



**Wissensbasierte Informationspräsentation:
Zwei Beiträge zum GI-Fachgespräch
Graphik und KI**

- 1. Ein planbasierter Ansatz zur Synthese
illustrierter Dokumente**
- 2. Wissensbasierte Perspektivenwahl für
die automatische Erzeugung von 3D-
Objektdarstellungen**

Elisabeth André, Thomas Rist

Mai 1990

**Deutsches Forschungszentrum für Künstliche Intelligenz
GmbH**

Postfach 20 80
D-6750 Kaiserslautern, FRG
Tel.: (+49 631) 205-3211/13
Fax: (+49 631) 205-3210

Stuhlsatzenhausweg 3
D-6600 Saarbrücken 11, FRG
Tel.: (+49 681) 302-5252
Fax: (+49 681) 302-5341

Deutsches Forschungszentrum für Künstliche Intelligenz

Das Deutsche Forschungszentrum für Künstliche Intelligenz (DFKI) mit Standorten in Kaiserslautern und Saarbrücken ist eine gemeinnützige GmbH, die 1988 von den Gesellschaftern ADV/Orga, AEG, IBM, Insiders, Fraunhofer Gesellschaft, GMD, Krupp-Atlas, Mannesmann-Kienzle, Nixdorf, Philips und Siemens gegründet wurde. Die am DFKI durchgeführten Projekte werden vom Bundesministerium für Forschung und Technologie, von den Gesellschaftern oder von externen Auftraggebern finanziert.

Am DFKI wird anwendungsorientierte Grundlagenforschung auf dem Gebiet der Künstlichen Intelligenz (KI) und angrenzender anderer Bereiche der Informatik durchgeführt. Ziel ist die Erstellung von "Intelligenten Fachsystemen", in denen ausgesuchte Probleme eines Anwendungsbereiches mit Hilfe von KI Methoden gelöst werden. Derzeitig werden am DFKI die folgenden Forschungsschwerpunkte verfolgt:

- Intelligente Ingenieur-Systeme
- Intelligente Benutzerschnittstellen
- Intelligente Kommunikations-Netzwerke
- Intelligente kooperative Systeme

Das DFKI strebt danach, seine wissenschaftlichen Ergebnisse der Forschungsgemeinschaft zugänglich zu machen. Aus diesem Grunde wurde ein Netzwerk von Kontakten zu in- und ausländischen Forschungseinrichtungen im akademischen und industriellen Bereich aufgebaut. Darüber hinaus veranstaltet das DFKI Wissenstransfer-Workshops, auf denen Gesellschafter und interessierte Dritte über den Forschungsstand informiert werden.

Seit seiner Gründung stellt das DFKI eine attraktive Arbeitsumgebung für namhafte Forscherinnen und Forscher aus dem In- und Ausland dar. Das Ziel ist, nach Beendigung der Aufbauphase ca. 100 Wissenschaftler zu beschäftigen.

Prof. Dr. Gerhard Barth
Technisch-wissenschaftlicher Geschäftsführer

**Wissensbasierte Informationspräsentation:
Zwei Beiträge zum GI-Fachgespräch Graphik und KI**

- 1. Ein planbasierter Ansatz zur Synthese
illustrierter Dokumente**
- 2. Wissensbasierte Perspektivenwahl für die
automatische Erzeugung von 3D-Objektdarstellungen**

Elisabeth André, Thomas Rist

DFKI-RR-90-07

Die beiden hier veröffentlichten Papiere erscheinen in K.Kansy and P. Wißkirchen (eds.): Graphik und KI, Springer-Verlag, 1990.

© Deutsches Forschungszentrum für Künstliche Intelligenz 1990

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.

Ein planbasierter Ansatz zur Synthese illustrierter Dokumente

Elisabeth André, Thomas Rist
Deutsches Forschungszentrum für Künstliche Intelligenz
Stuhlsatzenhausweg 3
D-6600 Saarbrücken 11

Zusammenfassung

Obwohl die Erzeugung illustrierter Dokumente in der KI-Forschung zunehmendes Interesse findet, werden in den meisten Systemen Text und Graphik weitgehend unabhängig voneinander aufgebaut und stehen daher beziehungslos nebeneinander. In dieser Arbeit* wird von der Überlegung ausgegangen, daß nicht nur die Erzeugung von Texten, sondern auch die Synthese illustrierter Dokumente als kommunikative Handlung zur Erreichung von Zielen aufgefaßt werden kann. Für die Realisierung eines Systems, das selbstständig illustrierte Dokumente erstellt, bietet sich daher ein planbasierter Ansatz an. Es wird zunächst gezeigt, daß die in der Textlinguistik gebräuchliche Unterscheidung zwischen Haupt- und Nebenhandlungen auch für Text-Bild-Kombinationen geeignet ist. Von dieser Unterscheidung ausgehend werden Strategien formuliert, die sich sowohl auf die Erzeugung von Text als auch auf den Aufbau von Bildern beziehen. Die gemeinsame Planung von Text und Bild wird als grundlegende Voraussetzung angesehen, die beiden Modi in einem Dokument aufeinander abzustimmen.

1 Einleitung

Intelligente Benutzerschnittstellen als Komponenten von Hilfesystemen, Expertensystemen oder intelligenten Betriebswarten der nächsten Generation müssen in der Lage sein, vorliegendes Wissen auf flexible Weise in unterschiedlichen Präsentationssituationen jeweils angemessen darzubieten (vgl. [39]). Dazu gehört vor allem auch die Fähigkeit, Verbalisierungs- und Visualisierungsergebnisse in einer multimodalen Ausgabe zu integrieren. Ohne Zweifel läßt sich in vielen Fällen Information gegenüber rein textuellen Darbietungen präziser und wirkungsvoller durch integrierte Text-Bild-

* Die vorliegende Arbeit entstand im Rahmen des vom BMFT unter dem Förderkennzeichen ITW8901 8 geförderten Projekts WIP (**W**issensbasierte **I**nformations**p**äsentation).

Präsentationen übermitteln. Beispiele sind etwa Gebrauchsanleitungen für technische Geräte oder Lehrbücher aus dem naturwissenschaftlichen Bereich, in denen zur Erklärung komplizierter Sachverhalte ausgiebig auf graphische Illustrationen zurückgegriffen wird. Es wäre allerdings falsch, hieraus zu folgern, daß der gleichzeitige Einsatz der Modi Text und Bild zwangsläufig zum Gelingen des Kommunikationsprozesses beiträgt. So ist zu berücksichtigen, daß die Art der Verwendung eines Bildes nicht notwendigerweise aus dem Bildinhalt allein hervorgeht. Wittgenstein weist beispielsweise darauf hin, daß das Bild eines Boxers in Kampfstellung verwendet werden kann, um zu zeigen, wie sich ein Boxer halten sollte, wie er sich nicht halten soll oder wie ein bestimmter Mann irgendwo gestanden hat usw. (vgl. [43]).

Um sicherzustellen, daß ein Dokument verstanden wird, sind Text und Bild in geeigneter Weise aufeinander abzustimmen. Eine solche Abstimmung erfordert Kenntnisse über die Funktionen textueller und bildhafter Dokumententeile sowie Kenntnisse über die zwischen ihren Funktionen bestehenden Zusammenhänge. Dies gilt insbesondere dann, wenn wie in unserem Fall illustrierte Dokumente maschinell entworfen und realisiert werden sollen.

2 Stand der Forschung

Während die Verwendung von Text-Bild-Kombinationen innerhalb der KI-Forschung gerade erst zum Forschungsgegenstand wird, liegen zu dieser Problemstellung bereits mehrere Arbeiten aus den Disziplinen Philosophie, Linguistik und Psychologie vor, die in diesem Zusammenhang von Bedeutung sind.

Die Grundlage philosophischer und linguistischer Ansätze bildet meist ein Vergleich von Bild und Sprache. Goodman (vgl. [18]) setzt sich mit dem Gegenstandsbezug von Bildern auseinander, indem er Bilder mit Namen und Kennzeichnungen vergleicht. Bennett (vgl. [6]) untersucht die Frage, inwiefern man Bildern, ähnlich wie Prädikationen, einen Wahrheitswert zuordnen kann. Vergleiche von Sprache und Bild, bezogen auf ihre Verwendung, finden sich in Wittgensteins Bildtheorie (vgl. [43]) und in Übertragungen der Sprechakttheorie auf Bilder (vgl. [23], [33] und [24]). Bei dem Entwurf einer Gebrauchstheorie für Bild und Sprache setzt sich Muckenhaupt für eine wechselseitige Betrachtungsweise ein: nicht nur Konzepte des sprachlichen Bereichs sollen als Vergleichsobjekte für den Bildbereich verwendet werden, sondern auch umgekehrt (vgl. [32]). Eine in vielen Arbeiten vertretene Auffassung ist die, daß ähnlich zur Sprache auch mit Bildern kommunikative Handlungen vollzogen werden können.

Gegenstand psychologischer Untersuchungen ist zum einen die Frage, inwieweit Bilder zum Verständnis von Texten beitragen (vgl. u.a. [10] und [41]), zum anderen das dazu duale Problem, nämlich der Einfluß von Texten auf das Verstehen von Bildern (vgl. u.a. [20] und [21]). Durch solche Untersuchungen werden Erkenntnisse über kognitive Verarbeitungsprozesse gewonnen, die dann insbesondere auch bei der Gestaltung illustrierter Dokumente Berücksichtigung finden können. Ein wichtiges Ergebnis ist die Feststellung, daß Text-Bild-Kombinationen nur dann von Vorteil sind,

wenn Text und Bild sich gegenseitig ergänzen, d.h. zueinander komplementär sind. Bleibt hingegen die Beziehung zwischen Text und Bildern unklar, ist kein positiver Effekt auf die Verstehensprozesse beobachtbar (vgl. [5]). Weitere Arbeiten beschäftigen sich mit der Frage, welche Information eher durch Bilder, welche eher durch Text wiedergegeben werden sollte. Den Untersuchungen in [7] und [42] zufolge wird beispielsweise in Anweisungsdokumenten räumliche Information schneller erfaßt, wenn sie durch Bilder vermittelt wird, andererseits unterlaufen den Versuchspersonen in diesem Fall mehr Fehler beim Ausführen der Anweisungen.

Die in jüngster Zeit abgehaltenen Workshops über intelligente Benutzerschnittstellen (vgl. [38] und [40]) verdeutlichen, daß in der KI-Forschung zwar die Entwicklung von Systemen angestrebt wird, die Graphik und Text in Dokumenten integrieren, daß jedoch die beiden Modi noch weitgehend unabhängig voneinander behandelt werden. Eine naheliegende Vorgehensweise ist dabei, auf bestehende Ansätze zur Text- und Graphikgenerierung zurückzugreifen.

Einen Überblick über die wichtigsten Ergebnisse aus dem Bereich der Textgenerierung bieten u.a. [22] oder [45]. Ein auch für die vorliegende Arbeit relevantes Teilproblem ist die Festlegung von Inhalt und Struktur von Texten. Neben schemabasierten Ansätzen (vgl. [29], [28] und [34]) werden heute vor allem planbasierte Strategien (vgl. [3], [19] und [31]) als erfolgversprechend angesehen. Bei planungsbasierten Verfahren wird davon ausgegangen, daß ein Autor beim Erzeugen von Texten Handlungen vollzieht, um bestimmte Ziele zu erreichen (vgl. [36]). Von besonderem Interesse für die vorliegende Arbeit sind die Ansätze von Hovy, Moore und Swartout, da sich diese nicht nur mit der Generierung von Einzelsätzen, sondern vor allem auch mit der Generierung kohärenter Texte befassen. Dazu greifen sie auf die von Mann und Thompson vorgeschlagene RST-Theorie (vgl. [27]) zurück, die den Aufbau von Texten mit Hilfe sogenannter rhetorischer Relationen, wie z.B. *Elaboration* oder *Motivation*, beschreibt.

Die bisherigen Ansätze zur Graphikgenerierung lassen sich nach den zugrundeliegenden Zielsetzungen unterscheiden. Wichtige Arbeiten sind die Darstellung relationaler Information durch Business-Graphiken (vgl. [26]), die Synthese von Ikonen (vgl. [16]), die graphische Präsentation von Handlungsabfolgen (vgl. [15]), die Visualisierung natürlichsprachlicher Beschreibungen (vgl. [8] und [1]), die Generierung von mentalen Bildern (vgl. [25]) oder die Repräsentation des Wissens über graphische Präsentationstechniken durch semantische Netzwerke (vgl. [44], [17] und [4]).

Einer der wenigen Ansätze, bei denen eine Integration von Text und Graphik angestrebt ist, wird im RADIO-Projekt verfolgt (vgl. [14]). Nach der Dekomposition des Dialogziels in Teilziele entscheidet eine Koordinationskomponente, welche Ziele durch Text und welche durch Graphik wiedergegeben werden sollen. Einer nahtlosen Verzahnung der beiden Modi steht allerdings die Tatsache entgegen, daß derzeit die Teilziele noch unabhängig voneinander bearbeitet werden.

3 Die Handlungsstruktur von Text-Bild-Präsentationen

Ziel unserer Arbeit ist die Entwicklung eines Systems, das Dokumente erzeugt, in denen Text und Bild aufeinander abgestimmt sind. Ausgangspunkt unserer Überlegungen ist dabei die Sichtweise, daß nicht nur die Erzeugung von Texten, sondern auch die Erstellung von multimodalen Dokumenten als kommunikative Handlung aufgefaßt wird, durch deren Ausführung bestimmte Ziele erreicht werden sollen. Oft sind Ziele nicht unmittelbar durch die Ausführung einer Handlung zu erreichen, da die dazu notwendigen Voraussetzungen erst durch entsprechende untergeordnete Handlungen geschaffen werden müssen. Darüberhinaus kann es sinnvoll sein, neben der eigentlichen Handlung weitere auszuführen, die zwar nicht unbedingt notwendig wären, jedoch unterstützende Wirkung haben. In Anlehnung an textlinguistische Arbeiten unterscheiden wir im folgenden zwischen *Haupt-* und *Nebenhandlungen*.¹ Nebenhandlungen unterscheiden wir weiter danach, ob sie zur Erreichung notwendiger Bedingungen dienen oder optionale Ergänzungen darstellen. Hierarchische Handlungsstrukturen ergeben sich dadurch, daß sich Nebenhandlungen wiederum aus Haupt- und Nebenhandlungen zusammensetzen können. Die Struktur eines Dokuments wird jedoch nicht nur durch die Handlungshierarchie bestimmt, sondern auch durch die funktionalen Beziehungen zwischen Haupt- und Nebenhandlungen und den Inhalten, die durch sie vermittelt werden. Beispielsweise kann man ein Bild durch Text erklären oder eine Folge von Ereignissen durch eine Bildsequenz wiedergeben. Zur Verdeutlichung der eben eingeführten Begriffe betrachten wir den in Abb. 1 dargestellten Dokumentausschnitt.

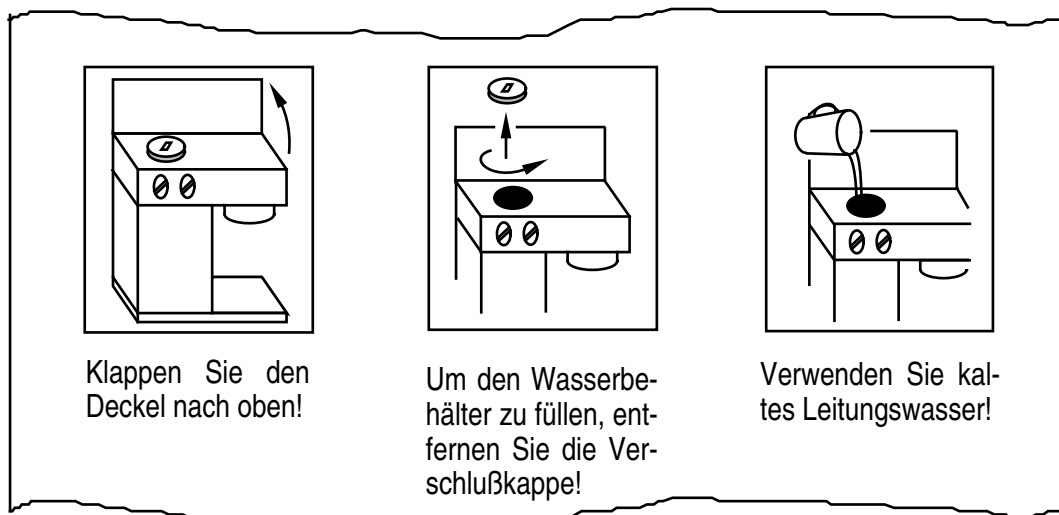


Abb. 1: Ausschnitt aus einer Gebrauchsanleitung einer Espressomaschine²

¹ Die Unterscheidung zwischen Haupt- und Nebenhandlungen entspricht im wesentlichen den Unterscheidungen zwischen *global speech acts* und *subsidiary speech acts* bei Searle (vgl. [36]), *main speech acts* und *subordinate speech acts* bei van Dijk (vgl. [13]), *dominierenden* und *subsidiären Handlungen* bei Brandt et al. (vgl. [9]) sowie zwischen *nucleus* und *satellites* bei Mann und Thompson (vgl. [27]).

² Als Vorlage für den hier gezeigten Dokumentausschnitt wurde eine Gebrauchsanleitung für die Philips Espressomaschine HD 5649 zugrundegelegt.

Mit diesem Dokumentausschnitt wird das Ziel verfolgt, den Adressaten darüber zu informieren, wie der Wasserbehälter der Espressomaschine zu füllen ist. Dieses Ziel wird erreicht, indem man ihm mitteilt, welche Teilschritte er ausführen soll. Zunächst wird dem Adressaten explizit gesagt (Haupthandlung), daß der Deckel hochgeklappt werden soll, und mit einem Bild näher spezifiziert (Nebenhandlung), wie diese Handlung auszuführen ist. Bei der Herstellung des Bildes läßt sich ebenfalls unterscheiden zwischen einer Haupthandlung, nämlich dem Zeigen der Position des Deckels nach dem Hochklappen, und einer Nebenhandlung, der Darstellung des gesamten Geräts, um dem Adressaten die Orientierung zu erleichtern. Als nächstes wird der Adressat gebeten (Haupthandlung), den Verschluß des Wasserbehälters zu entfernen. Um diese Bitte zu motivieren, wird das Ziel, das Füllen des Wasserbehälters, genannt (Nebenhandlung). Außerdem wird im dazugehörigen Bild gezeigt (Nebenhandlung), wie der Verschluß entfernt werden soll. Aus dem Bild geht hervor, daß der Deckel nach rechts gedreht und nach oben abgehoben werden soll. Der letzte Teil des Dokumentausschnitts ist ein Beispiel dafür, daß man auch jemanden implizit auffordern kann, etwas zu tun. Falls der Adressat weiß, daß die dargestellte Handlung zur Erreichung seines Ziels notwendig ist, genügt es offenbar, mit einem Bild zu zeigen (Nebenhandlung), wie das Einfüllen auszuführen ist. Der Text dient hier lediglich zur Ergänzung des Bildes. Er weist auf bildlich nicht darstellbare Details hin (Nebenhandlung), nämlich kaltes Leitungswasser einzufüllen. Die Haupthandlung, den Adressaten aufzufordern, mit einem Gefäß Wasser einzufüllen, folgt implizit aus dem situativen Kontext und den Nebenhandlungen.³ Ferner wird, wie auch schon beim zweiten Bild, außer den für das Öffnen bzw. Füllen des Wasserbehälters notwendigen Objekten wie Wasserbehälter, Verschlußkappe bzw. Gefäß und Wasser zusätzlich ein Teil der Espressomaschine dargestellt (Nebenhandlung). Die hierarchische Handlungsstruktur des obigen Dokumentausschnitts ist in Abb. 2 zusammengefaßt.

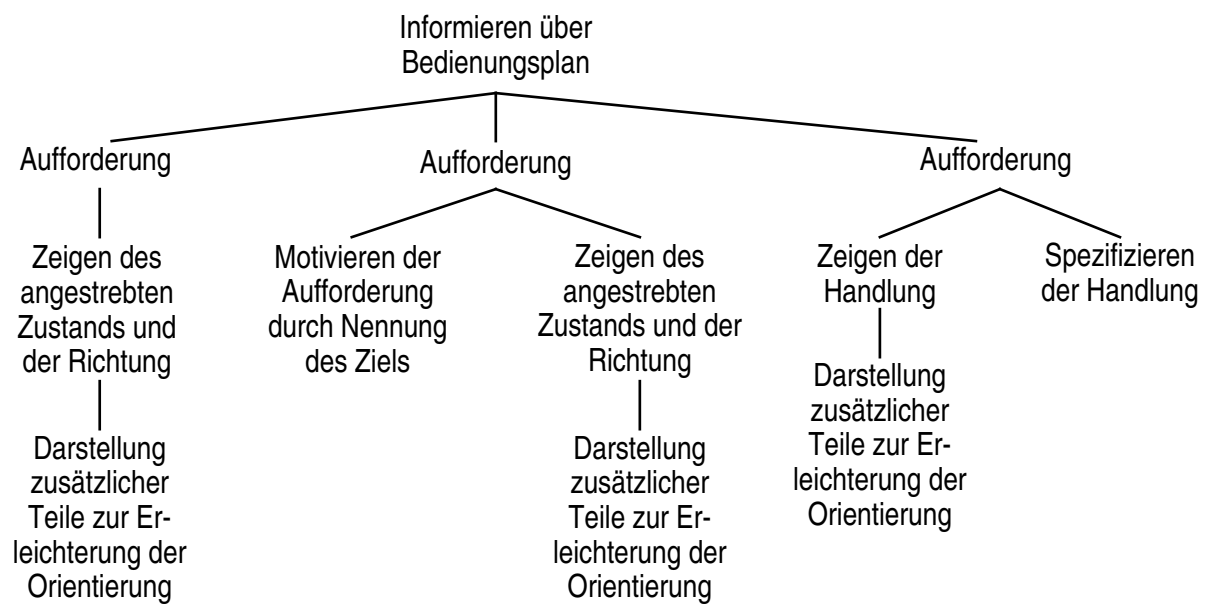


Abb. 2: Handlungsstruktur des Dokumentausschnitts aus Abb. 1

³ Auf das Auftreten nicht explizit ausgeführter Haupthandlungen wird auch in [9] hingewiesen.

Im folgenden wird der Frage nachgegangen, wie zu einem vorgegebenen Präsentationsziel ein verständliches und der jeweiligen Situation angepaßtes Dokument aufgebaut werden kann.

4 Strategien zum Aufbau von Text-Bild-Präsentationen

Wir gehen davon aus, daß der Autor eines Dokuments über ein Repertoire von *Präsentationsstrategien* verfügt, die er den Aufgaben entsprechend auswählen und miteinander kombinieren kann. Den Kern solcher Präsentationsstrategien bilden dabei die im vorangegangenen Abschnitt eingeführten Haupt- und Nebenhandlungen.

Bei der Repräsentation der Präsentationsstrategien folgen wir dem in [30] und [31] vorgestellten Ansatz zur Operationalisierung der RST-Theorie. In Anlehnung an textlinguistische Arbeiten (vgl. [9] und [35]) lassen wir jedoch auch Strategien zu, die keine Haupthandlung enthalten, und sehen vor, mehrere Haupthandlungen zu einer kommunikativen Einheit zusammenzufassen, falls keine dieser Handlungen untergeordnet ist oder nur zur Unterstützung einer anderen Handlung dient.

Handlungen werden durch die Angabe des Effekts oder durch die direkte Angabe einer Strategie spezifiziert. Der Effekt einer Strategie kann für eine funktionale Beziehung (z.B. *Elucidation* oder *Supplement*) stehen oder ein Präsentationsziel (z.B. der Adressat kennt das Ergebnis einer Handlung) bezeichnen (vgl. [30]). Präsentationsziele werden über mentalen Zuständen oder Prozessen eines Adressaten formuliert. Wir verwenden hierzu die Modaloperatoren GOAL, BEL und MB (vgl. [11] und [12]). Der Ausdruck (GOAL P p) steht für: P hat p zum Ziel. Durch (BEL P p) wird ausgedrückt: P glaubt, daß p erfüllt ist. (MB P A p) bedeutet: P und A ist gegenseitig bekannt, daß p erfüllt ist. Um auszudrücken, daß P und A gegenseitig bekannt ist, für welches eindeutig bestimmte x eine Formel p erfüllt ist, verwenden wir die abkürzende Schreibweise (MBR P A p). Ob eine Präsentationsstrategie eingesetzt werden kann, ist situationsabhängig. Bedingungen, die angeben, unter welchen Umständen eine Strategie angewandt wird, bezeichnen wir als *Anwendbarkeitsbedingungen*.

Als Beispiele werden im folgenden einige Präsentationsstrategien aufgeführt, die sich auch zum Aufbau des in Abb. 1 gezeigten Dokumentausschnitts verwenden lassen.

- (1) Strategie:
(EXPLICIT-REQUEST P A ?act)
Effekt:
(MB P A (GOAL P (DONE A ?act)))
Anwendbarkeitsbedingungen:
(GOAL P (DONE A ?act))
Haupthandlungen:
(REQUEST P A ?act)
Nebenhandlungen:
(MOTIVATE P A ?act)
(ENABLE P A ?act)

Strategie 1 kann von einem Präsentator P dazu verwendet werden, um einen Adressaten A explizit dazu aufzufordern, eine Handlung auszuführen. Die erste Nebenhandlung dient zur Motivation der Handlung, die zweite vermittelt dem Adressaten Information, die ihm die Durchführung der Handlung ermöglicht. Zur Präsentation dieser Information kann Strategie 2 angewandt werden. Sie sieht vor, das Ergebnis der Handlung graphisch darzustellen (SHOW-STATE) und einen Pfeil vom Anfangs- zum Endzustand zu zeichnen (SHOW-DIRECTION).

- (2) Strategie:
 (ENABLE-BY-ILLUSTRATION P A ?act)
 Effekt:
 (ENABLE P A ?act)
 Anwendbarkeitsbedingungen:
 (BEL P (ISA ?act MOTION))
 Haupthandlungen:
 (SHOW-STATE P A (RESULT ?act) ?picture)
 (SHOW-DIRECTION P A (DIRECTION ?act) ?picture)
 Nebenhandlungen:
 (ATTRACT-ATTENTION P A ?picture)
 (ELUCIDATE P A (MODE ?act) ?picture)

Durch die Haupthandlungen in Strategie 2 allein ist jedoch nicht sichergestellt, daß das angestrebte Ziel erreicht wird. Verwendet man beispielsweise ein an anderer Stelle erzeugtes Bild, so muß man dafür Sorge tragen, daß der Adressat das Bild fokussiert. Dies kann dadurch geschehen, daß man den Adressaten explizit auffordert, das Bild zu betrachten.

- (3) Strategie:
 (ATTRACT-ATTENTION-BY-REQUEST P A ?picture)
 Effekt:
 (ATTRACT-ATTENTION P A ?picture)
 Haupthandlungen:
 (MB P A (GOAL P (DONE A (LOOK-AT A ?picture))))

Selbst wenn der Adressat das betreffende Bild im richtigen Augenblick betrachtet, ist nicht unbedingt gewährleistet, daß er auch dessen Funktion erkennt; d.h., daß er weiß, welche Information er dem Bild entnehmen soll. Es ist allerdings nicht immer einfach, vorherzusehen, ob er dazu in der Lage sein wird. Eine entscheidende Rolle spielt hierbei das Wissen des Adressaten über die eingesetzten Präsentationstechniken. Verwendet der Präsentator beispielsweise ein Bildlexikon, so kann davon ausgegangen werden, daß die damit verbundene Intention erkannt wird. Wir nehmen zur Vereinfachung an, daß die Funktion eines Bildes erkannt wird, falls es ein Textstück gibt, innerhalb dessen das Bild fokussiert wird, und falls diesem Textstück und dem Bild ein gemeinsames Ziel, z.B. Darstellung eines Objekts, zugeordnet werden kann.⁴ Bei der Formulierung dieser Bedingung wird davon ausgegangen, daß die Funktion eines Textes in jedem Fall aus dem Text selbst hervorgeht und daß aus dieser die Funktion des Bildes hergeleitet werden kann. Ist die Bedingung nicht erfüllt, so

⁴ Eine schwächere Bedingung findet sich in Kjørups Bildakttheorie (vgl. [24]).

besteht Grund zur Annahme, daß die Funktion des Bildes nicht erkannt wird. In diesem Fall kommt nachstehende Präsentationsstrategie zum Einsatz.

- (4) Strategie:
(ELUCIDATE-BY-REQUEST P A ?info ?picture)
Effekt:
(ELUCIDATE P A ?info ?picture)
Anwendbarkeitsbedingungen:
(BEL P (CONTAINS ?info ?picture))
Haupthandlungen:
(MB P A (GOAL P (DONE A (EXTRACT A ?info ?picture))))

Kehren wir nun zurück zu der in Strategie 2 auftretenden Haupthandlung, die in zwei Teilhandlungen aufgeteilt ist, nämlich die graphische Darstellung des Ergebnisses der vom Adressaten auszuführenden Handlung und dem Zeichnen eines Pfeils vom Anfangs- zum Endzustand. Das Ergebnis einer Handlung kann dadurch beschrieben werden, daß man die räumlichen Beziehungen zwischen Objekten in einem Bild darstellt. Eine räumliche Relation wird dargestellt, indem man das Subjekt und die Bezugsobjekte abbildet und ihre Abbilder entsprechend positioniert. Anstatt näher auf die entsprechenden Strategien einzugehen - sie sind in [2] zu finden - konzentrieren wir uns auf die Darstellung von Objekten (vgl. 5).

- (5) Strategie:
(SHOW-OBJECT P A ?x ?picture)
Effekt:
(IDENTIFIED-REFERENT A ?x ?px ?picture)
Haupthandlungen:
(DEPICT-OBJECT P A ?x ?px ?picture)
Nebenhandlungen:
(IDENTIFIABLE A ?x ?px ?picture)
(*Res* (SPACE-AVAILABLE-P) (FACILITATE-LOC P A ?x ?px ?picture))

Hierbei ist sicherzustellen, daß der Adressat Welt- und Bildobjekte in gleicher Weise einander zuordnet wie der Präsentator. Es wird davon ausgegangen, daß der Adressat die Verbindung zwischen einem Objekt ?x und seinem Abbild ?px erkennt, falls keine Weltobjekte fokussiert werden, die dem Abbild von ?x bezüglich der dargestellten visuellen Merkmale gleich ähnlich oder sogar ähnlicher sehen als ?x. Gibt es ein weiteres Objekt, das zu dem Abbild ?px ebenso große Ähnlichkeit aufweist wie ?x, dann ist davon auszugehen, daß der Adressat nicht in der Lage ist, ?px und ?x einander zuzuordnen. In diesem Fall können die Strategien 6 und 7 angewandt werden, um ?x als Teil eines komplexeren Objekts darzustellen.

- (6) Strategie:
(ENABLE-IDENTIFICATION P A ?x ?px ?picture)
Effekt:
(IDENTIFIABLE A ?x ?px ?picture)
Anwendbarkeitsbedingungen:
(BEL P (PART-OF ?x ?z))
Haupthandlungen:
(SUPPLEMENT P A ?z ?pz ?picture)

- (7) Strategie:
 (GRAPHICAL-SUPPLEMENT P A ?x ?px ?picture)
 Effekt:
 (SUPPLEMENT P A ?x ?px ?picture)
 Haupthandlungen:
 (DEPICT-OBJECT P A ?x ?px ?picture)
 Nebenhandlungen:
 (IDENTIFIABLE A ?x ?px ?picture)

Die durch *Res* gekennzeichnete optionale Nebenhandlung in Strategie 5 wird nur dann ausgeführt, wenn noch genügend Platz zur Verfügung steht. Sie führt zur Darstellung zusätzlicher Objekte, um dem Adressaten die räumliche Einordnung eines Objekts ?x zu erleichtern. Je nachdem, ob sich ?x im visuellen Fokus befindet oder nicht, werden benachbarte Objektteile angedeutet oder ?x als Teil eines übergeordneten Objekts dargestellt (vgl. [2]).

5 Aufbau von Text-Bild-Präsentationen

Für den automatischen Entwurf eines Dokuments werden die vorgestellten Präsentationsstrategien als Operatoren eines Planungssystems aufgefaßt. Beim Planungsprozeß wird dann nach Strategien gesucht, deren Anwendbarkeitsbedingungen erfüllt sind und deren Ziele mit dem Präsentationsziel instantiierbar sind. Kommen mehrere Strategien in Frage, erfolgt die Auswahl regelbasiert. Bei der Formulierung von Auswahlregeln bietet es sich an, auf die in Abschnitt 2 erwähnten psychologischen Untersuchungen zuzugreifen, in denen die Wirksamkeit der Präsentationsmodi in Abhängigkeit vom Informationstyp untersucht wird. Beispielsweise könnte man durch eine Regel sicherstellen, daß graphische Darstellungen zur Präsentation räumlicher Information bevorzugt zum Einsatz kommen (vgl. 8).

- (8) IF (GOAL P (MBR P A (LOC ?loc ?obj)))
 THEN (DOBEFORE *graphics-strategies* *text-strategies*)

Nach der Auswahl einer Strategie sind die entsprechenden Neben- und Haupthandlungen durchzuführen. Ob eine Nebenhandlung vor oder nach der Haupthandlung ausgeführt wird, richtet sich danach, ob die Nebenhandlung unabhängig von der Haupthandlung ausführbar ist oder nicht. Soll beispielsweise durch eine Nebenhandlung sichergestellt werden, daß der Adressat die in einer Haupthandlung gezeigten Objektdarstellungen Weltobjekten zuordnen kann, setzt dies Kenntnis über das Bild voraus. In diesem Fall ist zunächst die Haupthandlung, nämlich das Erzeugen eines Bildes, auszuführen. Erst danach kann in einer Nebenhandlung durch Antizipation der Verstehensprozesse des Adressaten⁵ das Bild überprüft werden. Stellt sich heraus, daß dieses unverständlich ist, so ist das Dokument entsprechend zu ergänzen, z.B. durch Darstellung zusätzlicher Objekte (vgl. Abschnitt 4).

⁵ zur Antizipation der Verstehensprozesse eines Hörers siehe auch [37]

In Abb. 3 sind einige der zum Aufbau des ersten Teils der Gebrauchsanweisung aus Abschnitt 3 durchzuführenden Handlungen aufgeführt. Dabei sind Haupthandlungen durch HH, Nebenhandlungen durch NH gekennzeichnet.

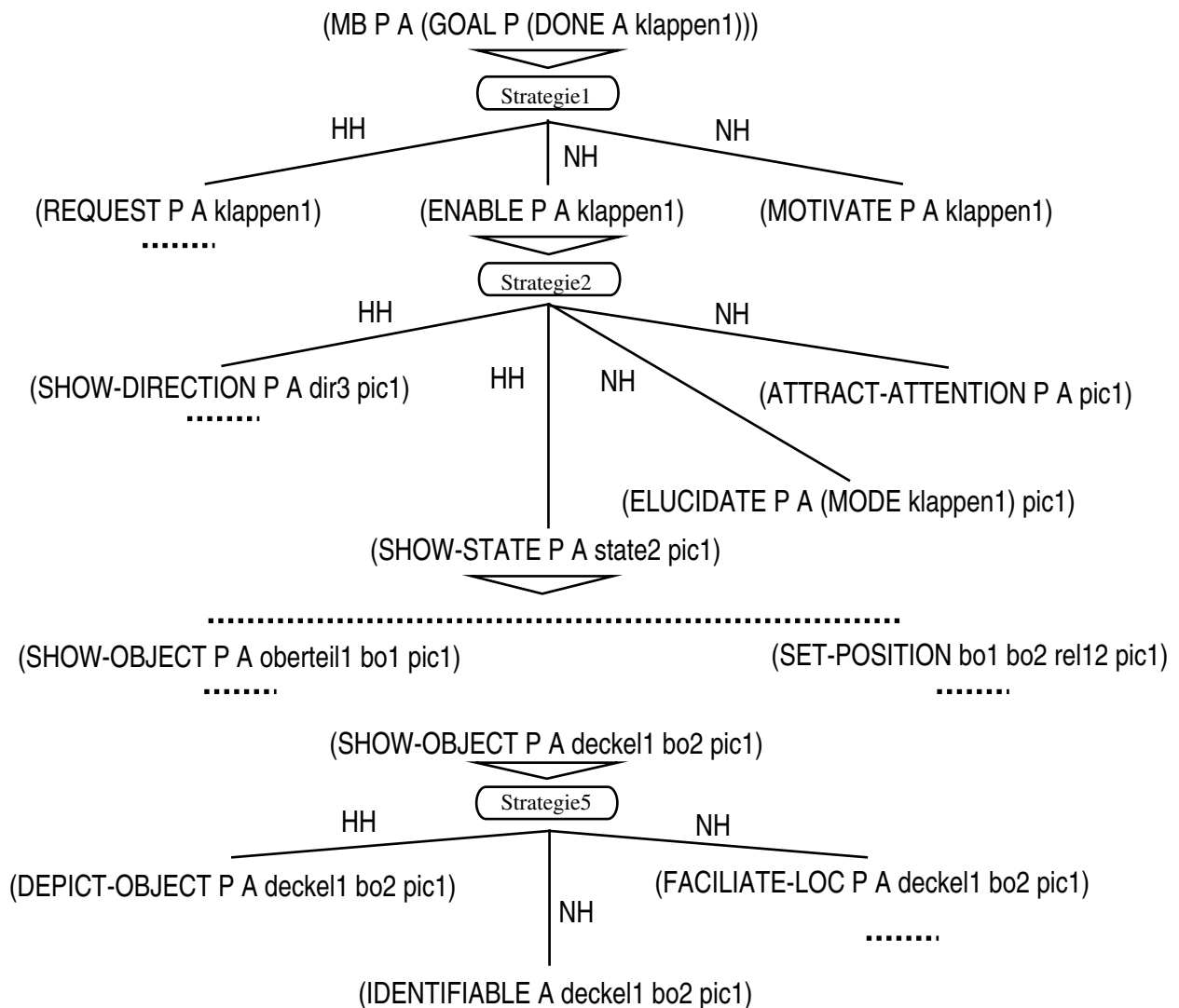


Abb. 3: Planausschnitt zur Generierung einer Aufforderung

6 Schlußbemerkung

In dieser Arbeit wurde dargelegt, daß nicht nur Texte, sondern auch Bilder und Text-Bild-Kombinationen als Mittel zur Erreichung von Zielen aufgefaßt werden können. Es wurden Strategien vorgestellt, die Wissen über Präsentationstechniken repräsentieren. In Anlehnung an textlinguistische Arbeiten wurde zwischen Haupt- und Nebenhandlungen unterschieden. Diese können sich sowohl auf den Aufbau einer Graphik als auch auf die Erzeugung von Text beziehen und ermöglichen somit

eine gemeinsame Planung von Text und Bild. Der vorgestellte Ansatz ist vor allem in bezug auf die folgenden Punkte von Vorteil:

○ *Komplementarität*

Da jedes Ziel genau einmal realisiert wird, ist gewährleistet, daß Text und Bild zueinander komplementär sind. Allerdings ist bei der Definition der Präsentationsstrategien zu berücksichtigen, daß diese das Komplementaritätsprinzip nicht von vornherein verletzen.

○ *Kohärenz*

Ein Dokument erscheint umso kohärenter, je einfacher sich eine Beziehung zwischen den einzelnen Dokumentteilen herleiten läßt. In unserem Ansatz besteht immer eine solche Beziehung, da der Planer nur Dokumentteile (Bilder und Textpassagen) generiert, die durch die in den Präsentationsstrategien spezifizierten funktionalen Beziehungen miteinander verbunden sind.

○ *Kohäsion*

Die vorgeschlagenen Präsentationsstrategien geben explizit an, welchen Zweck ein Dokumentteil zu erfüllen hat. Neben Verweisen auf den Bildinhalt können somit insbesondere auch Verweise generiert werden, die sich auf die Funktion eines Bildes beziehen.

Da es in vielen Fällen schwierig ist, Dokumente a priori so zu planen, daß diese verständlich sind, ist es nicht zuletzt aus Effizienzgründen sinnvoller, Dokumentteile vorab zu erzeugen und dann anschließend auf ihre Verständlichkeit hin zu überprüfen und gegebenenfalls zu revidieren. Unser Ansatz erlaubt es, die Verstehensprozesse des Adressaten in einer nach der Haupthandlung auszuführenden Nebenhandlung zu antizipieren.

Nachdem die Grundkonzeption des Präsentationsplaners feststeht, konzentrieren sich die derzeitigen Bemühungen auf dessen Implementierung. Parallel dazu werden Komponenten zur Text- und Graphikgenerierung entwickelt, die dann die in den Präsentationsstrategien aufgeführten elementaren Handlungen wie z.B. INFORM oder DEPICT-OBJECT ausführen. In einer weiteren Ausbaustufe soll der vorgestellte Ansatz auf weitere Modi wie Gestik und Animation erweitert werden.

Danksagung

Die Autoren bedanken sich bei Wolfgang Wahlster für wertvolle Hinweise und Anregungen zu diesem Papier.

Literatur

1. G. Adorni, M. DiManzo und F. Giunchiglia. Natural Language Driven Image Generation. In: COLING'84, S. 495-500 (1984)

2. E. **André**. Strategien zur Generierung multimodaler Dokumente. Memo, Deutsches Forschungszentrum für Künstliche Intelligenz, Saarbrücken, i. Vorb..
3. D.E. **Appelt**. Planning English Sentences. Cambridge University Press: London (1985)
4. Y. **Arens**, L. **Miller**, S.C. **Shapiro** und N.K. **Sondheimer**. Automatic Construction of User-Interface Design. In: Proceedings of the 7th National Conference of the American Association for Artificial Intelligence, S. 808-813 (1988)
5. S.-P. **Ballstaedt**, H. **Mandl**, W. **Schnotz** und S.-O. **Tergan**. Texte verstehen, Texte gestalten. Urban & Schwarzenberg: München, Wien, Baltimore (1981)
6. J.G. **Bennett**. Depiction and Convention. In: The Monist 58, S. 255-269 (1974)
7. G.R. **Bieger** und M.D. **Glock**. Comprehending Spatial and Contextual Information in Picture-Text Instructions. The Journal of Experimental Education 54(4), S. 181-188 (1986)
8. L.C. **Bogges**. Computational Interpretation of English Spatial Prepositions. Technical Report No. T-75, Coordinated Science Laboratory, University of Illinois (1978)
9. M. **Brandt**, W. **Koch**, W. **Motsch** und I. **Rosengren**. Der Einfluß der kommunikativen Strategie auf die Textstruktur - dargestellt am Beispiel des Geschäftsbriefes. In: I. Rosengren (Hrsg.), Sprache und Pragmatik, Lunder Symposium 1982. Almqvist & Wiksell: Stockholm, S. 105-135 (1983)
10. J. **Bransford** und M. **Johnson**. Contextual Prerequisites for Understanding: Some Investigations of Comprehension and Recall. In: Journal of Verbal Learning and Verbal Behavior 11, S. 717-726 (1972)
11. P.R. **Cohen**. On Knowing What to Say: Planning Speech Acts. Technical Report No. 118, University of Toronto (1978)
12. P.R. **Cohen** und H.J. **Levesque**. Speech Acts and Rationality. In: Proceedings of the 23rd Annual Meeting of the ACL, S. 49-59 (1985)
13. T. A. **van Dijk**. Textwissenschaft. dtv: München (1980)
14. M. **Elhadad**, D.D. **Seligmann**, S. **Feiner** und K.R. **McKeown**. A Common Intention Description Language for Interactive Multi-Media Systems. Proc. of the IJCAI-89 Workshop 'A New Generation of Intelligent Interfaces', Detroit, Michigan (1989)
15. S. **Feiner**. An Architecture for Knowledge-Based Graphical Interfaces. In: Proceedings of the Workshop on Architectures of Intelligent Interfaces: Elements & Prototypes, S. 129-140 (1988)
16. M. **Friedell**. Automatic Synthesis of Graphical Object Descriptions. In: Computer Graphics (ACM) 18(3), S. 53-62 (1984)
17. J. **Geller** und C. **Shapiro**. Graphical Deep Knowledge for Intelligent Machine Drafting. In: Proceedings of the 10th International Joint Conference on Artificial Intelligence, S. 545-551 (1987)
18. N. **Goodman**. Replies. In: Erkenntnis 12, S. 153-179 (1978)
19. E. H. **Hovy**. Approaches to the Planning of Coherent Text. Papers from the 4th International Workshop on Text Generation, Catalina Island (1988)

20. **S. Jörg.** Der Einfluß sprachlicher Bezeichnungen auf das Wiedererkennen von Bildern. Huber: Bern (1978)
21. **S. Jörg und H. Hörmann.** Sentences before and after Pictures: How do Verbal Specifications Influence Recognition?. In: Psychological Research 45, S. 255-266 (1983)
22. **G. Kempen** (Hrsg.). Natural Language Generation: New Results in Artificial Intelligence, Psychology, and Linguistics. Nijhoff: Dordrecht, Boston, Lancaster (1987)
23. **S. Kjørup.** George Innes and the Battle at Hastings, or Doing Things with Pictures. In: The Monist 58, S. 216-235 (1974)
24. **S. Kjørup.** Pictorial Speech Acts. In: Erkenntnis 12, S. 55-71 (1978)
25. **S.M. Kosslyn.** Image and Mind. Harvard University Press: Cambridge, Massachusetts, London (1980)
26. **J. Mackinlay.** Search Architecture for the Automatic Design of Graphical Presentations. In: Proceedings of the Workshop on Architectures of Intelligent Interfaces: Elements & Prototypes, S. 129-140 (1988)
27. **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, Nijhoff: Dordrecht, Boston, Lancaster, S. 85-95 (1987)
28. **K.F. McCoy.** Highlightening a User Model to Respond to Misconceptions. In: A. Kobsa und W. Wahlster (Hrsg.), User Models in Dialog Systems, Springer: Berlin, Heidelberg, New York (1989)
29. **K.R. McKeown.** Text Generation. Cambridge University Press: London (1985)
30. **J.D. Moore und C.L. Paris.** Planning Text for Advisory Dialogues. In: Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics (1989)
31. **J.D. Moore und W.R. Swartout.** A Reactive Approach to Explanation. In: Proceedings of the 11th International Joint Conference on Artificial Intelligence (1989)
32. **M. Muckenhaupt.** Text und Bild. Gunter Narr: Tübingen (1986)
33. **D. Novitz.** Picturing. In: The Journal of Aesthetics and Art Criticism 34, S. 145-155 (1975)
34. **C.L. Paris.** The Use of Explicit User Models in a Generation System for Tailoring Answers to a User's Level of Expertise. In: A. Kobsa und W. Wahlster (Hrsg.), User Models in Dialog Systems, Springer: Berlin, Heidelberg, New York (1989)
35. **I. Rosengren.** Die Textstruktur als Ergebnis strategischer Überlegungen des Senders. In: I. Rosengren (Hrsg.), Sprache und Pragmatik, Lunder Symposium 1982, Almqvist & Wiksell: Stockholm, S. 157-191 (1983)
36. **J.R. Searle.** Speech Acts: An Essay in the Philosophy of Language. Cambridge University Press: Cambridge, MA (1969)
37. **J.R.J. Schirra.** Ein erster Blick auf ANTLIMA: Visualisierung statischer räumlicher Reaktionen. In: Metzging (Hrsg.), GWAI-89, Springer: Berlin, Heidelberg, S. 301-311 (1989)

38. **J.W. Sullivan** und **S.W. Tyler** (Hrsg.). *Architectures for Intelligent User Interfaces: Elements and Prototypes*. Addison-Wesley (1989)
39. **W. Wahlster**, **E. André**, **M. Hecking** und **T. Rist**. *WIP: Knowledge-based Presentation of Information*. Report WIP-1, Deutsches Forschungszentrum für Künstliche Intelligenz, Saarbrücken (1989)
40. *Proceedings of the IJCAI-89 Workshop 'A New Generation of Intelligent Interfaces'*, Detroit, Michigan (1989)
41. **D.M. Willows** und **H.A. Houghton**. *The Psychology of Illustration - Basic Research*. Springer: Berlin, Heidelberg, New York (1987)
42. **M. Wintermantel**, **L. Laux** und **U. Fehr**. *Anweisung zum Handeln: Bilder oder Wörter*. Bericht Nr. 2, Psychologisches Institut der Universität Heidelberg (1989)
43. **L. Wittgenstein**. *Tractatus logico-philosophicus*. Nachgedruckt in: *Werkausgabe: L. Wittgenstein*. Band 1, Suhrkamp: Frankfurt (1988)
44. **F. Zdybel**, **N. Greenfeld** und **M. Yonke**. *An Information Presentation System*. In: *Proceedings of the 7th International Joint Conference on Artificial Intelligence*, S. 978-984 (1981)
45. **M. Zock** und **G. Sabah** (Hrsg.). *Advances in Natural Language Generation*. Pinter: London (1988)

Wissensbasierte Perspektivenwahl für die automatische Erzeugung von 3D-Objektdarstellungen

Thomas Rist, Elisabeth André
Deutsches Forschungszentrum für Künstliche Intelligenz
Stuhlsatzenhausweg 3
D-6600 Saarbrücken 11

Zusammenfassung

Aus welcher Perspektive ein Objekt gezeigt werden soll, ist eine der elementaren Fragen, die sich bei der automatischen Erzeugung von 3D-Darstellungen stellt, die aber in den wenigen Systemen, die graphische Objektdarstellungen selbstständig planen, bisher vernachlässigt wurde. Ziel der vorliegenden Arbeit ist es, aufzuzeigen, wie sich Wissen über Objekte und Darstellungstechniken verwenden läßt, um die Menge der möglichen Perspektiven, aus denen ein Objekt gesehen und gezeigt werden kann, sinnvoll einzuschränken. Als Grundlage zur Perspektivenwahl schlagen wir ein Bezugssystem vor, das eine Einteilung der Perspektiven in 26 Klassen nahelegt und das darüberhinaus Vorteile bietet, wenn gewählte Perspektiven natürlichsprachlich zu beschreiben sind. Anschließend führen wir einige für die Perspektivenwahl relevante Kriterien an. Diese Kriterien werden dann zur Formulierung von Regeln herangezogen, die wir dazu verwenden, um in einer konkreten Präsentationssituation eine geeignete Perspektive zu bestimmen.

1. Problemstellung

Neben Sprache wird in zahlreichen Dokumenten auf graphische Objektdarstellungen als Medium zur Informationsvermittlung zurückgegriffen. Sie werden beispielsweise in Beschreibungen für technische Geräte und Anlagen dazu verwendet, um die Identifikation oder die räumliche Einordnung von Objekten zu erleichtern. Ziel unserer Arbeit ist die Entwicklung eines wissensbasierten Systems¹, das in der Lage ist, Information der jeweiligen Präsentationssituation angepaßt zu präsentieren, wobei

¹ Der vorgestellte Ansatz zur Perspektivenwahl wurde für eine Graphikdesignkomponente entworfen, die derzeit im Rahmen des vom BMFT unter dem Förderkennzeichen ITW8901 8 geförderten WIP-Projekts am Deutschen Forschungszentrum für Künstliche Intelligenz in Saarbrücken entwickelt wird.

unterschiedliche Präsentationsmodi (Text, Graphik) aufeinander abgestimmt zum Einsatz kommen. In bezug auf die Erzeugung graphischer Objektdarstellungen muß ein solches System u.a. die folgenden Entscheidungen treffen:

- Arrangieren der darzustellenden Objekte
- Auswahl einer Projektionstechnik
- Auswahl einer geeigneten Perspektive

Bevor ein Abbildungsprozeß² ausgeführt wird, sollte eine günstige Vorlage geschaffen werden, indem man die darzustellenden Objekte in eine geeignete Ausgangsposition bringt. Die Auswahl einer geeigneten Projektionstechnik wird sich vor allem danach richten, für welchen Zweck und für welche Zielgruppe eine Abbildung bestimmt ist. Geht es darum, Objekte so darzustellen, wie sie erscheinen, so wird in der Regel eine perspektivische Projektion gewählt. Für technische Zeichnungen werden hingegen Parallelprojektionen bevorzugt, da diese Form und Ausmaß weniger stark verzerren. Wie gut eine Objektabbildung die ihr zugeordnete Funktion erfüllen kann, hängt schließlich auch davon ab, aus welcher Perspektive das Objekt gezeigt wird.

Während die Ausführung von Projektionen zum Standard heutiger Graphiksysteme zählt, werden die oben angeführten Entscheidungen nicht automatisch getroffen. Bei interaktiven Graphiksystemen, etwa aus dem CAD-Bereich, liegt die Verantwortung darüber, wie Objekte auf einem Ausgabemedium erscheinen, beim Benutzer. In den wenigen Systemen, die graphische Objektdarstellungen selbstständig planen (vgl. [4], [5] und [10]), wird anstelle eines vom System durchzuführenden Entscheidungsprozesses bislang auf vordefinierte Standardperspektiven zurückgegriffen, die fester Bestandteil der jeweils zugrundeliegenden geometrischen Objektrepräsentationen sind.

Eine Beschränkung auf Standardperspektiven führt allerdings in vielen Präsentationssituationen zu ungeeigneten Darstellungen. Rein geometrisch repräsentierte Perspektiven sind u.U. schwierig in natürlicher Sprache zu beschreiben. Solche Beschreibungen sind aber bei einigen graphischen Darstellungen erforderlich, um ihr Verstehen zu gewährleisten, etwa dann, wenn ein Objekt aus einer 'ungewöhnlichen' Perspektive gezeigt werden muß. In diesem Beitrag schlagen wir als Grundlage zur Perspektivenwahl ein Bezugssystem vor, das eine Einteilung der Perspektiven in 26 Klassen nahelegt und das darüberhinaus einer natürlichsprachlichen Beschreibung entgegenkommt. Zum anderen beschäftigen wir uns mit der Frage, wie Wissen über Objekte und Darstellungstechniken dazu eingesetzt werden kann, um in einer konkreten Präsentationssituation eine geeignete Perspektive auszuwählen.

2. Grundperspektiven

In einer Abhandlung über Sprache und Raum stellen Miller und Johnson-Laird (vgl. [9]) heraus, daß Menschen dazu tendieren, Objekten intrinsische Seiten zuzusprechen, und zwar maximal drei

² Abgebildet werden in unserem Fall Objektmodelle, die zur Repräsentation des Wissens über Geometrie und Topologie der Domänenobjekte dienen.

Seitenpaare; Ober- und Unterseite, Vorder- und Hinterseite sowie linke und rechte Seite. Diese Seiten werden nach dem Vorbild des menschlichen Körpers oder der menschlichen Raumorientierung vergeben (vgl. [12]). Als einen 'maximalen Prototyp' kann man sich ein mit sechs intrinsischen Seiten ausgezeichnetes, quaderförmiges Objekt vorstellen, das entweder wie in Abb. 1 oder wie in Abb. 2 dargestellt organisiert ist, je nachdem, ob die Zuweisung der Seiten nach dem *Spiegelbildprinzip* oder nach dem *Koinzidenzprinzip* erfolgt ist.

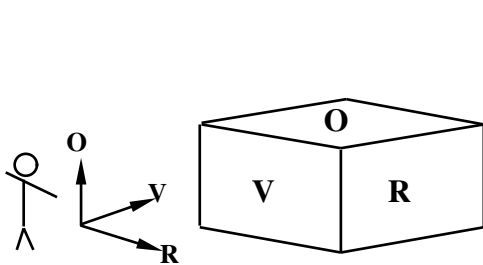


Abb.1: Zuordnung intrinsischer Objektseiten nach dem Spiegelbildprinzip

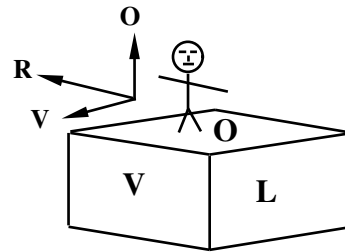


Abb.2: Zuordnung intrinsischer Objektseiten nach dem Koinzidenzprinzip

Betrachten wir nun die Situation, in der ein Beobachter einem wie in Abb. 1 organisierten quaderförmigen Objekt frontal gegenübersteht. Von dieser Situation ausgehend konstruieren wir ein dreidimensionales Bezugssystem, indem wir Achsen senkrecht durch die Mittelpunkte der Quaderflächen legen. Wenn wir in diesem Bezugssystem einen weiteren Beobachter annehmen, der ebenfalls aus genügender Entfernung den Quader betrachtet, macht es Sinn zu sagen, ob dieser in bezug auf den ersten Beobachter *von oben*, *von vorne*, *von vorne rechts* usw. auf den Quader schaut (vgl. Abb. 3). Die möglichen Perspektiven lassen sich somit nach den Raumrichtungen klassifizieren, aus denen der Quader betrachtet wird. Mit den Abkürzungen O, U, V, H, L, R für *von oben*, *von unten*, *von vorne*, *von hinten*, *von links* und *von rechts* ergeben sich die folgenden 26 Klassen³ von Perspektiven, die wir im folgenden als *Quader-Grundperspektiven* bezeichnen:

V, H, O, U, L, R
 VO, VU, VL, VR, OH, OL, OR, UL, UR, UH, HL, HR
 VOL, VOR, VUL, VUR, HOL, HOR, HUL, HUR

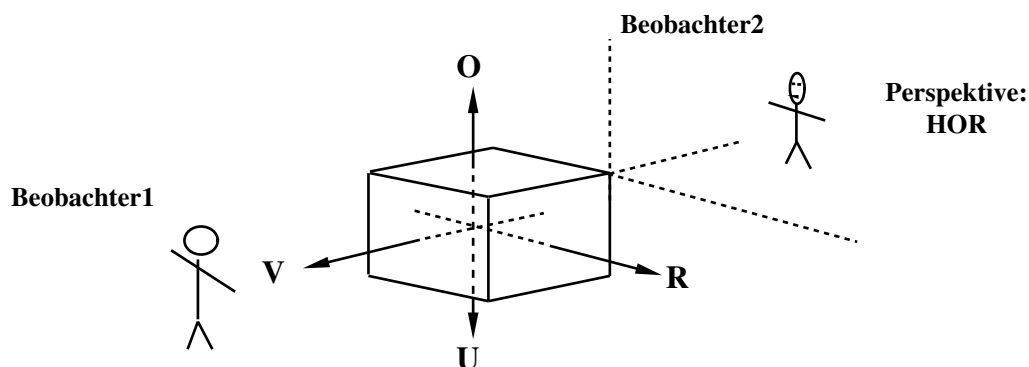


Abb.3: Beobachter2 betrachtet den Quader hinsichtlich des Bezugssystems von Beobachter1 *von hinten oben rechts*.

³ Weitere Klassen erhält man, wenn man Hohlkörper zulässt, die von innen betrachtet werden. Dort können auch Perspektiven gewählt werden, aus denen vier oder fünf Quaderseiten sichtbar sind.

Vorausgesetzt, der betrachtete Quader ist wie in Abb. 1 organisiert, spiegeln die eingeführten Grundperspektiven die jeweils sichtbaren Quaderseiten wider. Für den Fall, daß Objekte betrachtet werden, deren intrinsische Seiten nach dem Koinzidenzprinzip zugeordnet sind, zeigen die linke und rechte Seite in die entgegengesetzten Richtungen des Bezugssystems. Wird beispielsweise ein Auto aus einer Perspektive *von rechts* betrachtet, so ist dessen intrinsische *linke* Seite sichtbar. Unabhängig von der Zuweisung intrinsischer Seiten ist die Tatsache, daß ein Perspektivenwechsel nur dann zu einer qualitativen Änderung der Quaderdarstellung führt (d.h. neue Kanten und Flächen kommen zum Vorschein und/oder bisher sichtbare werden verdeckt), wenn damit auch ein Klassenwechsel verbunden ist (vgl. Abb. 4). Verschiebungen der Perspektive innerhalb einer Klasse bedeuten hingegen quantitative Veränderungen bezogen auf Kantenlänge, Flächeninhalt und Winkelgröße.

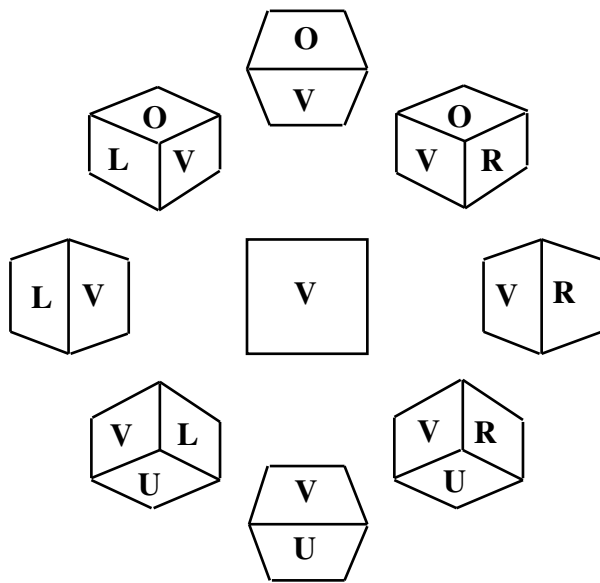


Abb.4: Quader-Grundperspektiven, aus denen die vordere Quaderseite sichtbar ist.

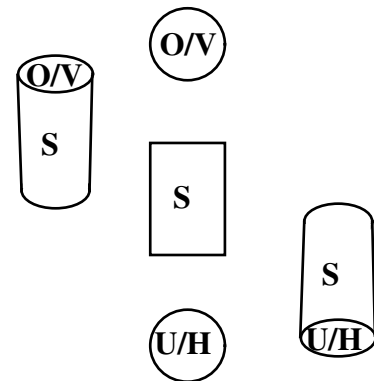


Abb.5: Die fünf Zylinder-Grundperspektiven.

Mit dem in Abb. 3 vorgestellten Bezugssystem ist es uns möglich, zwischen Arrangement und Perspektivenwahl zu unterscheiden. Eine Perspektive wählen bedeutet, eine bestimmte Beobachterposition in diesem Bezugssystem einzunehmen. Arrangieren bedeutet, ein Objekt im Bezugssystem zu plazieren. Eine 'Standard-Plazierung' erhalten wir, indem wir das abzubildende Objekt so in das in Abb. 3 dargestellte Bezugssystem einpassen, daß dessen intrinsische Seiten an den entsprechenden Raumrichtungen ausgerichtet sind. Für Objekte mit sechs intrinsischen Seiten, deren Form von der eines Quaders abweicht, greifen wir dabei auf *umschreibende* Quader zurück. Weisen Objekte nur zwei intrinsische Seiten auf, also nur Ober- und Unterseite (z.B. Pilz, Litfaßsäule) oder nur Vorder- und Rückseite (z.B. Bleistift, Nadel), dann können einige der 26 Quader-Grundperspektiven zusammengefaßt werden, da sie sich hinsichtlich der aus ihnen jeweils sichtbaren Objektseiten nicht unterscheiden. Beispielsweise ist die Seitenansicht eines an der Unten/Oben-Achse ausgerichteten Pilzes davon unabhängig, ob er aus einer der Perspektiven V, H, L, R oder aus einer ihrer Kombinationen betrachtet wird. Anstelle eines umschreibenden Quaders verwenden wir bei diesen Objekten umschreibende Zylinder und beschränken uns bei der Perspektivenwahl auf die fünf

in Abb. 5 dargestellten *Zylinder-Grundperspektiven*. Für die Darstellung eines Pilzes ergeben sich beispielsweise die Grundperspektiven: O, U, S, OS und US, wobei die Abkürzung S für die nicht weiter differenzierten Objektseiten steht. Für Objekte, die überhaupt keine intrinsischen Seiten aufweisen (z.B. Ball, Stein), ist eine Unterscheidung der Perspektiven hinsichtlich der sichtbaren intrinsischen Seiten nicht möglich.

Ob ein Objekt intrinsische Seiten besitzt oder nicht, kann entweder explizit in einer Wissensbasis vermerkt sein oder muß erschlossen werden, z.B. mit Hilfe der Regeln, die von Miller und Johnson-Laird aufgestellt wurden, um Objekte hinsichtlich ihrer intrinsischen Teile zu kategorisieren (vgl. [9]). Kann nicht festgestellt werden, ob ein Objekt über intrinsische Seiten verfügt, weil zu wenig Information über das Objekt vorliegt, ist es in vielen Fällen plausibel, wenigstens eine Unterseite zu bestimmen, etwa dadurch, daß man das Objekt hinsichtlich des Bezugssystems in eine stabile Lage bringt ⁴.

3 Kriterien zur Perspektivenauswahl

In Handbüchern für Graphiker und Designer findet man gewöhnlich den Hinweis, Objekte so darzustellen, wie sie unter `normalen Sehbedingungen´ erscheinen (vgl. u.a. [2] und [7]). Betrachtet man etwa Darstellungen von Pilzen (z.B. in Lehrbüchern oder Lexika), so fällt auf, daß dort meistens Seitenansichten bevorzugt werden, obwohl Pilze aufgrund ihrer geringen Größe von Menschen in der Regel von oben gesehen werden. Das Kriterium der `normalen Sehbedingung´ erscheint zu vage und im Hinblick auf unterschiedliche Präsentationssituationen auch keinesfalls ausreichend, um die Zahl der möglichen Perspektiven sinnvoll einzuschränken. Vielmehr ist davon auszugehen, daß dieser Entscheidungsprozeß durch eine Vielzahl von Faktoren determiniert wird. Welche Faktoren im einzelnen zu berücksichtigen sind, hängt davon ab, ob man vorrangig daran interessiert ist, wie menschliche Zeichner dieses Problem bewältigen, etwa im Hinblick auf die Entwicklung eines entsprechenden Simulationsmodells, oder ob der performative Aspekt, nämlich in einer gegebenen Präsentationssituation angemessene Darstellungen automatisch zu erzeugen, im Vordergrund steht.

Einige der für die Perspektivenwahl maßgeblichen Faktoren, die bei einem Simulationsansatz zu berücksichtigen wären, sind nach von Sommers eher technischen Charakters (vgl. [11]). Wie ein Objekt dargestellt wird, ist seiner Meinung nach auch davon abhängig, an welcher Stelle der Zeichenunterlage mit dem Zeichnen begonnen wird (nicht zuletzt aus anatomischen Gründen ist das bevorzugt der linke obere Teil der Zeichenunterlage) und was zuerst vom Objekt gezeichnet wird. Als Paradebeispiel führt er die stets als von links nach rechts wehend gezeichneten Flaggen an. Von diesen wird in der Regel zuerst mit dem Mast auf der linken Papierhälfte begonnen; die Flagge muß dann zwangsläufig nach rechts ergänzt werden. Interessant in bezug auf ein Simulationsmodell wäre auch die häufig vertretene These, daß menschliche Zeichner, konfrontiert mit der Aufgabe, ein Objekt

⁴ Hierzu kann z.B. auf die von Adorni formulierten Gesetze der naiven Statik zurückgegriffen werden (vgl. [1]).

abzuzeichnen, auf erlernte Schemata zurückgreifen, die sie dann solange abändern, bis eine mehr oder weniger große Ähnlichkeit zwischen Abbild und Vorlage erreicht ist (vgl. [6]). Zahlreiche Anleitungen für Zeichner zielen im wesentlichen auf die Vermittlung solcher Schemata ab. So wird etwa gelehrt, wie aus einfachen geometrischen Grundformen konventionelle Darstellungen für Blumen, Katzen, Vögel oder Schiffe aufgebaut werden (vgl. z.B. [8]). Vor allem bei ungeübten Zeichnern ist daher anzunehmen, daß die Wahl einer Perspektive auch davon abhängen wird, welche Schemata diese kennen bzw. wie leicht neue Schemata zu erlernen sind.

Da es uns allerdings weniger um die Simulation des menschlichen Zeichnens geht, bleiben Faktoren, wie sie eben genannt wurden, unberücksichtigt. Stattdessen konzentrieren wir uns auf einen Aspekt, der sowohl bei einem Simulations- als auch bei einem Performanzansatz eine entscheidende Rolle spielt, nämlich auf die Konsequenzen, die sich durch die Präsentationssituation und die sich daraus ergebenden Ziele für die Perspektivenwahl bei Objektdarstellungen ergeben. Im folgenden sind einige Beispiele angeführt.

○ **Funktionalität/Zugänglichkeit hervorheben**

In engem Zusammenhang mit dem Kriterium der `normalen Sehbedingungen´ steht vermutlich die von einem Betrachter gegenüber einem Objekt typischerweise eingenommene Betrachterposition. Diese spiegelt sich aber gerade in der Zuordnung der intrinsischen Objektseiten wider. Bei technischen Geräten wird beispielsweise häufig diejenige Seite als intrinsische Vorderseite angesehen, an der sich wichtige Bedienungs- und Funktionsteile befinden. Nicht zugängliche Objektflächen, etwa Auflageflächen, werden hingegen zu Unter- oder Rückseiten. Sie sind aus der Standard-Betrachterposition meist nicht zu sehen.

○ **Aktuelle Beobachterposition berücksichtigen**

Wird in einem Dokument, z.B. einer Konstruktionsanleitung, dazu aufgefordert, Objektmanipulationen (Drehungen usw.) durchzuführen, so ist zu beachten, daß dem Betrachter andere als die beispielsweise aus der Standard-Betrachterposition sichtbaren Objektseiten zugänglich werden. Die Verständlichkeit solcher Dokumente kann dann im allgemeinen erhöht werden, wenn diese Operationen in nachfolgenden Objektdarstellungen Berücksichtigung finden.

○ **Bewegungsrichtung betonen**

Soll herausgestellt werden, daß sich ein darzustellendes Objekt in eine bestimmte Richtung bewegt, werden üblicherweise Ansichten gewählt, die zur Bewegungsrichtung parallele Objektseiten zeigen. Dabei scheint es keine Rolle zu spielen, ob sich das darzustellende Objekt in der Horizontalen oder in der Vertikalen bewegt. Die typischsten Beispiele sind hier Darstellungen von Pfeilen. Ähnliches gilt aber auch für Schiffe, Autos, Fische, Raketen oder Regentropfen. Van Sommers (vgl. [11]) vertritt sogar die Auffassung, daß Trompeten bevorzugt von der Seite gezeichnet werden, um die Fortpflanzungsrichtung des Schalls anzudeuten.

○ **Räumlichkeit vermitteln**

Räumliche Tiefe kommt besser zum Ausdruck, wenn das Objekt aus einer Perspektive dargestellt wird, bei der in den Raum hineingehende Flächen deutlich erkennbar sind. Wird nur eine Oberseite dargestellt, verschwindet das Volumen gänzlich. Insbesondere ist zu berücksichtigen, daß im Falle einer ungünstig gewählten Perspektive formgebende Kanten und Linienschnittpunkte aufeinander fallen und somit nicht mehr zu unterscheiden sind. Während die linke Figur in Abb. 6 gut als Darstellung eines quaderförmigen Objektes zu erkennen ist, bereitet die rechte Schwierigkeiten. Durch das Zusammenfallen zweier Schnittpunkte bleibt unklar, ob es sich um die Darstellung eines dreidimensionalen Objekts handelt oder um ein Sechseck mit eingezeichneten Diagonalen.



Abb. 6: Günstige und ungünstige Darstellung eines quaderförmigen Drahtgitters

○ **Ein bestimmtes Objektteil zeigen**

Oft werden Objektdarstellungen verwendet, um Information über ein bestimmtes Objektteil darzubieten, etwa um zu zeigen, daß ein Objekt das Teile X besitzt, und/oder um zu zeigen, wo sich das Teil X befindet. In diesen Fällen müssen dann Perspektiven gewählt werden, aus denen das betreffende Teil sichtbar ist. Besteht das Präsentationsziel beispielsweise darin, auf die Aufhängung eines technischen Gerätes aufmerksam zu machen, wird man das Gerät zweckmäßigerweise nicht von vorne, sondern von hinten zeigen.

○ **Identifizierung erleichtern**

Variationen in der Perspektive führen zu unterschiedlichen Abbildungen ein und desselben Objektes. Das Ergebnis einer Projektion wird allerdings nur dann als Abbild eines Objektes erkennbar sein, wenn ausreichend viele visuell wahrnehmbare Eigenschaften vorhanden sind, die mit den visuell wahrnehmbaren Eigenschaften der Vorlage identifizierbar sind. Je mehr diskriminierende Eigenschaften eine graphische Objektdarstellung beinhaltet, desto geringer sind die Verwechslungsmöglichkeiten mit anderen Objekten. Ein Pilz ist beispielsweise aus der Vogelperspektive schlechter zu identifizieren als aus einer Seitenansicht. Geht man davon aus, daß großflächige Objektseiten gegenüber kleineren eher wahrgenommen werden, ist es oft günstig, solche ausgeprägten Seiten möglichst unverzerrt darzustellen. Beispielsweise fällt die Identifikation eines aus der Frontalperspektive gezeigten Bleistifts in der Regel wesentlich schwerer als eine Darstellung in Seitenansicht.

○ **Wirkungen erzielen**

Vor allem in Werbeanzeigen wird häufig versucht, Objekte so darzustellen, daß die Betrachter ihnen bestimmte wertende Merkmale (angenehm, wertvoll etc.) zuordnen. Die Wahl der Perspektive ist dabei ein wesentliches gestalterisches Mittel. In diesem Zusammenhang verweist Espe (vgl. [3]) auf eine

Faustregel, die besagt, daß ein aus der Vogelperspektive gezeigtes Motiv als dem Betrachter unterlegen, ein aus der Froschperspektive gezeigtes hingegen als dem Betrachter überlegen empfunden wird.

Durch Angabe solcher Präsentationsziele läßt sich repräsentieren, welche kommunikative Funktion eine Objektdarstellung in einer konkreten Präsentationssituation zu erfüllen hat. Als Standardperspektiven bezeichnen wir die Perspektiven, die sich durch die Vorgabe von Standardpräsentationszielen ergeben. Hierzu gehört die Vermittlung von Räumlichkeit, die Erleichterung der Identifizierung, bei statischen Objekten die Hervorhebung der Funktionalität/Zugänglichkeit und bei dynamischen die Betonung der Bewegungsrichtung.

4. Regelgesteuerte Perspektivenauswahl

Ausgehend von den eingeführten Grundperspektiven können wir mit den oben genannten Kriterien feststellen, was in einer konkreten Präsentationssituation für bzw. gegen eine bestimmte Perspektive spricht. Eine Entscheidung ist schließlich durch Abwägen der für- und widersprechenden Argumente herbeizuführen. Wir gehen dabei jedoch nicht davon aus, daß es stets nur eine Perspektive gibt, die den gestellten Anforderungen gerecht wird. Wir wollen vielmehr ungeeignete oder weniger geeignete Perspektiven ausschließen. Um Kriterien möglichst einfach zu verfeinern und neue Kriterien hinzunehmen zu können, wählen wir einen regelbasierten Ansatz. Die Regeln werden wie folgt abgearbeitet. Im Prämissenteil wird geprüft, ob die Regel in der vorgegebenen Präsentationssituation anwendbar ist; im Aktionsteil wird eine Bewertung der zur Auswahl stehenden Perspektiven vorgenommen. Dabei können einzelne Perspektiven durch ein Minuszeichen blockiert oder durch ein Pluszeichen als besonders geeignet gekennzeichnet werden. Einige solcher Regeln sind nachfolgend skizziert:

- (R1) Funktionalität betonen =>
blockiere alle Perspektiven, die die funktionale Seite nicht enthalten, befürworte Frontalansicht, der durch die Funktion gegebenen Seite
- (R2) Unterseite identisch mit Auflagefläche =>
blockiere alle Perspektiven, die die Unterseite zeigen
- (R3) Räumlichkeit betonen =>
blockiere Frontalansichten, befürworte Perspektiven, die drei Seiten zeigen
- (R4) Identifizierung erleichtern =>
blockiere Frontalansichten extrem kleiner Seiten

Wie solche Bewertungsregeln zur Perspektivenwahl eingesetzt werden, demonstrieren wir anhand eines Beispiels aus der Domäne des WIP-Projekts. In dieser Domäne geht es darum, eine illustrierte Gebrauchsanleitung für eine Espressomaschine unter Berücksichtigung verschiedener Generierungsparameter wie Präsentationssituation, Zielgruppe usw. automatisch zu erzeugen. Wir betrachten die Situation, in der zu Beginn einer Gebrauchsanleitung die zu beschreibende Espressomaschine erstma-

lig erscheinen soll. Spezielle Präsentationsziele seien nicht vorgegeben, so daß nur die mit den Standardzielen verbundenen Regeln (R1 bis R4) herangezogen werden. Ausgehend von der durch funktionale Teile (Einschaltknopf usw.) gegebenen Vorderseite und der durch die Auflagefläche bestimmten Unterseite werden der espressomaschine gemäß dem Spiegelbildprinzip (vgl. Abb. 1) sechs intrinsische Seiten zugeordnet. Es sind daher alle 26 Quader-Grundperspektiven in Betracht zu ziehen. Nach Anwendung von R1, R2 und R3 werden die Grundperspektiven VOL und VOR gefunden. Da die espressomaschine keine extrem kleinen intrinsischen Seiten aufweist, ist R4 nicht anwendbar. Ohne weitere Restriktionen erscheinen VOL und VOR gleichermaßen geeignet. Das Ergebnis der Bewertung durch die anwendbaren Regeln ist in nachstehender Tabelle zusammengefaßt.

	V	H	O	U	L	R	VO	VU	VL	VR	OH	OL	OR	UL	UR	UH	HL	HR	VOL	VOR	VUL	VUR	HOL	HOR	HUL	HUR	
(R1)	+	-	-	-	-	-					-	-	-	-	-	-	-	-					-	-	-	-	
(R2)				-			-							-	-							-	-			-	-
(R3)	-	-	-	-	-	-													+	+	+	+	+	+	+	+	

Abb.7: Ergebnis der Bewertung durch die Regeln R1 bis R3

Nicht auszuschließen ist allerdings, daß nach der Bewertung alle zur Auswahl stehenden Perspektiven blockiert sind. Soll in unserem Beispiel etwa noch die sich auf der Rückseite befindliche Aufhängung der espressomaschine dargestellt werden, dann gibt es keine Perspektive, die mit allen Präsentationszielen vereinbar ist. In solchen Fällen sind dann entweder mehrere Abbildungen zu erzeugen, oder die nicht dargestellte Information muß auf anderem Wege, z.B. verbal, mitgeteilt werden.

5. Schlußbemerkung

Vorgestellt wurde ein Ansatz, bei dem der prinzipiell unendliche⁵ Suchraum möglicher Perspektiven zunächst auf eine geringe Anzahl sich qualitativ unterscheidender Grundperspektiven eingeschränkt wurde, um aus diesen dann unter Rückgriff auf Wissen über Objekte und Darstellungstechniken eine der jeweiligen Präsentationssituation angemessene auszuwählen. Wir gehen davon aus, daß diese Grundperspektiven eine gute Basis für die Erzeugung von Darstellungen eines breiten Spektrums an Objekten bilden. Sicher ist richtig, daß für manche Objektklassen in bestimmten Präsentationssituationen bessere Ergebnisse erzielt werden können, wenn der Perspektivenwahl ein größerer Suchraum zugrundegelegt wird. Sollen beispielsweise mehrere Teile eines Objektes in einer Abbildung sichtbar sein, stellt sich die Wahl einer geeigneten Perspektive in erster Linie als geometrisches Problem. Wenn dann aber mit natürlicher Sprache auf die Abbildung Bezug

⁵ Selbst wenn nur ganzzahlige Koordinatentupel in Frage kommen und nur Positionen betrachtet werden, die einen bestimmten Abstand vom Objekt nicht überschreiten, verhält sich die Anzahl möglicher Perspektiven immerhin kubisch zu diesem Abstand.

genommen werden soll, kann wiederum auf die Grundperspektiven zurückgegriffen werden, um Perspektiven natürlichsprachlich zu beschreiben.

Noch offen ist die Frage, inwieweit sich der Ansatz für die Darstellung von Szenen eignet. Eine erste Überlegung zielt darauf ab, Szenen selbst als zusammengesetzte Objekte aufzufassen. Schwierigkeiten treten dann auf, wenn unterschiedliche Präsentationsziele nur durch geschicktes Arrangieren der Szenen-objekte zu erreichen sind.

Weitere Untersuchungen sind auch hinsichtlich der Kriterien zur Perspektivenwahl erforderlich. Zum einen können die aufgeführten Präsentationsziele verfeinert werden, etwa das Präsentationsziel 'Wirkung erzielen', zum andern lassen sich gewiß Objektklassen bestimmen, für die speziellere Kriterien zu besseren Ergebnissen führen. Wichtige Hinweise auf notwendige Modifikationen und Erweiterungen erwarten wir insbesondere auch durch die praktische Anwendung des Ansatzes.

Literatur

1. **G. Adorni, M. DiManzo** und **F. Giunchiglia**. Natural Language Driven Image Generation. In: COLING'84, S.495-500, 1984.
2. **J.N. Arnold**. Introductory Graphics. New York, Toronto, London: McGraw-Hill,1958.
3. **H. Espe**. Zum Einfluß von Kamerawinkeln auf die Beurteilung von Portraitfotografien - eine Erkundungsstudie. In: Semiotische Studien zur Kommunikation, Band 2, H. Espe (Hrsg.): Visuelle Kommunikation: Empirische Analysen. Hildesheim, Zürich, New York: Verlag G.Olms, S.166-182, 1986.
4. **S. Feiner**. APEX: An Experiment in the Automated Creation of Pictorial Explanations. IEEE Computer Graphics and Applications 5:11, S.129-140, 1985.
5. **M. Friedell**. Automatic Synthesis of Graphical Object Descriptions. In: Computer Graphics (ACM) 18(3), S.53-62, 1984.
6. **E.H. Gombrich**. Kunst und Illusion. Stuttgart, Zürich: Belser, 1978.
7. **E. Holder**. Design: Darstellungstechniken; ein Handbuch. Wiesbaden, Berlin: Bauerverlag, 1987.
8. **T. Meilhamer**. Grundkurs: Tiere zeichnen. Ravensburg: Otto Maier, 1987.
9. **G.A. Miller** und **P.N. Johnson-Laird**. Language and Perception. Cambridge: Cambridge University Press, 1976.
10. **D.D. Seligmann** und **S. Feiner**. Specifying Composite Illustrations with Communicative Goals. To appear in Proc. ACM UIST'89, Williamsburg, VA, 1989.
11. **P. van Sommers**. Drawing and Cognition. Cambridge, MA: Cambridge University Press, 1984.
12. **D. Wunderlich**. Raum, Zeit und das Lexikon. In: H. Schweizer (Hrsg.): Sprache und Raum: Psychologische und linguistische Aspekte der Aneignung und Verarbeitung von Räumlichkeit. Stuttgart: Metzler, S.66-89, 1985.



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 Research Reports

RR-92-43

Christoph Klauck, Jakob Mauss: A Heuristic driven Parser for Attributed Node Labeled Graph Grammars and its Application to Feature Recognition in CIM
17 pages

RR-92-44

Thomas Rist, Elisabeth André: Incorporating Graphics Design and Realization into the Multimodal Presentation System WIP
15 pages

RR-92-45

Elisabeth André, Thomas Rist: The Design of Illustrated Documents as a Planning Task
21 pages

RR-92-46

Elisabeth André, Wolfgang Finkler, Winfried Graf, Thomas Rist, Anne Schauder, Wolfgang Wahlster: WIP: The Automatic Synthesis of Multimodal Presentations
19 pages

RR-92-47

Frank Bomarius: A Multi-Agent Approach towards Modeling Urban Traffic Scenarios
24 pages

RR-92-48

Bernhard Nebel, Jana Koehler: Plan Modifications versus Plan Generation: A Complexity-Theoretic Perspective
15 pages

RR-92-49

Christoph Klauck, Ralf Legleitner, Ansgar Bernardi: Heuristic Classification for Automated CAPP
15 pages

RR-92-50

Stephan Busemann: Generierung natürlicher Sprache
61 Seiten

RR-92-51

Hans-Jürgen Bürckert, Werner Nutt: On Abduction and Answer Generation through Constrained Resolution
20 pages

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.

RR-92-52

Mathias Bauer, Susanne Biundo, Dietmar Dengler, Jana Koehler, Gabriele Paul: PHI - A Logic-Based Tool for Intelligent Help Systems
14 pages

RR-92-53

Werner Stephan, Susanne Biundo: A New Logical Framework for Deductive Planning
15 pages

RR-92-54

Harold Boley: A Direkt Semantic Characterization of RELFUN
30 pages

RR-92-55

John Nerbonne, Joachim Laubsch, Abdel Kader Diagne, Stephan Oepen: 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-93-02

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, Bernhard 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-93-33*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-93-45*Rainer Hoch:* On Virtual Partitioning of Large

Dictionaries for Contextual Post-Processing to Improve Character Recognition

21 pages

DFKI Technical Memos**TM-91-14***Rainer Bleisinger, Rainer Hoch, Andreas Dengel:*

ODA-based modeling for document analysis

14 pages

TM-91-15*Stefan Busemann:* Prototypical Concept Formation An

Alternative Approach to Knowledge Representation

28 pages

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

DFKI Documents**D-92-19**

Stefan Dittrich, Rainer Hoch: Automatische, Deskriptor-basierte Unterstützung der Dokument-analyse zur Fokussierung und Klassifizierung von Geschäftsbriefen
107 Seiten

D-92-21

Anne Schauder: Incremental Syntactic Generation of Natural Language with Tree Adjoining Grammars
57 pages

D-92-22

Werner Stein: Indexing Principles for Relational Languages Applied to PROLOG Code Generation
80 pages

D-92-23

Michael Herfert: Parsen und Generieren der Prolog-artigen Syntax von RELFUN
51 Seiten

D-92-24

Jürgen Müller, Donald Steiner (Hrsg.):
Kooperierende Agenten
78 Seiten

D-92-25

Martin Buchheit: Klassische Kommunikations- und Koordinationsmodelle
31 Seiten

D-92-26

Enno Tolzmann:
Realisierung eines Werkzeugauswahlmoduls mit Hilfe des Constraint-Systems CONTAX
28 Seiten

D-92-27

Martin Harm, Knut Hinkelmann, Thomas Labisch:
Integrating Top-down and Bottom-up Reasoning in COLAB
40 pages

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-\$).