

# Neue 'Neue Medien'?

## Von Multimedia zu Intellimedia-Präsentationssystemen<sup>1</sup>

Elisabeth André, Thomas Rist, Gerd Herzog<sup>2</sup>

### 1 Einleitung

Multimediasysteme, in denen unterschiedliche mediale Darbietungsformen 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 Multimediasysteme in ihre Angebotspalette aufnehmen und eingeführte Produkte um gewisse Multimediafunktionen erweitern. Solche Funktionen umfassen die Darstellung, Speicherung, Verbreitung bis hin zur Erzeugung und Verarbeitung multimedialer Dokumente. Für die Informationspräsentation bietet die Unterstützung mehrerer Darbietungsformen nicht nur Wahlmöglichkeiten, sondern vor allem auch Kombinierbarkeit. Rechnergestützte Präsentationen lassen sich darüber hinaus auch noch interaktiv als Hypermediapräsentation gestalten und eröffnen damit innovative Formen der Informationsdarbietung. So sind Multimedia-/Hypermediasysteme besonders dann attraktiv, wenn sich die gewünschte Information erst durch eine geschickt gewählte Kombination mehrerer Medien und Präsentationsstile effektiv darbieten läßt, wobei unter Effektivität zu verstehen ist, daß eine Präsentation - je nach ihrem Zweck - leicht verständlich, schnell wahrnehmbar, einprägsam und ansprechend ist. Im hier vorgelegten Beitrag soll versucht werden, in naher Zukunft zu erwartende Entwicklungen im Bereich der Nutzung Neuer Medien zu skizzieren und dadurch implizit auch den Begriff 'Neue Medien' sozusagen in die Zukunft zu vervollständigen.

---

<sup>1</sup> Die hier beschriebenen Forschungsarbeiten wurden teilweise von der EU im Rahmen des Programms "Telematikanwendungen", Projekt TE 2006: FLUIDS, und dem BMBF unter den Förderkennzeichen ITW 8901 8, ITW 94007 und 9701 0 gefördert.

<sup>2</sup> {andre,rist,flint}@dfki.de

Intelligente Präsentationssysteme als Komponenten von Nutzerschnittstellen zu Hilfesystemen, Expertensystemen oder intelligenten Betriebswarten der nächsten Generation zielen darauf ab, vorliegendes Wissen nutzerspezifisch und der jeweiligen Präsentationssituation angemessen auszuwählen und darzubieten. Dies setzt ein hohes Maß an Flexibilität bei der Informationspräsentation voraus; unterschiedliche Nutzer in unterschiedlichen Situationen erfordern auch immer unterschiedliche Präsentationen. Weil es im allgemeinen extrem aufwendig – wenn nicht gar unmöglich – ist, potentielle Nutzer 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 Präsentationswissen über grundlegende kommunikative Fähigkeiten verfügt. Ein solches Präsentationssystem wäre sowohl für Nutzer als auch für Ersteller einer Anwendung vorteilhaft. Während Nutzer auf sie individuell abgestimmte Präsentationen erwarten dürfen, werden Ersteller wesentlich entlastet, da sie sich auf die eigentlichen Probleme der Anwendung konzentrieren können.

Intellimedia-Präsentationssysteme (IMP-Systeme) stehen für den Versuch, Multimedia- und Hypermedia-Technologie für eine intelligente Informationspräsentation zu nutzen. IMP-Systeme gehen in ihrer Funktionalität weit über die heutigen Multimedia-/Hypermedia-Umgebungen hinaus, da nicht nur die Darstellung von Präsentationen geleistet wird, sondern auch deren Entwurf. Die Entwicklung ausgereifter IMP-Systeme ist ein interdisziplinäres und aufwendiges Unterfangen – interdisziplinär, weil es auf Ergebnisse unterschiedlicher Fachrichtungen angewiesen ist, insbesondere der Künstlichen Intelligenz, der Sprachphilosophie, der Textlinguistik, der Kommunikationswissenschaft und der Kognitionspsychologie, aufwendig, weil dieses Gebiet gerade erst als Forschungsgegenstand entdeckt wurde und es noch erheblicher Anstrengungen bedarf, bis diese ehrgeizige Vision umgesetzt ist.

Die Bedeutung von IMP-Systemen resultiert unmittelbar aus der Tatsache, daß solche Systeme immer dann von Nutzen sind, wenn Information nutzerspezifisch zu präsentieren ist. Im Grunde genommen ist dies bei nahezu jeder Anwendung wünschenswert. Im folgenden werden einige Anwendungen skizziert, die am Deutschen Forschungszentrum für Künstliche Intelligenz in Saarbrücken (DFKI) entwickelt wurden.

## **2 Aufgaben eines IMP Systems**

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

### **2.1 Festlegung und Strukturierung des Inhalts**

In einigen Fällen gibt das Anwendungssystem den Inhalt mehr oder weniger vor, und das Präsentationssystem hat lediglich irrelevante Details zu beseitigen und den Rest zu organisieren. Häufig ist es jedoch Aufgabe des Präsentationssystems, die mitzuteilende Information ausgehend von einem kommunikativen Ziel aus einer Wissensbasis auszuwählen. Gemäß den konversationellen Maximen von Grice (1975) sollte Nutzern einerseits notwendige Information vermittelt werden, andererseits ist Redundanz zu vermeiden. Inhaltsorganisation bezieht sich nicht nur auf die medienspezifische Anordnung und Gruppierung von Information. Das Präsentationssystem muß darüber hinaus sicherstellen, daß die Strukturen von Präsentationsteilen, die in unterschiedlichen medialen Darbietungsformen vorliegen, miteinander verträglich sind. Des weiteren sollte ein flexibles System in der Lage sein, seine Inhaltsselektions- und -organisationsprozesse an die jeweilige Situation anzupassen. Liegt beispielsweise die mitzuteilende Information zu Beginn der Präsentation vollständig vor, so bietet sich der Einsatz unterschiedlicher Sortierungstechniken (z.B. zeitbasiert vs. ortszentriert) an. Solche Techniken sind jedoch weniger geeignet, wenn die zu übermittelnden Daten erst zur Präsentationszeit geliefert werden. In einer solchen Situation steht die Aktualität der Beschreibung im Vordergrund; ob die Beschreibung zusammenhängend ist, spielt hingegen eine eher untergeordnete Rolle.

### **2.2 Verteilung der Information auf unterschiedliche Medien und Modi**

Die große Zahl schlecht konzipierter multimedialer Dokumente zeigt, daß die Kombination mehrerer Medien nicht automatisch zur besseren Präsentation führt. Beispiele hierfür sind konventionelle Papierdokumente, in denen Bilder überflüssig sind oder den Leser sogar verwirren (siehe hierzu auch die in Levie, 1987, bzw. Molitor, Ballstaedt & Mandl, 1989, aufgeführten empirischen Studien), ebenso wie die vielen aufwendig gestalteten Computerpräsentationen, die ihre Nutzer letztendlich einer nicht zu bewältigenden Reizüberflutung aussetzen. Um die Stärken unterschiedlicher Medien optimal nutzen zu können, müssen geeignete Kriterien gefunden werden, nach denen die mitzuteilende Information auf unterschiedliche Darbietungsformen, z.B. Text und Graphik, verteilt wird. Hierbei gilt es auch, das richtige Maß an Komplementarität und Redundanz zwischen den verwendeten Präsentationsformen 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 Nutzer sich nach kurzer Zeit auf ein Medium konzentrieren und möglicherweise Information übersehen. Steht das Medium fest, so ist zu entscheiden, durch welchen Modus die mitzuteilende Information zu realisieren ist. Dies hängt u.a. davon ab, welches Medium zur Übertragung der Information gewählt wurde. Beispielsweise ist in Verbindung mit Videosequenzen gesprochene Sprache gegenüber geschriebener Sprache zu bevorzugen, da nach Möglichkeit mehrere Perzeptionsorgane angesprochen werden sollten. Schließlich ist noch zu entscheiden, welches Hilfsmittel für die Ausgabe zu verwenden ist. So kann geschriebene Sprache sowohl auf Papier als auch auf einem Bildschirm ausgegeben werden.

### **2.3 Herstellung von Bezügen zwischen Präsentationsteilen in unterschiedlichen Medien**

Eine Multimediapräsentation erscheint nur dann als kohärent, wenn Nutzer erkennen, welche Präsentationsteile sich auf die gleichen Objekte beziehen. In einem multimedialen Kontext müssen dabei nicht nur Verbindungen zwischen Präsentationsteilen innerhalb eines Mediums, sondern auch Verbindungen zwischen Präsentationsteilen in unterschiedlichen Medien hergestellt werden. Eine Möglichkeit besteht darin, explizit auf Präsentationsteile in anderen Medien zu verweisen, beispielsweise durch natürlichsprachliche Äußerungen der Form: "der Schalter oben links im Bild". Darüber hinaus sollten in allen Medien nach

Möglichkeit die gleichen Begriffe gewählt werden. Wird beispielsweise ein Objekt im Bild annotiert, so sollte dieser Begriff auch im Text auftauchen. Wesentlich schwieriger ist die Herstellung solcher Verbindungen, wenn dynamische Medien verwendet werden. Wird Information beispielsweise durch animierte Darstellungen und gesprochene Sprache übertragen, so sollten Objekte möglichst dann sichtbar sein, wenn auf sie verwiesen wird.

## **2.4 Medienspezifische Inhaltsrealisierung**

In vollautomatisierten IMP-Systemen wird die Realisierung der mitzuteilenden Information von medienspezifischen Generierungskomponenten, z.B. einem Text- und einem Graphikgenerator, übernommen. Diese verfügen über spezielles Designwissen und stellen Mechanismen zum automatischen Aufbau von audio-visuellen Elementen bereit. Allerdings können solche Komponenten nicht als unidirektionale Ausgabekomponenten betrieben werden. Da die Generierungsergebnisse so aufeinander abzustimmen sind, daß sich die Medien in ihrer Wirkung gegenseitig verstärken, muß jeder Generator einerseits auf die Ergebnisse anderer Generatoren zurückgreifen können, andererseits eine explizite Repräsentation seiner Enkodierungen bereitstellen. Dies ist u.a. dann erforderlich, wenn auf die Ergebnisse anderer Generatoren Bezug genommen wird. Selbst wenn auf vorgefertigte Präsentationselemente zurückgegriffen wird, kann auf automatisierte Verfahren zur Inhaltsrealisierung meist nicht verzichtet werden. So kann es beispielsweise notwendig sein, kohärenzstiftende Elemente einzufügen.

## **2.5 Integration der Generierungsergebnisse in einer Ausgabe**

Liegen die Ergebnisse unterschiedlicher Generatoren vor, müssen diese unter Berücksichtigung räumlicher und zeitlicher *Constraints* in einer multimedialen Ausgabe integriert und Nutzern präsentiert werden. 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 Doku-

mentseiten 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. Wird Nutzern Information mit Hilfe dynamischer Medien vermittelt (z.B. gesprochene Sprache, Zeigegesten und animierte Darstellungen), so ist die Ausgabe der Generatorergebnisse nicht nur räumlich, sondern auch zeitlich zu koordinieren. Unter anderem ist festzulegen, zu welchem Zeitpunkt eine bestimmte Zeigegeste beginnen soll und wieviel Zeit sie in Anspruch nehmen darf.

### **3 WIP: Ein System zur automatischen Dokumentation technischer Geräte**

Ein geradezu ideales Einsatzfeld für ein IMP-System ist die Generierung von Anleitungen für die Handhabung, Wartung und Bedienung von technischen Geräten. Für Nutzer solcher Anleitungen ist es zweifellos vorteilhaft, wenn diese nicht pauschalisiert, sondern auf individuelle Bedürfnisse abgestimmt sind. Die Möglichkeit, Dokumentvarianten vor Ort nach Bedarf generieren zu können, umgeht das Problem, Anforderungen und Bedürfnisse aller potentiellen Nutzer in einer unendlichen Menge von Präsentationssituationen vorab antizipieren zu müssen. Aber auch für Produzenten technischer Dokumentation ergeben sich handfeste Vorteile. Während heute bereits eine kleine Produktänderung oft erheblichen Aufwand für die manuell durchzuführende Aktualisierung zur Folge hat, wäre bei Einsatz eines IMP-Systems nur die Gerätespezifikation auf den neuesten Stand zu bringen.

Ein erster Schritt in diese Richtung wurde mit dem an der DFKI entwickelten WIP-System (Wissensbasierte Informationspräsentation) geleistet, das automatisch Bedienungsanleitungen für technische Geräte generiert (André et al., 1993). Als Grundlage zur automatischen Generierung multimedialer Präsentationen entwickelten wir zunächst Beschreibungsprinzipien zur Erfassung von deren Struktur (André 1995). Im wesentlichen beruhen diese Prinzipien auf (1) *einer Verallgemeinerung der sogenannten Sprechakttheorie* (Searle, 1980). Idee dabei ist, daß sich auch die Durchführung nicht-verbaler Präsentationshandlungen klassifizieren läßt, stets intentionsgetrieben ist und daß sich komplexere Präsentationshandlungen in elementarere zerlegen lassen. So ergibt sich beispielsweise auch eine hierarchische Strukturbeschreibung von Bildern.

Zum anderen stellen sie (2) *eine Erweiterung der sogenannten Rhetorical Structure Theory* (RST, Mann & Thompson, 1987) dar. Von der Feststellung ausgehend, daß sich der strukturelle Aufbau von Multimediapräsentationen ähnlich beschreiben läßt wie reine Texte, lag eine Orientierung an bekannten Methoden zur Planung des Inhalts und der Struktur von Texten nahe.

Zur Operationalisierung des Generierungsprozesses wurde daher ein planbasiertes Verfahren entwickelt. Als Eingabe erhält diese Planungskomponente ein auf hohem Abstraktionsniveau formuliertes Präsentationsziel. Dieses Ziel wird dann in einer Reihe von Dekompositionsschritten zerlegt und schrittweise in medienspezifische Präsentationshandlungen überführt. Dieser 'Top-Down'-Planungsansatz kann dabei durch die Vorgabe von Generierungsparametern weiter gesteuert werden. Solche Parameter sind beispielsweise von Nutzern geäußerte Medien- oder Formatpräferenzen, aber auch Betriebsmittelbeschränkungen wie beispielsweise die vorhandene Größe des Bildschirms. Als Planungsoperatoren dienen dabei sogenannte Präsentationsstrategien. Diese enkodieren Wissen darüber, wie komplexe Präsentationsaufgaben in elementarere zu zerlegen sind, welche Medien bzw. Medienkombinationen zur Bewältigung der ausgearbeiteten Teilaufgaben in Frage kommen sowie medienspezifisches Designwissen, etwa wie ein Bedienungshandlung graphisch illustriert werden kann.

Abbildung 1 zeigt zwei Präsentationsvarianten für Nutzer mit unterschiedlichen Medienpräferenzen. In beiden Fällen bestand das Präsentationsziel darin, Nutzern zu erläutern, wie man ein Modem auf Empfang von Daten schaltet. Während im ersten Fall die gesamte Information bildlich wiedergegeben wird, wurde im zweiten Fall ausschließlich auf Text zurückgegriffen.

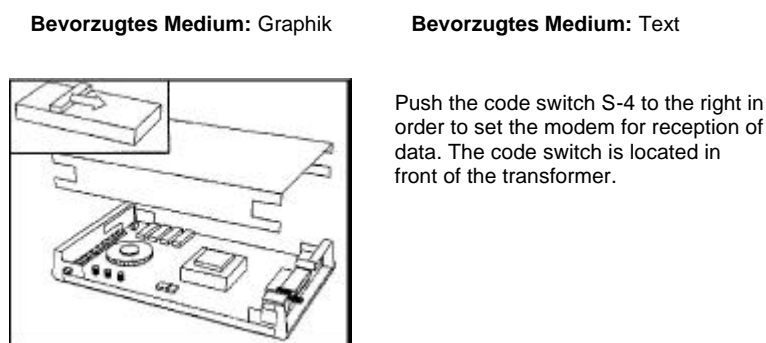


Abbildung 1: Einfluß von Medienpräferenzen in WIP

#### **4 FLUIDS: Multimediale Präsentation sensorischer Daten**

Da die Durchführung von Beobachtungs- und Überwachungsaufgaben oft ermüdend und zeitraubend ist, bietet der Einsatz eines IMP-Systems auch im Bereich der Präsentation sensorischer Daten erhebliche Vorteile. Durch die Kopplung eines IMP-Systems mit einem Leitstand können beispielsweise Entscheidungsträger über alle relevanten Ereignisse eines Beobachtungsraumes informiert werden, ohne daß diese selbst die Beobachtungen durchführen müssen. Bei dieser Anwendung kommt auch die Fähigkeit eines IMP-Systems zum Tragen, dem Nutzer visuelle Information in einem auf ihn zugeschnittenen Format zu präsentieren. Das Spektrum der Möglichkeiten zur Wiedergabe wahrgenommener Ereignisse reicht hier von rein sprachlichen Beschreibungen über illustrierte Berichte bis hin zu kommentiertem Videomaterial. Der Aufwand, der mit der manuellen Erzeugung solcher Präsentationen verbunden ist, steht jedoch dem Wunsch nach individuell auf die jeweiligen Nutzer zugeschnittenen Präsentationen entgegen. Ein IMP-System stellt hingegen eine kostengünstige Möglichkeit dar, dieselbe Information in Abhängigkeit von Generierungsparametern (wie z.B. simultane versus retrospektive Beschreibung) zielgruppenspezifisch zu präsentieren.

In dem von der Europäischen Union geförderten Projekt FLUIDS (Future Lines of User Interface Decision Support) wurde ein Rahmen geschaffen, der es erlaubt, bestehende Echtzeitmanagementanwendungen mit einer wissensbasierten Komponenten zur Entscheidungsunterstützung sowie einer leistungsfähigen multimedialen Nutzerschnittstelle aufzurüsten (Herzog, André, Baldes, & Rist, 1997). Der FLUIDS-Ansatz wurde anhand konkreter Verkehrsüberwachungssysteme, die sich in den Städten Turin und Madrid im Einsatz befinden, ausgearbeitet und getestet. Die Aufgabe der Informationspräsentation in FLUIDS läßt sich beschreiben als Abbildung von anwendungsspezifischen Informations- und Datenelementen auf Präsentationen, die aus räumlich und zeitlich angeordneten Medienobjekten aufgebaut sind. Die dargebotenen Schaubilder sind Karten mit unterschiedlichen Abstraktionsgraden sowie spezielle Diagrammtypen, etwa zur Darstellung einzelner Buslinien (siehe Abbildung 2). Jedes Schaubild besteht primär aus einem Hintergrund (z.B. einer Karte) und einer Menge von graphischen Objekten, die auf diesem Hintergrund plaziert werden. Zu nennen sind hier Ikonen für die bildliche Referenz auf Domänenobjekte (z.B. Ikonen für Fahrzeuge), Ikonen, die bestimmte Ereignisse oder Handlungen symbolisieren, sowie einige metagraphische Symbole zur Fokussierung (z.B. Pfeile und einrahmende Kreise). Im Gegensatz zu statischen Schaubildern ändern bei Anima-

tionen die erscheinenden Ikonen ihre Lokationen in einem Zeitintervall. Typisches Beispiel ist hier die Visualisierung einer Fahrzeugtrajektorie, bei der das Fahrzeugikon auf der Kartendarstellung seinen Weg auf einer bestimmten Route, also seine Trajektorie abfährt. In der vorliegenden Anwendungsdomäne kommt jedoch einer anderen Form der Animation eine weitaus größere Bedeutung zu, und zwar der teilanimierten Graphik. Hierunter verstehen wir solche graphischen Darstellungen, bei denen ein komplexes Schaubild vor den Augen des Betrachters sukzessive komplettiert wird. Obwohl es sich auch hier um eine Folge von Bildern handelt, besteht ein wesentlicher Unterschied zu dem zuvor genannten Animationstyp darin, daß das letzte Einzelbild als Enkodierung der gesamten Information zu verstehen ist. Die sukzessive Vervollständigung dient vornehmlich dazu, den Betrachter durch das Schaubild zu führen. Insbesondere in Kombination mit gesprochener Sprache lassen sich so komplexe Inhalte übersichtlich und einprägsam vermitteln.

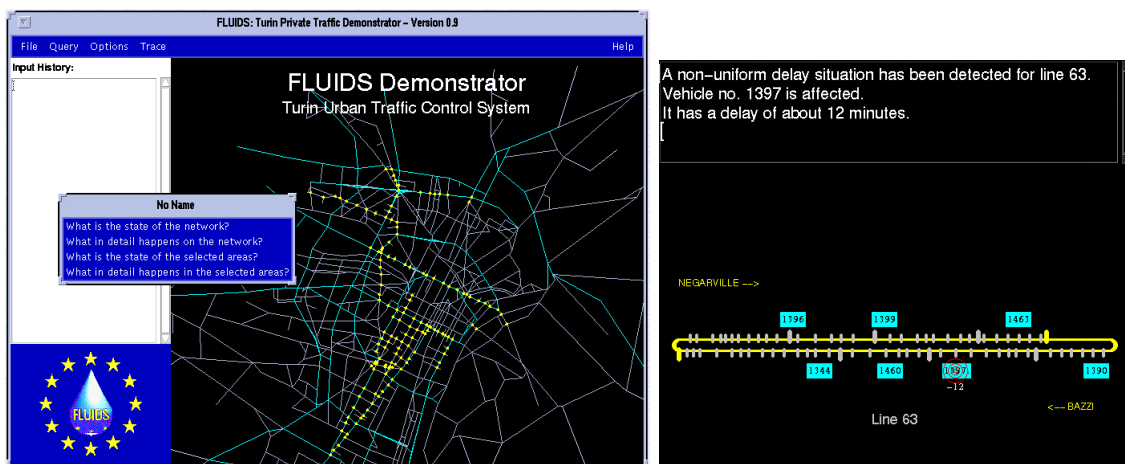


Abbildung 2: In FLUIDS verwendete Graphiktypen: Straßennetzkarte und Busliniendiagramm

In der Operationalisierung wird die Zuordnung von Informationseinheiten zu Medienobjekten wie bereits in WIP durch ein planbasiertes Generierungsverfahren bewältigt. Im Gegensatz zu WIP kommen in FLUIDS jedoch auch dynamische Medien zum Einsatz, deren Darbietung zeitlich zu koordinieren ist. Um dies zu gewährleisten, wurde der Generierungsansatz um zwei Aspekte erweitert: (1) *Beschreibungskonstrukte zur Notation qualitativer und quantitativer zeitlicher Constraints in den Präsentationsstrategien* wurden eingeführt. In Anlehnung an Allens Zeitlogik (Allen, 1987) können 13 verschiedene qualita-

tive Beziehungen zwischen zwei Zeitintervallen ausgedrückt werden. Hingegen werden quantitative zeitliche Beziehungen als metrische Gleichheit bzw. Ungleichheit ausgedrückt. Weiterhin wurde (2) *ein Mechanismus zur Berechnung möglicher zeitlicher Präsentationsabläufe gemäß der vorgegebenen zeitlichen Constraints* implementiert.

Um dies zu ermöglichen, wurde die Planungskomponente mit einer Komponente für zeitliches Schließen kombiniert, die auf dem von Kautz und Ladkin (1991) entwickelten Metric-Allen-Time-System (MATS) basiert (vgl. André & Rist, 1996).

## **5 PPP und AiA: Informationspräsentation durch animierte Agenten**

Animierte Agenten sei es in Form von Cartoonfiguren oder synthetisierter 3D-Humanoiden, können erheblich dazu beitragen, Information in einer für Nutzer leicht verständlichen und ansprechenden Form darzubieten. Inspiriert von der Beobachtung, daß ein gut gehaltener Vortrag oftmals ansprechender ist als das eigene Durcharbeiten von statischem Lern- und Dokumentationsmaterial, wurde am DFKI in den Projekten PPP (Personalized Plan-Based Presenter) und AiA (Adaptive Communication Assistant for Effective Infobahn Access) mit der Entwicklung eines animierten Präsentationsagenten begonnen. Aufgabe dieses Agenten ist es, Präsentationsmaterial, wie Bilder, textuelle Erläuterungen, Video- und Animationsclips, ansprechend darzubieten. Konkret lassen sich durch die Verwendung von animierten Präsentationsagenten folgende Vorteile erzielen:

- *Erweitertes Repertoire an Präsentationstechniken*

Ein Agent kann als zusätzliche Präsentationsmodalität betrachtet werden, durch die die Möglichkeiten eines Generierungssystems zur Informationsdarbietung erweitert werden. Ziel ist es dann, diese Modalität dort einzusetzen, wo sie gegenüber anderen Vorteile aufweist. So ermöglichen die mit Armen ausgestatteten Agenten effektivere Referenzhandlungen, da es in der Regel wesentlich anschaulicher ist, einen Querverweis durch eine Zeigegeste des Agenten zu realisieren als etwa durch einen sprachlich formulierten Verweis.

- *Navigationshilfe beim Erfassen der Information*

Agenten erlauben es, den zu Instruierenden so durch die Präsentation zu führen, daß die Inhalte in der optimalen zeitlichen Reihenfolge assimiliert werden. Dies stellt sich insbesondere bei multimedialen Präsentationen als Vorteil heraus, da dort oftmals Inhalte parallel über verschiedene Fenster ausgegeben werden, und der Präsentationskonsument eine zeitliche Einordnung der Teile selbst rekonstruieren muß.

- *Instruieren durch Vormachen*

Insbesondere wenn eine virtuelle 3D-Umgebung als Präsentationsumfeld gewählt wird, können Agenten die Durchführung von Handlungsanweisungen verständlich durch Vormachen vermitteln (Rickel & Johnson, 1999). Text-Bild-Kombinationen oder gar rein textuell verfaßte Anleitungen erfordern hier wesentlich höheren kognitiven Aufwand, da der räumliche Kontext sowie der zeitliche Ablauf vom Lernenden mental rekonstruiert werden muß.

Zu erwähnen ist schließlich noch, daß die Interaktion mit einem animierten Agenten oftmals als unterhaltsam empfunden wird und sich so die Lernmotivation der Nutzer steigern läßt (Lester et al., 1997). Voraussetzung ist jedoch, daß sie ein für Nutzer glaubhaftes Verhalten aufweisen und diese bei der Lösung ihrer Aufgabe tatsächlich unterstützen. Übertriebene Lebhaftigkeit sowie stumpfsinnig repetierendes Verhalten führen andererseits schnell dazu, daß Nutzer lieber auf einen anderen Präsentationsmodus umschalten. Unsere eigenen Untersuchungen haben ergeben, daß Nutzer Präsentationen mit animierten Agenten als weniger schwierig ansehen (vgl. van Mulken, André & Müller, 1998).

Die Aufgabe eines agentenbasierten Präsentationssystems ist vergleichbar mit der eines Sprechers, der einen audio-visuellen Vortrag zunächst vorbereitet, um ihn dann vor einem Publikum zu halten. Neben der Auswahl und Beschaffung bzw. Neuerstellung des zu präsentierenden Materials sind auch Präsentationshandlungen zu planen und zeitlich zu koordinieren.

Unsere Systemkonzeption folgt dem Bild eines mehr oder weniger vielseitig einsetzbaren Präsentationsagenten (André, Rist & Müller, 1999). Um anwendungsspezifische Präsentationen durchzuführen, muß dem Agenten ein Präsentationsskript übermittelt werden, das ähnlich einem Drehbuch die einzelnen Schritte benennt. Zum Aufbau solcher Skripte wird das bereits in FLUIDS verwendete Planungsverfahren eingesetzt.

Das Verhalten des Agenten ist durch dieses Skript weitgehend – jedoch nicht vollständig - festgelegt. Die Definition des Gesamtverhaltens folgt vielmehr der Gleichung:

*Gesamtverhalten = Präsentationsskript + Eigenverhalten*

Den Agenten mit einem spezifischen Eigenverhalten auszustatten, ist konzeptionell wünschenswert, da dies eine Entkoppelung von Generierungsphase (what-to-say) und Präsentationsphase (how-to-say) darstellt. Daß es sich hier um eine sinnvolle Trennung handelt, wird klar, wenn man bedenkt, daß es bei einer Vortragsvorbereitung im allgemeinen wenig Sinn macht, genau zu planen, ob man sich nun 40 oder 50 cm vom Folienprojektor aufstellt, nach 3 Minuten und 25 Sekunden den linken Fuß einen Schritt nach vorne stellt usw. Damit der synthetisierte Agent einigermaßen belebt und glaubhaft wirkt, sind solche nicht vorab geplanten Handlungen jedoch ebenfalls in die Präsentation einzustreuen. Im einzelnen unterscheiden wir zwischen:

- *Navigationshandlungen*

Diese dienen dazu, den Agenten in eine zur Durchführung gewisser Präsentationshandlungen (insbesondere Zeigehandlungen) günstige Ausgangslage zu bringen. Zu beachten ist, daß die Realisierung von Navigationshandlungen auf die gewählte Agentenmetapher abzustimmen ist. Wird wie in unserem Fall eine menschenähnliche Figur eingesetzt, so sind gehen, laufen und springen plausible Realisierungen, während Microsofts animierter Papagei Peedy (Ball et al., 1997) von einer Bildschirmposition zur nächsten fliegt.

- *Reflexhandlungen*

Die Glaubhaftigkeit eines Agenten läßt sich erhöhen, wenn dieser auf gewisse externe Ereignisse wie beispielsweise Nutzerinteraktionen unmittelbar reagiert. Dies ist wiederum vergleichbar mit einer natürlichen Vortragssituation, in der der oder die Vortragende etwa durch eine Zwischenfrage unterbrochen wird.

- *Leerlaufhandlungen*

Ein Präsentationsagent, der zwischen einzelnen Präsentationshandlungen stets in eine Art Winterstarre verfällt, wirkt im allgemeinen wenig ansprechend. Kleinere Pausen sind jedoch häufig einzustreuen, etwa um Präsentationskonsumenten hinreichend Zeit einzuräumen, einzelne Präsentationselemente zu erfassen, wie beispielsweise Textzitate oder Videoclips. Um den Agenten „am Leben zu erhalten“ bietet es sich an, ihn gewisse Leerlaufhandlungen durchführen zu lassen, die zwar von Nutzern bemerkt werden, sie jedoch nicht von der Inspektion anderer Präsentationsteile ablenken.

Das Eigenverhalten des Agenten wird in einer deklarativen Spezifikations-sprache definiert. Aus Verhaltensspezifikationen wird dann automatisch ein Kontrollmodul für die Animationskomponente des Agenten generiert.

In Abbildung 3 werden zwei Anwendungsbeispiele für animierte Präsentationsagenten vorgestellt. Links in Abbildung 3 erläutert eine animierte Persona den Nutzern die Funktionsweise eines Modems. Während das frühere WIP System auf Beschriftungen zurückgreift, um eindeutig auf Objekte zu verweisen, ermöglicht die Persona natürliche Zeigegesten, wie sie auch in der zwischenmenschlichen Kommunikation vorkommen. In der Abbildung zeigt sie beispielsweise auf den Transformator und teilt Nutzern - unter Verwendung eines Sprachsynthesizers - mit: "Das ist der Transformator". Das Beispiel zeigt auch, wie Gesichtsausdrücke und Kopfbewegungen eingesetzt werden, um den visuellen Fokus zu begrenzen. Dadurch daß die Persona in die Richtung des Zielobjekts schaut, wird der Blick von Nutzern automatisch dorthin gelenkt. Rechts in Abbildung 3 ist eine Präsentation mit drei animierten Agenten zu sehen, die die Nutzer durch einen elektronischen Produktkatalog führen, Empfehlungen aussprechen und auf Produktmerkmale hinweisen, die für sie von Interesse sind.

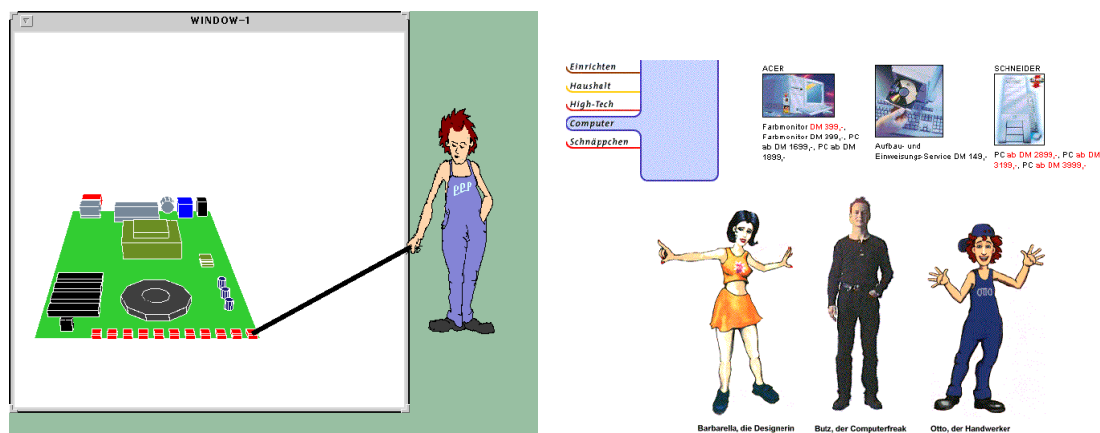


Abbildung 3: Anwendungsbeispiele für animierte Agenten

## 6. Schlußbemerkung

In der hier vorgelegten Arbeit wurde ein planbasierter Ansatz als Paradigma für die multimediale Informationspräsentation vorgestellt und skizziert, wie sich ein solcher Ansatz an unterschiedliche Anwendungsklassen, wie zum Beispiel technische Dokumentation und entscheidungsunterstützende Systeme, anpassen läßt.

Ein besonderer Vorteil des Ansatzes besteht darin, daß er die Möglichkeit bietet, Information in Abhängigkeit von unterschiedlichen Generierungsparametern zu vermitteln. Im Gegensatz zu WIP wird in FLUIDS, PPP und AiA nicht nur das darzubietende Multimediamaterial, sondern auch dessen Präsentation auf die jeweiligen Nutzer und die aktuelle Situation zugeschnitten, wobei in den letzten beiden Systemen den Nutzern das Präsentationssystem - das Intellimediasystem, wie wir es nennen - in Form animierter Agenten entgegentritt. Intellimediasysteme sind so nicht mehr einfache Medienkoppelungssysteme, sondern können durch ihre dynamische Eigentätigkeit als Meta-Medien und in diesem Sinne neue 'Neue Medien' gesehen werden.

## 7. Literatur

Allen, J.F. (1983). Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11), 832-843.

André, E. (1995). Ein planbasierter Ansatz zur Generierung multimedialer Präsentationen, DISKI 108, infix.

André, E. & Rist, T. (1996). Coping with temporal constraints in multimedia presentation planning. *Proceedings of AAAI-96*, 1, 142-147.

André, E., Rist, T. & Müller, J. (1999). Employing AI methods to control the behavior of animated interface agents. *Applied Artificial Intelligence* 13, 415-448.

Ball, G., Ling, D., Kurlander, D., Miller, J., Pugh, D., Skelly, T., Stankosky, A., Thiel, D., van Dantzich, M. & Wax, T. (1997). Lifelike computer characters: The persona project at Microsoft. In J.M. Bradshaw (Ed.), *Software agents* (pp. 191-222). Menlo Park, CA: AAAI/The MIT Press.

Grice, H.P. (1975). Logic and conversation. In P. Cole & J.L. Morgan (Eds.), *Syntax and semantics: Speech acts*, Volume 3 (pp. 41-58). New York: Academic Press.

Herzog, G., André, E., Baldes, S. & Rist, T. (1998). Zur multimedialen Präsentation von Handlungsalternativen in Verkehrsmanagement- und Verkehrssimulationssystemen. In P. Lorenz & B. Preim (Eds.), *Simulation und Visualisierung '98* (pp. 235-248). Delft:

- Kautz, H.A. & Ladkin, P.B. (1991). Integrating metric and qualitative temporal reasoning, *AAAI 91*, Anaheim, CA, pp. 241-246.
- Lester, J.C., Converse, S.A., Kahler, S.E., Barlow, S.T., Stone, B.A. & Bhogal, R.S. (1997). The persona effect: Affective impact of animated pedagogical agents. In S. Pemberton (Ed.), *Human factors in computing systems, CHI'97 conference proceedings* (pp. 359–366). New York: ACM Press.
- Levie, W.H. (1987). Research on pictures: A guide to the literature. In D.M. Willows & H.A. Houghton (Eds.), *The Psychology of Illustration, Basic Research*, VOLUME 1 (pp. 1-50). New York: Springer.
- Mann, W.C. & Thompson, S.A. (1987). Rhetorical Structure Theory: Description and construction of text structures. In G. Kempen (Ed.), *Natural language generation: New results in artificial intelligence, psychology, and linguistics* (pp. 85-95). Dordrecht: Nijhoff.
- Molitor, S., Ballstaedt, S.-P. & Mandl, H. (1989). Problems in knowledge acquisition from text and pictures. In H. Mandl & J.R. Levin (Eds.), *Knowledge acquisition from text and pictures* (pp. 3-35). Amsterdam: North Holland.
- Rickel, J. & Johnson, L. (1999). Animated agents for procedural training in virtual reality: Perception, cognition, and motor control employing AI methods to control the behavior of animated interface agents. *Applied Artificial Intelligence*, 13, 443–483.
- Searle, J.R. (1969). *Speech Acts: An essay in the philosophy of language*. Cambridge, MA: Cambridge University Press.
- van Mulken, S., André, E. & Müller, J. (1999). The persona effect: How substantial is it? In H. Johnson, L. Nigay & C. Roast (Eds.), *People and computers XIII* (Proceedings of HCI-98, pp. 53-66. Berlin: Springer.