE International Conference on Communications (ICC2006), Istanbul, Türkei, 11.-15. Juni 2006.

Peter Ibach

Live Präsentationen von MagicMap:

- Lange Nacht der Wissenschaften, WiWi Fakultät
- Tag der Informatik, Erwin Schrödinger Zentrum, Berlin-Adlershof
- Fraunhofer Fokus, Berlin-Spandau
- Siemens AG, Berlin-Spandau
- Universität Bonn
- Seehotel Zeuthen
- Live-Test Lagerhalle Lufapak GmbH, Neuwied
- Internetökonomie Tagung, Maritim, Stuttgart
- Bus-Betriebshof, Verkehrsbetriebe Potsdam

Diplomarbeiten

BIRGER SCHMIDT: *Dynamisches Traffic Engineering in einem drahtlosen IP-Backbone*, März 2005.

EMMANUEL NGANKAM WELANGABE: *Untersuchung von Quality of Service (QoS) Parametern in Netzwerkkomponenten*, März 2005.

JOCHEN WITTE: *Implementierung einer NFS-Schnittstelle im Kontext eines Grid-Management-Systems*, März 2005.

GREGOR ZEITLINGER: *User Profile Integration using Ontologies*, März 2005.

MAREN LENK: *Heuristische Komposition abstrakter Maschinen*, Juni 2005.

Lehr- und Forschungseinheit

Signalverarbeitung und Mustererkennung

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

Leiterin

PROF. DR. BEATE MEFFERT

Tel.: (030) 2093 3043

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

Sekretariat

SABINE DZIWISZ

Tel.: (030) 2093 3044

Fax: (030) 2093 3045

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

Mitarbeiter

DIPL.-INF. ROMAN BLASCHEK

DR.-ING. MANFRED GÜNTHER

LOTHAR HEESE

DR.-ING. OLAF HOCHMUTH

DIPL.-INF. UWE KNAUER

DIPL.-ING. THOMAS MORGENSTERN

DIPL.-INF. AXEL WEISS

DR.-ING. FRANK WINKLER

DIPL.-ING. KARL-HEINRICH HAUPTVOGEL

DR.-ING. MICHAEL RITZSCHKE

Doktorandinnen und Doktoranden

DIPL.-ING (FH) CARSTEN BÜTTNER

DIPL.-ING. MATHIAS HAASE

DIPL.-INF. ILLYA KOMAROV

DIPL.-INF. DAVID KRUTZ

DIPL.-ING. THOMAS LOEWEL

DIPL.-ING. HOLGER MAASS

MOHAMMED A. MEGEED SALEM, M. SC.

DIPL.-INF. KATJA WORM

Tutorinnen und Tutoren

DAVID DAMM

MARCUS EHRIG

ERIK FISCHER

ALEXANDRA JULIUS

THOMAS KUHN

ENRICO MAY

MATTHIAS PREISS

RAINER SCHNABEL
ROBERT STEIN
ANNE WEGERICH
JÜRGEN WOHLFEIL

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.

Schwerpunkt der Forschungsaktivitäten des Jahres 2005 war die Entwicklung von Hardwarelösungen für die Signalverarbeitung unter Echtzeitbedingungen u. a. 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.

Auch die Analyse von Biosignalen zur Unterstützung der Diagnostik und Therapiekontrolle ist 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, A. WEISS)
- Mikroprozessorenentwurf - Proseminar (F. WINKLER)
- Praktische Informatik 1 - Übung (M. RITZSCHKE)

Hauptstudium (Halbkurse)

- Grundlagen der Signalverarbeitung (B. MEFFERT, O. HOCHMUTH, U. KNAUER)
- Signalverarbeitung (B. MEFFERT, O. HOCHMUTH)
- Schaltkreisentwurf (F. WINKLER)
- Prozessinformatik (K.-H. HAUPTVOGEL, M. RITZSCHKE)

Projekte und Seminare

- Biometrie im Kontext (P. BITTNER, B. MEFFERT)
- Spezialgebiete der Signalverarbeitung (B. MEFFERT)
- Programmierung von digitalen Signalprozessoren (A. WEISS)
- Parallele Signalverarbeitung mit FPGA (F. WINKLER)

Ringvorlesung „Multidisziplinäre Umweltanalyse und -bewertung“

(in Zusammenarbeit mit dem Geographischen Institut)

- Signalverarbeitung (B. MEFFERT)

Ringvorlesung „Schwerpunkte der Informatik - Themen der Informatik im historischen Kontext“

- Signale und Muster (B. MEFFERT)

Forschung

Projekt: Entwurf eines integrierten optischen Systems für die Bildanalyse

Ansprechpartner: PROF. DR. BEATE MEFFERT, PROF. DR. RALF REULKE

Beteiligte Mitarbeiter: DIPL.-INF. UWE KNAUER, DIPL.-INF. ROMAN BLASCHEK, DR.-ING. FRANK WINKLER, DR.-ING. MICHAEL RITZSCHKE

Zusammenarbeit: Deutsches Zentrum für Luft- und Raumfahrt DLR, Berlin-Adlershof

Forschungsförderung: Investitionsbank Berlin (IBB)

Das mehrjährige Projekt soll dazu beitragen, die Anwendungsmöglichkeiten optischer Systeme erheblich zu erweitern und ihnen neue Perspektiven für die wirtschaftliche Verwertung zu eröffnen. Als Testbeispiel dient die für die Verkehrsforschung relevante Verkehrsszenenbeurteilung. Die Vielzahl der in einem solchen System zu integrierenden Hardware- und Softwarekomponenten erfordert einerseits die Entwicklung von Algorithmen für die sensornahe Vorverarbeitung der Signale und Daten und andererseits Untersuchungen zur sinnvollen Partitionierung der Hard- und Software des Gesamtsystems. Optimierungskriterien sind dabei die Arbeitsgeschwindigkeit (Echtzeitanforderungen), die Robustheit und die Zuverlässigkeit. Im Jahre 2005 waren Schwerpunkte in der Bearbeitung des Projekts u.a.:

- Installation und Inbetriebnahme der Hardwarebasis, insbesondere des Alpha-Data-FPGA-Boards und der zugehörigen Entwicklungsumgebung; Einrichtung eines Demonstrations- und Entwicklungsplatzes
- Systematisierung und Analyse von favorisierten Algorithmen zur Erfassung verkehrlicher Parameter; z.B. mehrdimensionale Farbhistogrammanalyse zur Fahrzeugdetektion und -erkennung
- Durchführung mehrerer Messkampagnen an verschiedenen Berliner Standorten zur Bereitstellung realer Datensätze für die Algorithmentestung
- permanente Verkehrsbeobachtung an einem Adlershofer Standort
- Entwicklung einer Plattform zum Entwurf signalverarbeitender Systeme (freeSP)
- Arbeiten zur Erhöhung der Zuverlässigkeit des integrierten optischen Systems

Projekt: Sicherheitsmanagement mit Bildauswertung

Ansprechpartner: PROF. DR. BEATE MEFFERT

Beteiligte Mitarbeiter: DIPL.-INF. UWE KNAUER, SEBASTIAN HINZ, MATTHIAS PREISS, DAVID DAMM, DR.-ING. OLAF HOCHMUTH, DR.-ING. THOMAS MORGENSTERN, DR.-ING. MICHAEL RITZSCHKE

Zusammenarbeit: Lüth & Dümchen Automatisierungsprojekt GmbH, Berlin

Forschungsförderung: Investitionsbank Berlin (IBB)

Das Projekt wurde erfolgreich abgeschlossen. Die Ergebnisse sind in einem Projektbericht festgehalten. Beim Projektpartner Lüth & Dümchen erfolgte eine Abschlussdemonstration. Ein Prototyp der biometrischen Zugangskontrolle ist als Exponat auf der Cebit 2005 vorgestellt worden.

Projekt: Bienenmonitoring

Ansprechpartner: DIPL.-INF. UWE KNAUER

Beteiligte Mitarbeiter: MICHAEL HIMMELSBACH, DR.-ING. FRANK WINKLER, PROF. DR. BEATE MEFFERT

Zusammenarbeit: Länderinstitut für Bienenkunde, Hohen Neuendorf

Von den 900.000 Bienenvölkern in Deutschland haben 300.000 den Winter 2002/2003 nicht überstanden. Ein wesentlicher Grund ist die seit den 70er Jahren in Deutschland vorkommende Varroamilbe. Dem soll mit der Züchtung varroaresistenter Bienen begegnet werden. Das Merkmal „Ausräumen varroaparasitierter Brut“ soll als Selektionskriterium verwendet werden. Mit einer am Länderinstitut entwickelten Methode (Infrarot-Videobeobachtung) kann dieses Verhalten an individuell markierten Einzelbienen beobachtet werden.

Gegenstand des gemeinsamen Projekts ist eine automatische Auswertung der aufgezeichneten Videos. Schwerpunkt ist die Entwicklung von Algorithmen für eine automatische Erkennung und Verfolgung der markierten Bienen sowie für die Erkennung von geöffneten Brutzellen.

Projekt: Algorithmen und Implementierungen zur stereoskopischen Bildanalyse

Ansprechpartner: DIPL.-INF. ROMAN BLASCHEK

Beteiligte Mitarbeiter: DIPL.-INF. UWE KNAUER, DR.-ING. OLAF HOCHMUTH, PROF. DR. BEATE MEFFERT, PROF. DR. RALF REULKE

Zusammenarbeit: Deutsches Zentrum für Luft- und Raumfahrt, Berlin-Adlershof

Für die Vermessung von Untersuchungsobjekten finden zunehmend bildverarbeitende Systeme Verwendung. Sie bieten den Vorteil, große räumliche Bereiche erfassen zu können, ohne dass dies eines komplizierten Versuchsaufbaus bedarf. Speziell die Erfassung der räumlichen Beziehungen zwischen den Objekten einer beobachteten Szene kann durch Kombination zweier Kameras mit unterschiedlichem Blickwinkel auf den Versuchsaufbau – der stereoskopischen Bildanalyse – gelöst werden. In diesem Projekt wurde exemplarisch ein stereoskopischer Messplatz entworfen, der es gestattet, die zu beobachtende Szene nicht nur als Einzelbilder zu erfassen, sondern als Video-Bildfolge. Dabei können die Bildsequenzen aus bis zu vier Blickwinkeln gleichzeitig erfasst werden. Weiterhin wurden verschiedene Algorithmen zur Kalibrierung der Kameras untersucht, welche es gestatten, die Bilder der Kameras zu entzerren und automatisiert die relative Lage der Kameras zueinander zu bestimmen. Die Funktion des Messplatzes wird derzeit an einer Außenbereichsszene „RUD25“ erprobt.

Projekt: Entwicklung, Implementierung und Test von parallelen Hardwarearchitekturen für drahtlose Hochgeschwindigkeitsnetzwerke**Ansprechpartner:** DR.-ING. FRANK WINKLER**Beteiligte Mitarbeiter:** PROF. DR. BEATE MEFFERT, ANNE WEGRICH, THOMAS KUHN**Zusammenarbeit:** IHP GmbH - Innovations for High Performance Microelectronics, Frankfurt (Oder)

Die Arbeiten sind in das vom BMBF geförderte Forschungsprojekt WIGWAM (Wireless Gigabit with Advanced Multimedia Support) eingeordnet. Ziel des Vorhabens sind konzeptionelle Untersuchungen, Modellierungsansätze und Implementierungsvarianten für Codierungsverfahren, die in drahtlosen Netzwerken (Wireless LAN) bis 1 Gbit/s Datenrate verwendet werden sollen. Zur näheren Auswahl werden LDPC,- Turbo- und Faltungscodes betrachtet, der besondere Schwerpunkt liegt auf der prototypischen Implementierung von LDPC-Codes (Low Density Parity Check Code). Die Modellierung erfolgt mit Matlab, es werden zunächst VHDL-Lösungsansätze für FPGA-Strukturen erarbeitet und bewertet. Die Ergebnisse werden danach anhand eines Chipentwurfes verifiziert. Dazu erfolgt eine Synthese mit Herstellerbibliotheken, der Entwurf eines Testkonzeptes und die Simulation der Implementierung anhand des Testkonzeptes. Aus den Untersuchungen wird die Eignung des LDPC-Codes für zukünftige Gigabit-Wireless-LAN-Netzwerke bestimmt.

Projekt: Spezifikation, Verifikation und Synthese global-asynchroner lokal-synchroner (GALS) Systeme und Schaltungen**Ansprechpartner:** DR.-ING. FRANK WINKLER**Beteiligte Mitarbeiter:** PROF. DR. BEATE MEFFERT, PROF. DR. WOLFGANG REISIG, DIPL.-INF. CHRISTIAN STAHL, ALEXANDRA JULIUS**Zusammenarbeit:** IHP GmbH - Innovations for High Performance Microelectronics, Frankfurt (Oder)

Anforderungen an Systeme zur mobilen Kommunikation führen zu immer komplexeren Schaltungen. Es wird zunehmend schwieriger, diese Schaltungen mit herkömmlichen Methoden des synchronen Schaltungsentwurfs zu implementieren. Insbesondere wachsen die Probleme der Taktsynchronisation, der elektromagnetischen Interferenz und des Energieverbrauchs. Zur Lösung bieten sich asynchrone Schaltungstechniken an; insbesondere global-asynchrone, aber lokal-synchrone Systeme. In diesem Projekt werden effektive Verfahren zum durchgängigen Entwurf und zur Implementierung von GALS-Schaltungen entwickelt. Schwerpunkt der Forschungen am Lehrstuhl sind die GALS-Schaltungssynthese, die Simulation und die Implementierung als integrierte Schaltungen. Die Leistungsfähigkeit der Entwurfsverfahren wird an exemplarischen Entwürfen nachgewiesen.

Projekt: Algorithmen und Implementation von Bildverarbeitungsfunktionen auf FPGA**Ansprechpartner:** DR.-ING. FRANK WINKLER**Beteiligte Mitarbeiter:** PROF. DR. BEATE MEFFERT, DR.-ING. OLAF HOCHMUTH, DIPL.-INF. UWE KNAUER**Zusammenarbeit:** Deutsches Zentrum für Luft- und Raumfahrt DLR, Berlin-Adlershof

Für die Steuerung und Kopplung von mehreren Bildsignalquellen 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 Entwurfsmethodik 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 Projektes 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 Entwurfs wird durch Simulation, Funktionstests im Labor und im Feldversuch nachgewiesen.

Projekt: Entwicklung eines Demonstrators zur Korrelation von Wireless-LAN-Signalen mit einer für Positionsbestimmungen geeigneten Zeitauflösung

Ansprechpartner: DR.-ING. FRANK WINKLER

Beteiligte Mitarbeiter: DR.-ING. MANFRED GÜNTHER, LOTHAR HEESE

Zusammenarbeit: IHP GmbH - Innovations for High Performance Microelectronics, Frankfurt (Oder)

Zur Positionsbestimmung mobiler drahtloser 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 im Bluetooth-Frequenzband zur Beurteilung der Leistungsfähigkeit ermöglicht. Aufbauend auf diesen Ergebnissen erfolgte eine Integration in 0,25- μ m-CMOS-TSMC-Technologie und die Realisierung eines Funktionsmusters einer Basisstation zur Positionsbestimmung. Mit diesem Aufbau wurden erfolgreich Messungen durchgeführt und das Messprinzip und die Auswertalgorithmen entwickelt. Für den Aufbau eines Demonstrators wurde eine Anpassung an die Schnittstellen eines 32-bit-Prozessors (LEON) und an die CMOS-Technologie des IHP Frankfurt vorgenommen. Diese Komponenten bilden einen Demonstrator, mit dem ein für 2006 geplanter Feldversuch mit einem drahtlosen Netzwerk (Wireless LAN, 5 GHz, gemäß IEEE802.11a) und einem Bluetooth-Netz (2,4 GHz) durchgeführt wird.

Projekt: Biosignalanalyse zur Ermittlung von Entspannung unter Snoezelen-Bedingungen

Bearbeiterin: PROF. DR. BEATE MEFFERT

Beteiligte Mitarbeiter: DR.-ING. OLAF HOCHMUTH

Zusammenarbeit: PROF. DR. KRISTA MERTENS, Abteilung Körperbehindertenpädagogik des Instituts für Rehabilitationswissenschaften der HU Berlin

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 hat 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 Probanden gewonnen worden, die unter-

schiedlichen akustischen Reizen ausgesetzt wurden, um einen möglichen Zusammenhang zwischen diesen Reizen und dem emotionalen Zustand der Versuchspersonen aufzudecken. Gemessen wurden der Hautleitwert, die Durchblutung und die Herz- und Atemfrequenz. Die Auswertung von Fragebögen ergänzt das Methodenspektrum. Die Auswertung der Experimente dauert an.

Projekt: Messsystem für den Betonfertigteiltbau

Ansprechpartner: DR.-ING. OLAF HOCHMUTH

Beteiligte Mitarbeiter: PROF. DR. BEATE MEFFERT, MICHAEL OTTO

Zusammenarbeit: Gesellschaft zur Förderung angewandter Informatik e.V.

Ziel ist die Entwicklung eines bildanalytisch arbeitenden Messsystems, das in der Lage ist, für den Montageprozess von Bauten aus Betonfertigteilen wichtige Daten zu erfassen. Dazu sollen alle Daten für eine passgenaue Montage eines jeweils neu einschwebenden Teils automatisch ermittelt werden. Diese Daten (beispielsweise Normalenrichtungen, Spaltenbreiten, Abstände) dienen zum einen dazu, die momentane reale Situation grafisch mit überlagerten Messwerten anzuzeigen, zum anderen sollen sie für die Steuerung eines Montagekrans verfügbar gemacht werden, damit dieser automatisch sein Steuerregime für ein sanftes und passgenaues Einschweben ableiten kann.

Das Lösungsprinzip basiert auf dem zielgerichteten Einbringen von Informationen in die Szene, hier in Form von mehreren (drei) projizierten Laserebenen, evtl. Laserebenenbündeln. Bei dem Messsystem wird es sich daher um einen Stativaufbau handeln, der mit einer Kamera und den Lasern mit Linienoptiken bestückt ist. Die räumliche Lage der Fertigteile soll dann mittels Lasertriangulation bestimmt werden. Notwendige Voraussetzung für eine exakte Lage der Fertigteile ist die Kalibrierung des Gesamtaufbaus.

Projekt: Beseitigung der Potentialasymmetrien beim EKG-Mapping-Verfahren

Ansprechpartner: DR.-ING. OLAF HOCHMUTH

Beteiligte Mitarbeiter: PROF. DR. BEATE MEFFERT, ATHANASSIOS 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. GERHARD KRENZKE)

Die Untersuchungen unterstützten die Forschungsarbeiten im Rahmen des DFG-Forschungsprojektes, das von einem interdisziplinären Wissenschaftlerteam bearbeitet wird. Mit Hilfe eines EKG-Mapping-Systems kann das elektrische Feld des Herzens ermittelt und auf eine Kugeloberfläche projiziert werden. Durch die Messung der Herzpotentiale auf der realen Thoraxoberfläche treten Abweichungen zum Kugelmodell auf. Insbesondere werden die Messwerte durch Potentialschwächungen auf Grund des Lungengewebes unsymmetrisch. Es sind Algorithmen entwickelt worden, die diese Potentialasymmetrien korrigieren können. Die Methoden zur Korrektur des Lungeneinflusses wurden implementiert und konnten somit erstmalig von der DFG-Forschergruppe bewertet werden.

Projekt: Neue Methoden zur Analyse der nichtlinearen Dynamik in der Herzratenvariabilität

Ansprechpartner: DR.-ING. OLAF HOCHMUTH

Beteiligte Mitarbeiter: PROF. DR. BEATE MEFFERT, HEIKO SCHARFF, ALEXANDER SUHRBIER

Zusammenarbeit: Forschungszentrum Karlsruhe GmbH, Universität Potsdam, Franz-Volhard-Klinik Berlin-Buch

Gegenstand der Zusammenarbeit ist Untersuchung von Regulationsvorgängen bei der menschlichen Herztätigkeit. Speziell sind Ansätze zur Analyse der spontanen Barorezeptorsensitivität zu untersuchen. Darüber hinaus sollen Wortverteilungen, ihre Kullback-Leibler-Entropie und ihre Eignung zur Patientenklassifikation untersucht werden. Die Untersuchungen werden ausgeweitet auf die Kopplungsanalyse von Herzrate und Blutdruck.

Projekt: Einsatz eines neu entwickelten Roboters KUKA KR 3-SI an einer Bearbeitungszelle

Ansprechpartner: DR.-ING. MICHAEL RITZSCHKE

Beteiligte Mitarbeiter: PROF. DR. BEATE MEFFERT, HANNES HÜBNER

Zusammenarbeit: FKS Maschinenbau GmbH

Im Rahmen des Projektes wurden die Möglichkeiten des Einsatzes eines Industrieroboters als Assistenzsystem in einer Montagezelle untersucht. Im Ergebnis dieser Untersuchungen entstand ein inzwischen realisierter Lösungsvorschlag, der eine sichere Mensch-Roboter-Kooperation gewährleistet. Weitere Aspekte waren die Optimierung der Handlungsabläufe zur Zykluszeitreduzierung, die Lösung konstruktiver Details zur Bereitstellung der Montageteile und die Entwicklung des Steuerungsprogramms.

Veröffentlichungen

D. BEMMANN: *IP Lookup on a Platform FPGA - a Comparative Study*. 19th International IEEE Parallel and Distributed Processing Symposium, Colorado, April 2005, Proceedings, p. 166a.

R. BLASCHEK; B. MEFFERT: *Automatisierte Korrespondenzsuche für die Stereobildverarbeitung*. 3D-NordOst, Berlin, Dezember 2005, Tagungsband, S. 5-12.

E. GRASS, F. WINKLER, M. KRSTIC, A. JULIUS, C. STAHL, M. PIZ: *Enhanced GALS Techniques for Datapath Applications*. 15th International Workshop, PATMOS 2005, Leuven/Belgium, September 2005. In: PALIOURAS, V.; VOUNCKX, J.; VERKEST, D.: *Integrated Circuit and System Design. Power and Timing Modeling, Optimization and Simulation*. Springer, ISBN: 3-540-29013-3, pp. 581-590.

H. F. HARMUTH, B. MEFFERT: *Dogma of the Continuum and the Calculus of Finite Differences in Quantum Physics*. Vol. 137, Advances in Imaging and Electron Physics. Elsevier, 2005, ISBN 0-12-014779-3.

U. KNAUER, M. HIMMELSBACH, F. WINKLER, F. ZAUTKE, K. BIENEFELD, B. MEFFERT: *Application of an Adaptive Background Model for Monitoring Honeybees*. 5th IASTED International Conference on Visualization, Imaging & Image Processing, Benidorm/Spain, September 2005, Proceedings, pp. 46-50.

U. KNAUER, R. REULKE, B. MEFFERT: *Fahrzeugdetektion und -erkennung mittels mehrdimensionaler Farbhistogrammanalyse*. Farbbildverarbeitung 2005, Gesellschaft zur Förderung angewandter Informatik, Berlin, Oktober 2005, Proceedings S. 93-100.

B. MEFFERT; R. BLASCHEK, U. KNAUER, R. REULKE, A. SCHISCHMANOW, F. WINKLER: *Monitoring traffic by optical sensors*. Second International Conference on Intelligent Computing and Information Systems (ICICIS 2005), Cairo/Egypt, März 2005, Proceedings, pp. 9-14.

B. MEFFERT; K. MERTENS, G. SCHNEIDER: *Forschendes Lernen - Begleitung diagnostischer und therapeutischer Verfahren in der Rehabilitationspädagogik*. Berlin, Humboldt-Spektrum 12 (2005) Heft 3, S. 54-58.

B. MEFFERT: *Snoezelen und seine Wirkung auf Biosignale*. In: Mertens, K.; Verheul, A. (Hrsg.): *Snoezelen - Anwendungsfelder in der Praxis*. ISNA 2005, S. 89-98.

D. SCHÄFER, G. VON WAGNER, S. HEY, K.-D. MÜLLER-GLASER, B. MEFFERT: *Comperative Study of Different Methods and Constraints in Heart Rate Variability Spectrum Estimation*. Proceedings, 39. Kongress für Biomedizinische Technik, Nürnberg, September 2005, S. 143-144.

H. SCHARFF, N. WESSEL, O. HOCHMUTH, H. MALBERG: *Vergleich von Methoden zur Analyse der spontanen Barorezeptorsensitivität*. 29. Wissenschaftlicher Kongress Hypertonie, Berlin, November 2005. In: *Deutsche Medizinische Wochenschrift* 130 (2005) Suppl. 4, S. 148.

A. SUHRBIER, H. MALBERG, O. HOCHMUTH, A. SCHIRDEWAN, N. WESSEL: *Different Approaches to Heart Rate Variability Analysis by Symbolic Dynamics*. 39. Kongress für Biomedizinische Technik, Nürnberg, September 2005, Proceedings, S. 145-146.

A. SUHRBIER, H. MALBERG, O. HOCHMUTH, A. SCHIRDEWAN, N. WESSEL: *Bivariate Symbolische Dynamik für die Herzraten- und Blutdruck-Analyse bei Dilatativer Kardiomyopathie*. 29. Wissenschaftlicher Kongress Hypertonie, Berlin, November 2005. In: *Deutsche Medizinische Wochenschrift* 130 (2005) Suppl. 4, S. 155.

A. WEISS, B. MEFFERT: *freeSp - ein XML-basiertes Framework für die Softwareentwicklung verteilter signalverarbeitender Systeme*. Berliner XML-Tage, September 2005, Tagungsband, S. 203-211.

T. WENCKEBACH, H. LAMECKER, H.-C. HEGE: *Capturing Anatomical Shape Variability Using B-Spline Registration*. Information Processing in Medical Imaging 2005 (IPMI), Lecture Notes in computer Science 3565, Springer 2005, Proceedings, pp. 578-590.

F. WINKLER, E. FISCHER, E. GRASS, G. FISCHER: *A 60 GHz OFDM Indoor Localization System Based on DTDOA*. 14th IST Mobile & Wireless Communication Summit Dresden, June 2005, Tagungsband, online-Publikation unter: www.mobilesummit2005.org.

Vorträge

R. BLASCHEK, B. MEFFERT: *Automatisierte Korrespondenzsuche für die Stereobildverarbeitung*. 8. Anwendungsbezogener Workshop zur Erfassung, Modellierung, Verarbeitung und Auswertung von 3D-Daten, Gesellschaft zur Förderung angewandter Informatik, Berlin, Dezember 2005.

C. BÜTTNER, K. BOBEY, B. SCHLICHTING: *Parametrierung digitaler Farbkameras auf Basis einer spektralen Charakterisierung*. 11. Workshop Farbbildverarbeitung, Gesellschaft zur Förderung angewandter Informatik, Berlin, Oktober 2005.

U. KNAUER, M. HIMMELSBACH, F. WINKLER, F. ZAUTKE, K. BIENEFELD, B. MEFFERT: *Application of an Adaptive Background Model for Monitoring Honeybees*. 5th IASTED International Conference on Visualization, Imaging and Image Processing (VIIP 2005), Benidorm/Spain, September 2005.

U. KNAUER, R. REULKE, B. MEFFERT: *Fahrzeugdetektion und -erkennung mittels mehrdimensionaler Farbhistogrammanalyse*. 11. Workshop Farbbildverarbeitung an der Gesellschaft zur Förderung angewandter Informatik, Berlin, Oktober 2005.

B. MEFFERT: *Traffic monitoring*. Key note, Second International Conference on Intelligent Computing and Information Systems (ICICIS 2005), Cairo/Egypt, März 2005.

B. MEFFERT, K. MERTENS G. SCHNEIDER: *Snoezelen Research - Biosignal Acquisition and Analysis*. 3. Internationales Snoezelen Symposium der International Snoezelen Association ISNA, Berlin, September/Oktober 2005.

A. WEISS, B. MEFFERT: *Ein XML-basiertes Framework für die Softwareentwicklung verteilter signalverarbeitender Systeme*. Berliner XML-Tage, September 2005.

F. WINKLER: *VHDL-Simulation*. Miniworkshop GALS, HU Berlin, Januar 2005.

Poster

M. HIMMELSBACH, U. KNAUER, F. WINKLER, B. MEFFERT, F. ZAUTKE, K. BIENEFELD: *Automatische Auswertung des Ausräumverhaltens der Honigbiene*. 19. Tagung der deutschsprachigen Sektion der Internationalen Union zum Studium sozialer Insekten (IUSI) gemeinsam mit der 52. Jahrestagung der Arbeitsgemeinschaft der Institute für Bienenforschung, Halle (Saale), März 2005.

H. SCHARFF, N. WESSEL, O. HOCHMUTH, H. MALBERG: *Vergleich von Methoden zur Analyse der spontanen Barorezeptorsensitivität*. 29. Wissenschaftlicher Kongress Hypertonie, Berlin, November 2005.

A. SUHRBIER, H. MALBERG, O. HOCHMUTH, A. SCHIRDEWAN, N. WESSEL: *Bivariate Symbolische Dynamik für die Herzraten- und Blutdruck-Analyse bei Dilatativer Kardiomyopathie*. 29. Wissenschaftlicher Kongress Hypertonie, Berlin, November 2005.

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
- FKS Maschinenbau GmbH, Berlin
- Forschungszentrum Karlsruhe GmbH
- Franz-Volhard-Klinik Berlin-Buch
- Geographisches Institut der HU Berlin
- Gesellschaft zur Förderung angewandter Informatik e.V., Berlin

- IHP GmbH - Innovations for High Performance Microelectronics, Frankfurt (Oder)
- Infineon AG, München
- Institut für Rehabilitationswissenschaften der HU Berlin
- Länderinstitut für Bienenkunde, Hohen Neuendorf
- Lüth & Dümchen Automatisierungsprojekt GmbH, Berlin
- PicoQuant GmbH, Berlin
- Siemens Dematic Postdienst-Automatisierung GmbH, Berlin
- Universität Potsdam, Hasso-Plattner-Institut für Softwaresystemtechnik
- Universität Stuttgart, Institut für Straßen- und Verkehrswesen

Sonstige Aktivitäten

Prof. Dr. Beate Meffert

- 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
- Mitglied des Forschungsverbundes Naturwissenschaft, Abrüstung und internationale Sicherheit (FONAS)
- Moderation des Arbeitskreises „Wissenschaftliche Arbeitsmethoden in den Natur- und Technikwissenschaften. Workshop für Stipendiatinnen des Berliner Programms zur Förderung der Chancengleichheit für Frauen in Forschung und Lehre. Berlin, 2. Juni 2005

Diplomarbeiten

GUJO, OLEG: *Modellierung signalverarbeitender Systeme durch Petri-Netze*, Februar 2005.

KOMAROV, ILLYA: *Modellierung signalverarbeitender Systeme durch Petri-Netze*, Februar 2005.

PREISS, MATTHIAS: *VHDL-Entwurf eines kaskadierten Detektors für die Objekterkennung in Bildern*, Dezember 2005.

RÄTZEL, MARCUS: *Entwurf einer allgemeinen Speicherschnittstelle für FIFO-gestützte Videodatenerfassung auf FPGA-Boards*, Dezember 2005

STEIN, ANDREJ: *Bestimmung von Fledermausarten anhand ihrer akustischen Signale*, März 2005.

TRIANAFYLLOU, ATHANASSIOS: *Signalverarbeitungsalgorithmen zur Darstellung des elektrischen Herzfeldes und Untersuchungen zur Beseitigung der Potentialasymmetrien beim EKG-Mapping-Verfahren*, März 2005.

Lehr- und Forschungseinheit

Computer Vision

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

Leiter

PROF. DR. RALF REULKE

Tel.: (030) 2093 3044

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

Sekretariat

SABINE DZIWSZ

Tel.: (030) 2093 3044

Fax: (030) 2093 3045

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

Doktoranden

DIPL.-ING. FREDERIK MEYSEL

DIPL.-ING. THOMAS DÖRING

Diplomanden

SASCHA BAUER

ARTHUR LIPPOK

JÜRGEN WOHLFEIL

Das Fachgebiet „Computer Vision“, vertreten durch die gleichnamige Professur innerhalb der Technischen Informatik, befasst sich in Lehre und Forschung mit Stereo-Bildverarbeitung, Szenenbeschreibung und Visualisierung, Datenfusion sowie Sensor-design und Verifikation. Die Professur steht in enger Verbindung zum Deutschen Zentrum für Luft- und Raumfahrt / Institut für Verkehrsforschung.

Schwerpunkt der Forschungsaktivitäten des Jahres 2005 war die Entwicklung von Verfahren zur automatisierten Erfassung des motorisierten, aber auch nicht-motorisierten Verkehrs mit optoelektronischen Sensoren, die Beschreibung und Geocodierung sowie Speicherung und Visualisierung von Verkehrsobjekten.

Lehre und Forschung erfolgt in enger Zusammenarbeit mit der Lehr- und Forschungseinheit Signalverarbeitung und Mustererkennung von Frau Prof. Dr.-Ing. Beate Meffert.

In der Lehre werden Lehrveranstaltungen zur Stereo-Bildverarbeitung angeboten.

Lehre

Hauptstudium (Halbkurs)

- Stereo-Bildverarbeitung (R. REULKE)

Forschung

Projekt: Objekterfassung und Objektverfolgung für die Verkehrsbeobachtung

Ansprechpartner: PROF. DR. RALF REULKE

Beteiligte Mitarbeiter: Sascha Bauer, Björn Piltz, Robert Spangenberg, Jürgen Wohlfeil, Thomas Döring

Zusammenarbeit: Lehr- und Forschungseinheit Signalverarbeitung und Mustererkennung, Deutsches Zentrum für Luft- und Raumfahrt, Berlin-Adlershof

Unter Verkehrsbeobachtung versteht man hier die Erfassung und Verfolgung von Verkehrsobjekten von ruhenden und bewegten Plattformen. Die abgeleiteten Daten sollen Eingang in die Verkehrsmodellierung und Simulation finden. Ziel ist eine verbesserte Verkehrssteuerung. Im Kontext der Verkehrsbeobachtung ist die Erfassung von Form, Farbe und Position der interessierenden Verkehrsobjekte. Weiteren Informationsgewinn ergibt die Analyse und Interpretation von Bildfolgen oder Videosequenzen. Hieraus können zusätzliche Daten, beispielsweise Geschwindigkeit und Beschleunigung und damit generelle dynamische Eigenschaften von Objekten gewonnen werden.

Für die flächendeckende und zeitversetzte Verfolgung von Verkehrsobjekten in großräumigen Gebieten mit verschiedenen positionierten optoelektronischen Sensoren erfolgt die Kombination der abgeleiteten Daten eines Sensors mittels Verfolgungsalgorithmen (Tracking). Zunächst wurde ein Kalmanfilter als Trackingansatz zur Verfolgung mehrerer Fahrzeuge innerhalb *einer* Videoszene gewählt. Dieses Kalman-Filter wurde anschließend für den Einsatz mehrerer Kameras angepasst.

Für Multikamerasysteme ist die Bestimmung der gegenseitigen Orientierung der Kameras von besonderer Bedeutung. Als Beispiel soll ein Projekt zur Verfolgung von Personen auf einem Campus dienen. Dazu müssen die Orientierungsparameter von mehreren Kameras so bestimmt werden, dass die von ihnen gemessenen Weltkoordinaten im Überlappungsbereich nur geringe Abweichungen aufweisen. Dies wurde auf der Grundlage von Passpunkten, die aus Stereoluftbildern gewonnen wurden, realisiert.

Ansätze für die Bildverarbeitung von bewegten Plattformen unterscheiden sich in wesentlichen Teilen von denen, die oben beschrieben wurden. Ziel ist hier die automatische Erkennung von Fahrbahn und Fahrzeugen aus einem fahrenden Fahrzeug heraus. Dazu müssen zuerst die verkehrsaktiven Flächen identifiziert werden. So werden Fluchtpunkt und Lage der Fahrbahnen durch die Analyse der linienartigen Kanten extrahiert. Ein erster Schritt zur Objektverfolgung aus dem Fahrzeug heraus ist die Erkennung, die Zusammenfassung und die Verfolgung von charakteristischen Strukturen.

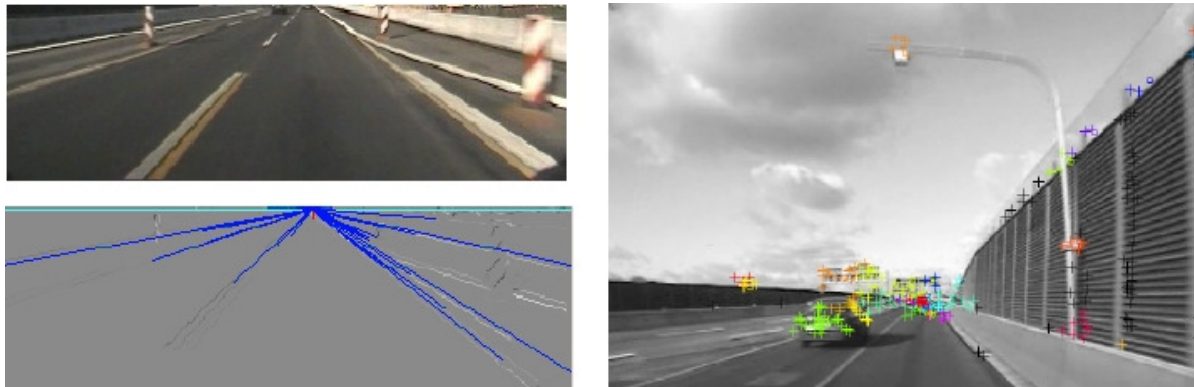


Abbildung 1: Fahrbahnerkennung (links) und Verfolgung interessanter Strukturen

Projekt: Multi-Sensor-Datenfusion zur Interpretation komplexer Situationen

Ansprechpartner: FREDERIK MEYSEL

Beteiligte Mitarbeiter: PROF. DR. RALF REULKE

Zusammenarbeit: Deutsches Zentrum für Luft- und Raumfahrt DLR, Berlin-Adlershof

Ziel ist die Beobachtung und die automatische Interpretation dynamischer Situationen mit vielen unabhängigen Akteuren. Zum Verständnis einer Szene gehört dabei nicht nur die Rekonstruktion der dreidimensionalen Zusammenhänge zwischen sich bewegenden Objekten, sondern auch die Bewertung der Vorgänge im Hinblick auf atypische Aspekte, die ggf. als Gefährdung eingestuft werden müssen. Solche Systeme können zum Einsatz kommen bei der Verkehrsüberwachung, in Luft- und Raumfahrt, aber auch zur Beurteilung belebter Plätze, Fußgängerzonen oder auch zur Sicherung von Großereignissen wie z.B. der Fußballweltmeisterschaft.

Ausgegangen wird dabei immer von der multisensoriellen Erfassung einer Situation, vor allem mit optischen Sensoren (Kameras) unterstützt durch Infrarot und LIDAR-Systeme. Den verschiedenen Typen von Sensoren und deren unterschiedlichen Erfassungsbereichen und Standorten muss durch geeignete Fusionsverfahren Rechnung getragen werden.

Schließlich werden Eigenschaften der Szene extrahiert und Objekte verfolgt (Tracking). Das eigentliche Ziel, die Erkennung von atypischen Ereignissen bzw. Entwicklungen in der Szene ist so komplex, dass es zunächst strukturiert werden muss. Im Zusammenhang mit dieser Arbeit wurde eine dreigliedrige Unterteilung in globale, objektbezogene und situationsbezogene Maße zur Bewertung der Situation vorgenommen. So können in erster Näherung bereits bestimmte Messungen ohne weitere Untersuchungen als atypisch bewertet werden, zum Beispiel eine zu hohe Geschwindigkeit eines Fahrzeugs. Bei objektbezogenen Verfahren werden die durch das Tracking ermittelten Trajektorien von Objekten auf ungewöhnliche Merkmale untersucht. Zum Beispiel können Schleifen im Gang eines Fußgängers auf Ziellosigkeit oder unlautere Absichten, evtl. aber auch auf seine Hilflosigkeit hindeuten. Es sollen weiterhin bezogen auf die Trajektorien Merkmale wie „Schlendern“ oder „Zielstrebigkeit“ zugeordnet werden können. Die letzte Variante der situationsbezogenen Maße stellt die schwierigste, aber auch interessanteste Möglichkeit zum Verständnis einer Situation bereit. Hierunter fallen einfache Verfahren, die z.B. die Geschwindigkeit eines Objekts in Abhängigkeit von seinem Ort zu betrachten. So kann ein zügig beschleunigendes Objekt auf einer gerade Straße ganz normal sein, nicht aber in einer sehr engen

Kurve. Auch komplexe Analysen wie z.B. die Bewertung von Trajektorien mehrerer miteinander interagierender Objekte in einer Szene gehören in diesen Bereich.

Projekt: Aufnahme und Visualisierung von komplexen Szenen

Ansprechpartner: THOMAS DÖRING

Beteiligte Mitarbeiter: PROF. DR. RALF REULKE

Zusammenarbeit: Deutsches Zentrum für Luft- und Raumfahrt, Berlin-Adlershof

Im Rahmen der Verkehrsforschung gibt es eine Reihe von Fragestellungen, die eine Visualisierung von Verkehrsobjekten erforderlich machen. So werden Kreuzungen mittels optoelektronischer Systeme (z.B. Kameras, die im sichtbaren oder im thermalen infraroten Spektralbereich messen) überwacht. Mit Bildverarbeitungsmethoden werden Objektinformationen (Ort, Typ, Größe, etc.) abgeleitet. Diese Bildinformation wird an den Kreuzungsrechner weitergeleitet und in einer Verkehrszentrale verarbeitet und visualisiert.

Für die Erstellung einer „realistischen virtuellen Umgebung“ sind Oberflächenmodelle und die Texturierung dieser 3D-Objekte notwendig. Die Messung dieser Informationen kann terrestrisch und von Flugzeugplattformen erfolgen. Zum Teil existieren diese Daten, andere müssen neu erhoben werden. Das entsprechende Sensorpaket besteht aus einer hochauflösenden Kamera (z.B. die Panoramakamera), einem Laserscanner und einem Inertialnavigationssystem.

Mehrere Experimente mit einem solchen Sensorpaket wurden im Jahre 2005 durchgeführt. Auf der Grundlage der gewonnenen Daten wurde zunächst ein mathematisches Modell zur Online-Kalibrierung einer Zeilenkamera entwickelt. Auf der Grundlage von Bildaufnahmen einer solchen Kamera und aufgezeichneten Positionierungsdaten ermöglicht ein solches Modell insbesondere die Bestimmung von Verzeichnungsparametern. Dazu werden im Gegensatz zur klassischen Photogrammetrie mit Flächenbilddaten lediglich wenige bekannte Objektpunkt-Koordinaten in einem *beliebigen* Koordinatensystem benötigt. Dies ermöglicht unter anderem die Kalibrierung einer Zeilenkamera unter realen Aufnahmebedingungen. Die erhaltenen Algorithmen wurden in *RSI IDL* umgesetzt und auf der Grundlage von generierten Idealdaten verifiziert.



Abbildung 2: Rohdaten einer Zeilenkameraaufnahme (links) und korrigierte Daten (rechts)

Weitere Untersuchungen befassen sich insbesondere mit der Automatisierung des Auswerteprozesses und der Genauigkeit der Koregistrierung und der Messungen.

Veröffentlichungen

R. REULKE, U. KNAUER: *International Archives Of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume XXXVI-5/W8, Proceedings of the ISPRS working group V/5, 'Panoramic Photogrammetry Workshop', Berlin, Germany, 24-25 February 2005, ISSN 1682-1750.

R. KLETTE, R. REULKE: *Modeling 3D Scenes: Paradigm Shifts in Photogrammetry, Remote Sensing and Computer Vision*, Invited talk, International Conference on Systems and Signals(ICSS), Kaohsiung, Taiwan on April 28-29, 2005.

R. REULKE, K. SCHEIBE, A. WEHR: *Integration of digital panoramic camera and laser scanner data*, International Workshop on Recording, Modeling and Visualization of Cultural Heritage, Ascona, 2005.

Vorträge

RALF REULKE: *Panoramic Data Acquisition and Evaluation*, Invited talk, ICOB 2005, Workshop On Immersive Communication And Broadcast Systems, 27./28. October 2005, Fraunhofer HHI, Germany.

R. SCHIFFEL, R. REULKE: *Berlin-Intermodal“ Integrierte Echtzeitinformation für einen verbesserten Verkehrsfluss*, Tag der Verkehrswirtschaft, Wissenschaft, Wirtschaft und Betreiber im Dialog, 07. Dezember 2005.

Wissenschaftliche Kooperationen

- Abteilung Vermessung und Geoinformation , Fachhochschule Basel
- Chair of Information Technology, The University of Auckland
- Computer Vision and Remote Sensing, Technische Universität Berlin
- Department of Cybernetics, Czech Technical University in Prague
- Institut für Verkehrsführung und Fahrzeugtechnik, DLR
- Institut für Robotik & Mechatronik, DLR
- Institut für Methoden der Fernerkundung, DLR
- Institut für Photogrammetrie, Universität Stuttgart
- Institut für Photogrammetrie und Fernerkundung, TU Dresden
- Institut für Photogrammetrie und Geoinformation, Universität Hannover
- Imaging, Robotics, and Intelligent Systems Laboratory, The University of Tennessee
- Laboratory for Atmospheric and Space Physics, University of Colorado

Dissertationen

THOMAS DÖRING: „Verfolgung von Verkehrsobjekten aus verschiedenen Kameraperspektiven“, Diplomarbeit

MELANIE SELL: „Implementierungsaspekte von Bildverarbeitungsoperatoren in einer FPGA-Umgebung, Diplomarbeit

V. „Mehr Frauen in die Informatik!“



*

**Studienwahl
Studium
Beruf**

**am Institut für Informatik der Humboldt-Universität zu Berlin
Roberta-RegioZentrum Berlin HU**

Dr. Márta Gutsche, Projektleiterin
Humboldt-Universität zu Berlin,
Institut für Informatik
Unter den Linden 6
10099 Berlin

Sitz: Rudower Chaussee 25
Haus IV, Raum 108
12489 Berlin

URL: www.informatik.hu-berlin.de/~gutsche/ideen-werkstatt/index.htm

Tel: +49 / 30 / 2093-5468

Fax: +49 / 30 / 2093-3045

Email: gutsche@informatik.hu-berlin.de

Bei der Förderung der Chancengleichheit hat sich das Institut für Informatik mit dem Projekt „Mehr Frauen in die Informatik!“ den Schwerpunkt auf die Steigerung des Anteils der Frauen in den Studiengängen der Informatik zum Ziel gesetzt.

An der Humboldt-Universität wurden im Jahr 2003 für alle Informatikstudiengänge zusammen 207 Studierende zugelassen, darunter 31 Studentinnen (14,9%). Unter 138 Studienanfängern des Diplomstudiengangs Informatik waren nur 9 Frauen (6,5%). Diese Zahlen sprechen für sich und sind allein ein guter Grund für die Maßnahmen des Projekts.

ErstsemestlerInnen am Institut für Informatik

	insgesamt	davon weiblich	davon weiblich in %
WS 03/04	207	31	14,9
WS 04/05	159	18	11,3
WS 05/06	149	30	20,1

Das wichtigste Anliegen aller Projektmaßnahmen bleibt weiterhin die Steigerung des Anteils weiblicher Studierender in allen Studiengängen der Informatik. Dies kann nur gelingen, wenn die gegenwärtigen Angebote für Schülerinnen und die neuen Formen der Zusammenarbeit mit den Gymnasien aufrecht erhalten werden. Steigt die Zahl der Studentinnen, so wird auch die Chance größer, dass mehr Frauen in dieser zukunftsorientierten und interessanten Wissenschaftsdisziplin tätig werden und als Vorbild (auch für Schülerinnen) fungieren können.

In das so entstehende Netzwerk können Schülerinnen, Studentinnen, Absolventinnen, Praktikantinnen und Lehrkräfte des Instituts eingebunden werden. Das Netzwerk bietet dem weiblichen Nachwuchs der Informatik und Naturwissenschaften eine verbesserte Beratungsleistung, mehr Informationen und wichtige Kontaktmöglichkeiten.



Frau Tempelhoff, Lehrerin der Hugo-Gaudig-Oberschule (unsere Absolventin) mit ihren Schülerinnen beim RoboCup Junior in Osaka (Japan) und ihre Roboter „Clärchen und Paulchen“



Sie haben ihren ersten Roboter in unserer Ideen-Werkstatt gebaut und programmiert.

„Mehr Frauen in die Informatik!“
Ideenwerkstatt zu Studienwahl, Studium und Beruf *
am Institut für Informatik der Humboldt-Universität zu Berlin
2005

Projektmaßnahmen	in Zahlen
Arbeitsgemeinschaften (AG) für Schülerinnen ab Kl. 9 unter Leitung von Studentinnen „Roberta-Roboter-AG“, „3D-AG“ „Jawa-AG, „Der Pinguin lernt laufen“-AG 2 Präsentation-Kurse	4 AGs halbjährig (wöchentlich) 36 Schülerinnen 1 Schüler 3 Studentinnen 1 Student 1 Mitarbeiterin 2 vierwöchige Kurse 34 Schülerinnen, 5 Schüler 1 Studentin
Betriebspraktika mit Betreuung am Institut	3 Schülerinnen, 3 Schüler
Roberta-KursleiterInnen-Schulung in Zusammenarbeit mit dem Fraunhofer-Institut- AIS (gendergerechte Didaktik)	15 KursleiterInnen (2 Studenten, 6 Lehrerinnen, 5 Sozialarbeiterinnen, 1 Informatiker, 1 Mitarbeiter)
Roberta-Roboter-Kurse	92 Schülerinnen, 22 Schüler 46 Roboter wurden konstruiert und programmiert
Projektwochen mit Schulen und Präsentationen zum Informatikstudium. Besuche von Schülerinnengruppen / Schulklassen am Institut	64 Schülerinnen, 30 Schüler
Tage der offenen Tür für Schülerinnen bei SIW	11 Schülerinnen, 34 Schüler
Girl's Day	42 Schülerinnen 5 Studentinnen, 1 Student
Tage der Forschung beim AIBO-Team	20 Schülerinnen, 26 Schüler
Lange Nacht der Wissenschaften	4 Schülerinnen, Flyer, Poster
Homepage-Tutorium für StudentInnen Studentische Leitung: 2 Studenten	12 Studentinnen 5 Studenten
Studentische Impressionen- Informatik-Studium in Adlershof (Fotoprojekt)	1 Student

**Gefördert durch das Programm zur Förderung der Chancengleichheit für Frauen in Forschung und Lehre
Förderschwerpunkt: Steigerung des Anteils der Frauen in naturwissenschaftlichen und technischen Studiengängen, gemeinsam mit dem Institut für Informatik der Humboldt-Universität zu Berlin*

VI. 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

FRANK WOZOBULE

Die Rechnerbetriebsgruppe ist dem Lehrstuhl Systemarchitektur zugeordnet.

Im Jahr 2005 verringerte sich die Zahl der durch die Rechnerbetriebsgruppe betreuten Nutzer erstmals seit dem Jahre 1990. Zum Jahresende waren mehr als 2100 Nutzer bei der Informatik registriert. Dies sind 50 Nutzer weniger als im Vorjahr. Die studentischen Rechnerpools in Adlershof wurden in diesem Jahr gleichmäßig von Montag bis Freitag genutzt. Trotz einer weiterhin hohen Zahl von Studenten in den ersten Studienjahren kam es während der betreuten Praktikumszeiten zu keinen nennenswerten Wartezeiten für Poolarbeitsplätze.

Entwicklung der rechentechnischen Ausstattung des Instituts 2005

Wesentlichste Neuerung im zentralen Bereich des Instituts im Jahre 2005 war die Umstellung der Mailserver des Instituts. Die beiden bisher genutzten SUN Netra t1 wurden durch zwei SUN Fire V240 mit je 2 Prozessoren und 2 GByte Hauptspeicher ersetzt.

Weiterhin wurden die Filespeicherkapazitäten für die Mitarbeiter erhöht. So stehen jetzt zusätzlich 900 GByte für Mitarbeiter im SAN zur Verfügung.

Die rechentechnische Ausstattung in einigen LFGs konnte verbessert werden. Wesentliche Neuerungen sind:

- Ausstattung des LFG Rechnerorganisation und -kommunikation mit einem Computer-Server SUN V890
- Ausstattung des LFG Künstliche Intelligenz mit einem Fileserver SUN V490.
- Anschluss des LFG Künstliche Intelligenz an das SAN des CMS (500 GByte)
- Anschluss des LFG Komplexität und Kryptografie an das SAN des CMS (1 TByte)
- Ausstattung des LFG Rechnerorganisation und Kommunikation mit neuen Arbeitsplatzrechnern für die Mitarbeiter
- Erweiterung des PC-Clusters des LFG Algorithmen und Komplexität um 8 Knoten
- Beschaffung von diversen Notebooks für Mitarbeiter

Entwicklung der Netzwerkinfrastruktur des Instituts 2005

Schwerpunkt beim Ausbau der Netzwerkinfrastruktur war auch im Jahre 2005 die Erweiterung der Gigabit Portkapazität im Backbone-Switch-Bereich. Im Gegensatz zu den Vorjahren wurden in diesem Jahr vorrangig Lehrstuhlserver an das Gigabit Backbone angeschlossen. Gegen Ende des Jahres wurde mit der vollständigen Erneuerung und Erweiterung der Etagen-Switche begonnen. Dabei wird die bisherige Fastethernet-Technik durch Gigabit-Technik abgelöst.

Auch in diesem Jahr konnten die Netzwerk-Kommunikations-Dienste ganzjährig stabil zur Verfügung gestellt werden. Alle notwendigen Systemupdates und Wartungsarbeiten (Erweiterungen) konnten ohne nennenswerte Unterbrechungen des laufenden Betriebs durchgeführt werden.

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.

Beim Austausch der Mailserver wurde gleichzeitig deren Software erneuert, so dass jetzt sowohl die SPAM-Filterung als auch die Virenfilterung der Mail vor der eigentlichen Annahme der Mail erfolgt. Dadurch konnte der Durchsatz der Mailserver wesentlich erhöht werden. Gleichzeitig wurden Firewalls für die Mailserver aktiviert. Die Umrüstung der zentralen Mailserver führte nicht zu längeren Unterbrechungen bzw. Datenverlusten. Das tägliche Mailaufkommen liegt bei 45.000 bis 55.000 Mails. Außerdem werden täglich 24.000 bis 30.000 Mails, die nicht standardkonform sind, zurückgewiesen. Diese Werte liegen auf dem Niveau des Jahres 2004. Die eingesetzten SPAM-Filter klassifizieren ca. 60% der eingehenden Mail als SPAM. Durch die zentralen SPAM-Filter ist die Güte der SPAM-Filterung etwas zurückgegangen, da die individuellen SPAM-Filtereinstellungen der Nutzer nicht bei Mail-Listen berücksichtigt werden können. Dies wird jedoch durch die zusätzlichen Rückweisungsmechanismen für nicht standardkonforme Mails kompensiert, so dass der ausgefilterte SPAM-Anteil bei ca. 80-90% der eintreffenden Mail liegt. Jede eingehende Mail wird durch einen Virenscanner geprüft. Diese Mails werden ebenfalls zurückgewiesen. Das Aufkommen an Mailviren ist in den letzten Monaten leicht steigend.

Nach der Beseitigung der Kapazitätsprobleme ist die Menge der ausgelieferten Daten des FTP-Servers während des letzten Jahres auf über 1,2 TByte monatlich gewachsen. Zur Spiegelung von wichtigen FTP-Servern wurde eine neue Software installiert, die stabiler läuft.

Die Software des NEWS-Servers wurde erneuert, so dass jetzt jeder Angehörige des Instituts News von jedem beliebigen Rechner aus lesen kann (Authentifizierung wird unterstützt). Dabei wurde die Auswahl der angebotenen Newsgruppen aktualisiert. Die Zugriffszahlen für den NEWS-Server sind bei erweitertem Angebot nach der Umstellung signifikant gestiegen, von zuvor maximal 5000 gelesenen Artikeln monatlich auf über 100.000 gelesene Artikel monatlich.

Die WWW-Zugriffe auf den Institutserver liegen semesterabhängig zwischen 150.000 und 200.000 täglich. Damit haben sich die Zugriffszahlen im Vergleich zum Vorjahr etwas erhöht. Die transportierte Datenmenge hat sich dabei kontinuierlich vergrößert und innerhalb eines Jahres verdoppelt und liegt jetzt bei fast 10 GByte täglich. Das Problem des Mail-Zugriffs über den WWW-Server hat sich durch den Einsatz der SUN V480 ent-

schärft. Eine Verlangsamung der Zugriffszeiten wegen Ausschöpfung der Prozessorkapazität durch das MAIL-Tool ist auch während der Spitzenbelastungszeiten kaum spürbar.

Die Account-Beantragung mittels WWW hat sich bewährt. Dadurch konnte 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 ebenfalls bewährt.

Weiterhin rückläufig sind die Zugriffszahlen für die Remote-Einwahl. Hier sind die Ursachen sicher in der preiswerten Einwahlmöglichkeit des CMS über [DFN@HOME](#) zu sehen. Der Terminalserver für die analoge Einwahl wurde endgültig abgeschaltet. Es wird nur noch der Terminalserver für ISDN-Einwahl (16 Leitungen) betrieben.

Mailverkehr 2005

Mailserver 1 (mail)				
<i>Monat</i>	<i>Gesamt</i>	<i>Empfangen extern</i>	<i>Versendet</i>	<i>Lokale E-mail</i>
Januar	894396	514.439	69.995	309.962
Februar	841.158	467.760	76.710	296.688
März	843.328	188.027	78.179	576.122
April	1.560.207	815.350	484.342	260.515
Mai	901.271	68.164	11.188	821.919
Juni	885.160	386.532	48.685	449.943
Juli	897.175	419.127	51.971	426.077
August	729.670	345.439	46.965	337.266
September	629.102	241.159	31.586	356.357
Oktober	775023	296.766	35.770	442.487
November	790.704	321.267	48.943	420.494
Dezember	867.728	376.237	39.404	452.087

Mailserver 2 (mailslv1)				
<i>Monat</i>	<i>Gesamt</i>	<i>Empfangen extern</i>	<i>Versendet extern</i>	<i>Lokale E-mail</i>
Januar	768.465	380.036	151.477	236.952
Februar	707.680	324.743	143.877	239.060
März	687.149	317.545	158.188	211.416
April	795.126	338.381	296.021	159.824
Mai	822.376	342.075	156.106	324.195
Juni	647.041	307.212	143.301	196.528
Juli	677.311	360.830	129.553	186.928
August	570.518	308.867	115.528	146.123
September	508.647	197.732	33.555	277.360
Oktober	619.344	226.888	37.752	354.704
November	655.452	235.949	38.379	381.124
Dezember	766,428	294.390	75.258	396.780

SPAM-mail

<i>Monat</i>	<i>SPAM- mail</i>	<i>Filterung</i>		<i>SPAM- mailslv1</i>	<i>Filterung</i>	
	<i>geprüft</i>	<i>gefunden</i>	<i>%</i>	<i>geprüft</i>	<i>Gefunden</i>	<i>%</i>
Aug.	309.271	177.170	57	-	-	-
Sept.	203.608	89.917	44	58.472	43.178	74
Okt.	255.017	111.455	44	62.413	43.683	70
Nov.	285.786	131.262	46	71.078	53.488	75
Dez.	304.665	172.693	56	93.659	73.915	79

Virenmail und Zurückweisungen

<i>Monat</i>	<i>Mails mit Viren</i>		<i>Zurückweisung</i>
	<i>mail</i>	<i>mailsl v1</i>	<i>mail</i>
Aug.	10.269	-	62.454
Sept.	1.778	654	255.106
Okt.	2.026	625	305.482
Nov.	2.566	523	249.154
Dez.	5.091	2.028	284.236

Zugriffe zum WWW-Server im Jahr 2005

<i>Monat</i>	<i>Transfer in MByte</i>	<i>Zugriffe</i>
Januar	137.459	6.000.647
Februar	128.289	6.005.530
März	143.230	6.914.822
April	138.883	5.636.429
Mai	132.471	5.220.702
Juni	118.721	4.900.329
Juli	166.231	5.964.417
August	179.543	4.972.896
September	244.397	4.167.656
Oktober	371.100	5.213.990
November	338.094	5.427.055
Dezember	282.928	4.729.245

Zugriffe zum FTP-Server im Jahr 2005

<i>Monat</i>	<i>Dateien</i>	<i>MByte</i>	<i>Systeme</i>	<i>Dateien tägl.</i>	<i>MByte tägl.</i>
Januar	395.137	1.418.255	11.802	12.348	44.320
Februar	314.773	1.362.828	10.567	10.854	46.004
März	262.511	1.389.705	10.584	8.203	43.428
April	286.573	1.234.579	8.168	9.244	39.825
Mai	201.774	1.331.954	7.818	6.305	41.623
Juni	219.134	1.943.516	7.590	6.848	60.734
Juli	335.094	1.289.059	7.617	10.472	40.283
August	301.686	1.266.583	8.353	9.428	39.580
September	309.349	1.382.307	8.310	9.947	44.590
Oktober	715.314	1.552.450	10.162	22.354	48.514
November	568.280	1.221.900	10.322	15.106	39.416
Dezember	345.675	953.649	9.402	10.802	29.801

Zugriffe zum News-Server im Jahr 2005

<i>Monat</i>	<i>Empfangene Artikel</i>	<i>Gesendete Artikel</i>	<i>Gelesene Artikel</i>
Januar	2.785.598	1.342.079	2.799
Februar	2.328.832	1.121.804	2.083
März	2.800.994	2.029.597	2.943
April	3.922.179	6.129.182	8.202
Mai	1.980.538	2.992.508	6.622
Juni	6.190.215	10.412.265	242.481
Juli	3.367.505	5.557.999	168.148
August	3.782.072	3.945.126	252.869
September	3.715.605	4.119.820	86.404
Oktober	3.783.707	4.134.759	111.593
November	3.870.306	10.450.282	125.238
Dezember	3.586.337	7.384.532	533.544

Zugriffe zu den Terminal-Servern im Jahr 2005

<i>Monat</i>	<i>Nutzer ISDN</i>	<i>Verbind.</i>	<i>Dauer in Stunden</i>
Januar	51	806	195:40
Februar	46	767	178:27
März	49	725	187:55
April	40	767	147:42
Mai	46	1135	299:09
Juni	22	537	83:39
Juli	15	579	82:03
August	20	834	120:15
September	17	388	103:17
Oktober	21	320	87:29
November	26	410	90:24
Dezember	19	488	143:47

VII. Lehrveranstaltungen

SOMMERSEMESTER 2005

GRUNDSTUDIUM

Diplomstudiengang (D)

2. Semester

Praktische Informatik 2

Vorlesung	4 Std.	H.-D. BURKHARD
Übung	4 Std.	M. MINOR
Übung	4 Std.	G. LINDEMANN-v. TREZBIATOWSKI
Übung	4 Std.	K. SCHMIDT
Übung	4 Std.	D. GÖHRING
Übung	2 Std.	I. DROST
Praktikum		K. AHRENS, B. HOHBERG

Technische Informatik 1

Vorlesung	2 Std.	F. WINKLER
Übung	2 Std.	F. WINKLER,
Übung	2 Std.	A. WEIß
Praktikum		K.-H. HAUPTVOGEL, M. GÜNTHER, F. WINKLER, M. RITZSCHKE

Mathematik 2

Vorlesung	4 Std.	A. GRIEWANK
Übung	4 Std.	H.-J. LANGE
Übung	4 Std.	J. RIEHME
Übung	2 Std.	H. WAGNER

4. Semester - Diplom

Technische Informatik 2

Vorlesung	4 Std.	M. MALEK
Übung	4 Std.	S. SOMMER
Übung	2 Std.	J. RICHLING

Theoretische Informatik 3

Vorlesung	2 Std.	J. KÖBLER
Übung	4 Std.	O. BEYERSDORFF
Übung	2 Std.	C. SCHWARZ

Informatik und Gesellschaft

Vorlesung	2 Std.	J. KOUBEK
-----------	--------	-----------

Magister – 2. Hauptfach (M)**Praktische Informatik 2**

Vorlesung	4 Std.	H.-D. BURKHARD
Übung	4 Std.	G. LINDEMANN-V. TRZ.
Übung	4 Std.	K. SCHMIDT
Übung	4 Std.	M. MINOR
Übung	4 Std.	D. GÖHRING
Übung	2 Std.	I. DROST
Praktikum		K. AHRENS, B. HOHBERG

Technische Informatik 2

Vorlesung	4 Std.	M. Malek
Übung	4 Std.	S. SOMMER
Übung	2 Std.	J. RICHLING

Magister-Nebenfach (N)**Praktische Informatik 2**

Vorlesung	4 Std.	H.-D. BURKHARD
Übung	4 Std.	G. LINDEMANN-V. TRZ.
Übung	4 Std.	K. SCHMIDT
Übung	4 Std.	M. MINOR
Übung	4 Std.	D. GÖHRING
Übung	2 Std.	I. DROST
Praktikum		K. AHRENS, B. HOHBERG

Lehramt (L)**Praktische Informatik 2**

Vorlesung	4 Std.	H.-D. BURKHARD
Übung	4 Std.	G. LINDEMANN-V. TRZ.
Übung	4 Std.	K. SCHMIDT
Übung	4 Std.	M. MINOR
Übung	4 Std.	D. GÖHRING
Übung	2 Std.	I. DROST
Praktikum		K. AHRENS, B. HOHBERG

Vorbereitung, Analyse, Planung von Informatikunterricht

Seminar	2 Std.	E. LEHMANN
---------	--------	------------

Unterrichtspraktikum/Blockpraktikum

Praktikum	CH. DAHME
-----------	-----------

Bachelor-Kombinationsstudiengang – Kernfach und Zweitfach (B.A.)**Praktische Informatik 2 (Kernfach)**

Vorlesung	4 Std.	H.-D. BURKHARD
Übung	4 Std.	G. LINDEMANN-V. TRZ.
Übung	4 Std.	K. SCHMIDT
Übung	4 Std.	M. MINOR
Übung	4 Std.	D. GÖHRING
Übung	2 Std.	I. DROST
Praktikum		K. AHRENS, B. HOHBERG

Technische Informatik 2 (Zweifach)

Vorlesung	4 Std.	M. Malek
Übung	4 Std.	S. SOMMER
Übung	2 Std.	J. RICHLING

Vorbereitung, Analyse, Planung von Informatikunterricht (Kernfach)

Seminar	2 Std.	E. LEHMANN
---------	--------	------------

Proseminare (D, M, B.A.)**Softwarespezifikation mit UML**

Proseminar	2 Std.	J. FISCHER
------------	--------	------------

Wissenschaftliches Publizieren mit LATEX

Proseminar	2 Std.	M. PIEFEL
------------	--------	-----------

Wurzeln der Informatik

Proseminar	2 Std.	R. KUBICA
------------	--------	-----------

Informationelle Selbstbestimmung

Proseminar	2 Std.	P. BITTNER
------------	--------	------------

Klassische Probleme und Algorithmen der Bioinformatik

Proseminar	2 Std.	U. LESER
------------	--------	----------

Beauty is our Business

Proseminar	2 Std.	W. REISIG
------------	--------	-----------

Die Grenzen der Berechenbarkeit

Proseminar	2 Std.	M. GROHE
------------	--------	----------

HAUPTSTUDIUM

Kurse/Halbkurse (D, N, L, M)Praktische und angewandte Informatik**Objektorientierte Analyse, Design und Implementation (K)**

Vorlesung	8 Std.	J. FISCHER, K. AHRENS
Praktikum	2 Std.	M. PIEFEL

Mobile Kommunikation (HK)

Vorlesung	4 Std.	J.-P. REDLICH
Praktikum	2 Std.	J. GÖLLER

Prinzipien moderner Betriebssysteme (HK)

Vorlesung	4 Std.	J.-P. REDLICH, J.-P. BELL
Übung	2 Std.	A. ZUBOW

Unix Development Tools und UNIX Systemadministration (HK-2. Teil)

Vorlesung	2 Std.	J.-P. BELL
-----------	--------	------------

Security-Engineering - Sichere Systeme (HK)

Vorlesung	4 Std.	W. MÜLLER
Übung	2 Std.	A. ZUBOW

Software-Engineering 1: Software-Entwicklung (HK)

Vorlesung	4 Std.	K. BOTHE
Vorlesung/Übung	2 Std.	K. BOTHE, K. SCHÜTZLER
Übung	4 Std.	K. SCHÜTZLER

Parallelsierende Compiler (HK)

Vorlesung	4 Std.	R. HÄNISCH, K. BOTHE
-----------	--------	----------------------

Informationstechnik für Geschäftsprozesse (HK-2. Teil)

Vorlesung	BLOCK	J.-M. HASEMANN
-----------	-------	----------------

Konzepte für die frühen Phasen der Softwareentwicklung (HK)

Vorlesung	4 Std.	CH. DAHME
-----------	--------	-----------

DBS II: Implementierung vom Datenbanksystemen (HK)

Vorlesung	4 Std.	U. LESER
Praktikum		HEIKO MÜLLER, FRANK HUBER

Verteilte Systeme: Konzepte und Basisalgorithmen (HK)

Vorlesung	2 Std.	A. REINEFELD, A. ANDRZEJAK
Übung	2 Std.	A. REINEFELD, A. ANDRZEJAK
Projekt	BLOCK	A. REINEFELD, A. ANDRZEJAK

Maschinelle Sprachverarbeitung (HK)

Vorlesung	4 Std.	T. SCHEFFER
Übung	2 Std.	T. SCHEFFER

Werkzeuge der empirischen Forschung (HK)

Vorlesung	4 Std.	W. KÖSSLER
Übung	2 Std.	W. KÖSSLER
Praktikum		W. KÖSSLER

Methoden und Modelle des Systementwurfs

Vorlesung	4 Std.	W. REISIG
Übung	2 Std.	W. REISIG

Theoretische Informatik**Parametrische Algorithmen und Komplexitätstheorie (HK)**

Vorlesung	4 Std.	M. GROHE
Übung	2 Std.	S. KREUTZER

Logik und Komplexität (HK)

Vorlesung	4 Std.	N. SCHWEIKARDT, S. KREUTZER
Übung	2 Std.	N. SCHWEIKARDT, S. KREUTZER

Lineare Optimierung (HK – auch Math. Erg.fach)

Vorlesung	4 Std.	L. POPOVA-ZEUGMANN
Übung	2 Std.	L. POPOVA-ZEUGMANN

Graphen und Algorithmen 2 (HK)

Vorlesung	4 Std.	A. COJA-OGHLAN
Übung	2 Std.	M. BODIRSKY

Approximationsalgorithmen

Vorlesung	4 Std.	A. COJA-OGHLAN
-----------	--------	----------------

Graphen in der Ebene (HK)

Vorlesung	4 Std.	D. SCHLATTER, M. SCHACHT
Übung	2 Std.	D. SCHLATTER, M. SCHACHT

Algorithmen im Drug Design (HK)

Vorlesung	2 Std.	S. HOUGARDY
-----------	--------	-------------

Combinatorics and its Applications

Vorlesung	4 Std.	M. KANG
-----------	--------	---------

Schaltkreiskomplexität (HK)

Vorlesung	4 Std.	J. KÖBLER
Übung	2 Std.	J. KÖBLER

OpenSSL – Kryptologie in C (HK-2. Teil)

Vorlesung	2 Std.	E. G. GIESSMANN
-----------	--------	-----------------

Technische Informatik**Grundlagen der Signalverarbeitung (HK)**

Vorlesung	2 Std.	B. MEFFERT
Übung	4 Std.	O. HOCHMUTH
Praktikum	2 Std.	U. KNAUER

Signalverarbeitung (HK)

Vorlesung	2 Std.	B. MEFFERT
Übung	2 Std.	O. HOCHMUTH
Praktikum	n. V.	O. HOCHMUTH

Stereo-Bildverarbeitung (HK-2. Teil)

Vorlesung	2 Std.	R. REULKE
-----------	--------	-----------

Prozessinformatik (HK – auch Praktische Informatik)

Vorlesung	2 Std.	M. RITZSCHKE
Übung	2 Std.	K.-H. HAUPTVOGEL
Übung	2 Std.	M. RITZSCHKE
Praktikum		M. RITZSCHKE, K.-H. HAUPTVOGEL

Spezielle Techniken der Rechnerkommunikation (HK)

Vorlesung	2 Std.	S. SOMMER
Projekt	2. Std.	S. SOMMER

Mathematisches Ergänzungsfach (D)**Lineare Optimierung (auch HK Theoretische Informatik)**

Vorlesung	4 Std.	L. POPOVA-ZEUGMANN
Übung	2 Std.	L. POPOVA-ZEUGMANN

Pflichtfach (L)**Multimedia in der Schule**

Seminar	2 Std.	J. KOUBEK
---------	--------	-----------

Informatik und Gesellschaft

Vorlesung	2 Std.	J. KOUBEK
Projekt	n. V.	J. KOUBEK

Spezialvorlesungen (D, M)**Data Warehouses**

Vorlesung	2 Std.	U. LESER
-----------	--------	----------

XML und Datenbanken

Vorlesung	2 Std.	F. NAUMANN
-----------	--------	------------

Stochastische Methoden der globalen Optimierung

Vorlesung	2 Std.	E. RÖDEL
Übung	2 Std.	E. RÖDEL
Projekt	2 Std.	E. RÖDEL

Seminare (D, M, L)*Praktische und Angewandte Informatik***Metamodellierung und Spezifikationsprachen**

Seminar	2 Std.	J. FISCHER
---------	--------	------------

Simulation von Rechnernetzen mit ns-2

Seminar	2 Std.	J. FISCHER
---------	--------	------------

Compilergeneratoren

Seminar/ Proseminar	2 Std.	A. KUNERT
---------------------	--------	-----------

Selbstorganisierte Middleware für mobile Systeme

Projekt	2 Std.	J.-P. REDLICH
---------	--------	---------------

Softwaresanierung für Neueinsteiger

Seminar	2 Std.	K. BOTHE
---------	--------	----------

Requirements Engineering für Entwicklungs- und Reproduktionsprozesse

Seminar	2 Std.	K. SCHÜTZLER
---------	--------	--------------

Modelle für Wahrnehmung

Seminar	2 Std.	H.-D. BURKHARD
---------	--------	----------------

Ressourcen-Management im Cluster und Grid

Seminar	BLOCK	A. REINEFELD, T. RÖBLITZ
---------	-------	--------------------------

Internet und Politik

Seminar	BLOCK	P. BITTNER
---------	-------	------------

Geschichte der Kryptologie

Seminar	2 Std.	R. KUBICA
---------	--------	-----------

Leben und Werk John von Neumanns

Projekt	2 Std.	R. KUBICIA
---------	--------	------------

Fortgeschrittene algorithmische Bioinformatik

Seminar	BLOCK	2 Std.	U. LESER
---------	-------	--------	----------

Syntaxanalyse von Genomdaten

Seminar	BLOCK	2 Std.	S. HEYMANN
---------	-------	--------	------------

Theoretische Informatik**Systementwurf**

Seminar	2 Std.	W. REISIG
---------	--------	-----------

Algorithmische Spieltheorie und Grundlagen der Internetökonomie

Seminar	2 Std.	M. GROHE
---------	--------	----------

Komplexität der Kryptologie

Seminar	2 Std.	J. KÖBLER
---------	--------	-----------

Elektronische Bezahlssysteme

Seminar	2 Std.	J. KÖBLER, M. SCHWAN
---------	--------	----------------------

Stochastische Methoden der Globalen Optimierung

Projekt	2 Std.	E. RÖDEL
---------	--------	----------

Technische Informatik**Biometrie im Kontext**

Seminar	2 Std.	P. BITTNER, B. MEFFERT
---------	--------	------------------------

Spezialgebiete der Signalverarbeitung

Projekt 2 Std. B. MEFFERT

Parallele Signalverarbeitung mit FPGA

Projekt 2 Std. F. WINKLER

Programmierung von digitalen Signalprozessoren

Projekt 2 Std. A. WEIB

NOMADS – Networks of Mobile Adaptive Dependable Systems

Seminar 2 Std. M. MALEK, P. IBACH

CERO – CE Robots Community

Seminar 2 Std. M. MALEK, P. IBACH, J. RICHLING

Spezielle Probleme von Echtzeitsystemen

Seminar 2 Std. J. RICHLING

Studium generale**Ringvorlesung**

Vorlesung 2 Std. N.N.

Forschungsseminare/Kolloquium**Forschungsseminare, 2stündig**

Geleitet von: K. BOTHE, H.-D. BURKHARD, W. COY, J. FISCHER, J.-C. FREYTAG, M. GROHE, J. KÖBLER, U. LESER, M. MALEK, B. MEFFERT, F. NAUMANN, H.-J. PRÖMEL, A. REINEFELD, W. REISIG, E. RÖDEL, T. SCHEFFER

Oberseminar Theoretische Informatik

Kolloquium A. COJA-OGLAN, M. GROHE, J. KÖBLER,

WINTERSEMESTER 2005/2006

GRUNDSTUDIUM

Diplomstudiengang (D)

1. Semester

Praktische Informatik 1

Vorlesung	4 Std.	K. BOTHE
Übung	4 Std.	G. LINDEMANN-V. TRZEBIATOWSKI,
Übung	4 Std.	P. MASSUTHE,
Übung	2 Std	A. ZUBOW
Übung	2 Std	K. SCHÜTZLER
Praktikum		K. AHRENS

Theoretische Informatik 1

Vorlesung	4 Std.	A. GROHE
Übung	8 Std.	L. POPOVA-ZEUGMANN
Übung	6 Std.	M. GROHE, M. GRÜBER, S. KREUTZER, L. POPOVA-ZEUGMANN, N. SCHWEIKARDT, M. WEYER
Übung	2 Std.	M. GROHE

Mathematik 1

Vorlesung	4 Std.	E. HERRMANN
Übung	4 Std.	E. HERRMANN
Übung	6 Std.	H.J. LANGE

3. Semester

Praktische Informatik 3

Vorlesung	4 Std.	J.-C. FREYTAG
Praktikum	6 Std.	A. KUNERT
Praktikum		R. ECKSTEIN
Praktikum		R. HEESE

Theoretische Informatik 2

Vorlesung	4 Std.	A. COJA-OGHLAN
Übung	8 Std.	M. BODIRSKY
Übung	4 Std.	M. BODIRSKY

Technische Informatik 1

Vorlesung	2 Std.	F. WINKLER
-----------	--------	------------

Übung	2 Std.	F. WINKLER
Übung	2 Std.	K.H. HAUPTVOOGEL, M. RITZSCHKE
Praktikum	2 Std.	M. GÜNTHER, K.-H. HAUPTVOGEL, M. RITZSCHKE

Mathematik 3

Vorlesung	3 Std.	A. GRIEWANK
Übung	4 Std.	H.-D. NIEPAGE
Übung	2 Std.	J. RIEHME

Magister – 2. Hauptfach (M)

Praktische Informatik 1 (M, N)

Vorlesung	4 Std.	T. SCHEFFER
Übung	4 Std.	W. KÖSSLER
Übung	2 Std.	M. KURTH
Praktikum		M. RITZSCHKE

Theoretische Informatik 1 (M)

Vorlesung	4 Std.	M. GROHE
Übung	8 Std.	L. POPOVA-ZEUGMANN
Übung	4 Std.	M. GRÜBER
Übung	2 Std.	M. GROHE
Übung	2 Std.	M. WEYER

Mathematik 1 (M)

Vorlesung	4 Std.	E. HERRMANN
Übung	6 Std.	H.-J. LANGE
Übung	4 Std.	E. HERRMANN

Magister – Nebenfach (N)

Praktische Informatik 1

Vorlesung	4 Std.	T. SCHEFFER
Übung	4 Std.	W. KÖSSLER
Übung	2 Std.	M. KURTH
Praktikum		M. RITZSCHKE

Einführung in die Technische Informatik (Rechnerorganisation/Betriebssysteme)

Vorlesung	2 Std.	K. WOLTER
Übung	2 Std.	S. SOMMER

Lehramt (L)

Praktische Informatik 1

Vorlesung	4 Std.	T. SCHEFFER
Übung	4 Std.	W. KÖSSLER
Übung	2 Std.	M. KURTH
Praktikum		M. RITZSCHKE

Einführung in die Technische Informatik (Rechnerorganisation/Betriebssysteme)

Vorlesung	2 Std.	K. WOLTER
Übung	2 Std.	S. SOMMER
Praktikum/Projekt	2 Std.	K. WOLTER

Einführung in die Fachdidaktik

Seminar	2 Std.	W. COY
---------	--------	--------

Unterrichtspraktikum/Blockpraktikum

Praktikum	BLOCK	CH. DAHME
-----------	-------	-----------

Bachelor-Kombinationsstudiengang – Kernfach und Zweitfach (B.A.)

1. Fachsemester

Praktische Informatik 1 (Kernfach)

Vorlesung	4 Std.	T. SCHEFFER
Übung	4 Std.	W. KÖSSLER
Übung	4 Std.	M. KURTH
Praktikum	n. V.	M. RITZSCHKE

Mathematik 1 / Analysis I (Kernfach)

Vorlesung	4 Std.	B. KUMMER
Übung	4 Std.	R. BÖTTCHER
Übung	4 Std.	L. RECKE
Übung	2 Std.	J. HEERDA
Übung	2 Std.	N.N.

Theoretische Informatik 1 (Zweifach)

Vorlesung	4 Std.	M. GROHE
Übung	8 Std.	L. POPOVA-ZEUGMANN
Übung	4 Std.	M. GRÜBER
Übung	2 Std.	M. GROHE
Übung	2 Std.	M. WEYER

Einführung in die Fachdidaktik (Zweifach)

Seminar	2 Std.	W.COY
---------	--------	-------

3. Fachsemester**Theoretische Informatik 1 (Kernfach)**

Vorlesung	4 Std.	M. GROHE
Übung	8 Std.	L. POPOVA-ZEUGMANN
Übung	4 Std.	M. GRÜBER
Übung	2 Std.	M. GROHE
Übung	2 Std.	M. WEYER

Theoretische Informatik 2 (Kernfach)

Vorlesung	4 Std.	A. COJA-OGHLAN
Übung	8 Std.	N.N.
Übung	2 Std.	M. BODIRSKY

Mathematik für InformatikerInnen 2 (Kernfach)

Vorlesung	4 Std.	H. KURKE
Übung	4 Std.	L. WOTZLAW
Übung	4 Std.	N.N.
Übung	2 Std.	M. GRABITZ
Übung	2 Std.	J. LEITERER

Einführung in die Fachdidaktik (Kernfach)

Seminar	2 Std.	W. COY
---------	--------	--------

Praktische Informatik (Zweifach)

Vorlesung	4 Std.	T. SCHEFFER
Übung	4 Std.	W. KÖSSLER
Übung	4 Std.	M. KURTH
Praktikum	n. V.	M. RITZSCHKE

Proseminare (D, M, B.A.)**Beauty is our Business**

Proseminar	2 Std.	W. REISIG
------------	--------	-----------

Urheberrecht zwischen digitaler Rechtekontrolle und Wissensfreiheiten

Proseminar	2 Std.	C. KURZ
------------	--------	---------

The Weakest Link – Sicherheit von Online –Systemen

Proseminar	2 Std.	J.-M. LOEBEL
------------	--------	--------------

Mit Mathematica spielend durch die Informatik

Proseminar 2 Std. K.-P. Neuendorf

Kryptografische Algorithmen und Protokolle

Proseminar 2 Std. J. KÖBLER, O. BEYERSDORFF

Fakultative Lehrveranstaltungen (D, M, N, L, B)**UNIX Crash Kurs**

Vorlesung BLOCK J. BELL

Einführung in C- als Vorveranstaltung zu Praktische Informatik 3

Vorlesung BLOCK J. C. FREYTAG

HAUPTSTUDIUM

Kurse/Halbkurse (D, N, L, M)**Praktische und angewandte Informatik****Objektorientierte Modellierung, Spezifikation und Implementation (HK)**

Vorlesung	2 Std.	J. FISCHER
Vorlesung	2 Std.	K. AHRENS
Praktikum	n. V.	M. PIEFEL

Middleware-Plattformen (HK)

Vorlesung	4 Std.	J.-P. REDLICH
Praktikum	2 Std.	M. KURTH
Praktikum	2 Std.	A. ZUBOW

Informatik und Informationsgesellschaft I: Digitale Medien (HK)

Vorlesung	4 Std.	W. COY
Übung/Praktikum	2 Std.	J. KOUBEK

Kooperatives Prototyping (HK)

Vorlesung	4 Std.	CH. DAHME
-----------	--------	-----------

Einführung in die Datenbanken/DBS I (HK)

Vorlesung	4 Std.	J.-C. FREYTAG
Praktikum	6 Std.	N.N.

Algorithmische Bioinformatik (HK)

Vorlesung	4 Std.	U. LESER
Übung	2 Std.	J. HAKENBERG, J. BAUCKMANN

Informationsintegration (HK)

Vorlesung	4 Std.	F. NAUMANN
Praktikum	2 Std.	F. NAUMANN

Einführung in die Künstliche Intelligenz

Vorlesung	4 Std.	H.-D. BURKHARD
Übung	4 Std.	D. GÖHRING

Kognitive Robotik (HK)

Vorlesung	4 Std.	H.-D. BURKHARD
Übung	2 Std.	H.-D. BURKHARD
Praktikum		H.-D. BURKHARD

Software Engineering II: Eingebettete Systeme, Software-Qualität, Projektmanagement (HK)

Vorlesung 4 Std. H. SCHLINGLOFF

Theoretische Informatik

Verteilte Algorithmen (HK)

Vorlesung 4 Std. W. REISIG
 Übung 2 Std. W. REISIG

Computergestützte Verifikation (HK)

Vorlesung 4 Std. K. SCHMIDT
 Übung 2 Std. P. MASSUTHE

Logik in der Informatik

Vorlesung 4 Std. M. GROHE
 Übung 2 Std. M. WEYER

SAT-Solving und Constraint Satisfaction Probleme (HK)

Vorlesung 4 Std. S. KREUTZER
 Übung 2 Std. S. KREUTZER

Graphen und Algorithmen (HK oder K – 1. Teil)

Vorlesung 4 Std. S. HOUGARDY
 Übung 2 Std. S. HOUGARDY
 Praktikum M. BEHRSCHE

Randomisierte Algorithmen und Probabilistische Analyse (HK)

Vorlesung 4 Std. M. KANG

Kryptologie I (HK) 4 Std.

Vorlesung 4 Std. J. KÖBLER
 Übung 2 Std. O. BEYERDORFF

Technische Informatik

Grundlagen der Signalverarbeitung (HK)

Vorlesung 4 Std. B. MEFFERT
 Übung 4 Std. O. HOCHMUTH
 Praktikum 2 Std. U. KNAUER, M. SALEM

Grundlagen der Bildverarbeitung (HK)

Vorlesung 2 Std. B. MEFFERT
 Übung 4 Std. O. HOCHMUTH
 Praktikum O. HOCHMUTH

Schaltkreisentwurf (HK)

Vorlesung	4 Std.	F. WINKLER
Praktikum		F. WINKLER

Stereo-Bildverarbeitung (HK)

Vorlesung	4 Std.	R. REULKE
Praktikum		R. REULKE

Zuverlässige Systeme (HK)

Vorlesung	2 Std.	J. RICHLING
Vorlesung	2 Std.	K. WOLTER
Praktikum	n. V.	J. RICHLING

Eigenschaften von mobilen und eingebetteten Systemen (HK)

Vorlesung	4 Std.	J. RICHLING
Praktikum	n. V.	J. RICHLING

Grundlagen der Rechnerkommunikation (HK)

Vorlesung	4 Std.	S. SOMMER
Praktikum	2 Std.	S. SOMMER
Praktikum	n.V.	S. SOMMER

Mathematisches Ergänzungsfach (D)**Stochastik für InformatikerInnen**

Vorlesung	4 Std.	W. KÖSSLER
Übung	4 Std.	W. KÖSSLER

Pflichtfach (L)**Hauptseminar Fachdidaktik**

Seminar	2 Std.	C. KURZ
---------	--------	---------

Spezialvorlesungen (D, M)**Softwareentwicklung mit UML 2**

Vorlesung	2 Std.	J. FISCHER
-----------	--------	------------

Informatische Bildung

Vorlesung	2 Std.	J. KOUBEK
-----------	--------	-----------

Datenschutz für Informatiker

Vorlesung	2 Std.	P. BITTNER
Übung/Praktikum	2 Std.	P. BITTNER

Data Mining – Algorithmen und Anwendungen

Vorlesung	2 Std.	A. ANDRZEJAK
-----------	--------	--------------

Sichere Systemadministration

Vorlesung	2 Std.	J. BELL
Praktikum	2 Std.	W. MÜLLER

Seminare (D, M, L)Praktische Informatik**Protokollsimulation selbstorganisierender Systeme**

Seminar	2 Std.	J. FISCHER
---------	--------	------------

ULF (Unified Language Family)

Projekt	2 Std.	J. FISCHER
---------	--------	------------

Metamodellierung für Modellgetriebene Softwareentwicklung mit MDA und UML

Seminar	2 Std	M. PIEFEL
---------	-------	-----------

IT-Sicherheit

Seminar/Praktikum	BLOCK.	J.-P. REDLICH
-------------------	--------	---------------

Selbst-organiesierende Middleware für mobile Systeme

Seminar/Projekt	2 Std.	J.-P. REDLICH
-----------------	--------	---------------

Peer-to-Peer-Netze

Seminar	2 Std.	J.-P. REDLICH
---------	--------	---------------

Sicherheit in drahtlosen Netzen

Seminar	2 Std.	W. MÜLLER
---------	--------	-----------

Softwaresanierung für Neueinsteiger

Projekt	2 Std.	K. BOTHE
---------	--------	----------

Requirements Engineering für Entwicklungs- und Produktionsprozesse

Projekt	2 Std.	K. SCHÜTZLER
---------	--------	--------------

eHealth – Verdatung im Gesundheitswesen

Seminar	2 Std.	P. BITTNER
---------	--------	------------

Gesellschaftliche, technische und juristische Perpektiven von Filesharing-Netzen

Seminar	2 Std.	C. KURZ
---------	--------	---------

Der Fehler im System

Seminar 2 Std. R. KUBICA

3D-Grafik

Seminar BLOCK R. KUBICA

Graphmanagement in Datenbanken

Seminar BLOCK U. LESER

Rekonstruktion von Sprachstambäumen (interdisziplinär)

Seminar BLOCK U. LESER, A. LÜDELING, K. DONHAUSER

Text Mining und Anwendung

Seminar BLOCK T. SCHEFFER

Neuronale Netze und Autonome Roboter

Seminar 2 Std. M. HILD

Künstliche Neuronale Netze

Seminar 2 Std. G. KOCK

Parallele Algorithmen in der Bioinformatik

Seminar 2 Std. A. REINEFELD, T. STEINKE

FPGA-Programmierung für die Bioinformatik auf der Cray XD1

Seminar 2 Std. A. REINEFELD, T. STEINKE

Statische Programmanalyse

Seminar 2 Std. P. MASSUTHE

Theoretische Informatik**Systementwurf**

Seminar 2 Std. W. REISIG

Kommunikationskomplexität

Seminar 2 Std. N. SCHWEIKHARDT

Endliche Modelltheorie

Seminar BLOCK S. KREUTZER, N. SCHWEIKHARDT

Komplexität und Kryptologie

Seminar 2 Std. J. KÖBLER, O. BEYERSDORFF

Extremale Kobinatorik

Seminar 2 Std. M. SCHACHT

Interoperabilität und Sicherheit

Seminar 2 Std. E. G. GIESSMANN

Technische Informatik**Spezialgebiete der Signalverarbeitung**

Projekt 2 Std. B. MEFFERT

Biosignalanalyse

Projekt n. V. B. MEFFERT

Computermusik

Seminar 2 Std. M. HILD

Kolloquien**Oberseminar Theoretische Informatik**

Kolloquium A. COJA-OGLAN, M. GROHE, J. KÖBLER,
S. KREUTZER, N. SCHWEIKHARDT

Forschungsseminare**Forschungsseminare, 2stündig**

Geleitet von: K. BOTHE, H.-D. BURKHARD, W. COY, J. FISCHER, J.-C. FREYTAG, M. GROHE,
S. KREUTZER, J. KÖBLER, U. LESER, M. MALEK, B. MEFFERT, F. NAUMANN, H.-J. PRÖMEL,
J.-P. REDLICH, A. REINEFELD, W. REISIG, T. SCHEFFER, N. SCHWEIKHARDT