



Deutsches
Forschungszentrum
für Künstliche
Intelligenz GmbH

**Research
Report**
RR-94-02

Von Textgeneratoren zu Intellimedia-Präsentationssystemen

Elisabeth André, Thomas Rist

February 1994

**Deutsches Forschungszentrum für Künstliche Intelligenz
GmbH**

Postfach 20 80
67608 Kaiserslautern, FRG
Tel.: (+49 631) 205-3211/13
Fax: (+49 631) 205-3210

Stuhlsatzenhausweg 3
66123 Saarbrücken, FRG
Tel.: (+49 681) 302-5252
Fax: (+49 681) 302-5341

Deutsches Forschungszentrum für Künstliche Intelligenz

The German Research Center for Artificial Intelligence (Deutsches Forschungszentrum für Künstliche Intelligenz, DFKI) with sites in Kaiserslautern and Saarbrücken is a non-profit organization which was founded in 1988. The shareholder companies are Atlas Elektronik, Daimler-Benz, Fraunhofer Gesellschaft, GMD, IBM, Insiders, Mannesmann-Kienzle, SEMA Group, and Siemens. Research projects conducted at the DFKI are funded by the German Ministry for Research and Technology, by the shareholder companies, or by other industrial contracts.

The DFKI conducts application-oriented basic research in the field of artificial intelligence and other related subfields of computer science. The overall goal is to construct *systems with technical knowledge and common sense* which - by using AI methods - implement a problem solution for a selected application area. Currently, there are the following research areas at the DFKI:

- Intelligent Engineering Systems
- Intelligent User Interfaces
- Computer Linguistics
- Programming Systems
- Deduction and Multiagent Systems
- Document Analysis and Office Automation.

The DFKI strives at making its research results available to the scientific community. There exist many contacts to domestic and foreign research institutions, both in academy and industry. The DFKI hosts technology transfer workshops for shareholders and other interested groups in order to inform about the current state of research.

From its beginning, the DFKI has provided an attractive working environment for AI researchers from Germany and from all over the world. The goal is to have a staff of about 100 researchers at the end of the building-up phase.

Friedrich J. Wendl
Director

Die vorliegende Veröffentlichung ist urheberrechtlich geschützt durch die Deutsche Forschungsgemeinschaft (DFG).

Von Textgeneratoren zu Intellimedia-Präsentationssystemen

Elisabeth André, Thomas Rist

DFKI-RR-94-02

© Deutsche Forschungsgemeinschaft (DFG) 1994. Alle Rechte vorbehalten. Nachdruck, Vervielfältigung und Verbreitung, auch auszugsweise, ist ohne schriftliche Genehmigung der Deutschen Forschungsgemeinschaft. Die Deutsche Forschungsgemeinschaft ist eine eingetragene Stiftung des öffentlichen Rechts. Die DFG ist ein Mitglied der International Association of Scientific Academies (IASA). Die DFG ist ein Mitglied der International Association of Universities (IAU). Die DFG ist ein Mitglied der International Association of Agricultural Universities (IAAU). Die DFG ist ein Mitglied der International Association of Agricultural Universities (IAAU). Die DFG ist ein Mitglied der International Association of Agricultural Universities (IAAU).

Die vorliegende Veröffentlichung erschien auch in *Künstliche Intelligenz (KI) 2/93*.

Diese Arbeit wurde finanziell unterstützt durch das Bundesministerium für Forschung und Technologie (FKZ ITW-8901 8).

© Deutsches Forschungszentrum für Künstliche Intelligenz 1994

This work may not be copied or reproduced in whole or in part for any commercial purpose. Permission to copy in whole or in part without payment of fee is granted for nonprofit educational and research purposes provided that all such whole or partial copies include the following: a notice that such copying is by permission of Deutsches Forschungszentrum für Künstliche Intelligenz, Kaiserslautern, Federal Republic of Germany; an acknowledgement of the authors and individual contributors to the work; all applicable portions of this copyright notice. Copying, reproducing, or republishing for any other purpose shall require a licence with payment of fee to Deutsches Forschungszentrum für Künstliche Intelligenz.

Von Textgeneratoren zu Intellimedia-Präsentationssystemen*

Elisabeth André, Thomas Rist

Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI)

Stuhlsatzenhausweg 3, D-66123 Saarbrücken

Email: andre/rist@dfki.uni-sb.de

3. März 1994

Zusammenfassung:

Während Textgeneratoren ausschließlich auf das Medium Sprache zurückgreifen, nutzen Intellimedia-Präsentationssysteme die individuellen Stärken unterschiedlicher Medien, wie Text, Graphik, Gestik und Animation, für die Informationsdarbietung. Die aus kommunikationstheoretischer Sicht allgemeinere Aufgabe wirft einerseits neue interessante Probleme auf, etwa die Selektion und die Koordination von Medien, führt aber andererseits zu einer umfassenderen Behandlung von Fragestellungen, die bereits von der Textgenerierung her bekannt sind. Dieses Papier stellt die erste Generation von sprachverarbeitenden Intellimedia-Präsentationssystemen vor, skizziert die neuen Problemstellungen und beschäftigt sich insbesondere mit der Frage, inwiefern sich Methoden zur Textgenerierung verallgemeinern lassen, damit sie für die Informationspräsentation mit mehreren Medien anwendbar sind.

Abstract:

While text generators exclusively rely on a single medium, intellimedia presentation systems take advantage of the individual strength of several media, such as text, graphics, gestures and animations, to present information. On the one hand, new interesting problems arise in the broader context of multimedia communication, in particular the selection and the coordination of media. On the other hand, this research leads to a more general treatment of problems already known from text generation. The paper discusses the first generation of NL processing intellimedia presentation systems and sketches the new problems. Particular emphasis is given to the question of how to generalize methods for text generation in such a way that they become useful for the production of multimedia presentations, too.

*Die vorliegende Arbeit wurde im Rahmen des vom BMFT geförderten Projekts WIP durchgeführt.

Inhaltsverzeichnis

1	Einführung	3
2	Erste Protoypen	4
3	Zentrale Verarbeitungsprozesse	6
4	Ansätze zur Textgenerierung als methodologische Basis	7
4.1	Generierung von Text-Bild-Kombinationen als zielgerichtete Tätigkeit . . .	7
4.2	Ein erweiterter Kohärenzbegriff für Text-Bild-Kombinationen	8
4.3	Ansätze zur Strukturierung von Text-Bild-Kombinationen	10
5	Medienkoordination	11
5.1	Auswahl von Präsentationsmedien	12
5.2	Generierung von Querverweisen und multimedialen Referenzen	13
5.3	Räumliche und zeitliche Koordination der Ausgabe	15
6	Aufbau von IMP-Systemen	15
7	Schlußbemerkung	17

1 Einführung

Multimedia-Systeme, in denen unterschiedliche Medien wie beispielsweise Graphik, Sprache, Musik und Animation zur Informationspräsentation zur Verfügung stehen, sind seit Ende der 80er Jahre unaufhaltsam auf dem Vormarsch. Ein Gang über eine beliebige Computerfachmesse zeigt, daß immer mehr Anbieter Multimedia-Systeme in ihre Angebotspalette aufnehmen und eingeführte Produkte um gewisse Multimedia-Funktionen erweitern. Solche Funktionen umfassen die Darstellung, Speicherung, Verbreitung bis hin zur Erzeugung und Verarbeitung multimedialer Dokumente. Für die Informationspräsentation bietet die Unterstützung mehrerer Medien nicht nur Wahlmöglichkeit, sondern vor allem auch Kombinierbarkeit. Multimedia-Systeme sind besonders dann attraktiv, wenn sich Information erst durch eine geschickt gewählte Kombination mehrerer Medien effektiv darbieten läßt, wobei unter Effektivität leicht verständlich, schnell wahrnehmbar, einprägsam, ansprechend usw. zu verstehen ist - je nach Zweck einer Präsentation.

Intelligente Präsentationssysteme als Komponenten von Benutzerschnittstellen zu Hilfesystemen, Expertensystemen oder intelligenten Betriebswarten der nächsten Generation zielen darauf ab, vorliegendes Wissen benutzerspezifisch und der jeweiligen Präsentationssituation angemessen auszuwählen und darzubieten. Dies setzt ein hohes Maß an Flexibilität bei der Informationspräsentation voraus; unterschiedliche Benutzer in unterschiedlichen Situationen erfordern fast immer unterschiedliche Präsentationen. Weil es im allgemeinen extrem aufwendig - wenn nicht gar unmöglich - ist, potentielle Benutzer und Präsentationssituationen vollständig zu antizipieren und dafür entsprechende Präsentationen vorzubereiten, ist es naheliegend, diese erst bei Bedarf unter Berücksichtigung der aktuellen Gegebenheiten zu erzeugen. Dieser Gedanke führt zum intelligenten, im Idealfall universell einsetzbaren Präsentationsexperten, der neben fachspezifischem Präsentationwissen über grundlegende kommunikative Fähigkeiten verfügt. Ein solches Präsentationssystem wäre sowohl für den Benutzer als auch für den Ersteller einer Anwendung vorteilhaft. Während der Benutzer auf ihn individuell abgestimmte Präsentationen erwarten darf, wird der Ersteller wesentlich entlastet, da er sich auf die eigentlichen Probleme der Anwendung konzentrieren kann.

Intellimedia-Präsentationssysteme (IMP-Systeme) stehen für den Versuch, Multimedia-Technologie für eine intelligente Informationspräsentation zu nutzen. Dazu genügt es allerdings nicht, eine Workstation mit den technischen Notwendigkeiten für den Einsatz multipler Medien auszurüsten. Die Schwierigkeit besteht vielmehr darin, die individuellen Stärken der zur Verfügung stehenden Medien geschickt einzusetzen und unterschiedliche Medien in einer Multimedia-Präsentation in konsistenter und kohärenter Weise zu integrieren. Daß es sich hierbei keinesfalls um eine triviale Aufgabe handelt, läßt sich anhand zahlreicher mehr oder weniger mißglückter Multimedia-Präsentationen belegen. Zu den Beispielen zählen gedruckte illustrierte Gebrauchsanweisungen, in denen Bilder nichts zum Verständnis beitragen, wenn nicht gar Verwirrung stiften, ebenso wie die vielen, aufwendig gestalteten Computerpräsentationen, die den Benutzer letztendlich einer nicht zu bewältigenden Reizüberflutung aussetzen. Um die zur Verfügung stehenden Medien

optimal einsetzen zu können, muß ein IMP-System zum einen die medien-spezifische Enkodierung von Information beherrschen - tritt also sowohl als Textautor, Graphikdesigner, Drehbuchautor, Komponist usw. in Erscheinung. Zum anderen kommt ihm die Rolle des Medienkoordinators zu, der ähnlich einem Desktoppublisher verschiedene Medien in einem Dokument integriert.

Aus der Sicht der Textgenerierung sind IMP-Systeme aus zwei Gründen interessant. Zum einen macht Sprache als eines der wichtigsten Kommunikationsmedien Textgeneratoren zu unentbehrlichen Komponenten von IMP-Systemen. Zum anderen haben Verfahren zur Textgenerierung zunehmend an Reife gewonnen, so daß es erfolgversprechend ist, zu untersuchen, inwiefern sich die dort entwickelten Ansätze auch für den allgemeineren Fall der Informationspräsentation mit mehreren Medien übertragen lassen. In den letzten Jahren wurden in einer Reihe von Projekten ausgehend von Fragestellungen der Textgenerierung Systeme entwickelt, die neben Sprache auch Medien wie Graphik, Gestik und animierte Darstellungen heranziehen.

2 Erste Prototypen

Aus funktionaler Sicht läßt sich jedes der bisher entwickelten sprachverarbeitenden IMP-Systeme in das in Abbildung 1 gezeigte allgemeine Schema einpassen. Aufgabe eines IMP-Systems ist es, die von einer Anwendung bereitgestellte Information (Pfeil a) einem(r) Benutzer(gruppe) zu übermitteln (Pfeil b), wobei dem System unterschiedliche Medien zum Aufbau von Präsentationen zur Verfügung stehen.

Die einzelnen Systeme sind zunächst einmal hinsichtlich der vorgesehenen Anwendung voneinander zu unterscheiden. Zum einen gibt es Systeme, die als Informationssystem konzipiert sind. Hierzu zählen: ALFresco [Stock 91], das dem Benutzer Information über italienische Fresken vermittelt, das von Kerpedjiev (vgl. [Kerpedjiev 92]) entwickelte System zur Präsentation meteorologischer Daten, die Dokumentationssysteme EDWARD [Claassen 92], IDAS (Intelligent Documentation Advisory System, [Reiter et al. 92]) und IGiNG (Integrating Graphics and Natural Language Generation, [Dilley et al. 92]), die Auskunft über Dateien, Rechneranlagen oder Konferenzen geben, das kartographische Informationssystem I² (Integrated Interfaces, [Arens et al. 91]) sowie die Wegauskunftssysteme AIMI (vgl. [Maybury 91]) und VITRA-GUIDE (vgl. [Herzog et al. 93]). Zum anderen wurden Systeme als Schnittstellen zu Experten- oder Hilfesystemen entwickelt. XTRA (eXpert TRANslator, [Allgayer et al. 89]) ist ein Zugangssystem zu einem Expertensystem, das den Benutzer beim Ausfüllen eines Steuerformulars unterstützt. CUBRICON (the CUBRC Intelligent CONversationalist, [Neal & Shapiro 91]) ist als intelligente Benutzerschnittstelle zu einem militärischen Planungssystem konzipiert. COSPS [Marks 91] und SAGE (a System for Automatic and Graphical Explanation, [Roth et al. 91]) generieren Erklärungen für quantitative Modelle, z.B. im Bereich Projektmanagement oder im Finanzwesen. MMI² (A Multi-Modal Interface for Man Machine Interaction with Knowledge-Based Systems, [Wilson et al. 92]) ist eine

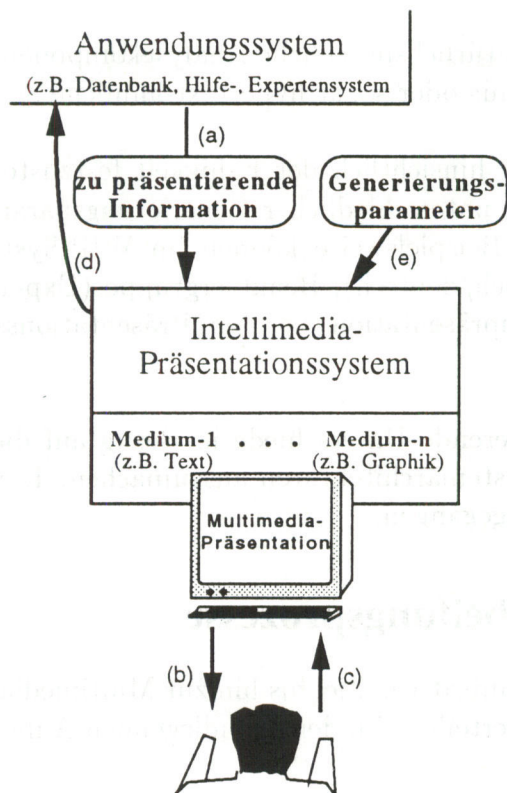


Abbildung 1: Schema für IMP-Systeme

Schnittstelle zu einem Konfigurierungssystem von Computernetzwerken. PEA (Program Enhancement Advisor, [Moore & Swartout 90]) ist mit einem Expertensystem gekoppelt, das Vorschläge zur Optimierung von Lispprogrammen liefert. COMET (COordinated Multimedia Explanation Testbed, [Feiner & McKeown 91]) und WIP (Wissensbasierte Informationspräsentation [Wahlster et al. 91; Wahlster et al. 92; Wahlster et al. 93; André et al. 93a]) übermitteln Instruktionen für die Montage, Bedienung und Reparatur technischer Geräte, wobei vorgesehen ist, daß die mitzuteilende Information von einem Diagnose- oder Hilfesystem bereitgestellt wird.

Ein augenscheinliches Vergleichskriterium sind die von den Systemen unterstützten Präsentationsmedien. Neben statischen Medien wie Text und Graphik, die in allen Systemen in irgendeiner Form zum Einsatz kommen, werden in einigen Systemen auch dynamische Medien unterstützt. Beispielsweise präsentiert ALFresco dem Benutzer Information über italienische Fresken nicht nur durch Einzelbilder, sondern auch durch kurze Videosequenzen. Die Generierung von animierten Darstellungen wird u.a. für die Systeme VITRA-GUIDE und eine Erweiterung von COMET angestrebt.

Je nachdem ob Interaktion zwischen Präsentationssystem und Benutzer bzw. einem Anwendungssystem vorgesehen ist, sind Pfeil c bzw. Pfeil d instantiiert (vgl. Abbildung 1). Die Systeme XTRA, CUBRICON, ALFresco, MMI², I², PEA und IDAS verfügen sowohl über eine Generierungs- als auch eine Analysekomponente. Um komplizierte Referenzprobleme bei der Analyse von natürlicher Sprache zu umgehen, wurde für die Systeme

I², PEA und IDAS keine natürlichsprachliche Analysekomponente entwickelt. Stattdessen werden dem Benutzer Menüs oder sogar hypertext-ähnliche Schnittstellen angeboten.

Unterschiede sind auch hinsichtlich der Fähigkeit festzustellen, ein und dieselbe Information in Abhängigkeit unterschiedlicher Generierungsparameter (Pfeil e) situationsgerecht zu präsentieren. Beispielsweise können im WIP-System Generierungsparameter wie Zielsprache (englisch/deutsch), Benutzergruppe (Experte/Laie) und Dokumenttyp (Handbuch/Bildschirmpräsentation) vor dem Präsentationssyntheseprozess eingestellt werden.

Desweiteren sind gravierende Unterschiede in Bezug auf die zugrundeliegenden Verarbeitungsprozesse und Systemarchitekturen auszumachen. Hierauf wird in den nachfolgenden Kapiteln näher eingegangen.

3 Zentrale Verarbeitungsprozesse

Der Weg von einem kommunikativen Ziel bis hin zur Multimedia-Präsentation läßt sich in mehrere Arbeitsschritte unterteilen. Zu den grundlegenden Aufgaben, die ein IMP-System zu bewältigen hat, zählen:

- Festlegung und Strukturierung des Inhalts

Hier ist zu unterscheiden, ob das Anwendungssystem den Inhalt mehr oder weniger vorgibt, so daß dieser "nur" noch zu strukturieren ist, oder ob ausgehend von einem kommunikativen Ziel die mitzuteilende Information aus einer Wissensbasis ausgewählt werden muß. Was präsentiert wird, hängt nicht zuletzt auch von dem gewählten Präsentationsmedium ab. Während Sprache zahlreiche Abstraktionsmöglichkeiten bietet, ist ein Präsentator bei der bildlichen Wiedergabe von Information oftmals gezwungen, diese zu konkretisieren. Ebenso wie die Inhaltsauswahl ist auch die Strukturierung einer Multimedia-Präsentation nicht unabhängig von dem gewählten Medium zu sehen.

- Koordinierte Verteilung der Information auf die verschiedenen Medien

Um unterschiedliche Medien so zu kombinieren, daß ihre individuellen Stärken optimal genutzt werden, müssen geeignete Kriterien gefunden werden, nach denen die mitzuteilende Information auf unterschiedliche Generatoren verteilt wird. Hierbei gilt es auch, das richtige Maß an Komplementarität und Redundanz zwischen den verwendeten Präsentationsmedien zu finden. Eine komplementär gestaltete Präsentation, die keinerlei redundant kodierte Information enthält, wirkt meist inkohärent. Wird dagegen zuviel in unterschiedlichen Medien "paraphrasiert", besteht die Gefahr, daß sich der Benutzer nach kurzer Zeit auf ein Medium konzentriert und möglicherweise Information übersieht.

- Medienspezifische Inhaltsrealisierung

Hierunter fällt die Realisierung der mitzuteilenden Information durch die unter-

schiedlichen Generatoren, wobei die Generierungsergebnisse so aufeinander abzustimmen sind, daß sich die Medien in ihrer Wirkung gegenseitig verstärken. Beispielsweise ist darauf zu achten, daß die Ergebnisse der einzelnen Generatoren strukturell verträglich sind und daß auf die Ergebnisse anderer Generatoren Bezug genommen wird.

- Integration der Generierungsergebnisse in einer Ausgabe
Während das Layout bei der Textgenerierung bisher weitgehend vernachlässigt wurde, wird es in IMP-Systemen als wichtiger Informationsträger genutzt. Liegen die Ergebnisse unterschiedlicher Generatoren vor, müssen diese unter Berücksichtigung räumlicher und zeitlicher Constraints in einer multimedialen Ausgabe integriert und dem Benutzer präsentiert werden.

4 Ansätze zur Textgenerierung als methodologische Basis

Obwohl die Generierung von Multimedia-Präsentationen neue Probleme aufwirft, fällt eine große Ähnlichkeit zu Fragestellungen auf, die bereits in mehreren Arbeiten zur Textgenerierung thematisiert wurden. Hierzu zählen Verfahren zur Planung des Inhalts und der Struktur von Texten.

4.1 Generierung von Text-Bild-Kombinationen als zielgerichtete Tätigkeit

Ausgehend von einer sprechakttheoretischen Sichtweise wird nicht nur die Generierung von Texten, sondern allgemein die Generierung von Multimedia-Präsentationen als kommunikative Handlung zur Erreichung von Zielen aufgefaßt. In [André & Rist 90b] wurde dieser Ansatz erstmals für die Generierung von Text-Bildkombinationen ausgearbeitet. Kommunikative Handlungen können sich wiederum aus anderen kommunikativen Handlungen zusammensetzen. Die Blätter der entstehenden hierarchischen Struktur sind elementare Sprechakte (vgl. [Searle 80]) oder Bildakte (vgl. [Kjorup 78]). Beispiele für Sprechakte sind *Behaupten*, *Auffordern* oder *Bitten*, während Handlungen wie *Annotieren*, *Darstellen* oder *Hervorheben* zu den Bildakten zählen. Durch die mit den einzelnen Handlungen assoziierten Ziele ergibt sich neben einer hierarchischen Handlungsstruktur auch eine hierarchische Struktur von Zielen (*intentionale Struktur*). Dabei tragen die Handlungen, die zur Erreichung eines untergeordneten Ziels dienen, zur Erreichung übergeordneter Ziele bei.

Abbildung 2 zeigt einen aus einer Bedienungsanleitung entnommenen Dokumentausschnitt und die ihm zugrundeliegende intentionale Struktur. Mit diesem Dokumentausschnitt soll erreicht werden, daß der Benutzer die Verschußkappe eines Wasserbehälters entfernt. Dieses Ziel spiegelt sich im wesentlichen durch die verbale Aufforderung "Entfernen Sie die Verschußkappe" wider. Aufgrund der Aufforderung weiß der Benutzer, daß der Präsentator möchte, daß er die Verschußkappe entfernt. Ob er jedoch dazu in der Lage ist oder den Wünschen des Präsentators nachkommen möchte, ist damit nicht unbedingt

sichergestellt. Diese beiden Ziele sollen durch das Bild und die verbale Äußerung "um den Wasserbehälter zu füllen" erreicht werden. Auch den einzelnen Bildteilen können bestimmte Ziele zugeordnet werden. Beispielsweise soll durch die beiden Pfeile sichergestellt werden, daß der Benutzer den Bewegungsverlauf der von ihm auszuführenden Handlungsschritte kennt.

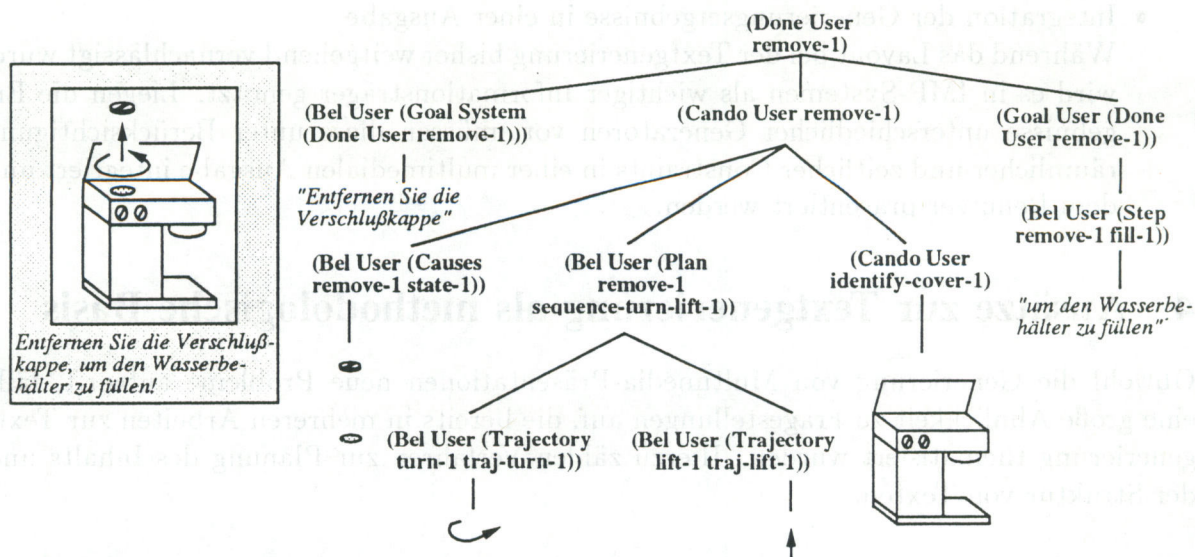


Abbildung 2: Intentionale Struktur eines Dokumentausschnitts

4.2 Ein erweiterter Kohärenzbegriff für Text-Bild-Kombinationen

Mehrere Textlinguisten haben Kohärenz durch Relationen zwischen einzelnen Diskurssegmenten charakterisiert (vgl. [Grimes 75; Hobbs 78; Mann & Thompson 87]). Ein Text wird nur dann als kohärent betrachtet, wenn er in Segmente zerlegt werden kann, die durch bestimmte Relationen miteinander verbunden sind. Beispielsweise haben Mann und Thompson ausgehend von einer umfangreichen Analyse der rhetorischen Struktur kohärenter Texte etwa 25 solcher Relationen identifiziert, die die Grundlage der von ihnen entwickelten *Rhetorical Structure Theory* (RST) bilden. Zu den RST-Relationen zählen beispielsweise *Enablement*, *Motivation*, *Condition* oder *Elaboration*. Im Hinblick auf einen verallgemeinerten Kohärenzbegriff stellt sich die Frage, welche Beziehungen zwischen den Dokumentteilen einer kohärenten multimedialen Präsentation bestehen.

In kognitionspsychologischen Arbeiten wurde untersucht, welche Rolle ein bestimmtes Bild in Bezug auf eine begleitende Textpassage spielt. Beispielsweise hat Levin [Levin et al. 87] fünf Grundfunktionen von Bildern erkannt: *Representation*, *Organization*, *Interpretation*, *Transformation* und *Decoration*. Bei der Untersuchung von Lehrbüchern schrieben Hunter und Kollegen (vgl. [Hunter et al. 87]) Bildern vor allem die Funktionen *Embellish*, *Reinforce*, *Elaborate*, *Summarize* und *Compare* zu. Bandyopadhyay [Bandyopadhyay 90]

unternahm den Versuch, die von Hobbs [Hobbs 78] vorgeschlagenen Kohärenzrelationen auf Bilder zu übertragen.

Leider sind die von den Text-Bild-Forschern vorgeschlagenen Definitionen sehr allgemein gehalten. Darüberhinaus wird nur die Funktion von ganzen Bildern betrachtet, d.h. es wird nicht untersucht, aus welchen Teilen sich ein Bild zusammensetzt und welche Rolle diese für die Gesamtinterpretation spielen. Um jedoch eine vollständige Beschreibung der gesamten Präsentationsstruktur zu erhalten, müssen auch die Beziehungen zwischen einzelnen Bildteilen und zwischen Bildteilen und Text explizit repräsentiert werden.

Untersuchungen illustrierter Dokumente (vgl. [André & Rist 93; Arens et al. 93b; Roth & Mattis 90]) zeigen, daß sich Beziehungen zwischen Bildern/Bildteilen untereinander und zwischen Bildern/Bildteilen und Text oft durch die von Textlinguisten gefundenen Relationen beschreiben lassen. Beispiele sind *Elaboration* bei Haupt- und Nebensbildern, *Sequence* bei Bildfolgen oder *Interpretation* bei einem Bild und der dazugehörigen Erklärung im Text. Andererseits gibt es Relationen, wie beispielsweise *Concession* oder *Condition*, die nicht ohne weiteres bildlich ausdrückbar sind. Schließlich treten Relationen auf, die in keiner der oben angegebenen Arbeiten zur Textkohärenz zu finden sind, etwa die Relation *Label* zwischen einem Text und seinem Titel oder zwischen Bildobjekten und ihrer Beschriftung.

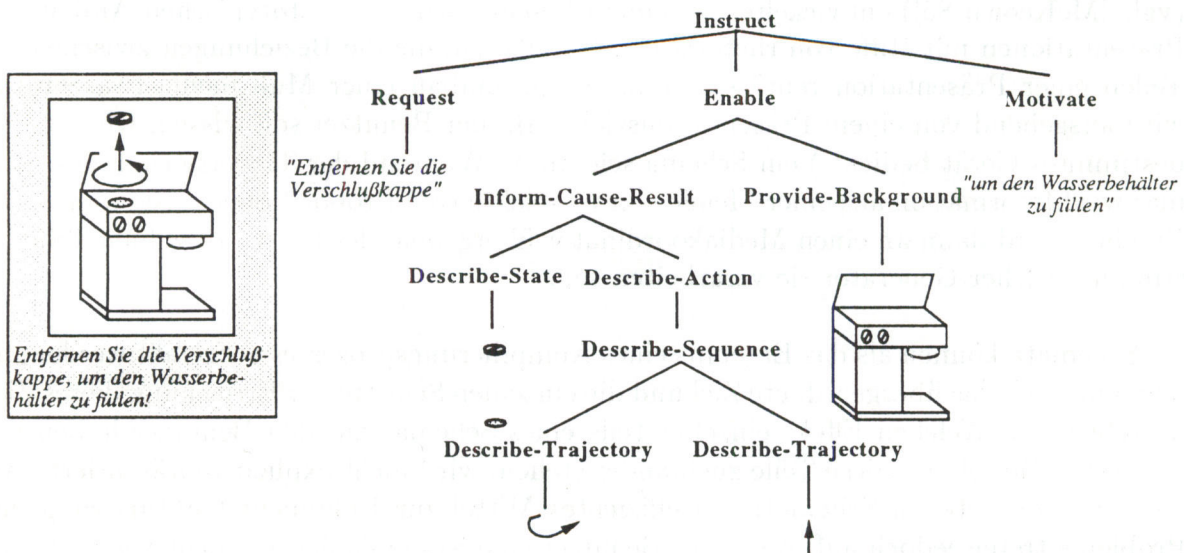


Abbildung 3: Rhetorische Struktur des Dokumentausschnitts aus Abbildung 2

Abbildung 3 zeigt die rhetorische Struktur des Dokumentausschnitts aus Abbildung 2. Das Dokument setzt sich zusammen aus einer Aufforderung, einem Motivationsteil und einem Teil, der die Ausführung der Handlung ermöglichen soll. Auch den einzelnen Bildteilen können unterschiedliche rhetorische Funktionen zugeordnet werden. Beispielsweise

dient die Darstellung der Espressomaschine als Hintergrund für den Rest des Bildes.

4.3 Ansätze zur Strukturierung von Text-Bild-Kombinationen

Wie in den vorangegangenen Abschnitten dargelegt, folgen Bilder und Text-Bild-Kombinationen ähnlichen Strukturierungskriterien wie Text. Insbesondere ist eine Text-Bild-Kombination charakterisiert durch ihre *intentionale Struktur*, die die Ziele des Präsentators widerspiegelt, und ihre *rhetorische Struktur*, die durch rhetorische Relationen bestimmt wird.

Es bietet sich daher an, Ansätze zur Textstrukturierung nicht nur zur Organisation der textuellen Teile einer multimedialen Präsentation, sondern zur Strukturierung der gesamten Präsentation zu verwenden. Ein wesentlicher Vorteil eines einheitlichen Strukturierungsansatzes ist darin zu sehen, daß nicht nur Beziehungen innerhalb eines Mediums, sondern auch Beziehungen zwischen Teilen in unterschiedlichen Medien explizit repräsentiert werden können. In Systemen wie SAGE [Roth et al. 91], in denen medienspezifische Verfahren zur Inhaltsstrukturierung zum Einsatz kommen, geht hingegen der Zusammenhang zwischen Text und Graphik verloren.

In COMET werden zur Strukturierung der Gesamtpräsentation Schemata verwendet (vgl. [McKeown et al. 90]), die McKeown ursprünglich zur Strukturierung von Texten (vgl. [McKeown 85]) entwickelte. Schemata beschreiben den prototypischen Aufbau von Präsentationen mit Hilfe von rhetorischen Prädikaten, die die Beziehungen zwischen den Teilen einer Präsentation repräsentieren. Beim Aufbau einer Multimedia-Präsentation wird ausgehend von einem Präsentationsziel (z.B. 'der Benutzer soll wissen, wie man ein bestimmtes Gerät bedient') ein Schema selektiert. Während der Traversierung des Schemas wird Information aus einer Menge von vorselektierten Propositionen ausgewählt. Das Ergebnis wird dann an einen Mediakoordinator übergeben, der für die einzelnen Teile bestimmt, welcher Generator sie verarbeiten soll.

Schemata können als das Ergebnis eines Kompilierungsprozesses aufgefaßt werden, bei dem nur noch das übergeordnete Ziel und die einzelnen Schritte zur Erreichung dieses Ziels übrigbleiben. Welchen Effekt einzelne Teile eines Schemas auf den Benutzer haben und in welcher Beziehung diese Teile zueinander stehen, wird nicht explizit repräsentiert. Aus diesem Grund stellen Schemata ein effizientes Mittel zur Dokumentstrukturierung dar. Probleme treten jedoch auf, wenn der Benutzer eine Systemäußerung nicht versteht. Auf Schemata basierende Systeme haben dann nur noch die Möglichkeit, ein anderes Schema auszuwählen. Sie sind nicht in der Lage, nur auf einen Teil des zum Schema korrespondierenden Textes oder Bildes näher einzugehen (siehe auch [Paris 91]).

Neben schemabasierten Ansätzen werden bei der Synthese multimedialer Dokumente planbasierte Ansätze verwendet, die auf Erweiterungen von Textplanungsansätzen (vgl. [Moore & Paris 89; Maybury 89]) beruhen. Um multimediale Präsentationen zu planen, faßt man die zu Beginn dieses Kapitels eingeführten kommunikativen Handlungen als

Operatoren eines Planungssystems auf. Beim Planungsprozeß wird ausgehend von einem kommunikativen Ziel, z.B. der Benutzer soll in der Lage sein, eine bestimmte Aktion auszuführen, nach Planungsoperatoren gesucht, deren Effekt dieses Ziel subsumiert. Wird ein solcher Planungsoperator gefunden, so werden die Ausdrücke im Rumpf als neue Teilziele bearbeitet. Handelt es sich dabei um elementare Sprechakte oder Bildakte, so werden diese an die medienspezifischen Generatoren übergeben.

Wie in [Moore & Paris 89] gezeigt wurde, ermöglichen es planbasierte Ansätze, sowohl die Intention, die ein Sprecher mit einer Äußerung verfolgt, als auch die rhetorischen Beziehungen zwischen Äußerungen aufzuzeichnen, vorausgesetzt es werden entsprechende Planoperatoren definiert, die intentionale mit rhetorischen Zielen in Verbindung bringen. Präsentationen lassen sich dadurch individuell auf einen Benutzer abstimmen. Beispielsweise könnte man in dem Dokumentausschnitt aus Abbildung 2 das Bild weglassen, falls sichergestellt ist, daß der Benutzer weiß, wo sich die Verschlussschleife befindet und wie sie zu entfernen ist. Darüberhinaus lassen sich Teile einer Präsentation nachbessern, falls sich herausstellt, daß diese für den jeweiligen Benutzer unverständlich sind. Teilt der Benutzer, nachdem ihm der Dokumentausschnitt aus Abbildung 2 präsentiert wurde, dem System mit, daß er die Verschlussschleife nicht finden kann, so könnte das System versuchen, das Objekt durch bestimmte Merkmale eindeutig zu charakterisieren, und z.B. einen Satz wie "Die Verschlussschleife ist die rote Schleife links unter dem Deckel" generieren. Bei einem schemabasierten Ansatz kann das System zwar ein neues Schema auswählen, um dem Benutzer das Auffinden der Verschlussschleife zu ermöglichen. Es kann jedoch passieren, daß wieder ein Bild generiert wird, da das System nicht in der Lage ist, zu erkennen, daß dieses Ziel bereits als Teilziel bearbeitet wurde und ein Bild nicht zum Erfolg führte.

Planungsbasierte Ansätze für die Synthese multimedialer Dokumente werden u.a. in den Systemen WIP [André & Rist 90b] und AIMI [Maybury 91] verwendet. Während jedoch in [Maybury 91] die Mediumselektion erst an den Blättern des entstehenden Baumes erfolgt, konnte in [André & Rist 93] gezeigt werden, daß es Planer dieses Typs ermöglichen, gleichzeitig Inhalte auszuwählen, zu strukturieren und auf die medienspezifischen Generatoren zu verteilen, wenn die Planoperatoren entsprechende Constraints für die Mediumselektion enthalten.

5 Medienkoordination

Um zu verhindern, daß die Ergebnisse der verschiedenen Generatoren beziehungslos nebeneinander stehen, dürfen die Generatoren nicht isoliert voneinander betrieben werden. Es reicht also nicht, unterschiedliche Generatoren zu einem IMP-System zusammenzufassen, man muß deren Arbeit auch koordinieren. Hierunter fällt die Verteilung der Information auf verschiedene Generatoren, die Verzahnung der einzelnen Generierungsprozesse sowie die Integration der Generierungsergebnisse in einer multimedialen Ausgabe.

5.1 Auswahl von Präsentationsmedien

Ausgehend von der Frage, welches Medium oder welche Medienkombination welche Art von Information am besten überträgt, wird in den meisten Ansätzen die zu präsentierende Information nach verschiedenen Gesichtspunkten klassifiziert.

Beispielsweise geht COMET von sechs Informationstypen (*location, physical attributes, abstract actions, connectives, simple and compound actions*) aus (vgl. [Feiner & McKeown 91]). Für räumliche Information und physikalische Attribute ist als Präsentationsmedium Graphik vorgesehen, während für abstrakte Aktionen und Konnektiva Text verwendet wird. Zur Präsentation von einfachen und zusammengesetzten Aktionen werden beide Medien herangezogen. Eine feinere Klassifikation wird im WIP System verwendet (vgl. [André & Rist 93]). Dort stehen insgesamt sieben Informationstypen (*concrete, abstract, spatial, covariant, temporal, quantification, negation*) mit bis zu zehn Untertypen zur Verfügung. Desweiteren wird zwischen zehn kommunikativen Funktionen (*attract-attention, compare, elaborate, enable, elucidate, label, motivate, evidence, background, summarize*) unterschieden, die ebenfalls die Mediumselektion beeinflussen.

Statt Informationstypen bestimmten Medien zuzuordnen, besteht die Möglichkeit, sowohl die Information als auch die zur Verfügung stehenden Medien und Präsentationstechniken nach bestimmten Merkmalen zu charakterisieren und die Auswahl der Medien und der Präsentationstechniken auf der Grundlage dieser Klassifikation durchzuführen. Diese Vorgehensweise wurde von Bertin [Bertin 67] für die Gestaltung von Graphiken gewählt, dann von Mackinlay [Mackinlay 86] und später von Roth [Roth et al. 91] für die automatische Generierung von Geschäftsgraphiken aufgegriffen. In [Arens et al. 93b] wird eine ähnliche Vorgehensweise nicht nur für die Auswahl von Präsentationstechniken, sondern auch für die Mediumselektion eines IMP-Systems vorgeschlagen. Die für die Auswahl relevante Information wird dabei in einem Und-Oder-Abhängigkeitsnetzwerk dargestellt. Aufgrund der wenigen Merkmale, anhand derer die Medien klassifiziert werden, benötigt aber auch dieser Ansatz bei der Generierung spezieller Auswahlregeln (vgl. [Vossers 91]) wie sie etwa in WIP und COMET Verwendung finden.

Während für die Auswahl zwischen Text und Graphik bisher nur einfache Heuristiken bekannt sind, liegen für die Auswahl zwischen mehreren Präsentationstechniken bereits auf tieferen Inferenzen beruhende Ansätze vor. Beispielsweise wird in dem System APT [Mackinlay 86] bei der Auswahl von Präsentationstechniken anhand von formalen Kriterien geprüft, welche Information durch eine spezielle Präsentationstechnik vermittelt werden kann (*Kriterium der Expressivität*) und wie wirkungsvoll diese Technik die zu übermittelnde Information präsentiert (*Kriterium der Effektivität*). Casner [Casner 91] beschreibt einen Ansatz, bei dem die Selektion von Präsentationstechniken auf der Grundlage von perzeptuellen Faktoren erfolgt. Das von ihm entwickelte System BOZ analysiert zunächst die vom Benutzer durchzuführende Aufgabe (z.B. Temperaturunterschiede zu erkennen) und formuliert eine entsprechende perzeptuelle Aufgabe (z.B. Längenunterschiede zu erkennen), indem die logischen Operatoren in der Aufgabenbeschreibung durch perzep-

tuelle Operatoren ersetzt werden. Anschließend wird eine graphische Darstellung entworfen, die die Daten in der Graphik so strukturiert, daß sie alle erforderlichen perzeptuellen Operatoren unterstützt und gleichzeitig die visuelle Suche minimiert.

Zusammenfassend läßt sich sagen, daß in den meisten Systemen die Mediumselektion in erster Linie vom Informationstyp abhängt. Weitere Faktoren sind die Ziele des Präsentators, Betriebsmittelbeschränkungen, Benutzerklasse und Benutzeraufgabe. Schwierig wird der Auswahlprozeß vor allem dann, wenn sich mehrere Faktoren gegenseitig beeinflussen. Beispielsweise stellten Bieger und Glock [Bieger & Glock 86] in einem Performanztest fest, daß Versuchspersonen Instruktionen zwar schneller verstehen, wenn räumliche Information graphisch dargestellt wird, den Versuchspersonen jedoch mehr Fehler unterlaufen als bei rein textuellen Präsentationen. In diesem Fall wäre zumindest eine Gewichtung der Ziele des Präsentators erforderlich.

5.2 Generierung von Querverweisen und multimedialen Referenzen

Um Präsentationen kohärent zu gestalten, müssen die einzelnen Generatoren ihre Ergebnisse aufeinander abstimmen. Wichtige Beispiele hierfür sind *Querverweise* und *multimediale Referenzen*.

Querverweise wie "der linke Bildteil" oder "Abbildung x" stellen Verbindungen zwischen Dokumentteilen in unterschiedlichen Medien her und tragen damit wesentlich zur Kohärenz einer Multimedia-Präsentation bei. Eine multimediale Referenz liegt vor, wenn auf mehr als ein Medium zurückgegriffen wird, um auf eine Entität der Diskurswelt zu verweisen. Typische Beispiele sind natürlichsprachliche Äußerungen der Form "die im Bild grün markierte Schraube" oder von natürlichsprachlichen Äußerungen begleitete Zeigegesten. Wird unter Verwendung verschiedener Medien auf die frühere Darstellung einer Entität Bezug genommen, so handelt es sich um einen *multimedialen Querverweis*. In den meisten Fällen wird von Text auf Graphik oder andere Textteile verwiesen, jedoch auch mit Hilfe graphischer Mittel lassen sich Querverweise realisieren. Beispiele hierfür sind Pfeile, die von einem Bildteil zum entsprechenden Text zeigen.

Multimediale Referenzausdrücke bieten die Möglichkeit, zusätzliche Attribute zur Abgrenzung eines Objekts von seinen Alternativen zu verwenden. Hierunter fallen Attribute des Objekts im Bild, deren Interpretation sowie die räumliche Lage des Objekts relativ zum Bild oder zu anderen Bildobjekten. Wird auf die Interpretation eines Bildattributs verwiesen, so muß sichergestellt sein, daß der Benutzer diese kennt. Beispielsweise wird ein Ausdruck der Form "der Fluß oben links auf der Karte" nur dann verstanden, wenn der Benutzer in der Lage ist, bestimmte Bildteile als Flußdarstellungen zu erkennen. Darüberhinaus sollte aus dem Kontext eindeutig hervorgehen, ob es sich bei den verwendeten Attributen um Welt- oder Bildattribute handelt. Verwirrung beim Benutzer könnte insbesondere dann entstehen, wenn sich Welt- und Bildattribute widersprechen.

Häufig werden bei der Generierung von Querverweisen und multimedialen Referenzen

räumliche Relationen verwendet. Da sich sowohl die Anordnung von Bildobjekten als auch die Anordnung von Textblöcken und Bildern nicht vorhersehen läßt, ist es unmöglich, räumliche Relationen vorab in einer Wissensbasis abzulegen. Einige IMP-Systeme wie COMET und WIP verfügen daher über Lokalisierungskomponenten, um die Lage von Bildobjekten absolut im Bild (*absolute Lokalisation*) oder relativ zu anderen Bildobjekten (*relative Lokalisationen*) zu berechnen. Neben *elementaren* (z.B. "der Knopf links im Bild") können mit diesen Komponenten auch *zusammengesetzte Lokalisationen* (z.B. "der Knopf oben links im Bild") generiert werden. Im Gegensatz zu COMET [McKeown et al. 92] werden in WIP Objekte auch in Bezug auf Bildteile (z.B. die Ecken) oder Objektgruppen lokalisiert (vgl. [Wazinski 91; Wazinski 92]).

Eine wichtige Voraussetzung für die Generierung von Querverweisen ist die explizite Repräsentation des sprachlichen und des piktorischen Kontextes. Im System CUBRICON [Neal & Shapiro 91] ist der sprachliche Kontext durch eine Fokusliste von Entitäten und Propositionen repräsentiert, auf die sprachlich oder mit Hilfe von Zeigegesten verwiesen wird. Zur Repräsentation des piktorischen Kontextes wird ein Display-Modell verwendet, das darüber Buch führt, welche Entitäten sichtbar sind, in welchem Fenster sie sich befinden und welche Fenster auf dem Bildschirm zu sehen sind. Auch XTRA repräsentiert nicht nur den sprachlichen Kontext, sondern verwaltet auch eine Datenstruktur für die Graphik, auf die während des Dialogs Bezug genommen wird (vgl. [Reithinger 92b]). In der Steuerdomäne ist dies eine Formularhierarchie, die die Positionen und Größen der Formularfelder enthält sowie deren geometrischen und logischen Zusammenhang. Darüberhinaus werden die Verbindungen zwischen Teilen des Formulars, z.B. *region437*, und den entsprechenden Konzepten in der Wissensbasis, z.B. *wohntort*, repräsentiert. Während die Systeme XTRA und CUBRICON auf unterschiedliche Kontextmodelle für den sprachlichen und den piktorischen Kontext zurückgreifen, verwendet das System EDWARD [Claassen 92] ein einheitliches Modell, das verschiedene Typen von Kontextfaktoren, wie z.B. Stellung im Satz oder Sichtbarkeit auf dem Bildschirm, berücksichtigt. WIP unterscheidet sich von den oben genannten Systemen dadurch, daß die Semantik von Bildinhalten nicht nur auf Objektebene (z.B. daß die Bildkonstituente *pic-ob-1* das Weltobjekt *w-obj-1* darstellt), sondern auch auf Merkmalsebene (z.B. daß die Eigenschaft 'ist-rot' eines Bildobjekts die Eigenschaft 'ist-defekt' eines Bauteils darstellt) beschrieben wird (vgl. [Rist & André 92b; Rist & André 92a]). Um die semantische Beziehung zwischen Bildteilen/Bildrelationen und Objekten/Relationen zu beschreiben, wird in Anlehnung an [Mackinlay 86] ein zweistelliges Prädikat der Form (Enkodiert <graphisches Mittel> <information>) verwendet. Eine Menge solcher Enkodierrelationen bildet dann eine semantische Beschreibung eines Bildes, auf deren Grundlage Inferenzprozesse durchgeführt werden können.

Da sich Bildattribute während des Generierungsprozesses ändern können, ist zu gewährleisten, daß der piktorische Kontext ständig aktualisiert wird. Wählt der Graphikgenerator beispielsweise aufgrund zusätzlicher Aufträge eine andere Perspektive, so ist zu prüfen, ob bereits berechnete räumliche Relationen noch aktuell sind.

5.3 Räumliche und zeitliche Koordination der Ausgabe

Eine weitere Koordinationsaufgabe ist die Integration der einzelnen Generatorergebnisse in einer multimedialen Ausgabe. Hierunter fällt die räumliche Anordnung von Textblöcken und Bildern durch eine Layoutkomponente. Einer solchen Komponente kommt nicht nur die Aufgabe zu, Text- und Bildteile in einer möglichst ansprechenden Form auf die zur Verfügung stehenden Dokumentseiten oder Bildschirmfenster zu verteilen (*künstlerisches Layout*), sondern auch inhaltliche Zusammenhänge zu vermitteln (*funktionales Layout*). Beispielsweise kann die rhetorische Struktur eines Dokuments mitunter sehr effektiv durch das Layout übertragen werden. Daher müssen neben linguistischen Constraints (Wortwahl, Festlegung syntaktischer Strukturen usw.) zur sprachlichen Wiedergabe von rhetorischen Relationen (vgl. [Scott & Sieckenius de Souza 90]) auch graphische Constraints (Plazierung, Abstand usw.) definiert werden (vgl. [Hovy & Arens 91; Graf 92]), die die rhetorische Struktur des Dokuments reflektieren. Für eine Reihe von rhetorischen Relationen sind bereits Designheuristiken bekannt. Beispielsweise werden Bilder, die durch 'Sequence' oder 'Contrast' zueinander in Beziehung stehen, vorzugsweise nebeneinander angeordnet. Faßt eine Graphik den Text zusammen, so sollte sie unter dem Text stehen, dient sie hingegen zur Einführung in ein Gebiet, so ist sie nach Möglichkeit über dem Text zu plazieren.

Wird dem Benutzer Information mit Hilfe dynamischer Medien vermittelt (z.B. gesprochene Sprache, Zeigegesten und animierte Darstellungen), so sind die Ausgabe der Generatorergebnisse nicht nur räumlich, sondern auch zeitlich zu koordinieren. Beispielsweise ist festzulegen, zu welchem Zeitpunkt eine bestimmte Zeigegeste beginnen soll und wieviel Zeit sie in Anspruch nehmen darf. Bisher beschäftigen sich erst wenige Arbeiten mit der Koordination dynamischer Medien. Erste Ideen zur zeitlichen Koordination gesprochener Sprache und animierter Darstellungen werden in [André et al. 87], [Feiner et al. 93] und [André et al. 93b] beschrieben.

6 Aufbau von IMP-Systemen

Beim Entwurf eines IMP-Systems ist eine adäquate Behandlung der Abhängigkeiten zwischen den durchzuführenden Entscheidungsprozessen zu gewährleisten.

Aus der Textgenerierung ist bekannt, daß es äußerst schwierig ist, Inhalte zu planen, ohne über sprachliches Wissen zu verfügen. Ein streng sequentielles Modell, in dem Information ausschließlich von der inhaltsfestlegenden Komponente zur inhaltsrealisierenden Komponente fließt, hat sich als unangemessen erwiesen, da die inhaltsfestlegende Komponente alle Entscheidungen der inhaltsrealisierenden Komponente antizipieren müßte (siehe auch [Reithinger 92a]). Dieses Problem verschärft sich, wenn wie in allen oben genannten Systemen für die Inhaltsrealisierung getrennte Komponenten verantwortlich sind (z.B. ein Text-, ein Graphik- und ein Gestengenerator), deren Entscheidungsprozesse die inhaltsfestlegende Komponente nicht antizipieren kann.

Ähnliche Probleme treten bei der Inhaltsplanung, der Inhaltsstrukturierung und der Mediumselektion auf. Einerseits hängt die Wahl eines geeigneten Präsentationsmedium weitgehend von der Art und der Struktur der zu präsentierenden Information ab (siehe auch [André & Rist 90a; Arens et al. 93b; Roth & Mattis 90]). Andererseits werden Inhaltsauswahl und Inhaltsstrukturierung durch frühere Entscheidungen bei der Wahl der Präsentationsmedien beeinflusst. Beispielsweise wird visuelle Information benötigt, um mit graphischen Mitteln auf ein Objekt zu verweisen; wird hingegen auf dasselbe Objekt textuell Bezug genommen, ist diese Information möglicherweise irrelevant.

Neben Abhängigkeiten zwischen Inhaltsfestlegung, Inhaltsrealisierung und Mediumselektion sind Abhängigkeiten zwischen medienspezifischen Generatoren zu berücksichtigen. Informationsfluß sollte daher nicht nur zwischen der inhaltsfestlegenden und den inhaltsrealisierenden Komponenten, sondern auch zwischen einzelnen Generatoren möglich sein.

Während über die Art der Entscheidungsprozesse eines IMP-Systems weitgehend Einigkeit herrscht, stehen unterschiedliche Verarbeitungsmodelle zur Diskussion.

In COMET wird beim Entwurf einer Präsentation zunächst eine baumartige Struktur aufgebaut, die den Aufbau der zu generierenden Präsentation widerspiegelt. Dieser Strukturbaum wird von den medienspezifischen Generatoren monoton erweitert - Revisionen aufgrund der Mediumselektion sind nicht vorgesehen. Den Abhängigkeiten zwischen Text- und Graphikgenerierung wurde mehr Aufmerksamkeit gewidmet. Beispiele hierfür sind die Abstimmung von Bild- und Satzgrenzen sowie die Generierung von Querverweisen. Um die Kommunikation zwischen Text- und Graphikgenerierung zu erleichtern, wird die mitzuteilende Information in einer einheitlichen Repräsentationssprache an die beiden Generatoren übergeben. Da beide Generatoren von derselben Repräsentationssprache ausgehen, können sie sich gegenseitig Information über ihre Entscheidungen und Ergebnisse mitteilen, indem sie die Inhaltsspezifikation mit Annotationen versehen. Jeder Generator hat damit Zugriff auf die Ergebnisse anderer Generatoren und kann seine eigenen Ergebnisse entsprechend darauf abstimmen.

Arens und Kollegen [Arens et al. 93a] beschäftigen sich mit Abhängigkeiten zwischen Inhaltsstrukturierung und Mediumselektion. Sie vertreten die Auffassung, daß die Bestimmung eines geeigneten Mediums erst dann möglich ist, wenn die Diskursstruktur vollständig vorliegt. Daher schlagen sie eine strikte Trennung von Planungs- und Mediumauswahlprozessen vor. Während des Planungsprozesses wird zunächst die Diskursstruktur des zu generierenden Dokuments vollständig aufgebaut. Diese ist im wesentlichen durch die kommunikativen Ziele des Präsentators und den mitzuteilenden Inhalt bestimmt. Anschließend erfolgt anhand spezieller Regeln [Arens et al. 93b] die Mediumauswahl. Nach der Mediumauswahl wird die Diskursstruktur von unten nach oben durchlaufen, wobei anhand von Transformationsregeln die Diskursstruktur in eine präsentationsnahe Struktur (*Präsentationsstruktur*) überführt wird. Problematisch an diesem Ansatz ist, daß diese Präsentationsstruktur offensichtlich keinerlei Einfluß auf die Diskursstruktur hat. Zwar wird die Auswahl des Mediums von der Diskursstruktur bein-

flußt, Inhalte werden allerdings unabhängig vom Medium ausgewählt.

Um Abhängigkeiten zwischen Inhaltsselektion, Inhaltsstrukturierung und Mediumselektion zu berücksichtigen, ist es jedoch notwendig, die einzelnen Prozesse enger als in den oben beschriebenen Systemen miteinander zu verzahnen. Wie bereits erläutert wird in WIP für alle drei Prozesse ein einheitlicher Planungsmechanismus verwendet. Er ermöglicht es, die Mediumselektion schon während der Inhaltsfestlegung durchzuführen und nicht erst danach wie beispielsweise in den obigen Systemen. Inhaltsrealisierung und Inhaltsfestlegung werden hingegen von separaten Komponenten, die auf unterschiedliche Wissensbasen zugreifen, übernommen. Die Modularisierung ermöglicht einerseits Parallelverarbeitung, macht jedoch andererseits Interaktion zwischen den Komponenten notwendig. Ähnlich wie in COMET erfolgt die Kommunikation zwischen den einzelnen Komponenten über den Designplan, auf den alle Komponenten Lese- und Schreibzugriff haben.

7 Schlußbemerkung

Die Verfügbarkeit neuer Medien eröffnet neue Wege der Informationspräsentation und wirft neue interessante Fragestellungen auf, wie beispielsweise die Verteilung der Information auf verschiedene Medien sowie die Koordination von Medien untereinander. Andererseits treten Probleme auf, die bereits von der Textgenerierung her bekannt sind. Eine zentrale Feststellung dieses Papiers ist, daß sich multimediale Präsentationen nach ähnlichen Prinzipien strukturieren lassen wie reine Texte. Konzepte, wie Sprechakte und rhetorische Relationen, die sich bereits für die automatische Generierung von Text als erfolgreich erwiesen haben, sind auch für die Generierung von multimedialen Präsentationen brauchbar. Aus diesem Grund können Methoden zur Planung des Inhalts und der Struktur von Texten im Hinblick auf die Generierung von multimedialen Präsentationen verallgemeinert werden. Tatsächlich hat der Forschungsbereich Textgenerierung einen wesentlichen Anteil an der Entwicklung der ersten Generation von IMP-Systemen. Inwiefern eine eingehendere Beschäftigung mit den zahlreichen noch unzureichend gelösten Aufgabenstellungen anregende Impulse für die Textgenerierung bringen wird, bleibt abzuwarten.

Literatur

- [Allgayer et al. 89] J. Allgayer, K. Harbusch, A. Kobsa, C. Reddig, N. Reithinger, und D. Schmauks. *XTRA: A Natural-Language Access System to Expert Systems*. International Journal of Man-Machine Studies, 31:161–195, 1989.
- [André & Rist 90a] E. André und T. Rist. *Synthesizing Illustrated Documents: A Plan-Based Approach*. In: Proceedings of InfoJapan '90, Vol. 2, S. 163–170, Tokyo, 1990. Auch als DFKI Research Report RR-91-06.

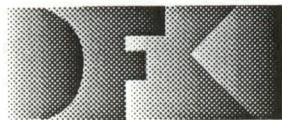
- [André & Rist 90b] E. **André** und T. **Rist**. *Towards a Plan-Based Synthesis of Illustrated Documents*. In: Proc. of the 9th ECAI, S. 25–30, Stockholm, July 1990. Auch als DFKI Research Report RR-90-11.
- [André & Rist 93] E. **André** und T. **Rist**. *The Design of Illustrated Documents as a Planning Task*. In: M. Maybury (Hrsg.), *Intelligent Multimedia Interfaces*, S. 94–116. AAAI Press, 1993. Auch als DFKI Research Report RR-92-45.
- [André et al. 87] E. **André**, T. **Rist**, und G. **Herzog**. *Generierung natürlichsprachlicher Äußerungen zur simultanen Beschreibung von zeitveränderlichen Szenen*. In: K. Morik (Hrsg.), *GWAI-87: 11th German Workshop on Artificial Intelligence*, S. 330–338. Berlin, Heidelberg: Springer, 1987.
- [André et al. 93a] E. **André**, W. **Finkler**, W. **Graf**, T. **Rist**, A. **Schauder**, und W. **Wahlster**. *WIP: The Automatic Synthesis of Multimodal Presentations*. In: M. Maybury (Hrsg.), *Intelligent Multimedia Interfaces*, S. 75–93. AAAI Press, 1993. Auch als DFKI Research Report RR-92-46.
- [André et al. 93b] E. **André**, W. **Graf**, J. **Heinsohn**, B. **Nebel**, H.-J. **Profitlich**, T. **Rist**, und W. **Wahlster**. *PPP - Personalized Plan-Based Presenter*. Document D-93-5, DFKI, Saarbrücken, 1993.
- [Arens et al. 91] Y. **Arens**, L. **Miller**, und N. **Sondheimer**. *Presentation Design Using an Integrated Knowledge Base*. In: J. W. Sullivan und S. W. Tyler (Hrsg.), *Intelligent User Interfaces*, S. 241–258. New York, NY: acm press, 1991.
- [Arens et al. 93a] Y. **Arens**, E. **Hovy**, und S. **van Mulken**. *Structure and Rules in Automated Multimedia Presentation Planning*. In: Proc. of the 13th IJCAI, Chambéry, France, 1993.
- [Arens et al. 93b] Y. **Arens**, E. **Hovy**, und M. **Vossers**. *Describing the Presentational Knowledge Underlying Multimedia Instruction Manuals*. In: M. Maybury (Hrsg.), *Intelligent Multimedia Interfaces*. AAAI Press, 1993.
- [Bandyopadhyay 90] S. **Bandyopadhyay**. *Towards an Understanding of Coherence in Multimodal Discourse*. Technical Memo TM-90-01, Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI), Saarbrücken, Germany, 1990.
- [Bertin 67] J. **Bertin**. *Sémiologie graphique*. Paris: Mouton & Cie, 1967.
- [Bieger & Glock 86] G.R. **Bieger** und M.D. **Glock**. *Comprehending Spatial and Contextual Information in Picture-Text Instructions*. *The Journal of Experimental Education*, 45(4):181–188, 1986.
- [Casner 91] S.M. **Casner**. *A Task-Analytic Approach to the Automated Design of Graphic Presentations*. *ACM Transactions on Graphics*, 10(2):111–151, April 1991.

- [Claassen 92] W. **Claassen**. *Generating Referring Expressions in a Multimodal Environment*. In: R. Dale, E. Hovy, D. Rösner, und O. Stock (Hrsg.), *Aspects of Automated Natural Language Generation: Proc. of the 6th International Workshop on Natural Language Generation*, S. 247–262. Berlin, Heidelberg: Springer, 1992.
- [Costabile et al. 92] M. F. **Costabile**, T. **Catarci**, und S. **Levialdi** (Hrsg.). *Advanced Visual Interfaces (Proceedings of AVI '92, Rome, Italy)*. Singapore: World Scientific Press, 1992.
- [Dilley et al. 92] S. **Dilley**, J. **Bateman**, U. **Thiel**, und A. **Tissen**. *Integrating Natural Language Components into Graphical Discourse*. In: Proc. of the Third Conference on Applied Natural Language Processing, S. 72–79, Trento, Italy, 1992.
- [Feiner & McKeown 91] S.K. **Feiner** und K.R. **McKeown**. *Automating the Generation of Coordinated Multimedia Explanations*. IEEE Computer, 24(10):33–41, 1991.
- [Feiner et al. 93] S. K. **Feiner**, D. J. **Litman**, K. R. **McKeown**, und R. J. **Passonneau**. *Towards Coordinated Temporal Multimedia Presentations*. In: M. Maybury (Hrsg.), *Intelligent Multimedia Interfaces*. AAAI Press, 1993.
- [Graf 92] W. **Graf**. *Constraint-Based Graphical Layout of Multimodal Presentations*. In: Costabile et al. [Costabile et al. 92]. Auch als DFKI Research Report RR-92-15.
- [Grimes 75] J.E. **Grimes**. *The Thread of Discourse*. The Hague, Paris: Mouton, 1975.
- [Herzog et al. 93] G. **Herzog**, W. **Maass**, und P. **Wazinski**. *VITRA GUIDE: Utilisation du langage naturel et de représentations graphiques pour la description d'itinéraires*. In: *Images et Langages: Multimodalité et Modélisation Cognitive*, Paris, April 1993.
- [Hobbs 78] J. **Hobbs**. *Why is a discourse coherent?* Technical Report 176, SRI International, Menlo Park, CA, 1978.
- [Hovy & Arens 91] E. H. **Hovy** und Y. **Arens**. *Automatic Generation of Formatted Text*. In: Proc. of AAAI-91, S. 92–97, Anaheim, CA, 1991.
- [Hunter et al. 87] B. **Hunter**, A. **Crismore**, und P.D. **Pearson**. *Visual Displays in Basal Readers and Social Studies Textbooks*. In: D.M. Willows und H. A. Houghton (Hrsg.), *The Psychology of Illustration, Basic Research, Band 2*, S. 116–135. New York, Berlin, Heidelberg: Springer, 1987.
- [Kerpedjiev 92] S. M. **Kerpedjiev**. *Automatic Generation of Multimodal Weather Reports from Datasets*. In: Proc. of the Third Conference on Applied Natural Language Processing, S. 48–55, Trento, Italy, 1992.
- [Kjorup 78] S. **Kjorup**. *Pictorial Speech Acts*. Erkenntnis, 12:55–71, 1978.

- [Levin et al. 87] J.R. **Levin**, G.J. **Anglin**, und R. N. **Carney**. *On Empirically Validating Functions of Pictures in Prose*. In: D.M. Willows und H. A. Houghton (Hrsg.), *The Psychology of Illustration, Basic Research, Band 1*, S. 51–85. New York, Berlin, Heidelberg: Springer, 1987.
- [Mackinlay 86] J. **Mackinlay**. *Automating the Design of Graphical Presentations of Relational Information*. *ACM Transactions on Graphics*, 5(2):110–141, April 1986.
- [Mann & Thompson 87] W. C. **Mann** und S. A. **Thompson**. *Rhetorical Structure Theory: Description and Construction of Text Structures*. In: G. Kempen (Hrsg.), *Natural Language Generation: New Results in Artificial Intelligence, Psychology, and Linguistics*, S. 85–95. Dordrecht: Martinus Nijhoff Publishers, 1987.
- [Marks 91] J. **Marks**. *Discourse Coherence and the Consistent Design of Informational Graphics*. In: *Workshop Notes from the Ninth National Conference on Artificial Intelligence*, S. 29–36, Anaheim, CA, 1991.
- [Maybury 89] M. T. **Maybury**. *Enhancing Explanation Coherence with Rhetorical Strategies*. In: *Proc. of the 4th EACL*, S. 168–173, Manchester, UK, 1989.
- [Maybury 91] M. T. **Maybury**. *Planning Multimedia Explanations Using Communicative Acts*. In: *Proc. of AAAI-91*, S. 61–66, Anaheim, CA, 1991.
- [McKeown et al. 90] K. **McKeown**, M. **Elhadad**, Y. **Fukumoto**, J. **Lim**, C. **Lombardi**, J. **Robin**, und F. **Smadja**. *Natural Language Generation in COMET*. In: R. Dale, C. Mellish, und M. Zock (Hrsg.), *Current Research in Natural Language Generation*, S. 103–139. London: Academic Press, 1990.
- [McKeown et al. 92] K. R. **McKeown**, S. K. **Feiner**, J. **Robin**, D. D. **Seligmann**, und M. **Tanenblatt**. *Generating Cross-References for Multimedia Explanation*. In: *Proc. of AAAI-92*, S. 9–16, San Jose, CA, 1992.
- [McKeown 85] K. R. **McKeown**. *Text Generation*. Cambridge, MA: Cambridge University Press, 1985.
- [Moore & Paris 89] J. D. **Moore** und C. L. **Paris**. *Planning Text for Advisory Dialogues*. In: *Proc. of the 27th ACL*, S. 203–211, 1989.
- [Moore & Swartout 90] J. D. **Moore** und W. R. **Swartout**. *Pointing: A Way Toward Explanation Dialogue*. In: *Proc. of AAAI-90*, S. 457–464, Boston, MA, 1990.
- [Neal & Shapiro 91] J. G. **Neal** und S. C. **Shapiro**. *Intelligent Multi-Media Interface Technology*. In: J. W. Sullivan und S. W. Tyler (Hrsg.), *Intelligent User Interfaces*, S. 11–43. New York, NY: acm press, 1991.
- [Paris 91] C. L. **Paris**. *Generation and Explanation: Building an Explanation Facility for the Explainable Expert Systems Framework*. In: C. L. Paris, W. R. Swartout, und W. C. Mann (Hrsg.), *Natural Language Generation in Artificial Intelligence and Computational Linguistics*, S. 49–82. Boston: Kluwer, 1991.

- [Reiter et al. 92] E. **Reiter**, C. **Mellish**, and J. **Levine**. *Automatic Generation of On-Line Documentation in the IDAS Project*. In: Proc. of the Third Conference on Applied Natural Language Processing, S. 64–71, Trento, Italy, 1992.
- [Reithinger 92a] N. **Reithinger**. *Eine parallele Architektur zur inkrementellen Generierung multimodaler Dialogbeiträge*. infix, 1992.
- [Reithinger 92b] N. **Reithinger**. *The Performance of an Incremental Generation Component for Multi-Modal Dialog Contributions*. In: R. Dale, E. Hovy, D. Rösner, und O. Stock (Hrsg.), Aspects of Automated Natural Language Generation: Proc. of the 6th International Workshop on Natural Language Generation, S. 263–276. Berlin, Heidelberg: Springer, 1992.
- [Rist & André 92a] T. **Rist** und E. **André**. *From Presentation Tasks to Pictures: Towards a Computational Approach to Graphics Design*. In: Proc. of the 10th ECAI, S. 764–768, Wien, 1992. Auch als DFKI Research Report RR-92-44.
- [Rist & André 92b] T. **Rist** und E. **André**. *Incorporating Graphics Design and Realization into the Multimodal Presentation System WIP*. In: Costabile et al. [Costabile et al. 92].
- [Roth & Mattis 90] S. F. **Roth** und J. **Mattis**. *Data Characterization for Intelligent Graphics Presentation*. In: Proc. of CHI-90, S. 193–200, 1990.
- [Roth et al. 91] S. F. **Roth**, J. **Mattis**, und X. **Mesnard**. *Graphics and Natural Language as Components of Automatic Explanation*. In: J. W. Sullivan und S. W. Tyler (Hrsg.), Intelligent User Interfaces, S. 207–239. New York, NY: acm press, 1991.
- [Scott & Sieckenius de Souza 90] D. R. **Scott** und C. **Sieckenius de Souza**. *Getting the Message Across in RST-Based Text Generation*. In: R. Dale, C. Mellish, und M. Zock (Hrsg.), Current Research in Natural Language Generation, S. 47–73. London: Academic Press, 1990.
- [Searle 80] J.R. **Searle**. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge, MA: Cambridge University Press, 1980.
- [Stock 91] O. **Stock**. *Natural Language and Exploration of an Information Space: The ALFresco Interactive System*. In: Proc. of the 12th IJCAI, S. 972–978, Sidney, Australia, 1991.
- [Vossers 91] M. **Vossers**. *To Write or to Draw? Automatic Generation of Formatted Text and Line Drawings*. Diplomarbeit, University of Nijmegen, Nijmegen, Niederlande, 1991.
- [Wahlster et al. 91] W. **Wahlster**, E. **André**, W. **Graf**, und T. **Rist**. *Designing Illustrated Texts: How Language Production is Influenced by Graphics Generation*. In: Proc. of the 5th EACL, S. 8–14, Berlin, Germany, 1991.

- [Wahlster et al. 92] W. **Wahlster**, E. **André**, S. **Bandyopadhyay**, W. **Graf**, and T. **Rist**. *WIP: The Coordinated Generation of Multimodal Presentations from a Common Representation*. In: A. Ortony, J. Slack, und O. Stock (Hrsg.), *Communication from an Artificial Intelligence Perspective: Theoretical and Applied Issues*, S. 121–144. Berlin, Heidelberg: Springer, 1992. Auch als DFKI Research Report RR-91-08.
- [Wahlster et al. 93] W. **Wahlster**, E. **André**, W. **Finkler**, H.-J. **Profitlich**, und T. **Rist**. *Plan-Based Integration of Natural Language and Graphics Generation*. *AI Journal*, 63:387–427, 1993. Auch als DFKI Research Report RR-93-02.
- [Wazinski 91] P. **Wazinski**. *Objektlokalisierung in graphischen Darstellungen*. Diplomarbeit, Universität des Saarlandes, Saarbrücken, 1991.
- [Wazinski 92] P. **Wazinski**. *Generating Spatial Descriptions for Cross-Modal References*. In: Proc. of the Third Conference on Applied Natural Language Processing, S. 56–63, Trento, Italy, 1992.
- [Wilson et al. 92] M. **Wilson**, D. **Sedlock**, J.-L. **Binot**, und P. **Falzon**. *An Architecture For Multimodal Dialogue*. In: Proc. of the Second Vencona Workshop for Multimodal Dialogue, Vencona, Italy, 1992.



Deutsches
Forschungszentrum
für Künstliche
Intelligenz GmbH

DFKI
-Bibliothek-
PF 2080
67608 Kaiserslautern
FRG

DFKI Publikationen

Die folgenden DFKI Veröffentlichungen sowie die aktuelle Liste von allen bisher erschienenen Publikationen können von der oben angegebenen Adresse oder per anonymem ftp von ftp.dfki.uni-kl.de (131.246.241.100) unter pub/Publications bezogen werden.

Die Berichte werden, wenn nicht anders gekennzeichnet, kostenlos abgegeben.

DFKI Publications

The following DFKI publications or the list of all published papers so far are obtainable from the above address or per anonymous ftp from ftp.dfki.uni-kl.de (131.246.241.100) under pub/Publications.

The reports are distributed free of charge except if otherwise indicated.

DFKI Research Reports

RR-92-54

Harold Boley:

A Direkt Semantic Characterization of RELFUN
30 pages

RR-92-55

John Nerbonne, Joachim Laubsch, Abdel Kader Diagne, Stephan Open: Natural Language Semantics and Compiler Technology
17 pages

RR-92-56

Armin Laux: Integrating a Modal Logic of Knowledge into Terminological Logics
34 pages

RR-92-58

Franz Baader, Bernhard Hollunder:
How to Prefer More Specific Defaults in Terminological Default Logic
31 pages

RR-92-59

Karl Schlechta and David Makinson: On Principles and Problems of Defeasible Inheritance
13 pages

RR-92-60

Karl Schlechta: Defaults, Preorder Semantics and Circumscription
19 pages

RR-93-01

Bernhard Hollunder:
An Alternative Proof Method for Possibilistic Logic and its Application to Terminological Logics
25 pages

RR-93-02

Wolfgang Wahlster, Elisabeth André, Wolfgang Finkler, Hans-Jürgen Profitlich, Thomas Rist:
Plan-based Integration of Natural Language and Graphics Generation
50 pages

RR-93-03

Franz Baader, Berhard Hollunder, Bernhard Nebel, Hans-Jürgen Profitlich, Enrico Franconi:
An Empirical Analysis of Optimization Techniques for Terminological Representation Systems
28 pages

RR-93-04

Christoph Klauck, Johannes Schwagereit:
GGD: Graph Grammar Developer for features in CAD/CAM
13 pages

RR-93-05

Franz Baader, Klaus Schulz: Combination Techniques and Decision Problems for Disunification
29 pages

RR-93-06

Hans-Jürgen Bürckert, Bernhard Hollunder, Armin Laux: On Skolemization in Constrained Logics
40 pages

RR-93-07

Hans-Jürgen Bürckert, Bernhard Hollunder, Armin Laux: Concept Logics with Function Symbols
36 pages

RR-93-08

Harold Boley, Philipp Hanschke, Knut Hinkelmann, Manfred Meyer: COLAB: A Hybrid Knowledge Representation and Compilation Laboratory
64 pages

RR-93-09

Philipp Hanschke, Jörg Würtz:
Satisfiability of the Smallest Binary Program
8 Seiten

RR-93-10

Martin Buchheit, Francesco M. Donini, Andrea Schaerf: Decidable Reasoning in Terminological Knowledge Representation Systems
35 pages

- RR-93-11**
Bernhard Nebel, Hans-Juergen Buerckert:
Reasoning about Temporal Relations:
A Maximal Tractable Subclass of Allen's Interval
Algebra
28 pages
- RR-93-12**
Pierre Sablayrolles: A Two-Level Semantics for
French Expressions of Motion
51 pages
- RR-93-13**
Franz Baader, Karl Schlechta:
A Semantics for Open Normal Defaults via a
Modified Preferential Approach
25 pages
- RR-93-14**
Joachim Niehren, Andreas Podelski, Ralf Treinen:
Equational and Membership Constraints for
Infinite Trees
33 pages
- RR-93-15**
*Frank Berger, Thomas Fehrle, Kristof Klöckner,
Volker Schölles, Markus A. Thies, Wolfgang
Wahlster:* PLUS - Plan-based User Support
Final Project Report
33 pages
- RR-93-16**
Gert Smolka, Martin Henz, Jörg Würtz: Object-
Oriented Concurrent Constraint Programming in
Oz
17 pages
- RR-93-17**
Rolf Backofen:
Regular Path Expressions in Feature Logic
37 pages
- RR-93-18**
Klaus Schild: Terminological Cycles and the
Propositional μ -Calculus
32 pages
- RR-93-20**
Franz Baader, Bernhard Hollunder:
Embedding Defaults into Terminological
Knowledge Representation Formalisms
34 pages
- RR-93-22**
Manfred Meyer, Jörg Müller:
Weak Looking-Ahead and its Application in
Computer-Aided Process Planning
17 pages
- RR-93-23**
Andreas Dengel, Ottmar Lutzy:
Comparative Study of Connectionist Simulators
20 pages
- RR-93-24**
Rainer Hoch, Andreas Dengel:
Document Highlighting —
Message Classification in Printed Business Letters
17 pages
- RR-93-25**
Klaus Fischer, Norbert Kuhn: A DAI Approach to
Modeling the Transportation Domain
93 pages
- RR-93-26**
Jörg P. Müller, Markus Pischel: The Agent
Architecture InteRRaP: Concept and Application
99 pages
- RR-93-27**
Hans-Ulrich Krieger:
Derivation Without Lexical Rules
33 pages
- RR-93-28**
*Hans-Ulrich Krieger, John Nerbonne,
Hannes Pirker:* Feature-Based Allomorphy
8 pages
- RR-93-29**
Armin Laux: Representing Belief in Multi-Agent
Worlds via Terminological Logics
35 pages
- RR-93-30**
Stephen P. Spackman, Elizabeth A. Hinkelman:
Corporate Agents
14 pages
- RR-93-31**
Elizabeth A. Hinkelman, Stephen P. Spackman:
Abductive Speech Act Recognition, Corporate
Agents and the COSMA System
34 pages
- RR-93-32**
David R. Traum, Elizabeth A. Hinkelman:
Conversation Acts in Task-Oriented Spoken
Dialogue
28 pages
- RR-93-33**
Bernhard Nebel, Jana Koehler:
Plan Reuse versus Plan Generation: A Theoretical
and Empirical Analysis
33 pages
- RR-93-34**
Wolfgang Wahlster:
VerbMobil Translation of Face-To-Face Dialogs
10 pages
- RR-93-35**
Harold Boley, François Bry, Ulrich Geske (Eds.):
Neuere Entwicklungen der deklarativen KI-
Programmierung — *Proceedings*
150 Seiten
Note: This document is available only for a
nominal charge of 25 DM (or 15 US-\$).
- RR-93-36**
*Michael M. Richter, Bernd Bachmann, Ansgar
Bernardi, Christoph Klauck, Ralf Legleitner,
Gabriele Schmidt:* Von IDA bis IMCOD:
Expertensysteme im CIM-Umfeld
13 Seiten

RR-93-38

Stephan Baumann: Document Recognition of Printed Scores and Transformation into MIDI
24 pages

RR-93-40

Francesco M. Donini, Maurizio Lenzerini, Daniele Nardi, Werner Nutt, Andrea Schaerf: Queries, Rules and Definitions as Epistemic Statements in Concept Languages
23 pages

RR-93-41

Winfried H. Graf: LAYLAB: A Constraint-Based Layout Manager for Multimedia Presentations
9 pages

RR-93-42

Hubert Comon, Ralf Treinen: The First-Order Theory of Lexicographic Path Orderings is Undecidable
9 pages

RR-93-43

M. Bauer, G. Paul: Logic-based Plan Recognition for Intelligent Help Systems
15 pages

RR-93-44

Martin Buchheit, Manfred A. Jeusfeld, Werner Nutt, Martin Staudt: Subsumption between Queries to Object-Oriented Databases
36 pages

RR-93-45

Rainer Hoch: On Virtual Partitioning of Large Dictionaries for Contextual Post-Processing to Improve Character Recognition
21 pages

RR-93-46

Philipp Hanschke: A Declarative Integration of Terminological, Constraint-based, Data-driven, and Goal-directed Reasoning
81 pages

RR-93-48

Franz Baader, Martin Buchheit, Bernhard Hollunder: Cardinality Restrictions on Concepts
20 pages

RR-94-01

Elisabeth André, Thomas Rist: Multimedia Presentations: The Support of Passive and Active Viewing
15 pages

RR-94-02

Elisabeth André, Thomas Rist: Von Textgeneratoren zu Intellimedia-Präsentationssystemen
22 pages

RR-94-05

Franz Schmalhofer, J. Stuart Aitken, Lyle E. Bourne jr.: Beyond the Knowledge Level: Descriptions of Rational Behavior for Sharing and Reuse
81 pages

DFKI Technical Memos**TM-92-01**

Lijuan Zhang: Entwurf und Implementierung eines Compilers zur Transformation von Werkstückrepräsentationen
34 Seiten

TM-92-02

Achim Schupeta: Organizing Communication and Introspection in a Multi-Agent Blocksworld
32 pages

TM-92-03

Mona Singh: A Cognitive Analysis of Event Structure
21 pages

TM-92-04

Jürgen Müller, Jörg Müller, Markus Pischel, Ralf Scheidhauer: On the Representation of Temporal Knowledge
61 pages

TM-92-05

Franz Schmalhofer, Christoph Globig, Jörg Thoben: The refitting of plans by a human expert
10 pages

TM-92-06

Otto Kühn, Franz Schmalhofer: Hierarchical skeletal plan refinement: Task- and inference structures
14 pages

TM-92-08

Anne Kilger: Realization of Tree Adjoining Grammars with Unification
27 pages

TM-93-01

Otto Kühn, Andreas Birk: Reconstructive Integrated Explanation of Lathe Production Plans
20 pages

TM-93-02

Pierre Sablayrolles, Achim Schupeta: Conflict Resolving Negotiation for COoperative Schedule Management
21 pages

TM-93-03

Harold Boley, Ulrich Buhrmann, Christof Kremer: Konzeption einer deklarativen Wissensbasis über recyclingrelevante Materialien
11 pages

TM-93-04

Hans-Günther Hein: Propagation Techniques in WAM-based Architectures — The FIDO-III Approach
105 pages

TM-93-05

Michael Sintek: Indexing PROLOG Procedures into DAGs by Heuristic Classification
64 pages

DFKI Documents

D-92-28

Klaus-Peter Gores, Rainer Bleisinger: Ein Modell zur Repräsentation von Nachrichtentypen
56 Seiten

D-93-01

Philipp Hanschke, Thom Frühwirth: Terminological Reasoning with Constraint Handling Rules
12 pages

D-93-02

Gabriele Schmidt, Frank Peters, Gernod Laufkötter: User Manual of COKAM+
23 pages

D-93-03

Stephan Busemann, Karin Harbusch(Eds.): DFKI Workshop on Natural Language Systems: Reusability and Modularity - Proceedings
74 pages

D-93-04

DFKI Wissenschaftlich-Technischer Jahresbericht 1992
194 Seiten

D-93-05

Elisabeth André, Winfried Graf, Jochen Heinsohn, Bernhard Nebel, Hans-Jürgen Profitlich, Thomas Rist, Wolfgang Wahlster: PPP: Personalized Plan-Based Presenter
70 pages

D-93-06

Jürgen Müller (Hrsg.): Beiträge zum Gründungsworkshop der Fachgruppe Verteilte Künstliche Intelligenz, Saarbrücken, 29. - 30. April 1993
235 Seiten

Note: This document is available only for a nominal charge of 25 DM (or 15 US-\$).

D-93-07

Klaus-Peter Gores, Rainer Bleisinger: Ein erwartungsgesteuerter Koordinator zur partiellen Textanalyse
53 Seiten

D-93-08

Thomas Kieninger, Rainer Hoch: Ein Generator mit Anfragesystem für strukturierte Wörterbücher zur Unterstützung von Texterkennung und Textanalyse
125 Seiten

D-93-09

Hans-Ulrich Krieger, Ulrich Schäfer: TDL ExtraLight User's Guide
35 pages

D-93-10

Elizabeth Hinkelman, Markus Vonerden, Christoph Jung: Natural Language Software Registry (Second Edition)
174 pages

D-93-11

Knut Hinkelmann, Armin Laux (Eds.): DFKI Workshop on Knowledge Representation Techniques — Proceedings
88 pages

D-93-12

Harold Boley, Klaus Elsbernd, Michael Herfert, Michael Sintek, Werner Stein: RELFUN Guide: Programming with Relations and Functions Made Easy
86 pages

D-93-14

Manfred Meyer (Ed.): Constraint Processing — Proceedings of the International Workshop at CSAM'93, July 20-21, 1993
264 pages

Note: This document is available only for a nominal charge of 25 DM (or 15 US-\$).

D-93-15

Robert Laux: Untersuchung maschineller Lernverfahren und heuristischer Methoden im Hinblick auf deren Kombination zur Unterstützung eines Chart-Parsers
86 Seiten

D-93-16

Bernd Bachmann, Ansgar Bernardi, Christoph Klauk, Gabriele Schmidt: Design & KI
74 Seiten

D-93-20

Bernhard Herbig: Eine homogene Implementierungsebene für einen hybriden Wissensrepräsentationsformalismus
97 Seiten

D-93-21

Dennis Drollinger: Intelligentes Backtracking in Inferenzsystemen am Beispiel Terminologischer Logiken
53 Seiten

D-93-22

Andreas Abecker: Implementierung graphischer Benutzungsoberflächen mit Tcl/Tk und Common Lisp
44 Seiten

D-93-24

Brigitte Krenn, Martin Volk: DiTo-Datenbank: Datendokumentation zu Funktionsverbgefügen und Relativsätzen
66 Seiten

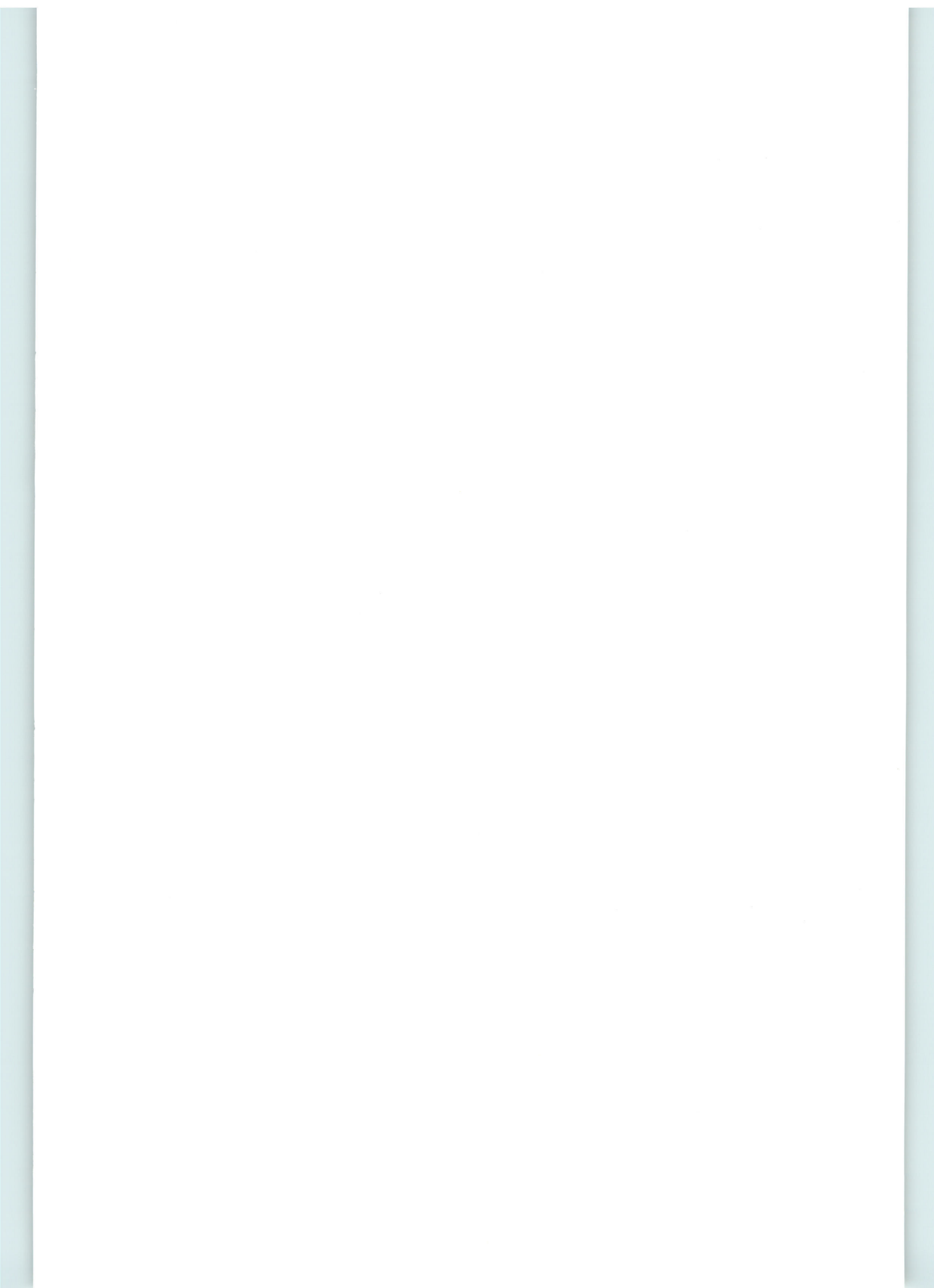
D-93-25

Hans-Jürgen Bürckert, Werner Nutt (Eds.): Modeling Epistemic Propositions
118 pages

Note: This document is available only for a nominal charge of 25 DM (or 15 US-\$).

D-93-26

Frank Peters: Unterstützung des Experten bei der Formalisierung von Textwissen
INFOCOM - Eine interaktive Formalisierungskomponente
58 Seiten



Von Textgeneratoren zu Intellimedia-Präsentationssystemen

Elisabeth André, Thomas Rist

RR-94-02

Research Report