Informatik Forsch. Entw. (1989) 4: 89-107

Informatik
Forschung und Entwicklung

TWRM-TOPOGRAPHIC

Ein wissensbasiertes System zur situationsgerechten Aufbereitung und Präsentation von Textinformation in graphischen Retrievaldialogen

R. Kuhlen, R. Hammwöhner, G. Sonnenberger und U. Thiel

Universität Konstanz, Postfach 5560, D-7750 Konstanz 1

Zusammenfassung. TWRM-TOPOGRAPHIC ist Teil eines neuartigen Informationssystems, das sich auf inhaltsorientierte Repräsentation von Volltexten stützt. Die beiden wesentlichsten Leistungsmerkmale von TWRM-TOPOGRAPHIC sind die *graphische Retrievaldialogführung* über ein flexibles, objektorientiert spezifiziertes *User-Interface-Management-System (UIMS)* und die flexible, situationsgerechte Aufbereitung und Präsentation von Textwissen: Die Dialogführung erlaubt dem Benutzer die direkte Navigation in den auf dem Bildschirm graphisch dargestellten Wissensstrukturen, die Selektion dargestellter Objekte zur Formulierung einer Suchfrage sowie das Wechseln des Abstraktionsniveaus der dargestellten Textinformation. Textwissen wird in unterschiedlichen Abstraktionsstufen präsentiert: von einer sehr generischen Ebene über *Wissensgraphen*, automatisch generierten *Abstracts* mit unterschiedlichen Themenschwerpunkten und variabler Ausführlichkeit bis zur diskursiven Form der *Textpassage*.

Schlüsselwörter: Automatisches Textkondensieren, interaktives graphisches Retrieval, pragmatischer Systemdesign

Abstract. TWRM-TOPOGRAPHIC is part of a novel type of information-systems, which is based on a semantic representation of thematic descriptions of fulltexts. TWRM-TOPOGRAPHIC employs graphical interaction to provide access to the knowledge bases (text and world knowledge) and presents information in a flexible, situation-specific way. The user may navigate directly within the knowledge structures depicted on the screen, select objects in order to formulate a query and vary the abstraction-level of the represented textual information. Textual knowledge is presented on variable layers of abstraction including a taxonomical layer, conceptual graphs, automatically generated abstracts (varying in detail and thematic focal point) and the discursive form of textpassages.

Key words: Automatic text condensation, interactive graphical retrieval, pragmatic system design

CR Subject Classifications: H.3.1, I.2.7, H.3.3, I.3.6, I.2.8

---

1 Entwurfsziele eines neuen Systemtyps: Prototypische Realisierung durch TWRM-TOPOGRAPHIC

Durch die fortschreitende Verwendung moderner Drucklegungstechniken, den vermehrten Einsatz dezentraler Arbeitsplatzrechner, durch *Mail- und Message-Systeme* und Formen des elektronischen Publizierens liegen in allen Bereichen der Fachkommunikation zunehmend mehr Texte maschinenlesbar (und damit maschinenverarbeitbar) vor. Diese fast schon flächendeckende Verbreitung elektronischer Volltexte korrespondiert aber keinesfalls mit entsprechend leistungsstarken Nachweis-, Verarbeitungs-/Aufbereitungs- oder Präsentationstechniken. Die Gefahr besteht, daß potentiell relevante Information in maschinenlesbaren Speichern verschwindet, ohne bekannt zu werden.

Angesichts des großen Aufwandes, Volltexte intellektuell über Indexierung oder durch Referieren inhaltlich zu erschließen, sollen nach Einschätzung vieler Informationspraktiker Volltext-Retrieval-Systeme die Aufgabe des Relevanznachweises über-

nehmen, und entsprechend haben Volltextdatenbanken von den zur Zeit ca. 3500 weltweit verfügbaren elektronischen Online-Informationsbanken die größten Zuwachsraten; innerhalb eines halben Jahres ist ihre Zahl von ca. 800 auf heute (4/88) 1100 angestiegen.

Den Fortschritten bei Speicherungstechniken und Zugriffsmöglichkeiten entsprechen aber nicht im gleichen Ausmaß Fortschritte bei der automatisierten Inhaltserschließung oder beim Retrieval. Bislang gibt es kaum eigens für die Verarbeitung von Volltexten entwickelte Retrieval-Systeme. Diese Defizite gelten für Inhaltserschließung und Suchtechniken (das eigentliche Retrieval) gleichermaßen.

1.1 Inhaltserschließung

Zwar werden in manchen Fällen zur Wahrung von Qualitätsstandards intellektuelle Erschließungstechniken unter Verwendung von Klassifikationssystemen oder Thesauri eingesetzt, bei den meisten Volltextsystemen wird aber eine Automatisierung der Inhaltserschließung angestrebt, und zwar dadurch, daß die Volltexte invertiert werden. Dieses Verfahren, bei dem also die Textwörter unverändert als Suchargumente beim Retrieval verwendet werden, kann man allerdings kaum als Inhaltserschließung bezeichnen. Zu wenig kann bei diesem Verfahren der Bandbreite morphologischer, syntaktischer und semantischer Varianten Rechnung getragen werden. Weitergehende Verfahren der *automatischen Indexierung*, soweit sie überhaupt zum Einsatz kommen, sind fast ausschließlich wort- oder partiell satzorientiert und beruhen entsprechend auf morphologischen oder syntaktischen Konzepten und einer schwachen und/oder statistisch basierten Semantik [28]. Textadäquate Analysetechniken kommen so gut wie gar nicht zum Einsatz (vgl. Literaturbericht [12]). Soweit die Inhaltserschließung nicht nur durch Indexate auf die Originaltexte verweist, sondern Textzusammenfassungen liefern soll, ist man auf intellektuell erstellte Referate angewiesen. Die zahlreichen, seit den sechziger Jahren laufenden Experimente des *automatischen Abstracting* auf statistischer und/oder oberflächenlinguistischer Basis (vgl. [21]) haben in keinem Fall zu einem realen Einsatz in Informationssystemen geführt. Die verschiedenen Ansätze aus der Künstlichen Intelligenz, die versuchen, aus Wissensstrukturen Textzusammenfassungen zu erzeugen (vgl. Abschn. 3.1), sind weniger auf eine praktische Anwendung in der Fachkommunikation ausgerichtet, sondern mehr an der Simulation einer kognitiv hochstehenden humanen Leistung interessiert.

1.2 Retrieval

Angesichts der Semantikdefizite bei der Inhaltserschließung realer Informationssysteme können natürlich keine leistungsfähigen Retrievalsysteme entstehen. Der methodische Stand der sechziger Jahre wird insgesamt kaum überschritten (vgl. [28]), so daß Veränderungen auf der Retrieval-Seite entweder formalsyntaktische Verbesserungen (Menütechnik, Ikonengraphik) sind oder aber die Idee der Kontextoperatoren, mit denen Kontext- bzw. Abstandsbedingungen in der Kombination von Fragewörtern definiert werden können, weiter verfeinern (z. B. Berücksichtigung des Satz- oder Absatzkontexts, Lokalisierung der Suchargumente an bevorzugten Stellen, z. B. am Anfang von Absätzen). Entsprechend ist sich die Forschung nach verschiedenen Evaluierungsstudien (z. B. [25]) einig, daß Volltextsysteme in ihrer jetzigen Ausprägung, obgleich sie immer stärkere Verwendung finden, nur ein unzureichendes Mittel der Volltextverarbeitung sind.

Nach unserer Einschätzung sind bisherige *(Volltext-)Informations-Retrieval-Systeme* eher „Konfektionsware" und als solche natürlich auch weiterhin unersetzlich; sollen jedoch „maßgeschneiderte" Systeme entstehen, müssen Inhaltserschließung, Retrieval und die Präsentation der Suchergebnisse mehr auf intelligente Verfahren abgestützt werden. *Informationssysteme* – und dies folgt aus unserem informationswissenschaftlichen Verständnis von Information, nach dem Information lediglich als die Teilmenge von Wissen verstanden werden sollte, die aufgrund konkreter Benutzerbedürfnisse und Bedarfssituationen aktuell gebraucht wird – verdienen erst dann ihren Namen, wenn sie über pragmatische Komponenten verfügen, also in der Lage sind, flexibel, d. h. benutzer- und situationsgerecht zu reagieren. Neben leistungsstarken Semantik-Komponenten und, im Falle von textorientierten Systemen, robusten und auf die Möglichkeiten der Wissensrepräsentation zugeschnittenen Textanalyseverfahren sind also auf der Retrieval-Seite flexible Ableitungs- und Darstellungsformen von Wissens- und Textstrukturen notwendig. Auf unterschiedliche Anforderungen müssen Retrieval-Systeme differenziert reagieren können.

Diese Überlegung, den Zugang zu immer größer werdenden Textmengen durch leistungsstärkere Verfahren offen zu halten, war Ausgangspunkt der seit 1982 durchgeführten Forschungsprojekte TOPIC[1],

[1] TOPIC (*T*ext *O*riented *P*rocedures for *I*nformation Management and *C*ondensation of Expository Texts, Projektträger: GID, Förderungskennzeichen: 10200160) wurde in C entwickelt. Die Software ist portierbar und läuft z. Zt. auf einem Cadmus 9200 unter UNIX™

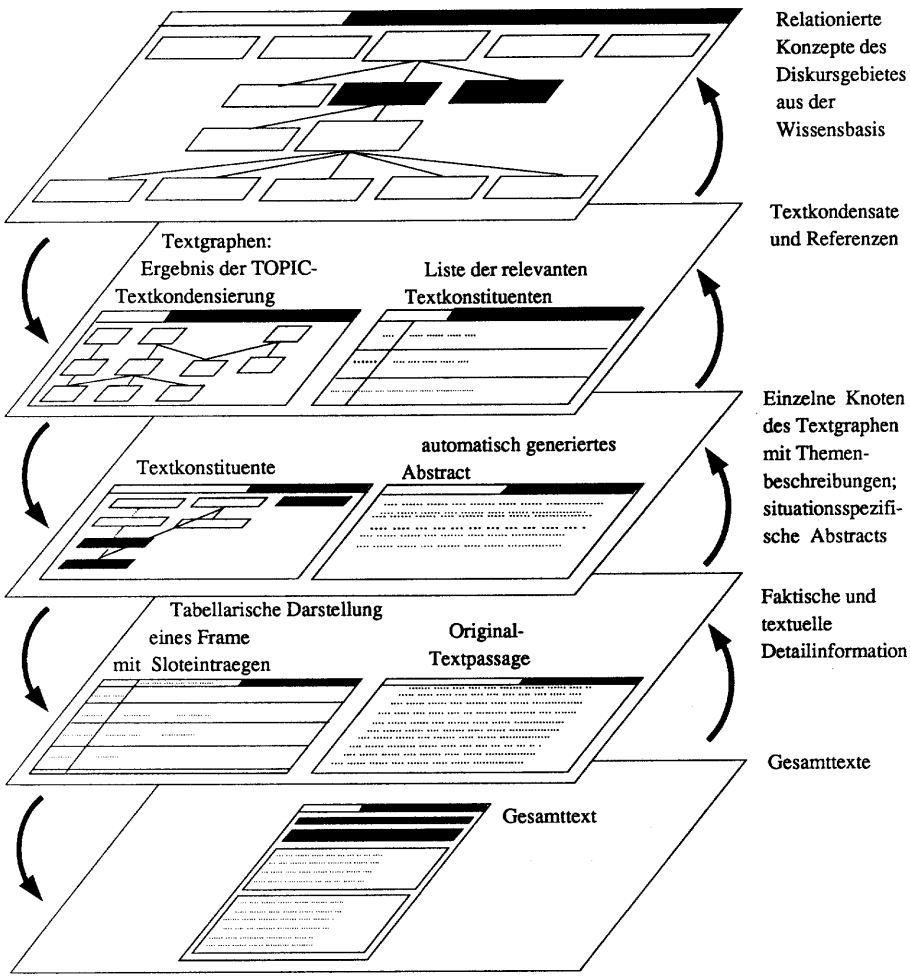


Abb. 1. Stufen der kaskadierten Kondensierung in TWRM-TOPOGRAPHIC

TOPOGRAPHIC[2] und TWRM-TOPOGRAPHIC[3]. Dabei haben wir das zum allgemeinen Gestaltungsprinzip gemacht, was uns die besondere Stärke rechnergestützter Systeme zu sein scheint: die Flexibilität der Verarbeitung und Darstellung. Anstatt primär mit der menschlichen Leistung zu konkurrieren, die in unserem Fall darin besteht, aus einem zehnseitigen Text eine zehnzeilige Zusammenfassung zu erstellen, sollte das Ziel verfolgt werden, aus den bei der Textanalyse erstellten Textwissensstrukturen variable „Kondensate" in unterschiedlichen, sich im Benutzerdialog entwickelnden Kaskaden zu präsentieren. Entsprechend ersetzen wir das Konzept des *automatischen Abstracting* durch das des *kaskadierten Kondensierens*.

Speziell bei dem Entwurf von TWRM-TOPOGRAPHIC sind wir davon ausgegangen, daß die erwünschte Flexibilität wesentlich über graphische Darstellungsformen erreicht werden soll (vgl. [11], [21]), wobei allerdings auch textuell ausgestaltete Kaskadierungsstufen als graphische Objekte aufgefaßt und präsentiert werden. TWRM-TOPOGRAPHIC steht also in der Tradition des graphikorientierten Retrieval. Insgesamt beruht die graphische Ausrichtung der Dialogführung und der Präsentation auf kognitiv-ergonomischen Prinzipien. Das System berücksichtigt die begrenzte Aufnahmekapazität von Benutzern und stellt die Bedeutung der zeitlichen Anordnung von Informationseinheiten für Wahrnehmung und Gedächtnisspeicherung in Rechnung.

Die Idee des kaskadierten Kondensierens wird in TWRM-TOPOGRAPHIC nach dem gegenwärtigen Stand über die folgenden graphisch realisierten Stufen verwirklicht (vgl. Abb. 1):

[2] TOPOGRAPHIC (*TOP*ic *O*perating with *GRAPH*ical *I*nteractive *C*omponents, Projektträger: GID, Förderungskennzeichen: 10200160) ist in C und IF-Prolog implementiert und z. Zt. auf einem Cadmus 9200 unter UNIX™ installiert

[3] TWRM-TOPOGRAPHIC (*T*ext*w*issens*r*eptions*m*echanismus TOPOGRAPHIC, Projektträger: GMD, Förderungskennzeichen: 10200181) ist als eine Erweiterung von TOPOGRAPHIC ebenfalls in C und IF-Prolog programmiert und läuft auf einem Cadmus 9200 unter UNIX™

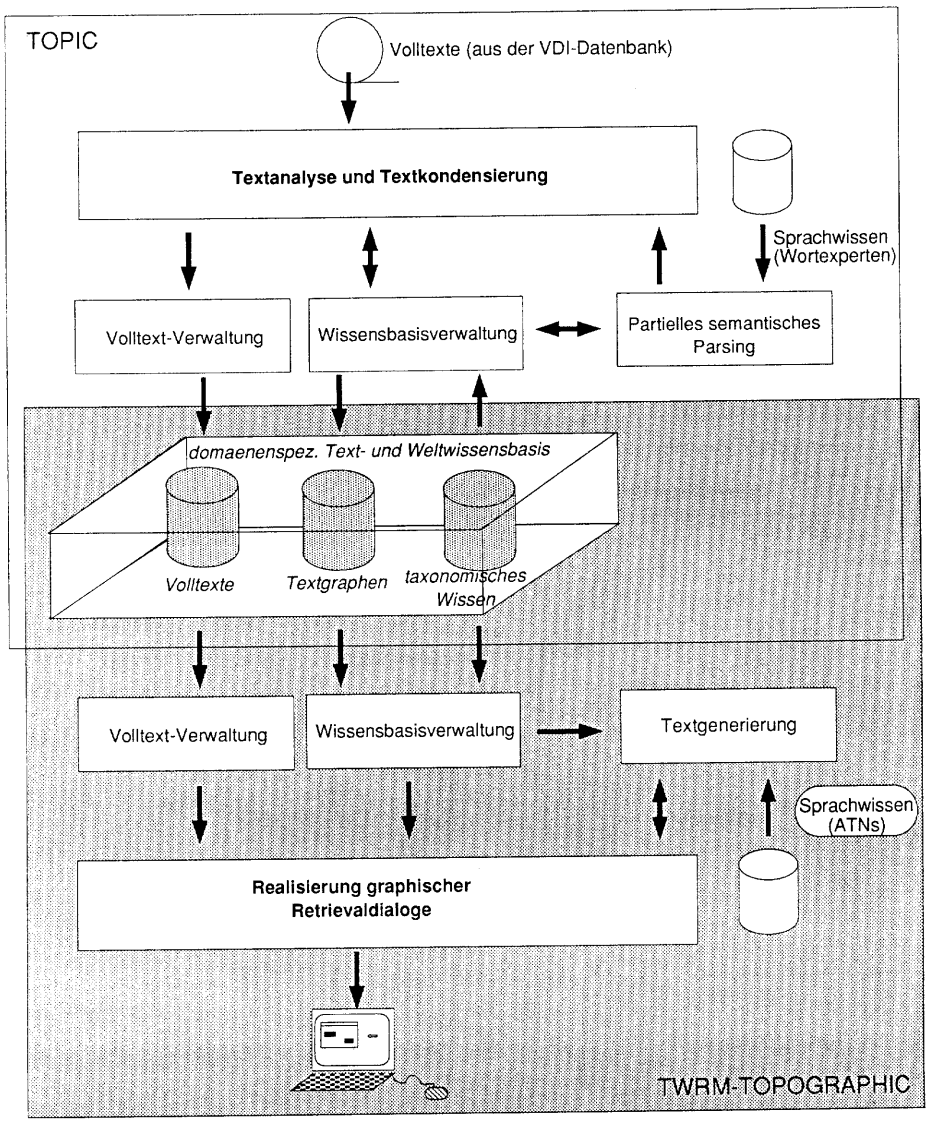


Abb. 2. Architektur des Informationssystems TOPIC/TWRM-TOPOGRAPHIC

• Taxonomische Informationen über den Diskursbereich eines Textes (bzw. einer Textmenge) werden über die durch Relationen verbundenen Konzepte der Weltwissensbasis bereitgestellt. Eine graphische Präsentation der Begriffshierarchie erlaubt eine Auswahl der relevanten Konzepte.
• Die semantische Struktur der jeweiligen Texte wird durch einen Textgraphen dargestellt, der als Ergebnis der TOPIC-Analyse bereitgestellt wird.
• Nach der prinzipiellen Relevanzentscheidung für einen Text wird die graphische Darstellung der thematischen Struktur der Textpassagen zugänglich. Sie erlaubt dem Benutzer,
- die Detailinformation zu den als einschlägig erkannten Konzepten zu betrachten, die in tabellarischer Form angeboten wird, oder
- die entsprechende Passage als Ganzes im Original zu lesen.
• Die in einem Textgraphen vorgegebene thematische und faktische Information bietet zusätzlich die Möglichkeit, nach situationsspezifischer Selektion relevanter Netzteile ein benutzerangepaßtes Abstract, also eine textuelle Kurzfassung, anzubieten.
• Nach Bedarf ist das ‚Blättern' im Gesamttext und die Anzeige von Grafiken im Text möglich.

Mit diesem Beitrag geben wir einen Überblick über die Systemkomponenten und den Leistungsstand des Gesamtsystems. Die Ergebnisse der Textanalyse, basierend auf einem *Frame-Modell* [26] und einem lexikalisch verteilten *Text-Parsing* [13], werden

kurz in Abschnitt 2 zusammengefaßt, allerdings nur insoweit, als es für das Verständnis der Ausgabe- bzw. Interaktionsleistungen erforderlich ist. Das Ergebnis wird - wie erwähnt - in einem konzeptuellen Textgraphen repräsentiert, welcher die Basis für die weitere, hier näher darzustellende Verarbeitung in TWRM-TOPOGRAPHIC ist. Abschnitt 3 stellt die Erweiterung der ursprünglich rein graphisch konzipierten Ausgabe durch einen Abstract-Generator dar, der aus den in den Textkonstituenten enthaltenen Wissensstrukturen natürlichsprachige flexible Zusammenfassungen erzeugt. Abschnitt 4 enthält eine Beschreibung des graphischen Kernstücks von TWRM-TOPOGRAPHIC und zeigt, wie syntaktisch definierte graphische Objekte im Dialog als informationelle Objekte interpretiert werden. Abschnitt 5 diskutiert die kognitiven und ergonomischen Prämissen, die dem Prinzip des kaskadierten Kondensierens und der flexiblen Dialogführung zugrunde liegen und demonstriert die Systemleistung an einem Beispieldialog. Abbildung 2 zeigt den Zusammenhang der wesentlichen Teile der beiden Systeme TOPIC und TWRM-TOPOGRAPHIC [15] sowie die Struktur der weiteren Darstellung.

Seit Beginn der Arbeit an den Projekten ist zumindest in der Forschung von Information-Retrieval-Systemen einiges in Bewegung gekommen. Wir haben den Eindruck, daß die seit ca. 15 Jahren festzustellende konzeptionelle Stagnation beim Entwurf von Volltextsystemen durch intensive Forschungsaktivitäten auf dem Gebiet intelligenter Information-Retrieval-Systeme überwunden werden kann (vgl. [4]). Auch das System TWRM-TOPOGRAPHIC, über das hier in erster Linie berichtet wird, steht im Zusammenhang des Entwurfs intelligenter Information-Retrieval-Systeme.

2 Textgraphen: Ergebnis der TOPIC-Textanalyse/Textkondensierung und Ausgangsbasis für den TWRM-TOPOGRAPHIC Retrievaldialog

TOPIC [14] analysiert deutschsprachige Texte, vollständige Zeitschriftenartikel aus dem Gebiet der Informations- und Kommunikationstechnologie (aus der VDI-Volltextdatenbank) und überführt thematisch zusammenhängende Textabschnitte in eine Themenbeschreibung, im weiteren auch Textkonstituente genannt. Eine solche Themenbeschreibung ist als hierarchisches Netz aufgebaut, dessen Knoten Frames und, je nach Spezifität, auch *Slots* und *Slot-Entries* zugeordnet sind. Diese repräsentieren die thematisch relevanten Konzepte eines Textabschnitts, deren semantische Beziehungen durch sie verbindende Kanten aufgezeigt werden. Die auftretenden *Frames* können durch die Ober-/Unterbegriff- und Prototyp/Instanz-Beziehung relationiert sein. Ausgehend von derartigen Textkonstituenten wird durch Ableitung weiterer konzeptueller Graphen, die in verallgemeinerter Form die Gemeinsamkeiten der beteiligten Textkonstituenten beschreiben, ein sogenannter Textgraph (das Textkondensat) gebildet, dessen Knoten die Textkonstituenten zugeordnet sind. Die am höchsten liegenden Knoten des Textgraphen enthalten die generalisiertesten und die Blattknoten die spezifischsten Themenbeschreibungen. Die Kanten des Textgraphen zeigen die Abstraktionsbeziehungen an, die zwischen den Textkonstituenten existieren. Folgende vier Kantentypen sind zwischen den Knoten und den ihnen zugeordneten Konstituenten definiert:

Is-a/Instance-Relation:
Eine Kante dieses Typs zwischen einem Knoten n und einem tiefer liegenden Knoten n′ zeigt an, daß die Textkonstituente des Knotens n′ eine Beschreibung eines Unterbegriffs (Is-a-Relation) bzw. einer Instanz (Instance-Relation) des in diesem Fall einzigen Konzepts der Themenbeschreibung des Knotens n enthält.

Slot/Slot-Entry-Relation:
Eine derartige Kante zwischen einem Knoten n und einem tieferen Knoten n′ signalisiert, daß die Themenbeschreibung des Knotens n in derjenigen des Knotens n′ enthalten ist und dessen Beschreibung durch einen zusätzlichen *Slot* bzw. *Slot-Entry* näher spezifiziert ist.

Zur weiteren Erläuterung soll Abb. 3 dienen, die einen Textgraphen-Ausschnitt zeigt, bestehend aus zwei Textkonstituenten, die jeweils einen thematisch zusammenhängenden Textabschnitt beschreiben, und zwei abgeleiteten Textkonstituenten, die die Gemeinsamkeiten der ihnen zugrunde liegenden Textkonstituenten in verallgemeinerter Form wiedergeben.

Die TOPIC-Textgraphen sind zusammen mit dem Volltext in einer Textwissensbasis abgelegt und bilden die Ausgangsbasis sowohl für das graphisch-interaktive Retrieval - aufgrund der Ähnlichkeit zwischen der Suchfrage des Benutzers und der Textrepräsentation werden diejenigen Repräsentationen ausgewählt, deren Textinhalte geeignet scheinen, die Suchfrage zu beantworten - als auch für die situationsspezifische Präsentation von Textwissen auf der vom Benutzer gewählten Kaskadierungsstufe.

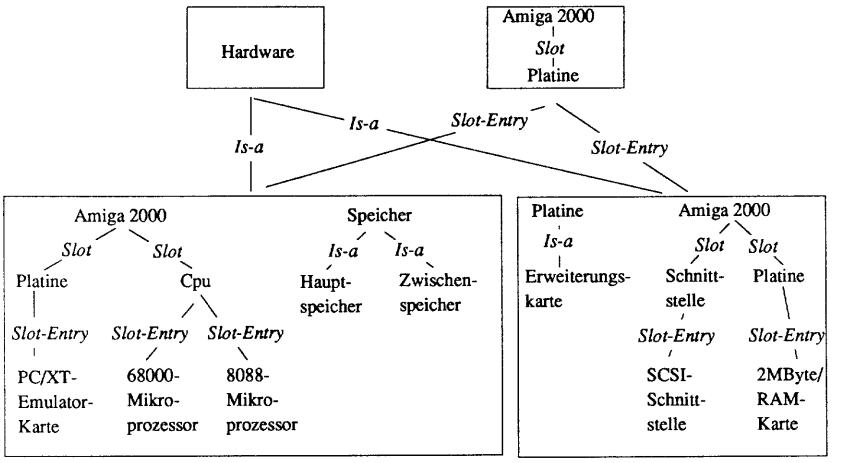


Abb. 3. Beispiel für eine Basis-Textkonstituente

3 Die informationslinguistische Komponente des Systems: Natürlichsprachige Präsentation von Textwissen

3.1 Situationsspezifische Generierung von natürlichsprachigen Abstracts aus den TOPIC-Textgraphen

Die gegenwärtige Beschäftigung mit der natürlichsprachigen Präsentation von Textwissen läßt einen entscheidenden Paradigmenwechsel insofern erkennen (vgl. [21]), als versucht wird, nicht mehr wie bei den früheren statistischen Verfahren aus den Texten direkt, sondern aus semantischen Repräsentationen der Texte Darstellungen der Textinhalte zu erstellen. Für TWRM-TOPOGRAPHIC stellen die vom Textanalyse-/Textkondensierungssystem TOPIC erstellten *semantischen Repräsentationsstrukturen* für Textinhalte (Textgraphen) die Basis für die Präsentation von Textwissen dar. Da die TOPIC-Textgraphen die thematischen Schwerpunkte eines Textes vorwiegend auf indikativem (lediglich anzeigendem) Niveau beschreiben, jedoch auch signifikante Fakteninformation beinhalten können, ist somit die Möglichkeit zur Erzeugung indikativer als auch indikativ-informativer Abstract (Zusammenfassungen)[4] gegeben. Die Hauptforderung, die an diese beiden Abstract-Typen gestellt wird, ist die, den wesentlichen Textinhalt anzuzeigen bzw., im Fall des indikativ-informativen Abstract, auch teilweise wiederzugeben.

Der Forderung, daß Länge, Komplexität und Abstraktionsniveau eines Abstracts sowie der angebotene Inhalt die jeweiligen Benutzerinteressen reflektieren sollen [8] [19] [20], konnte bisher nicht befriedigend entsprochen werden, da Abstracts in der Regel *einmalig* für *einen* bestimmten Zweck und *einen* angenommenen Benutzertyp angefertigt wurden. Mit der Generierungskomponente von TWRM-TOPOGRAPHIC können jedoch aus einem Text (bzw. dessen Repräsentation) *situationsspezifische* Abstracts mit unterschiedlichem Themenschwerpunkt und unterschiedlicher Ausführlichkeit produziert werden. Derart situationsspezifische Abstracts erfordern, im Gegensatz z. B. zum Abstracting-System SUSY [8], in dem der Benutzer Schemata angeben muß, die die Textanalyse und die Erstellung des Abstracts steuern, kein Eingreifen des Benutzers, sondern werden durch Auswertung der Vorgaben der Dialogführung von TWRM-TOPOGRAPHIC produziert.

3.2 Problematik der Generierung natürlichsprachiger Abstracts aus den Repräsentationsstrukturen der TOPIC-Textgraphen

Zwar liegt mit dem TOPIC-Textgraphen bereits eine kondensierte Textrepräsentation vor, doch sind im Textgraphen auch Themenbeschreibungen enthalten, die zwar für einen Textabschnitt, nicht aber für den gesamten Text von zentraler Bedeutung sind, da TOPIC die Relevanz eines Konzepts bezüglich eines Textabschnitts beurteilt. Ziel eines Abstracts dagegen ist, den wesentlichen Textinhalt (mit gegebenenfalls signifikanter Fakteninformation) bereitzustellen; aus diesem Grund müssen zunächst die zentralen Textthemen und die zugehörigen Konzepte identifiziert werden. Diese relevanten *Textthemen* sind je nach Interessenschwerpunkt und gewünschter Ausführlichkeit jedoch nicht für jeden Benutzer gleichermaßen relevant, so daß zur Bestimmung der

[4] Im Bereich der wissensbasierten Sprachverarbeitung wird *Zusammenfassung* meist für die Kurzwiedergabe der wesentlichen Handlung eines Narrativtexts gebraucht. Um Unklarheiten zu vermeiden, wird hier der Begriff *Abstract* bevorzugt

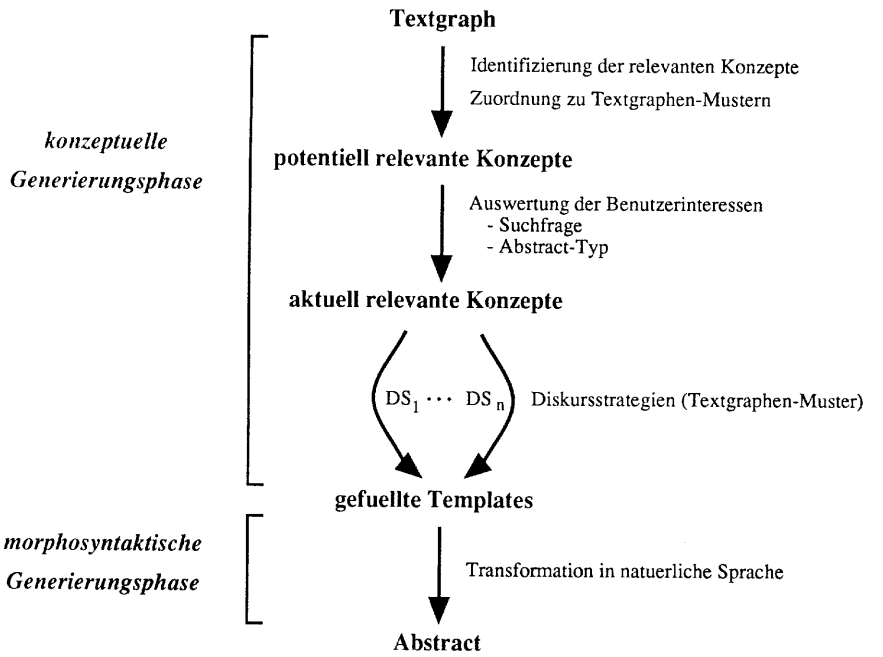


Abb. 4. Textgenerierungs-Modell

aktuell relevanten Konzepte eine weitere Bewertung durch Analyse der Vorgaben der Dialogführung (z. B. der Suchfrage des Benutzers) stattfinden muß.

Nachdem entschieden ist, welche Konzepte im Abstract ausgedrückt werden sollen, muß ein *Textplan* erstellt werden (vgl. zur skizzierten Problematik auch [3] [6] [22] [24]), der festlegt,

- welche Konzepte zusammengehören und in einem Satz ausgedrückt werden sollen,
- welche lexikalischen und syntaktischen Realisierungen am besten geeignet sind, die Beziehung zwischen den Konzepten auszudrücken,
- in welcher Reihenfolge die Sätze angeordnet werden sollen, wie sie zusammenhängen und wie dies verdeutlicht werden kann.

Es muß also die Generierung von zusammenhängendem Text gewährleistet werden, der sowohl die Expansion der einzelnen Textthemen und ihre Abgrenzung gegenüber anderen Themen *(Textkohäsion)* als auch die textuelle Relationierung der Themen *(Textkohärenz)* erkennen läßt (vgl. [17]) und darüber hinaus die speziellen Anforderungen, die an ein Abstract gestellt werden, erfüllt.

3.3 Textgenerierungs-Modell

Die angestrebte Funktionalität erfordert keine vollständig natürlichsprachige Generierung, so daß bei der Generierung vorgefertigte Satzmuster *(Templates)* verwendet werden können, in deren Lücken die ausgewählten Konzepte eingesetzt werden.

Die Textgenerierung wird in zwei Phasen, in eine konzeptuelle und in eine morphosyntaktische Phase, aufgeteilt (vgl. [24]). In der *konzeptuellen Phase* wird entschieden, welche Konzepte aus den Textkonstituenten des Textgraphen im Abstract ausgedrückt, wie sie gruppiert und angeordnet werden sollen, und welche lexikalischen und syntaktischen Realisierungen geeignet sind, diesen Inhalt auszudrücken. Aufgabe der *morphosyntaktischen Phase* ist die Transformation der erarbeiteten Struktur in natürliche Sprache.

Das hier kurz skizzierte Textgenerierungs-Modell legt den Schwerpunkt auf die Erstellung des Textplans, dementsprechend konzentriert sich die weitere Darstellung ganz auf die konzeptuelle Generierungsphase.

3.4 Konzeptuelle Generierungsphase

Im ersten Schritt der konzeptuellen Phase werden die zentralen Textthemen und die zugehörigen Konzepte identifiziert, also diejenigen Konzepte bestimmt, die aufgrund ihrer Bedeutung im Text *potentiell relevant* sind. Danach werden aus ihnen durch Auswertung der Vorgaben der Dialogführung, die die Benutzerinteressen reflektieren, diejenigen bestimmt, die für die gegebene Dialogsituation *aktuell relevant* sind.

Die ausgewählten relevanten Konzepte werden aufgrund ihrer thematischen Relationierung typischen Mustern zugeordnet. Entsprechende *Diskursstrategien* (vgl. z. B. [23] [24]) steuern sowohl die

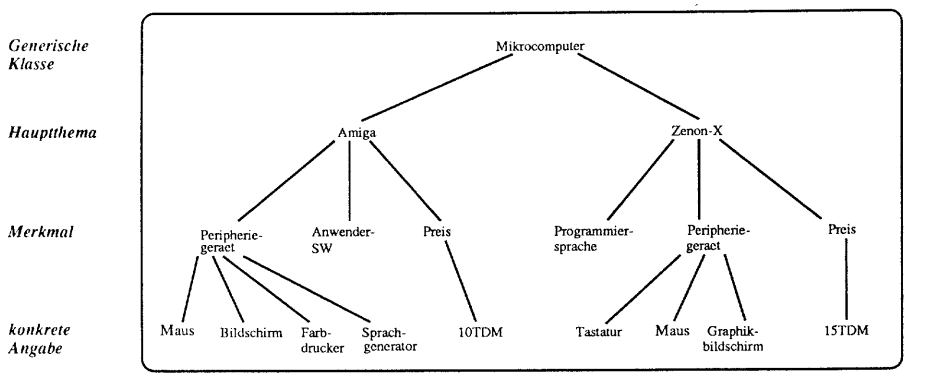


Abb. 5. Beispiel für die Relationierung der potentiell relevanten Konzepte im Textgraphen-Muster „Vergleichende Gegenüberstellung mehrerer verwandter Hauptthemen"

Auswahl und Anordnung der Konzepte im Text als auch das Einsetzen der Konzepte in adäquate Satzmuster.

Die Bestimmung der potentiell relevanten Konzepte beginnt mit der Identifizierung der zentralen Textthemen. Ein Konzept aus einer Themenbeschreibung des Textgraphen soll als ein zentrales Textthema gelten, wenn es im Text in mindestens zwei Abschnitten als thematisch relevant bewertet wurde (also nicht nur von lokaler Bedeutung ist), selbst durch andere thematisch relevante Konzepte näher spezifiziert wird und darüber hinaus genügend spezifisch ist, um Aussagekraft zu besitzen; eine formale Beschreibung der Bedingungen zur Bestimmung eines zentralen Themas enthält [30].

Anschließend werden die zu einem Hauptthema gehörigen Konzepte bestimmt und gemäß ihrer Relation zum Hauptthema unterteilt in:

- Generische Klasse (Prototyp oder Oberbegriff des Hauptthemas)
- Hauptthema
- Merkmal (ein *Slot* des Hauptthemas)
- konkrete Angabe zu einem Merkmal (*Slot-Entry* zu einem Merkmal)

Sind die potentiell relevanten Konzepte des Textgraphen bestimmt und thematischen Blöcken zugeordnet, werden aufgrund der Relationierung dieser Blöcke verschiedene, typische Textgraphen-Muster unterschieden, aus denen sich durch Kombination weitere, komplexere Muster ableiten lassen.

- *Einzelnes Hauptthema:*
Bei der Klassifizierung des Textgraphen wird ein einziges Hauptthema bestimmt, das durch Merkmale und konkrete Angaben zu den Merkmalen näher spezifiziert ist.

- *Mehrere, einfach linear relationierte Hauptthemen:*
Bei diesem Muster werden bei der Klassifizierung des Textgraphen mehrere Hauptthemen bestimmt. Dabei ist eines der näher spezifizierenden Konzepte des einen Hauptthemas ein anderes Hauptthema des Textgraphen, das wiederum näher spezifiziert wird. Gibt es mehr als zwei Hauptthemen, sind sie fortlaufend auf diese Weise relationiert.
- *Vergleichende Gegenüberstellung mehrerer verwandter Hauptthemen:*
Auch hier werden mehrere Hauptthemen bestimmt, die aber in diesem Fall alle derselben generischen Klasse angehören und darüber hinaus auch gemeinsame Merkmale besitzen. Aufgrund dieser thematischen Relationierung kann geschlossen werden, daß die Hauptthemen vergleichend gegenübergestellt werden (vgl. Abb. 5).

Um die für eine gegebene Dialogsituation aktuell relevanten Konzepte bestimmen zu können, werden die Suchfrage *(Query)* des Benutzers in Form semantisch relationierter Konzepte und die optionale Angabe über den gewünschten Abstract-Typ ausgewertet.

- *Abgleich mit der Suchfrage:*
Durch den Abgleich mit der Suchfrage werden aus den potentiell relevanten Konzepten diejenigen ausgewählt, die geeignet sind, die Suchfrage des Benutzers zu beantworten, die anderen Konzepte aber eliminiert. Damit es nicht zu einer Überbewertung der Relevanz des nachgewiesenen Textes kommt, muß jedoch auch der Teil des Textinhaltes angezeigt werden, der keinen direkten Bezug zur Suchfrage hat; d.h. es werden unter Beibehaltung der thematischen Relationierung die gesuchten Textthemen so detailliert wie möglich, die anderen nur so detailliert wie nötig beschrieben. Mit dieser Selektion wird der Grice'schen Quantitätsmaxime [10] entsprochen, so daß die Abstracts in den Benutzerdialog integriert

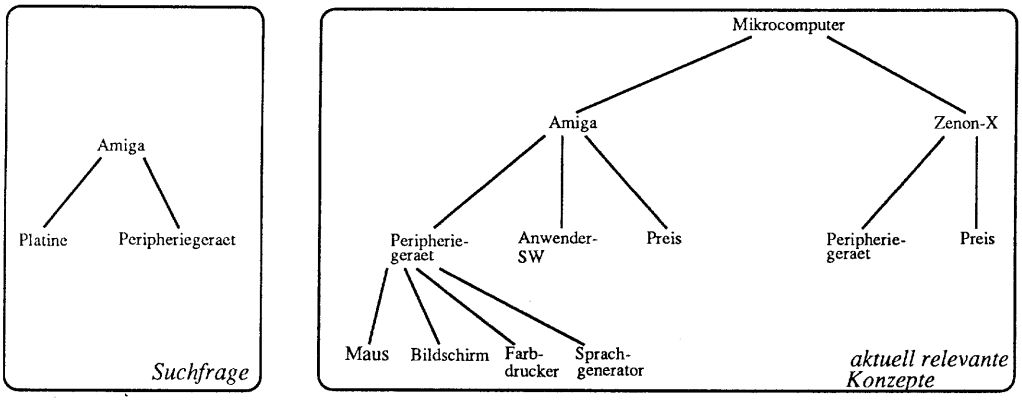


Abb. 6. Beispiel zur Bestimmung der aktuell relevanten Konzepte

werden können, ohne dessen topikalische Struktur zu verletzen.

Abbildung 6 zeigt, welche Konzepte aus den potentiell relevanten Konzepten von Abb. 5 durch Abgleich mit der angegebenen Suchfrage als aktuell relevant bestimmt werden.

• *Abstract-Typ:*
Im Anschluß an den Abgleich mit der Suchfrage erfolgt die Auswertung des Parameters ‚Abstract-Typ', für den die Werte ‚indikativ' oder ‚indikativ-informativ' angegeben werden können. Bei der Angabe ‚indikativ' soll der wesentliche Textinhalt angezeigt werden, ohne jedoch konkrete Informationen zu nennen. Dementsprechend werden aus den potentiell relevanten Konzepten diejenigen eliminiert, die konkrete Informationen zu einem Merkmal liefern. Wird als Typ ‚indikativ-informativ' vorgegeben, sind alle potentiell relevanten Konzepte tatsächlich relevant (eine formale Beschreibung zur Bestimmung der aktuell relevanten Konzepte enthält [30]).

Dem Auswahlmechanismus zur Bestimmung der relevanten Konzepte ist eine Ordnung der Konzepte inhärent, da die ausgewählten Konzepte verschiedenen Themenblöcken zugeordnet werden und die Themenblöcke aufgrund ihrer thematischen Relationierung typische Muster bilden. Mit der dadurch erarbeiteten Struktur ist die Basis für die Generierung von kohäsivem und kohärentem Text gegeben, da sowohl die Expansion der einzelnen Themen als auch ihre Relationierung erkennbar ist.

Daneš [5] hat die Anordnung und Relationierung von Themen in Texten untersucht und dabei drei Haupttypen der thematischen Progression erarbeitet, die ebenfalls kombiniert werden können. Diese drei Haupttypen der thematischen Progression korrespondieren mit den zuvor vorgestellten Textgraph-Mustern (s. Tabelle 1).

Tabelle 1

Thematisches Progressionsmuster:	**Textgraphen-Muster:**
konstantes Thema	einzelnes Hauptthema
einfache lineare Progression	mehrere, einfach linear relationierte Hauptthemen
abgeleitetes Thema	vergleichende Gegenüberstellung mehrerer verwandter Hauptthemen

Ursprünglich wurden die thematischen Progressionsmuster bei der Analyse bzw. Beschreibung von Texten verwendet; hier dienen sie dazu, Kriterien für eine sinnvolle lineare Anordnung der Themenblöcke im Text abzuleiten. Aus der Zuordnung der typischen Muster zu den thematischen Progressionsmustern von Daneš und der Operationalisierung der Konventionen zur Erstellung eines Abstracts, z. B., daß in indikativ-informativen Abstracts der indikative Teil vor dem informativen stehen soll, wurden Regeln in Form von Diskursstrategien abgeleitet, die die lineare Anordnung sowohl der Themenblöcke als auch der einzelnen Konzepte im Text steuern. Durch speziell auf die verschiedenen typischen Muster abgestimmte Diskursstrategien wird im letzten Schritt der konzeptuellen Phase die Generierung von zusammenhängendem Text realisiert. Mit jedem Schritt einer Diskursstrategie werden zusammengehörige Konzepte und ein dazu passendes Satzmuster ausgewählt, in dessen Lücken die Konzepte eingesetzt werden. Für das Textgraph-Muster „Vergleichende Gegenüberstellung mehrerer verwandter Hauptthemen" lautet z. B. die Diskursstrategie verbal formuliert:

1. Nenne die Hauptthemen des Textes und deren gemeinsame generische Klasse.
2. Nenne die Merkmale, die allen Hauptthemen gemeinsam sind.
3. Thematisiere ein gemeinsames Merkmal und gib die konkreten Angaben aus dem Textgraphen für eines dieser Hauptthemen wieder.
4. Nenne für ein anderes Hauptthema die konkreten Angaben zu diesem Merkmal.
5. Thematisiere ein anderes gemeinsames Merkmal und nenne die konkreten Angaben für die verschiedenen Hauptthemen.
6. Nenne für alle Hauptthemen deren eigene Merkmale.
7. Führe für die eigenen Merkmale eines der Hauptthemen die Schritte 3) und 4) der Diskursstrategie zur Behandlung des Textgraph-Musters „einzelnes Hauptthema" durch.

Haben die Hauptthemen des Textgraphen nur ein gemeinsames Merkmal, entfällt Schritt (5). Existieren mehr als zwei gemeinsame Merkmale, werden die Schritte (3) bis (5) wiederholt. Gibt es mehr als ein Hauptthema mit eigenen Merkmalen, wird Schritt (7) entsprechend oft wiederholt. Die Arbeitsweise der Diskursstrategien soll an nachfolgendem Beispiel verdeutlicht werden, das einen Schritt der oben genannten Strategie detaillierter beschreibt.

Beispiel: Durch Anwendung der Diskursstrategie zur Behandlung des Textgraph-Musters „vergleichende Gegenüberstellung mehrerer verwandter Hauptthemen" auf die in Abb. 6 dargestellten aktuell relevanten Konzepte wird im ersten Schritt das Satzmuster

„Der Artikel handelt über ((?ART_DET) (GENERIC_CLASS)) (MAIN_TOPICS)."

ausgewählt. Die Darstellung der Satzmuster ist leicht vereinfacht. Die Platzhalter des Satzmusters, „MAIN_TOPICS" und „GENERIC_CLASS" (die anzeigen, an welcher Stelle, mit welcher Kategorie von relevanten Konzepten und mit wievielen Konzepten die Lücken gefüllt werden können), werden durch die in diesem Schritt ausgewählten Konzepte „Mikrocomputer" (gemeinsame generische Klasse der Hauptthemen), „Zenon-X" und „Amiga" (Hauptthemen) ersetzt, so daß das Satzmuster anschließend lautet:

„Der Artikel handelt über ((?ART_DET)(Mikrocomputer))(Zenon-X, Amiga)."

Nach der Transformation der Satzmuster in natürliche Sprache durch die morphosyntaktische Phase ergibt sich folgendes Abstract (die aus dem Textgraphen ausgewählten Konzepte sind kursiv gesetzt):

„Der Artikel handelt über die *Mikrocomputer Zenon-X* und *Amiga*. Die *Preise* und die *Peripheriegeräte* der *Mikrocomputer* werden vergleichend gegenübergestellt. Für den *Amiga* gibt es die *Peripheriegeräte Maus, Bildschirm, Farbdrucker* und *Sprachgenerator*. Außerdem wird auf die *Anwendungssoftware* des *Amiga* eingegangen."

Zum Vergleich: Wären bei der Ausführung der Diskursstrategie die *potentiell* relevanten und nicht die *aktuell* relevanten Konzepte berücksichtigt, also die Benutzerinteressen vernachlässigt worden, würde das Abstract folgendermaßen lauten:

„Der Artikel handelt über die *Mikrocomputer Zenon-X* und *Amiga*. Die *Peripheriegeräte* und die *Preise* der *Mikrocomputer* werden vergleichend gegenübergestellt. Für den *Zenon-X* gibt es die *Peripheriegeräte Graphikbildschirm, Tastatur* und *Maus*. Der *Amiga* hat die *Peripheriegeräte Maus, Bildschirm, Farbdrucker* und *Sprachgenerator*. Als *Preis* wird für den *Zenon-X 15000 DM* und für den *Amiga 10000 DM* genannt. Außerdem wird auf die *Anwendungssoftware* des *Amiga* und auf die *Programmiersprache* des *Zenon-X* eingegangen."

Mit dem durch TOPIC erarbeiteten Textgraphen und den durch die Generierungskomponente erzeugten Abstracts liegen die informationellen Objekte fest, die der Benutzer während des Retrieval abfragen kann bzw. die ihm angeboten werden. Im folgenden Abschnitt werden die technischen Bedingungen der Umsetzung informationeller Objekte in graphische Objekte diskutiert, bevor im letzten Abschnitt der Dialog selber dargestellt wird.

4 Ein objektorientiertes User-Interface-Management-System zur flexiblen Definition graphischer Dialoge

4.1 Rahmenbedingungen für die Spezifikation eines UIMS

Der experimentelle Ansatz des Projekts TWRM-TOPOGRAPHIC, d.h. die Erprobung verschiedener Layout-Verfahren zur Darstellung von Repräsentationsstrukturen und von Interaktionsstilen zur Dialogsteuerung (bei evtl. gleichzeitiger Modifikation dieser Strukturen) macht ein flexibles User-Interface-Management-System erforderlich. Die Trennung zwischen Interface und Applikation erfolgt dabei dergestalt, daß die Inhalte der Applikation zugeordnet werden, während die Formen und die

Tabelle 2. Methoden informationeller Objekte im Retrievaldialog: Für den *select*-Operator werden die situationsspezifischen Wirkungen auf die manipulierten Objekte angegeben und die objekt-orientierten Navigationsoperatoren *zoom* und *browse* durch Spezifikation der situationsadäquaten Zielobjekte skizziert

Informationelle Objekte (Retrieval)	Graphische Darstellung	Select	Zoom	Browse
Frame	Beschrifteter Knoten (vgl. Abb. 7, 8)	Auswahl zum Retrieval	Interne Struktur	Kontext
expandierter Frame	Tabelle (vgl. Abb. 13)	Auswahl zum Retrieval	-	Kontext
Relationsverbindung	Kante (vgl. Abb. 7, 9)	Auswahl zum Retrieval	-	Verb. Beschr. der Relation
Wissensbasis	Baum (vgl. Abb. 7)	-	Query	-
Polyhierarchische Relationierung eines Frame	„zentrierter Knoten" (vgl. Abb. 8)	-	-	andere Relation
Query	Tabelle bzw. Netz (vgl. Abb. 9, 10)	-	Liste der relevanten Passagen	-
Passagenbeschreibung	Tabelleneintrag (vgl. Abb. 11)	-	Themenbeschreibung	Textgraph
Themenbeschreibung	Netz (vgl. Abb. 12)	Themenbeschr. als neue Query	Passage	weitere Themenbeschreibung
Passage	Text (vgl. Abb. 14)	-	Volltext	weitere Passage
Volltext	Text (vgl. Abb. 14)	Textgraph	-	weiteren relevanten Text
Textgraph	Netz	-	Volltext	-

Abbildung von Inhalten auf Formen durch das Interface verwaltet werden. So sind z. B. Texte und Textrepräsentation der Applikation, Bäume, Netze, Tabellen sowie die Darstellung einer Repräsentationsstruktur als Netzwerk dem Interface zugeordnet. Die Anforderungen an die Flexibilität der Schnittstelle wie auch die Notwendigkeit, prototypisch erstellte Dialogbausteine in unterschiedlichen Anwendungssituationen erproben zu können, ließen eine objektorientierte Spezifikation des UIMS als sinnvoll erscheinen (vgl. [29])[5].

4.2 Objektklassen im TWRM-TOPOGRAPHIC-System

Die oberste Spezifikationsebene des Interface stellen die *informationellen Objekte* dar. Diese Objekte kontrollieren die Abbildung von *Applikationsobjekten* auf *graphische Strukturen,* indem sie applikationsorientierte Methoden (z. B. Abfragen der von einem Frame ausgehenden Relationskanten) Methoden auf graphischen Strukturen zuordnen (z. B. Einfügen von Sohn-Knoten in einen Baum). Einem informationellen Objekt sind also je eine graphische Struktur und ein Applikationsobjekt zugeordnet, mit denen es Nachrichten austauscht.

Diese hybriden graphik- und applikationsorientierten Methoden werden entweder durch Nachrichten von anderen informationellen Objekten oder durch die ihnen zugeordnete graphische Struktur aktiviert.

Nachrichten an ein informationelles Objekt können von allen anderen informationellen Objekten gesendet werden, die durch unmittelbare Kommunikationslinien mit ihm verbunden sind und damit einen Dialogkontext bilden. Elemente des Dialogkontextes eines Objekts sind grundsätzlich:

- das Objekt, das seine Erzeugung veranlaßt hat,
- das zugehörige Aggregat,
- alle i-Objekte (im folgenden für informationelle Objekte), die das gleiche Applikationsobjekt abbilden (ansatzweise den *Perspectives* vergleichbar (vgl. [31]).

Die zugeordnete graphische Struktur kann Methoden des i-Objekts aktivieren, indem sie ein vom Benutzer gegebenes deiktisches Kommando, wie z. B. die Betätigung einer Maustaste in einem graphischen Objekt, weiterleitet. Derartige Kommandos genügen im allgemeinen, die von dem i-Objekt auszulösende Aktion unter Berücksichtigung des Dialogkontextes und der Dialoghistorie vollständig zu determinieren. Andernfalls veranlaßt das angesprochene i-Objekt die Objekte seines Dialogkontextes, ihr Verhalten bis auf Widerruf in einer der folgenden Weisen zu ändern:

- das Objekt reagiert nicht mehr auf Benutzeraktionen,
- die Reaktion auf Benutzeraktionen wird modifiziert,

[5] Die prinzipiellen Eigenschaften objektorientierter Systeme, wie *Vererbungs-* und *Message-Passing-Mechanismen* (Objekte kommunizieren durch Versenden von Nachrichten miteinander) wurden hier als bekannt vorausgesetzt [9]

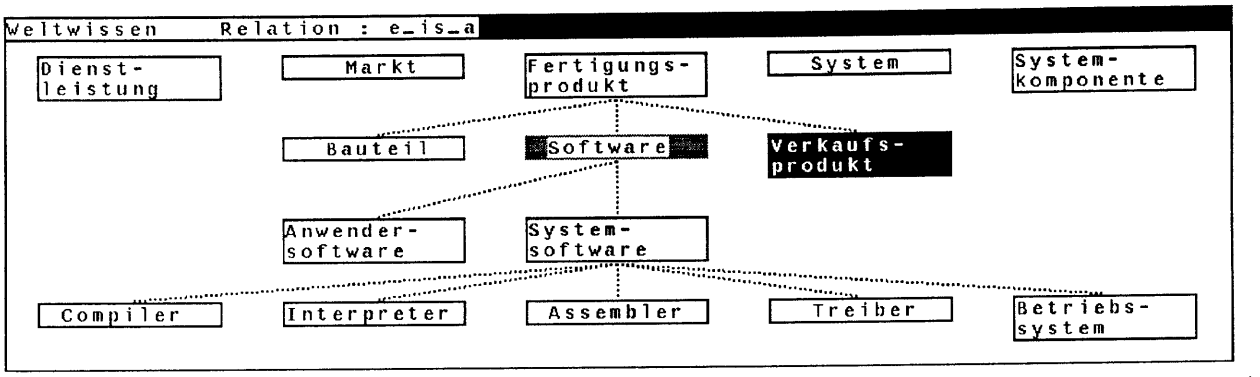


Abb. 7. Hierarchische Darstellung eines Wissensbasisausschnitts mit 2 aktivierten Frames (Verkaufsprodukt: positiv, Software negativ aktiviert).
Methoden auf den Teilobjekten:
select: *Frame* bzw. Relationskante selektieren/deselektieren
zoom: *Frame* als Tabelle darstellen (vgl. Abb. 13)
browse:
Frame: Kontext des *Frame* zeigen - durch Umordnung bzw. Erweiterung des Baums oder falls das nicht möglich ist durch Darstellung anderer Relationsverbindungen (vgl. Abb. 8)
Relationskante: Bedeutung der Relation verbal erläutern
Methoden auf dem Aggregat:
zoom: Präsentation der Query (vgl. Abb. 9 und 10)

- alle Benutzeraktionen werden an das initiierende i-Objekt weitergeleitet.

Damit sind die elementaren Voraussetzungen zur Realisierung graphischer Klärungsdialoge gegeben.

I-Objekte können in aggregierter Form vorliegen: so hat ein i-Objekt, das z. B. die baumorientierte graphische Darstellung einer Wissensbasis kontrolliert, als informationelle Teilobjekte diejenigen Objekte, die Frames und die Relationen zwischen ihnen als Knoten und Kanten dieses Baumes realisieren. Dabei sind auf aggregierten i-Objekten durchaus andere Methoden definiert als auf ihren Teilen. Gelangt aber an ein solches Teilobjekt eine Nachricht, die von diesem nicht verarbeitet werden kann, so wird sie an das Aggregatobjekt weitergeleitet. Aggregierte i-Objekte können oft nur partiell präsentiert werden. Da jedoch die ihnen zugeordneten graphischen Objekte als Fenster implementiert sind, können alle Teilobjekte durch Scrolloperationen erreicht werden.

Eine Zusammenstellung der informationellen Objekte findet sich in der ersten Spalte von Tabelle 2. Auf Beispiele der graphischen Realisierung wird in der zweiten Spalte von Tabelle 2 verwiesen. Die graphischen Darstellungen selber finden sich am Ende von Abschn. 5, wo sie in ein Dialogbeispiel eingebunden sind. Die in der dritten bis fünften Spalte zusammengestellten Operationen werden jeweils bei den Abbildungen am Schluß näher erläutert.

Auf der gleichen Spezifikationsebene wie die i-Objekte sind die *Kontrollobjekte* angesiedelt, die zur Kontrolle von Meta-Dialogen dienen. Zur Wahrnehmung ihrer Aufgaben kommunizieren sie nicht mit Applikationsobjekten, sondern mit i-Objekten, von denen sie die Informationen über Funktionalität, Dialoghistorie etc. erhalten, die sie benötigen. Die graphische Repräsentation von Kontrollobjekten erfolgt im allgemeinen durch Ikonen, z. B. werden *help* und *undo* über Piktogramme realisiert. Dies findet seine Begründung darin, daß so eine eingeschränkte Anzahl von wichtigen Funktionen, die zudem im allgemeinen parameterfrei sind, auf einsichtige Weise präsent gehalten werden kann. Eine Sonderstellung nimmt allerdings das Kommandoobjekt ein, das formalsprachige Eingaben behandelt. Diese werden einem Kommandointerpreter zugestellt, sodann werden aufgrund der interpretierten Struktur Nachrichten an i-Objekte gesendet, die das entsprechende Kommando ausführen. Formalsprachige Kommandos haben im TWRM-System nur abkürzende Wirkung, indem sie Folgen graphischer Kommandos ersetzen. So dient z. B. das Kommando *find(⟨Frame_ Name⟩)* als Abkürzung für eine Sequenz von *Browse*-Operationen im Weltwissen (vgl. Abb. 7). Eine Zusammenfassung der auf Kontrollobjekten definierten Methoden findet sich in Tabelle 3.

Häufig wiederkehrende *graphische Strukturen*, wie Bäume, Polyhierarchien, Tabellen, Texte als formatierte Zeichenfolgen, werden durch prototypische Aggregate realisiert, um den Spezifikationsaufwand zu minimieren. Jeder dieser Strukturen ist ein

Tabelle 3. Auf Kontrollobjekten definierte Methoden

Kontrollobjekte	Select	Zoom
Aktivierungs-gewicht	Gewicht erhöhen/ vermindern	-
Dialog	-	Dialog-historie darstellen
Element der Dialog-historie	Kommando erneut ausführen	-
Menü	Menü-element auswählen	-
Piktogramm	repräsentierte Funktion ausführen	-

Fensterobjekt zugeordnet, in dem ihre Elemente als graphische Elementarobjekte dargestellt werden. Die auf graphischen Strukturen definierten Methoden, wie z. B. Einfügen von Knoten in einen Baum oder von Einträgen in eine Tabelle, werden durch Nachrichten vom zugehörigen i-Objekt aktiviert. Nachrichten von den Elementarobjekten werden unbearbeitet an das i-Objekt weitergegeben. Die Bausteine zum Aufbau graphischer Strukturen sind einzelne *graphische Elemente*. Sie werden durch die Vorgabe eines graphischen Musters und einer Koordinatenangabe definiert.

Bildschirmobjekte stellen die Abbildung der graphischen Elemente auf Bildschirmkoordinaten und Elemente einer einfachen Graphiksprache dar, die vom UIMS aus nicht direkt verfügbar ist. Sie kommunizieren mit *Interaktionsobjekten,* die eine objektorientierte Sicht auf die Eingabetreiber ermöglichen.

4.3 Einige Aspekte der Implementation des UIMS

Die Implementation des UIMS erfolgte durch die Einbettung von Konstrukten objektorientierter Sprachen in Prolog. Das Message-Passing erfolgt dabei durch Variablen-Instanzen, während die einer Objektklasse zugeordneten Methoden durch Klauseln definiert sind. Zustände von Objekten (Belegung der Instanzvariablen) werden analog zum SPOOL-System [7] durch Fakten in der Prolog-Datenbasis repräsentiert. Anders als dort wird aber im TWRM-UIMS die Möglichkeit zur nicht-deterministischen Programmierung nicht eingeschränkt, da etwaige Seiteneffekte (z. B. Änderung der Wissensbasis) im Fall von *backtracking* rückgängig gemacht werden [16]. Daher können die auf TWRM-Objekten definierten Methoden auch als Transaktionen aufgefaßt werden, so daß nicht erfolgreich abgeschlossene Aktionen selbst beim Fehlen einer expliziten Fehlerbehandlung immer einen definierten Zustand erzeugen. Das ist insbesondere beim Editieren von Wissensbasen wichtig.

5 Die Interaktion mit dem Benutzer

TWRM-TOPOGRAPHIC realisiert das von Bates [1] vorgeschlagene *exploratorische Paradigma* des Information Retrieval, das den indirekten Zugriff auf die Inhalte einer Datenbank über formal- bzw. natürlichsprachige Abfragen ersetzt durch direkte Navigation in den auf dem Bildschirm graphisch präsentierten Wissensstrukturen. Im folgenden geben wir einen kurzen Überblick der Prämissen, die wir bei der Konzeption der Schnittstelle bezüglich des Informationsverarbeitungs- und Problemlösungsverhaltens eines durchschnittlichen Benutzers von TWRM-TOPOGRAPHIC berücksichtigt haben. Den Abschluß des Kapitels bildet dann ein Dialogbeispiel.

5.1 Kognitiv-ergonomische Gestaltungsprinzipien

Benutzer von Informationssystemen können nicht beliebig viele Daten gleichzeitig wahrnehmen. Also kommt zu den stilistischen Layout-Aufgaben der klassischen Ergonomie (angemessene Helligkeit, Kontrast, Schriftgröße etc.) die Auswahl der im aktuellen Dialogzustand zu präsentierenden Datenmenge hinzu, die kognitiv-ergonomisch zu begründen ist. TWRM-TOPOGRAPHIC erlaubt die gleichzeitige Präsentation von Textpassagen und graphisch dargestellten Textgraphfragmente. So kann der holistische Eindruck der Graphik zusammen mit einem diagonalen Lesen der Passage dem Benutzer Informationen vermitteln, die sonst nur durch einen intensiveren und damit ermüdenderen Leseprozeß erschlossen werden können.

Die Designphilosophie von TWRM-TOPOGRAPHIC trägt der Beschränkung gleichzeitig verarbeitbarer Daten[6] dadurch Rechnung, daß der Benutzer in einer gegebenen Dialogsituation nur einen kleinen Teil der vorhandenen, als relevant erkannten Daten auf der jeweils adäquaten Kondensierungsstufe gezeigt bekommt. Relativ große Informationsmengen können leichter verarbeitet werden, wenn sie aggregiert, also zu neuen Einheiten zusammengefaßt werden. Diesen in der kognitiven Psy-

[6] Eine Übersicht der quantitativen Beschränkungen des menschlichen Kurzzeitgedächtnisses, die für das Design von Benutzerschnittstellen relevant sind, geben [2]

chologie als *chunking* bezeichneten Effekt nutzt die Benutzerführung von TWRM-TOPOGRAPHIC aus, indem bei der Präsentation informationeller Objekte kognitive Strategien, die ein Benutzer beim Umgang mit Dokumenten erworben hat, möglichst unterstützt werden. So beginnt die Darstellung der Textinformation mit globalen, den Gesamttext charakterisierenden Angaben, die sukzessive durch immer detaillierter werdende informationelle Objekte ergänzt werden. Dieses Vorgehen orientiert sich am „selektiven Lesen" eines Dokuments, wobei man zunächst einschlägige Kapitel, dann Abschnitte, Passagen, Tabellen oder Abbildungen rezipiert.

Die Navigation in den Wissensstrukturen ist ein interaktiver Prozeß, bei dem von Seiten des Systems die in einer gegebenen Situation zu präsentierende Datenmenge zu bestimmen und graphisch aufzubereiten ist. Darüber hinaus sind auch die potentiellen Folgesitutionen zu determinieren, wie z. B. die Frage, welche Daten in welchem Umfang mit welchen Operationen zugänglich sein sollen. Eine wesentliche Rolle bei der Bestimmung der zeitlich-inhaltlichen Anordnung von Informationen spielt auch die begründete Annahme, daß neues Wissen in den Kontext des bereits vorhandenen leichter integrierbar ist, wenn Assoziationen mit bekannten Fakten möglich sind (vgl. [18]).

Die Benutzerführung darf allerdings dem Benutzer die Art der Textrezeption nicht aufoktroyieren, da dies für Personen, die gewohnt sind, ihren Bedürfnissen entsprechend die Reihenfolge angebotener Informationseinheiten zu ändern, eher ein Hindernis bei der Erschließung von Information darstellt. Den Benutzern von TWRM-TOPOGRAPHIC steht deshalb jederzeit die Möglichkeit offen, zu früheren Dialogsituationen und den damit verbundenen Informationseinheiten zurückzukehren und von dort aus andere Navigationspfade zu verfolgen.

5.2 Pragmatische Dialogstrukturen der Informationssuche

Beim Design innovativer Informationssysteme müssen Situation und Kenntnisstand des Benutzers berücksichtigt werden. Im Bereich des Referenzretrieval führt dieser *Primat der Pragmatik* [19] zu einer Integration der kognitiven und der pragmatischen Aspekte des Benutzerverhaltens (vgl. [27]): Die Informationssuche wird als Teil eines Problemlösungsprozesses begriffen. Präzisiert man diesen Ansatz in einem Szenario, wo (mehr oder weniger) konkrete Handlungsmuster definiert und postuliert werden können, so läßt sich die *Relevanzinformation,* also die Menge aller theoretisch relevanten Daten ermitteln.

Im Rahmen von TWRM-TOPOGRAPHIC beziehen sich diese Handlungsmuster auf die manipulative Navigation im Raum der informationellen Objekte, wobei zwei prinzipielle Bewegungsrichtungen unterschieden werden können, denen objektklassen-spezifische Methoden entsprechen, nämlich *Browsing* als Navigation innerhalb des Weltwissens, eines Textkondensats etc. und *Zooming* als Bewegung entlang den Stufen des kaskadierten Abstracting (vgl. Tabelle 2).

Zunächst kann der Benutzer die ihm auf einem gegebenen Abstraktionsniveau zugänglichen Objekte erforschen. Das objekt-orientierte *Browsing* ist durch semantische Kriterien definiert, d. h. es sind von einem Objekt ausgehend nur die „Nachbarn" zu erreichen, die in einer beim *Browse*-Kommando implizit (durch den Kontext) oder explizit (durch Menü-Selektion) spezifizierten Relation stehen. Während dieses Navigierens in konzeptuellen Strukturen kann der Benutzer durch wiederholte Auswahloperationen ein Suchprofil zusammenstellen (vgl. die *Select*-Methoden in Tabelle 2). Nach Bearbeitung der Suchfrage stellt die Textwissensverwaltung dem Benutzer eine exemplarische Textpassage zur Verfügung, deren thematische Struktur er zunächst betrachten kann, und präsentiert eine Liste weiterer relevanter Passagen. Entweder wird der Benutzer seine Anfrage aufgrund der bisher präsentierten Textinformationen verfeinern oder weitere Textfragmente anfordern[7].

Der Wechsel des Abstraktionsniveaus erfolgt durch die Anwendung des *Zoom*-Operators auf ein informationelles Objekt (vgl. Tabelle 2). Ist dieses in aggregierter Form dargestellt, so werden seine internen Strukturen sichtbar. So bringt z. B. die Anwendung des *Zoom*-Operators auf einen Frame-Knoten in einem Netz dessen *Slot*-Struktur als Tabelle zum Vorschein. Dagegen offenbart eine als komplexes graphisches Objekt dargestellte Struktur nur die vorher selektierten Teile in detaillierterer Form. Damit wird einerseits den ergonomischen Prinzipien entsprochen, andererseits ermöglicht diese Konzep-

[7] Ist die weitere Passage bereits in der Liste enthalten, so kann der Benutzer (evtl. nach einer Scroll-Operation) den entsprechenden Listeneintrag durch *Zooming* in eine graphische Themenskizze expandieren. Der Zugang zu anderen, evtl. nicht als relevant klassifizierten Fragmenten ist über die intratextuellen Beziehungen, die im Textgraph repräsentiert sind, realisiert: Eine *Browse*-Operation auf der aktuellen, d. h. als expandierte Konstituente präsentierten Passage resultiert zunächst in der Darstellung des zugehörigen Textgraphen. Dieser gibt als ein *strukturiertes Menü* einen Überblick über die Textstruktur und ermöglicht durch die Expansion einzelner Knoten den Zugriff auf die Konstituenten des Textes

tion die Anwendung von Relevanzkriterien. Beim Übergang von der taxonomischen in die Textebene ist dies vergleichbar mit der klassischen Retrievalfunktion: Ein Teil des durch die Suchfrage selektierten taxonomischen Wissens wird auf eine Textkollektion abgebildet, die nach absteigender Relevanz geordnet und als Liste präsentiert wird. Auch die weiteren Niveauwechsel erfolgen durch *Zooming*.

Entsprechend dem Paradigma der graphischen Interaktion sind die Operationen *Browse, Select* und *Zoom* durch direkte Manipulation mit Hilfe deiktischer Kommandos realisiert.

5.3 Ein Retrieval-Dialog mit TWRM-TOPOGRAPHIC

Nachdem nun die Prinzipien der graphischen Interaktion erläutert worden sind, sollen die Möglichkeiten von TWRM-TOPOGRAPHIC anhand eines leicht vereinfachten Retrievaldialoges erläutert werden. Dieser ist insbesondere darauf ausgerichtet, alle Darstellungsebenen, die während eines Dialogs durch Folgen von *Zoom*-Operationen durchlaufen werden können, zu erreichen. Der Wechsel der Darstellungsebene wird dabei jeweils durch Auswahloperatoren vorbereitet, die eine Fokussierung auf relevante Information ermöglichen. Die Erkundung der jeweiligen informationellen Ebene erfolgt durch *Browse*-Operationen, die einen Wechsel des Fokus verursachen, ohne dabei das Abstraktionsniveau zu verändern.

Reale Dialoge unterscheiden sich von dem Beispiel durch das weitaus weniger zielgerichtete Verhalten der Benutzer. Um getroffene Dialogentscheidungen revidieren zu können, muß deshalb zu jedem Zeitpunkt des Dialogs, statt zu immer spezifischerer Information fortzuschreiten, eine Rückkehr zu allgemeineren Kaskadierungsstufen möglich sein. Zudem erfolgt aus Gründen der Übersichtlichkeit der Bildschirmaufbau dergestalt, daß nur für den aktuellen Dialogzustand relevante Objekte gezeigt werden. Irrelevante Objekte werden gelöscht, Objekte, die in späteren Dialogsituationen wieder relevant werden können, werden komprimiert.

Der Dialog beginnt mit einer Darstellung der allgemeinsten Konzepte der Wissensbasis. Dies entspricht der obersten Hierarchieebene von Abb. 7, die den Benutzer über die Ausdehnung des Diskursbereichs informieren soll. Durch eine Folge von *Browse*-Operationen können nun die konzeptuellen Beziehungen erforscht werden (z. B. ‚Fertigungsprodukt', ‚Software', ‚Systemsoftware' in Abb. 7). Soll ein schon bekanntes Konzept, das zur Zeit nicht auf dem Bildschirm sichtbar ist, aufgesucht werden, so kann der Suchprozeß durch Verwendung des Kom-

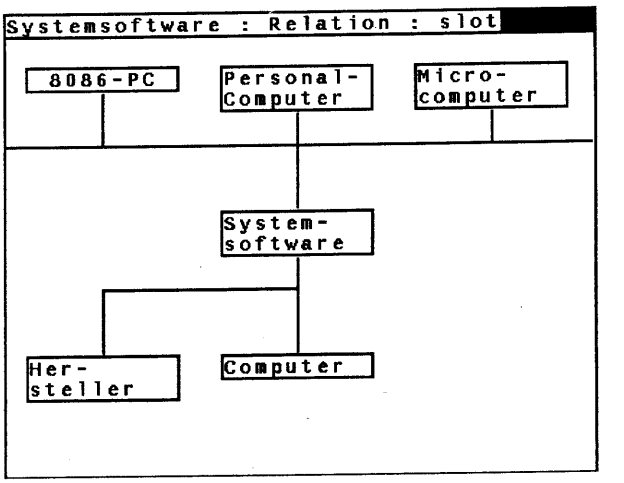


Abb. 8. Polyhierarchische Relationsverbindungen eines Konzepts
Methoden auf den Teilobjekten:
select: Frame bzw. Relationskante selektieren/deselektieren
zoom: Frame als Tabelle darstellen (vgl. Abb. 13)
browse:
Frame: Kontext des Frame zeigen (partielle Darstellung der Konzepthierarchie) (vgl. Abb. 7)
Relationskante: Bedeutung der Relation verbal erläutern
Methoden auf dem Aggregat:
browse: Vorgabe einer anderen Relationskante

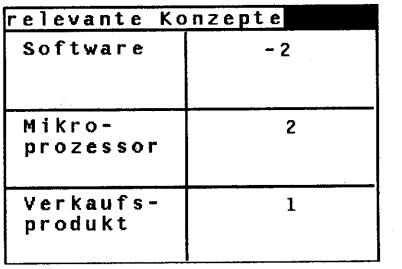


Abb. 9. Relevante Konzepte mit Aktivierungsgewichten
Methoden auf den Teilobjekten:
select: Aktivierungsgewicht modifizieren
Methoden auf dem Aggregat:
zoom: Liste der relevanten Konstituenten erzeugen (vgl. Abb. 11)

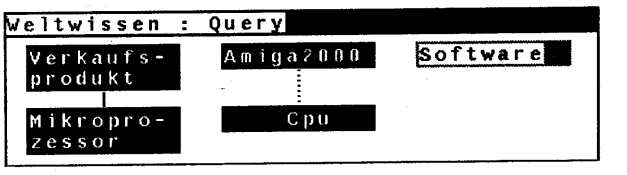


Abb. 10. Suchfrage (Query) als Graph
Methoden auf den Teilobjekten:
select: Frame bzw. Relationskante deaktivieren
Methoden auf dem Aggregat:
zoom: Liste der relevanten Konstituenten erzeugen (vgl. Abb. 11)

4 relevante Konstituenten aus 1 Text		
k1 : 36 (t1)	Autor/in : Loys Nachtmann Titel : Zwei Systeme in einem Quelle : Personal Computer Nr.3	
k7 : 36 (t1)	Soll der Amiga2000 mit MS-DOS arbeiten, so ist die sogenannte PC/XT-Emulatorkarte erforderlich. Im Prinzip befindet sich auf ihr ein selbstaendiger IBM-kompatibler Computer. Wie ueblich verrichtet	Autor/in : Titel : Zwe einem Quelle : Pe
k2 : 33 (t1)	Soll ein neuer Computer entwickelt werden, so kann man nur recht traege auf die momentane Marktsituation reagieren. Nicht nur ein neues Gehaeuse und andere Bauteile sind erforderlich, sondern	Autor/in : Titel : Zwe einem Quelle : Pe

Abb. 11. Liste der relevanten Textkonstituenten: Angabe von systeminternen Konstituentenbezeichnern, Relevanzmaß und dem Anfang der 1. Passage der Konstituente. Weitere Informationen, insbesondere bibliographische Angaben sind durch Scrolling in der Tabellenzeile erreichbar.
Methoden auf den Teilobjekten:
zoom: Themenbeschreibungsgraph präsentieren
browse: Textgraph darstellen

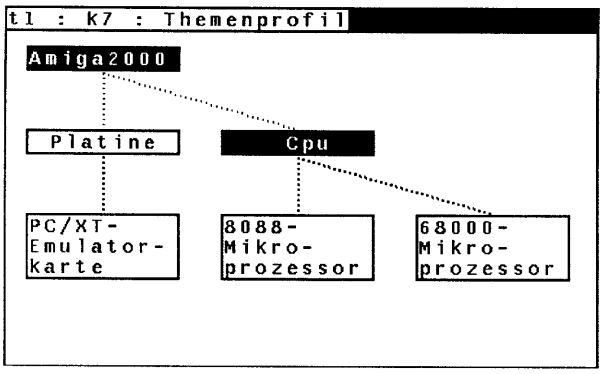


Abb. 12. Themenbeschreibungsgraph
Kantentypen:
durchbrochen (eng): Slot Relation
durchbrochen (weit): Slot Entry Relation
Methoden auf den Teilobjekten:
select: *Frame* bzw. Relationskante selektieren/deselektieren
zoom: *Frame* als Tabelle darstellen (vgl. Abb. 13)
browse:
Frame: Kontext des Frame zeigen (ermöglicht die Selektion von Frames zur Erweiterung einer Suchfrage)
Relationskante: Bedeutung der Relation verbal erläutern
Methoden auf dem Aggregat:
select: Themenbeschreibung wird neue Suchfrage *(Query by Example)*
zoom: Präsentation der Suchfrage (vgl. Abb. 9 und 10)
browse: Präsentation der Themenbeschreibung einer weiteren Konstituente, wobei diese nach in syntagmatisch, semantisch oder pragmatisch begründeter Beziehung zur Ausgangskonstituente stehen kann

Amiga2000 slots/entries		
Preis		
Schnitt-stelle	SCSI-Schnittste	Floppy-Port
Hersteller	Commodore	
Anwender-software	Computer-Spiel	

Abb. 13. Tabellarische Darstellung eines Frame
Methoden auf den Teilobjekten (nur bei non-terminalen Slots und Slot Entries, die selbst wieder als Frames repräsentiert sind):
select: Frame selektieren/deselektieren
zoom: Frame als Tabelle darstellen
browse: Kontext eines Frames zeigen

mandos „find", das durch das Prolog-Fenster eingegeben wird, abgekürzt werden. So ist z. B. das Kommando „find ('Systemsoftware')" äquivalent zu der oben genannten *Browse*-Sequenz. Zusätzlich zu der Spezialisierungsrelation („Is-a", „Instance") können andere inhaltliche Beziehungen zwischen Konzepten nach Bedarf in einem zusätzlichen Fenster gezeigt werden. Die *Slot*-Relation ordnet z. B. einem Konzept seine Merkmale zu (Abb. 8). Während der Erkundung des konzeptuellen Netzwerkes mit Hilfe von *Browse*- und *Find*-Operationen kann der Benutzer durch Anwählen *(Select)* von Begriffen oder Relationskanten ein Suchprofil zusammenstellen. Zur Verdeutlichung werden angewählte Begriffe invertiert abgebildet.

Die Konstruktion des Suchprofils wird abgeschlossen, indem der *Zoom*-Operator auf das Weltwissensobjekt angewendet wird. Zusätzlich zur Netzstruktur des Suchprofils *(Query)* wird eine Liste der bisher angewählten Konzepte ausgegeben, die mit Aktivierungsgewichten versehen sind und mit *Select* modifiziert werden können, was eine Gewichtung des Profils zur Folge hat (Abb. 9). Die Zuordnung eines negativen Aktivierungsgewichts bewirkt eine explizite Unterdrückung von Textthemen. Ein Beispiel dafür ist in Abb. 10 dargestellt, die negativ aktivierten Konzepte *(Software)* sind grau unterlegt dargestellt.

Mittels *Zoom* wird nun eine Liste der bezüglich der Anfrage relevanten Textkonstituenten unter Angabe bibliographischer Daten (Titel etc.) und eines kurzen Textausschnittes produziert, die nach einem Relevanzmaß sortiert ist, das sich aus dem Grad der Übereinstimmung von Suchprofil und thematischem Profil der Passage unter Berücksichtigung der Aktivierungsgewichte berechnet (Abb. 11). Gleichzeitig wird das Fenster, das den konzeptuellen Graphen (Wissensbasis) enthält, auf ein Minimum verkleinert, um die Aufmerksamkeit des Benutzers auf die in dieser Phase des Dialogs wichtigeren textorientierten Kaskadierungsstufen zu lenken.

Die zur Textpassage gehörigen Themenbeschreibungsgraphen können durch *Zoom* auf die Kurzbeschreibung der Textkonstituente gezeigt werden (Abb. 12). Von hier aus ergeben sich zwei Anwendungsmöglichkeiten für den *Zoom*-Operator: Wendet man ihn auf den gesamten Themenbeschreibungsgraphen an, so erhält man eine textuelle Version der Passage, während ein *Zoom* auf ein einzelnes Konzept des Themenprofils das bezüglich dieses Begriffs aus dem Text extrahierte Faktenwissen in einer Tabelle zur Darstellung (Abb. 13) bringt. Von der Passage ausgehend kann nun weitere Textinformation durch den Zugriff auf den Volltext bereitgestellt werden. Zur weiteren Relevanzentscheidung hat der Benutzer die Möglichkeit, ein situationsangepaßtes Abstract (indikativ oder indikativ-informativ) anzufordern, bevor er zum Gesamttext übergeht (Abb. 14). Mit der Darstellung des Volltextes

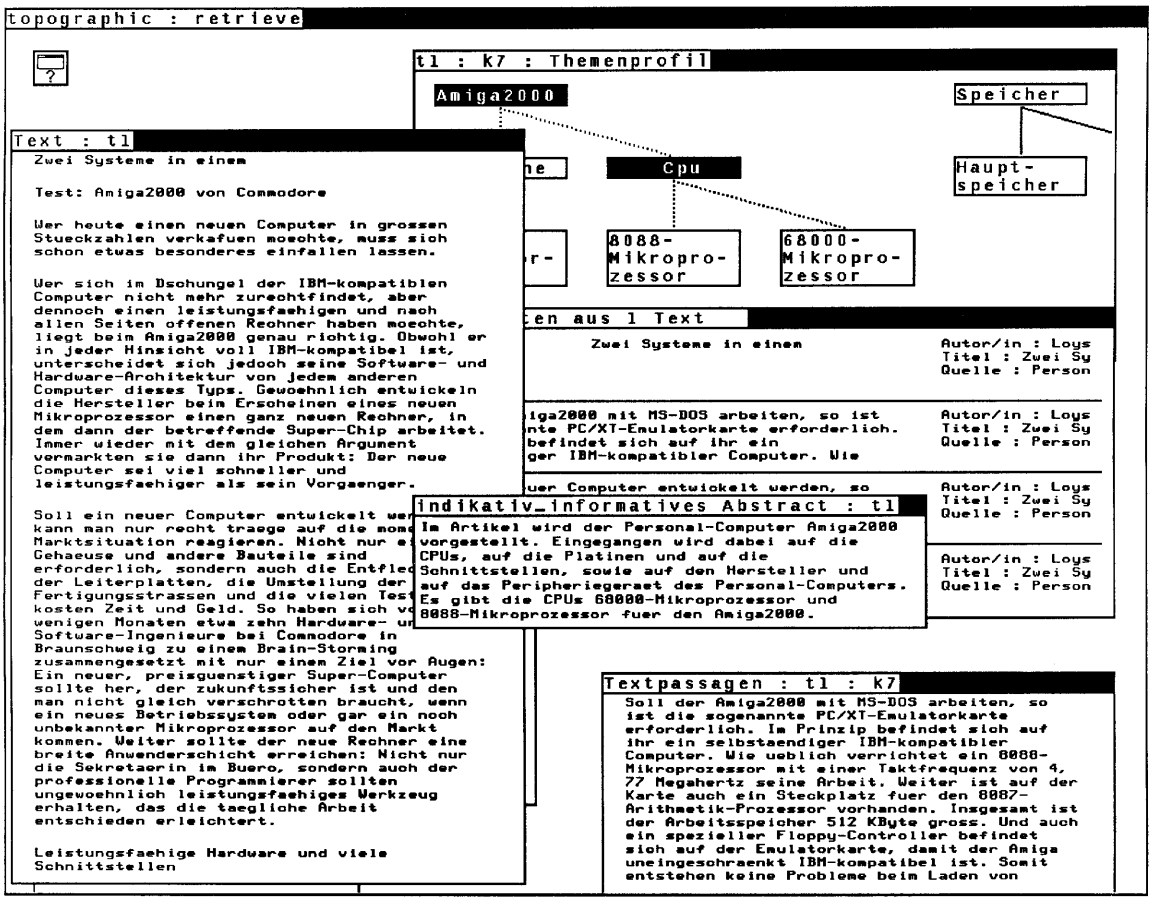


Abb. 14. Darstellung eines gesamten Bildschirms mit den Teilobjekten: Textpassage, Abstract und Volltext (im Hintergrund noch die Liste der relevanten Konstituenten und ein Themenprofil).
Methoden auf der Textpassage
zoom: Abstract bzw. Volltext präsentieren
Methoden auf dem Abstract
zoom: Volltext präsentieren
Methoden auf dem Volltext
zoom: Textgraph präsentieren
browse: weiteren relevanten Text darstellen

Mittels *Zoom* wird nun eine Liste der bezüglich der Anfrage relevanten Textkonstituenten unter Angabe bibliographischer Daten (Titel etc.) und eines kurzen Textausschnittes produziert, die nach einem Relevanzmaß sortiert ist, das sich aus dem Grad der Übereinstimmung von Suchprofil und thematischem Profil der Passage unter Berücksichtigung der Aktivierungsgewichte berechnet (Abb. 11). Gleichzeitig wird das Fenster, das den konzeptuellen Graphen (Wissensbasis) enthält, auf ein Minimum verkleinert, um die Aufmerksamkeit des Benutzers auf die in dieser Phase des Dialogs wichtigeren textorientierten Kaskadierungsstufen zu lenken.

Die zur Textpassage gehörigen Themenbeschreibungsgraphen können durch *Zoom* auf die Kurzbeschreibung der Textkonstituente gezeigt werden (Abb. 12). Von hier aus ergeben sich zwei Anwendungsmöglichkeiten für den *Zoom*-Operator: Wendet man ihn auf den gesamten Themenbeschreibungsgraphen an, so erhält man eine textuelle Version der Passage, während ein *Zoom* auf ein einzelnes Konzept des Themenprofils das bezüglich dieses Begriffs aus dem Text extrahierte Faktenwissen in einer Tabelle zur Darstellung (Abb. 13) bringt. Von der Passage ausgehend kann nun weitere Textinformation durch den Zugriff auf den Volltext bereitgestellt werden. Zur weiteren Relevanzentscheidung hat der Benutzer die Möglichkeit, ein situationsangepaßtes Abstract (indikativ oder indikativ-informativ) anzufordern, bevor er zum Gesamttext übergeht (Abb. 14). Mit der Darstellung des Volltextes

(Abb. 14) als letzter Kaskadierungsstufe kehren wir zum Ausgangspunkt zurück.

Wir waren davon ausgegangen, daß die zunehmende Verbreitung von maschinengespeicherten Volltexten Gefahr und Herausforderung zugleich ist. Der Gefahr, daß Wissen in maschinellen Speichern verschwindet, ohne zu Information zu werden, kann nur durch, allerdings aufwendige, wissensbasierte Textanalyse begegnet werden. Diese wiederum erlaubt eine reich strukturierte Retrievalkomponente, mit der flexibel auf unterschiedliche Bedürfnisse und Situationen eingegangen werden kann. Das Prinzip der kaskadierten Darstellung von Textwissen in TWRM-TOPOGRAPHIC ist eine mögliche Antwort auf die Volltext-Herausforderung.

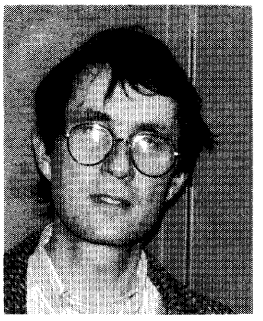
Danksagung. In diesem Beitrag sollte in erster Linie über die Ergebnisse des Projekts TWRM-TOPOGRAPHIC berichtet werden. Diese setzen auf den im früheren Projekt TOPIC von Dr. Udo Hahn (jetzt Universität Passau) und Dr. Ulrich Reimer (Universität Konstanz) erarbeiteten Leistungen auf. Ihnen sei herzlich für die „Vorarbeit" gedankt. Das Projekt hat Dipl.-Inform. (FH) Arno Weber über die ganze Zeit hinweg mit erheblichen Systemleistungen unterstützt, und bei der Texterstellung, nicht zuletzt bei der Graphik, haben wir auf den Einsatz und die Ideen von Frau Dr. Jutta Thellmann vertrauen können. Darüber hinaus sei den vielen Fachkollegen/innen gedankt, die im Laufe der Zeit das Projekt durch Kritik und Anregungen weitergebracht haben.

Literatur

1. Bates, M.J.: An exploratory paradigm for online information retrieval. In: Brooks, B.C. (ed.): Intelligent information systems for the information society. Proceedings of the 6th International Research Forum in Information Science (IRFIS 6), Frascati, Italy, 1985, pp.91-99, Amsterdam, Oxford, New York: North-Holland 1986
2. Card, S.K., Moran, T.P., Newell, A.: The psychology of human computer interaction. Hillsdale, N.J., London: Lawrence Erlbaum Ass. 1983
3. Cook, M.E., Lehnert, W.G., McDonald, D.D.: Conveying implicit content in narrative summaries. In: COLING-84: Proceedings of the 10th International Conference on Computational Linguistics. pp.5-7, Stanford 1984
4. Croft, W.B.: Approaches to intelligent information retrieval. In: Inf. Process. Manage. **23,** 249-254 (1987)
5. Daneš, F.: Functional sentence perspective and the organization of the text. In: Daneš, F. (ed.) Papers on functional sentence perspective, pp.106-128. Prague: Academia 1974
6. Danlos, L.: Conceptual and linguistic decisions in generation. In: COLING-84: Proceedings of the 10th International Conference on Computational Linguistics. pp.501-504, Stanford 1984
7. Fukunaga, K., Hirose, S.: An experience with a prolog-based object-oriented language. In: OOPSLA '86 Proceedings, pp.224-231 (1986)
8. Fum, D., Guida, G., Tasso, C.: Forward and backward reasoning in automatic abstracting. In: Horecky, J. (ed.): COLING-82: Proceedings of the 9th International Conference on Computational Linguistics, Prague, 1982. pp.83-88. Prague: Academia 1986
9. Goldberg, A., Robson, D.: Smalltalk-80. Reading: Addison-Wesley 1983
10. Grice, H.P.: Logic and conversation. In: Cole, R., Morgan, J.L. (eds.) Syntax and semantics. Vol.3: Speech Acts. pp.41-58. New York, Academic Press 1975
11. Hahn, U., Hammwöhner, R., Kuhlen, R., Reimer, U., Thiel, U.: TOPIC II/TOPOGRAPHIC II. Automatische Textkondensierung und text-orientiertes Informationsmanagement - Projektziele - State of the Art. Konstanz 1984 (Bericht TOPIC 12/84 & TOPOGRAPHIC-3/84)
12. Hahn, U.: Methoden der Volltextverarbeitung in Informationssystemen. Ein State-of-the-art-Bericht. In: Kuhlen, R. (ed.) Informationslinguistik. pp.195-216. Tübingen: Niemeyer 1986
13. Hahn, U.: Lexikalisch verteiltes Text-Parsing: Eine objektorientierte Spezifikation eines Wortexpertensystems auf der Grundlage des Aktorenmodells. Konstanz: Universität Konstanz, Sozialwissenschaftliche Fakultät, 1987 (Dissertation). Informatik-Fachberichte. Berlin, Heidelberg, New York: Springer 1989
14. Hahn, U., Reimer, U.: TOPIC essentials. In: COLING-86: Proceedings of the 11th International Conference on Computational Linguistics. pp.497-503, Bonn 1986
15. Hammwöhner, R., Kuhlen, R., Thiel, U.: TWRM-TOPOGRAPHIC: Automatische Textkondensierung mit flexiblem graphikgestütztem Retrieval und Wissensrepräsentation. Universität Konstanz, Informationswissenschaft 1987 (Bericht TOPOGRAPHIC-8/87)
16. Hammwöhner, R., Thiel, U.: Die Behandlung graphischer Seiteneffekte beim Backtracking in Prolog. In: Tagungsband der 16.Jahrestagung der Gesellschaft für Informatik in Berlin, Bd.1, pp.314-321. Berlin, Heidelberg, New York: Springer 1986
17. Hobbs, J.R.: Why is discourse coherent? In: Neubauer (ed.): Coherence in natural language texts, pp.29-70. Hamburg: Buske 1983
18. Kommers, P.A.M.: Webteaching as a design consideration for the adaptive presentation of textual information. In: Van der Veer, G.C., Tauber, M.J., Green, T.R.G., Gorny, P. (eds.): Readings on cognitive ergonomics - mind and computers. Proceedings of the 2nd European Conference, Gmunden, Austria, 1984. In: Lecture Notes in Computer Science, Vol.178, pp.161-169. Berlin, Heidelberg, New York: Springer 1984
19. Kuhlen, R.: Verarbeitung von Daten, Repräsentation von Wissen. Erarbeitung von Information, Primat der Pragmatik bei informationeller Sprachverarbeitung. In: Endres-Niggemeyer, B. (ed.): Sprachverarbeitung in Information und Dokumentation. Jahrestagung der Gesellschaft für Linguistische Datenverarbeitung (GLDV). Informatik Fachberichte, Vol.174, pp.1-22. Berlin, Heidelberg, New York: Springer 1985
20. Kuhlen, R.: Some similarities and differences between intellectual and machine text understanding for the purpose of abstracting. In: Kuhlen, R. (ed.): Informationslinguistik, pp.133-151. Tübingen: Niemeyer 1986
21. Kuhlen, R.: Information Retrieval: Verfahren des Abstracting. In: Lenders, W. (ed.): Computational linguistics. Handbücher zur Sprach- und Kommunikationswissenschaft. Berlin, New York: W.De Gruyter 1989
22. Mann, W.C., Moore, J.A.: Computer generation of multiparagraph english text. Am. J. Comput. Linguist. **7,** 17-29 (1981)
23. Mann, W.C.: Discourse structures for text generation. In: COLING-84: Proceedings of the 10th International Confer-

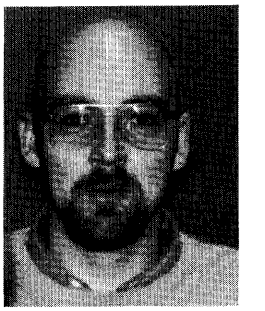
ence on Computational Linguistics, pp. 367–375, Stanford 1984
24. McKeown, K. R.: Language generation: applications, issues, and approaches. Proc. IEEE *74*, 961–968 (1986)
25. PADOK: Test und Vergleich von Texterschließungssystemen für das Deutsche Patent- und Fachinformationssystem. Endbericht. 1.1. 1985–31.3. 1986. Regensburg: Univ. Regensburg, Linguistische Informationswissenschaft 1986
26. Reimer, U.: FRM: Ein Frame-Repräsentationsmodell und seine formale Semantik. Zur Integration von Datenbank- und Wissensrepräsentationsansätzen. In: Informatik-Fachberichte, Vol. 198. Berlin, Heidelberg, New York: Springer 1989
27. Rouse, W. B., Rouse, S. H.: Human information seeking and design of information systems. Inf. Process. Manage. *20*, 129–138 (1984)
28. Salton, G., McGill, M. J.: Information Retrieval – Grundlegendes für Informationswissenschaftler. Hamburg, New York: McGraw Hill 1986 (engl. Erstausgabe 1983)
29. Sibert, J. L., Hurley, W. D., Bleser, T. W.: An object oriented user interface management system. Comput. Graphics *20*, 259–268 (1986)
30. Sonnenberger, G.: Flexible Generierung von natürlichsprachigen Abstracts aus Textrepräsentationsstrukturen. In: Trost, H. (ed.): 4. Österreichische Artificial-Intellience-Tagung. Wiener Workshop – Wissensbasierte Sprachverarbeitung, pp. 72–82. Berlin, Heidelberg, New York: Springer 1988
31. Stefik, M., Bobrow, D. G.: Object-oriented programming: themes and variations. In: The AI-Magazine *6*, 40–62 (1986)

Eingegangen am 21. Juli 1988
Angenommen am 17. Februar 1989

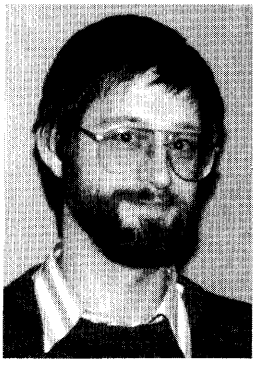


Rainer Kuhlen, geb. 1944 in Potsdam; ab 1964 Studium der Philosophie, Germanistik und Soziologie an der Universität Münster, 1969–1972 Assistent für Philosophie an der Universität Münster; Postgraduierten-Ausbildung in der Informationswissenschaft an der Zentralstelle für maschinelle Dokumentation, Frankfurt; 1976 Promotion an der Universität Regensburg auf dem Gebiet der Informationslinguistik über morphologische Reduktionsalgorithmen. Berufliche Tätigkeit: 1974–1979 Dozent am Lehrinstitut für Dokumentation in Frankfurt, zuständig für die Ausbildung wissenschaftlicher Dokumentare, speziell für Informationswissenschaft, Information Retrieval und Informationslinguistik; 1979/80 Vertretung eines Informatik-Lehrstuhls in Koblenz im Studiengang Angewandte Informatik mit Ausrichtung Linguistische Datenverarbeitung; seit 1980 Lehrstuhl für Informationswissenschaft an der Universität Konstanz. Arbeitsgebiete: Information Retrieval, Wissensbasierte Informationssysteme, Informationslinguistik, Ausbildungsforschung, Informationsverarbeitung in Organisationen, Informationsmanagement, Informationspolitik, Grundlagenprobleme der Informationswissenschaft.

Rainer Hammwöhner, geb. 1956; 1983 Diplom der Med. Informatik in Heidelberg/Heilbronn, ab 1983 wissenschaftlicher Angestellter am Lehrstuhl Informationswissenschaft in den Projekten TOPOGRAPHIC (1983–1984), TOPICII/TOPOGRAPHICII (1985–1986), TWRM-TOPOGRAPHIC (seit 1987), Arbeitsgebiete: Information Retrieval, Hypertextsysteme, graphische Benutzerschnittstellen, Diskursmodellierung für graphisches Retrieval bzw. Hypertexte.



Gabriele Sonnenberger, geb. 1957; 1982 Diplom der Informatik an der FH Konstanz, mehrjährige Berufserfahrung in der Systemprogrammierung, 1987 Diplom der Informationswissenschaft an der Universität Konstanz. Seit April 1987 wissenschaftliche Angestellte im Projekt TWRM-TOPOGRAPHIC. Arbeitsgebiet: Natürlichsprachige, situationsadäquate Präsentation von Textwissen. Interessengebiete: Textanalyse/Textverstehen; Repräsentation von Textwissen; Textgenerierung; Textlinguistik; Wissensrepräsentation, Information Retrieval.



Ulrich Thiel, geb. 1955; 1983 Diplom in Informatik an der Universität Dortmund, seit 1983 wissenschaftlicher Angestellter am Lehrstuhl Informationswissenschaft in den Projekten TOPOGRAPHIC (1983–1984), TOPICII/TOPOGRAPHICII (1985–1986) und TWRM-TOPOGRAPHIC (seit 1987); Arbeitsgebiete: Wissensbasierte Ansätze zur Mensch-Maschine-Interaktion, speziell Diskurs- und Benutzermodellierung für graphische Retrievalschnittstellen.