

Forschungsbericht 2005

Institut für Wissens- und Sprachverarbeitung



Otto-von-Guericke-Universität Magdeburg

Fakultät für Informatik

Institut für Wissens- und Sprachverarbeitung

Universitätsplatz 2, 39106 Magdeburg
Tel. +49 (0)391 67 18343, Fax +49 (0)391 67 12018
office@iws.cs.uni-magdeburg.de

1. Leitung

Prof. Dr. rer. nat. habil. Dietmar Rösner (geschäftsführender Leiter)

2. Hochschullehrer

Prof. Dr. rer. nat. habil. Jürgen Dassow
Prof. Dr. rer. nat. habil. Rudolf Kruse
Prof. Dr. rer. nat. habil. Dietmar Rösner
Prof. Dr. Andreas Nürnberger

3. Forschungsprofil

1. Wissensbasierte Systeme und Dokumentverarbeitung

- Teilprojekt Integration von Werkzeugen zur Wissensakquisition der DFG-Forschergruppe Informationsfusion
- Promotionsvorhaben Informationsextraktion aus wissenschaftlichen Abstracts
- XML-Technologie für Lehr-/Lernsysteme
- Aufbereitung von Dokumenten für das Semantic Web

2. Fuzzy-Systeme und Neuronale Netze

- Neuro-Fuzzy-Systeme
- Information Mining
- Fuzzy-Daten-Analyse

3. Formale Sprachen und Automaten

- Grammatiken mit gesteuerter Ableitung
- Kompetenz-basierte Kooperation in Grammatiksystemen
- Syntaktische Komplexität von Lindenmayer-Systemen
- Weitere Sprachklassen
- Spezifikation von XML-Dokumenten

4. Wissensentdeckung und Maschinelles Lernen

- Aktives Lernen
- Boosting-Ansätze für ILP-Verfahren
- ILP Lernverfahren für Spiele
- Transformierende ILP-Verfahren für große Datenbestände

4. Forschungsprojekte

Projektleiter: Prof. Dr. Rudolf Kruse
Projektbearbeiter: Christian Borgelt, Christian Döring
Förderer: Industrie; 01.01.2004 - 31.03.2005

Advanced Intelligent Data Analysis

Die Intelligent Systems Research Unit der British Telecom Research Laboratories (BT Labs) befasst sich mit der Entwicklung intelligenter Lösungen im Bereich der Telekommunikation. In dem Projekt werden gemeinsam verschiedene Datenanalyseprobleme bearbeitet. Unter anderem geht es um die Vorhersage der Dauer von Wartungsarbeiten, um die Wege der Techniker optimal planen zu können. Hierfür wurden statistische Verfahren und verschiedene Soft-Computing-Verfahren eingesetzt, u.a. neuronale Netze und Entscheidungs bzw. Regressionsbäume. Daneben wurde eine Software entwickelt, die interaktive, auf die vorliegenden Fragestellungen spezialisierte Datenanalysen ermöglicht.

Projektleiter: Prof. Dr. Rudolf Kruse
Förderer: Industrie; 01.06.2004 - 31.12.2005

Datanalyse von Finanz- und speziell Kreditwürdigkeitsdaten

Die Erstellung qualitativ hochwertiger Vorhersagen zur Kreditwürdigkeit zählt zu den bedeutenden Problemen im Bereich des Finanzwesens. Das Ziel dieses Projektes besteht in der Entwicklung von Werkzeugen, die eine umfassende Analyse und Auswertung der verfügbaren Datenquellen sowie die für Vorhersagen notwendige Modellgenerierung ermöglichen.

Im Rahmen dieses Projektes wird von der Arbeitsgruppe ein Datenanalysetool bereitgestellt und durch die Entwicklung zusätzlicher Module entsprechend der speziellen Anforderungen im gegebenen Problemfeld erweitert.

Projektleiter: Prof. Dr. Rudolf Kruse
Projektbearbeiter: Xiaomeng Wang
Förderer: Land (Sachsen-Anhalt); 01.07.2003 - 30.06.2005

Datanalyse in Kommunikationssystemen

In der Telekommunikation, speziell in Mobilfunknetzen, entsteht eine große Menge von Daten, die Alarme und Fehler beschreiben. Die effiziente Auswertung und Anzeige dieser Alarme ist eine wichtige Aufgabe, da die Sendestationen in Mobilfunknetzen natürlich an vielen Orten verteilt sind und es daher mit einigem Aufwand verbunden ist, einen Techniker zur Überprüfung und Wartung vor Ort zu bringen. Insbesondere müssen mehrere Alarme zusammengefasst werden, da jede Einheit auch dann einen Alarm auslöst, wenn eine ihrer Teileinheiten einen Alarm auslöst. Die Schwierigkeit besteht dabei darin, dass durch unterschiedliche Signallaufzeiten und ausfallende Übertragungsstrecken die Alarme in unterschiedlicher Reihenfolge und unvollständig ankommen. In diesem Projekt wird versucht über Clusteringverfahren und die Erzeugung unscharfer Regeln, Vorschriften für die automatische Zusammenfassung von Alarmen zu finden, die z.B. ... [mehr](#)

Projektleiter: Prof. Dr. Rudolf Kruse
Projektbearbeiter: Christian Döring, Frank Rügheimer

Förderer: Haushalt; 01.08.2003 - 01.08.2005

Datenanalyse Plattform "InfoMiner"

Gegenstand des Projektes ist die Fortführung der Ideen aus dem DFG Forschungsvorhaben (KR 521/4-1) und deren Implementierung in Form der flexiblen Data-Mining Plattform "InfoMiner". Die aus Vorarbeiten hervorgegangene Software ist komponentenbasiert und erlaubt in ihrer weiterentwickelten Form die Erstellung von Konfigurationen des Miners mit unterschiedlichem Funktionsumfang je nach Anwendungsfeld. Miner ist ein vertikales System: Die Menge von verfügbaren Datenzugriffs-, Vorverarbeitungs-, und Datenanalyseverfahren kann dem jeweiligen Sachbereich angepasst werden. Aktuelle Forschungsergebnisse auf dem Gebiet der Datenanalyseverfahren sowie den Sachbereichen angepasste Softwarelösungen können als neue Komponenten mit geringem Aufwand in die bestehende Architektur integriert und im Rahmen einer zugeschnittenen Konfiguration des Miners ausgeliefert werden.

Projektleiter: Prof. Dr. Rudolf Kruse

Projektbearbeiter: Christian Borgelt

Kooperationen: Michael Berthold, Altana Lehrstuhl für angewandte Informatik, Universität Konstanz

Förderer: Haushalt; 01.10.2003 - 01.10.2005

Finden relevanter Fragmente in Moleküldatenbanken

Viele Datenanalyseaufgaben in der Bioinformatik bestehen in der Analyse großer Sammlungen von Molekülen mit dem Ziel, Regelmäßigkeiten in den Molekülen einer bestimmten Klasse zu finden. Zum Beispiel möchte man in der Medikamentenentdeckung neue Kandidaten finden, indem man hunderttausende von Moleküle einem Screening unterzieht, um experimentell ihre Aktivität in Bezug auf eine bestimmte Krankheit festzustellen. In diesem Projekt wird versucht, Fragmente (Teilstrukturen) von Molekülen zu finden, die in den Aktiven häufig und in den Inaktiven selten vorkommen. Als Ansatz wird eine Übertragung des Eclat-Algorithmus zum Lernen von Assoziationsregeln auf attributierte ungerichtete Graphen entwickelt.

Projektleiter: Prof. Dr. Rudolf Kruse

Projektbearbeiter: Rudolf Kruse, Christian Borgelt, Frank Rügheimer

Förderer: Industrie; 01.01.2004 - 31.12.2005

Lernen von graphischen Modellen

In diesem Projekt werden Verfahren zur automatischen Gewinnung von Wissen aus Datenbanken (Knowledge Discovery in Databases) erforscht. Schwerpunkte sind das Lernen von probabilistischen (Bayesschen) und possibilistischen Schlussfolgerungsnetzen. Beim Lernen von Schlussfolgerungsnetzen wird eine mehrdimensionale Wahrscheinlichkeits- oder Possibilitätsverteilung in Verteilungen auf niedrigdimensionalen Teilräumen zerlegt, durch die die Gesamtverteilung möglichst gut angenähert wird. Schon eine solche Zerlegung allein kann interessante Aufschlüsse über in den Daten vorhandene Abhängigkeiten geben. Besonders vorteilhaft ist jedoch, dass es mit Hilfe einer solchen Zerlegung möglich ist, Schlussfolgerungen in dem zugrundeliegenden mehrdimensionalen Raum zu ziehen, ohne auf die Gesamtverteilung, die oft nicht im Speicher eines Rechners darstellbar ist, zurückgreifen zu müssen. ... [mehr](#)

Projektleiter: Jun. Prof. Dr. Andreas Nürnberger

Projektbearbeiter: Andreas Nürnberger, Korinna Bade, Ernesto William De Luca

Förderer: DFG; 01.05.2003 - 30.04.2007

Benutzeradaptive Schnittstellen zur Suche und Navigation in Dokumentendatenbanken

Das Ziel des Forschungsprojektes ist die Entwicklung interaktiver, adaptiver Benutzerschnittstellen für die Suche und Navigation in unstrukturierten Datenbanken (Information Retrieval Support Systems; IRS-Systeme). Hierbei sollen einige Hauptprobleme existierender Modelle behoben werden. Dies betrifft vor allem die unzureichende Einbindung von Benutzermodellen als auch die mangelhafte Integration der einzelnen Systemkomponenten in ein Gesamtsystem.

Im Unterschied zur Dokumentensuche ist das Ziel eines IRS-Systems nicht nur die Suche nach Dokumenten, die zu einer Anfrage passen, sondern eine umfassende Unterstützung eines Nutzers bei der Suche nach Informationen in einer Dokumentensammlung. Diese ist in der Regel nicht nach einem Schritt abgeschlossen, sondern besteht aus einem interaktiven Prozess. ... [mehr](#)

Projektleiter: Jun. Prof. Dr. Andreas Nürnberger

Projektbearbeiter: Andreas Nürnberger

Förderer: EU; 01.02.2005 - 31.01.2008

Nature-inspired Smart Information Systems (NiSIS)

NiSIS ist ein Co-ordinated Action (CA) Projekt der EU. Das Projekt hat zum Ziel Forschung an intelligenten Verfahren im Bereich von Informationssystemen zu koordinieren. Ziel ist insbesondere die Betrachtung von Modellen, die auf in der Natur vorkommenden Systemen beruhen oder durch natürliche Systeme motiviert werden. Hierbei werden sowohl natürliche Prozesse auf biologischer Ebene (micro) als auch auf Verhaltensebene (macro) einbezogen. Das Projekt hat 31 Partnerinstitutionen aus 13 Ländern.

Projektleiter: Prof. Dr. Dietmar Rösner

Projektbearbeiter: Dipl.-W-Inf. Mario Amelung

Förderer: Land (Sachsen-Anhalt); 01.05.2003 - 31.03.2005

XML-Technologie zur Unterstützung der Entwicklung und Wiederverwendung von Lehr- und Lernmaterialien

Ziel des Projektes ist, das Potential der XML-Technologie für die Bedürfnisse insbesondere von Lehrenden, die Lehr-/Lernsysteme entwickeln wollen, voll zur Nutzung zu bringen.

Grundgedanke ist dabei, den Lehrenden bei der Entwicklung von Lehr-/Lernmaterialien so zu unterstützen, dass er sich weitgehend auf die Inhalte und die inhaltliche Strukturierung konzentrieren kann und ihm Fragen der Nutzung, der Präsentation, der Interaktion usw. weitgehend abgenommen werden können.

Der methodische Ansatz dafür ist, wie bei XML üblich, die konsequente Trennung von Inhalt und Präsentation und die Annotation von Informationseinheiten mit maschinell interpretierbaren Metadaten. Über lehrgebietsbezogene Begriffssysteme (Ontologien) ergibt sich die Möglichkeit, strukturiert in den Dokumentbeständen mit Lehr- und Lernmaterialien zu navigieren.

Das bisher konventionell entwickelte Material soll für die Wiederverwendung in einem XML-basierten Rahmensystem ausgearbeitet werden. ... [mehr](#)

5. Veröffentlichungen

Originalartikel in internationalen Zeitschriften

Csuhaj-Varjß, Erzs bet (ext.); Dassow, Jürgen

A remark on evolutionary systems.

In: Discrete applied mathematics : DAM [Amsterdam] 146(2005), S. 111 - 115

[Imp.fact.: 0.557]

Dassow, Jürgen

Contextual grammars with subregular choice.

In: Fundamenta informaticae [Amsterdam] 64(2005), Nr. 1-4, S. 109 - 118

[Imp.fact.: 0.075]

Dassow, Jürgen; Holzer, Markus (ext.)

Language families defined by a ciliate bio-operation : hierarchies and decision problems.

In: International journal of foundations of computer science [Singapore] 16(2005), Nr. 4, S. 645 - 662

Klose, Aljoscha; Kruse, Rudolf

Semi-supervised learning in knowledge discovery.

In: Fuzzy sets and systems [Amsterdam] 149(2005), Nr. 1, S. 209 - 233

[Imp.fact.: 0.577]

Originalartikel in zeitschriftenartigen Reihen

Kruse, Peter M. (ext.); Naujoks, Andre (ext.); Rösner, Dietmar; Kunze, Manuela

Clever search : a WordNet based Wrapper for internet search engines.

In: Fisseni, Bernhard (Hrsg.) ; ... (Hrsg.): Sprachtechnologie, mobile Kommunikation und linguistische Ressourcen : Beiträge zur GLDV-Tagung 2005 in Bonn. Frankfurt am Main : Lang, 2005, S. 367 - 380 (Sprache, Sprechen, Computer 8)

Rösner, Dietmar; Amelung, Mario; Piotrowski, Michael

LlsChecker, ein CAA-System für die Lehre im Bereich Programmiersprachen.

In: Haake, Jörg M. (Hrsg.) ; Lucke, Ulrike (Hrsg.) ; Tavangarian, Djamshid (Hrsg.): DeLFI 2005 (3. Deutsche e-Learning Fachtagung Informatik der Gesellschaft für Informatik e. V., Rostock 13.-16. September 2005). - Tagungsband. Bonn : Köllen, 2005, S. 307 - 318 (GI-Edition : Proceeding 66)

Rösner, Dietmar; Piotrowski, Michael

Integration von E-Assessment und Content-Management.

In: Haake, Jörg M. (Hrsg.) ; Lucke, Ulrike (Hrsg.) ; Tavangarian, Djamshid (Hrsg.): DeLFI 2005 (3. Deutsche e-Learning Fachtagung Informatik der Gesellschaft für Informatik e. V., Rostock 13.-16. September 2005). - Tagungsband. Bonn : Köllen, 2005, S. 129 - 140 (GI-Edition : Proceeding 66)

Wissenschaftliche Monographien

Xiao, Chun

Corpus-based approaches for the processing of the English sublanguage of biomedical abstracts. München : Verl. Dr. Hut, 2005, XI, 115 S. - Magdeburg, Univ., Fak. für Informatik, Diss. Zugl.: Magdeburg, Univ., Fak. für Informatik, Diss.

Lehrbücher, Sachbücher

Dassow, Jürgen

Logik für Informatiker. Stuttgart : Teubner, 2005, 177 S

Buchbeiträge (einschließlich Lehrbuchbeiträge)

Kunze, Manuela; Rösner, Dietmar

Transforming business rules into natural language text.

In: Bunt, Harry (Hrsg.) ; Geertzen, Jeroen (Hrsg.) ; Thijse, Elias (Hrsg.): Computational semantics, IWCS- 6 (Sixth international workshop Tilburg, The Netherlands january 12-14, 2005). - proceedings. Tilburg : Univ., 2005, S. 387 - 389 (IWCS-6)

Hochschulschriften

Krogel, Mark-Andre

On propositionalization for knowledge discovery in relational databases. 2005, XIV, 183 S. Magdeburg, Univ., Fak. für Informatik, Diss., 2005

Xiao, Chun

Corpus-based approaches for the processing of the English sublanguage of biomedical abstracts. 2005, XI, 115 S. Magdeburg, Univ., Fak. für Informatik, Diss., 2005 Zugl. unter ISBN 3-89963-231-1