

SmartKom

Norbert Reithinger und Anselm Blocher

SmartKom - Multimodale Mensch-Technik-Interaktion

SmartKom – Multimodal Human Computer Interaction

Kurzzusammenfassung:

Eine der größten Herausforderungen für die Wissensgesellschaft ist derzeit die Schaffung intelligenter Benutzerschnittstellen, die den natürlichen Kommunikationsstil von Computerlaien akzeptieren und so unterstützen, dass eine für den Menschen intuitive und benutzergerechte Mensch-Technik-Interaktion entsteht. Im Leitprojekt SmartKom werden Konzepte für die Entwicklung völlig neuartiger Formen der Mensch-Technik-Interaktion erprobt. Das Ziel von SmartKom ist die Erforschung und Entwicklung einer selbsterklärenden, benutzeradaptiven Schnittstelle für die Interaktion von Mensch und Technik im Dialog. Das Vorhaben wird von der Vorstellung geleitet, die Vorteile sprachlich dialogischer Kommunikation zu verschmelzen mit den Vorteilen graphischer Bedienoberflächen und gestisch Ausdrucks. Das Ergebnis ist eine höherwertige Benutzerschnittstelle, die die natürlichen menschlichen Sinne in größerem Umfang als bisher berücksichtigt. SmartKom geht damit einen wichtigen Schritt auf dem Weg zur Entwicklung einer auf den Menschen zugeschnittenen Informationstechnologie.

Abstract:

One of the most important challenges facing the information society today is the creation of intelligent user interfaces which accept and support the natural communication style of computer lay-people so that a form of interaction between human and computer comes about which is natural and user friendly for the individual. In the project SmartKom, designs are being tested for the development of totally new forms of human-technology interaction. SmartKom's goals are the research and the development of a self-explanatory, user adaptive interface for the dialogic interaction of humans and technology. This project's intent is to merge the advantages of dialogic communication with the advantages of a mixture of graphic control surfaces and gestures. The result is a high value user interface, which accommodates the natural human senses to a greater degree than ever before. SmartKom thus takes an important step towards the development of information technology designed to meet human needs and abilities.

1. Die Leitvorstellung von SmartKom

Innovative Konzepte zur multimodalen Interaktion zwischen Mensch und Technik erforscht das Leitprojekt SmartKom (Wahlster, Reithinger und Blocher 2001a, SmartKom 2003) im Rahmen und mit Förderung des BMBF-Schwerpunktprogrammes Mensch-Technik Interaktion (MTI 2003).

In einem neuartigen Ansatz (Wahlster, Reithinger und Blocher 2001b) sollen in SmartKom die bestehenden Hemmschwellen von Computerlaien bei der Nutzung der Informationstechnologie abgebaut und damit ein Beitrag zur Nutzerfreundlichkeit und Nutzerzentrierung der Technik in der Wissensgesellschaft geliefert werden. Aufbauend auf den Erfahrungen mit dem BMBF-Leitvorhaben Verbmobil (1993-2000) zur multilingualen Verarbeitung von Spontansprache, hat SmartKom das Ziel der Erforschung und der Entwicklung einer selbsterklärenden, benutzeradaptiven Schnittstelle für die Interaktion von Mensch und Technik in einem multimodalen Dialog (Maybury und Wahlster 1998).

SmartKom verschmilzt die Vorteile sprachlich dialogischer Kommunikation einerseits mit den Vorteilen graphischer Bedienoberflächen andererseits sowie mit der Integration zusätzlicher Modalitäten wie Gestik und Mimik. Aus der Realisierung dieser Leitvorstellung geht eine höherwertige Benutzerschnittstelle hervor, die die natürlichen menschlichen Sinne in größerem Umfang als bisher berücksichtigt. Die Ergebnisse von SmartKom werden damit einen wichtigen Schritt zur Entwicklung einer auf den Menschen zugeschnittenen Informationstechnologie darstellen.

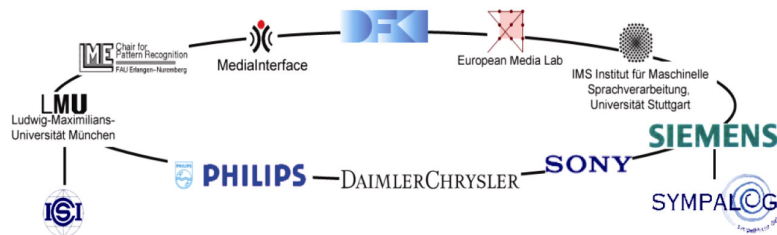


Bild 1: Die Konsortialpartner von SmartKom

Das SmartKom-Konsortium unter der Gesamtprojektleitung von Prof. Wolfgang Wahlster (DFKI) umfasst das Deutsche Forschungszentrum für Künstliche Intelligenz GmbH (DFKI) als Konsortialführer mit dem Forschungsbereich Intelligente Benutzerschnittstellen sowie die Partner DaimlerChrysler AG, European Media Laboratory, International Computer Science Institute, MediaInterface Dresden GmbH, PHILIPS GmbH, Siemens AG, Sony International (Europe) GmbH, Sympalog Speech Technologies AG, Universität München, Universität Erlangen-Nürnberg und Universität Stuttgart (siehe Bild 1). Die Laufzeit von SmartKom ist auf vier Jahre angelegt und endet im September 2003.

2. Eine Interaktionsmetapher - drei Szenarien

SmartKom ermöglicht komplexe dialogische Interaktionen, wobei sowohl der Benutzer als auch das System Interaktionen initiieren, Rück- und Klärungsfragen stellen, Verstehensprobleme signalisieren oder den Dialogpartner unterbrechen können. Dabei bricht SmartKom mit der traditionellen Desktop-Metapher: Anstelle des WIMP-Ansatzes (WIMP = Windows, Icon, Menus, Pointer) steht das situierte, delegationszentrierte Dialogparadigma im Zentrum. In dieser Interaktionsmetapher löst der Benutzer zusammen mit dem System, das durch einen animierten Agenten verkörpert wird, kollaborativ Aufgaben (siehe Bild 2). Dabei muss der Benutzer nicht alle Lösungsschritte vorgeben, sondern kann das Problem an das System weitergeben. Dieses bestimmt die Lösungsschritte, fragt unter Umständen zurück, und präsentiert die Lösung angepasst an die Aufgabenstellung. Die Objekte der Welt, wie z.B. ein Telefon, werden dabei in der virtuellen Welt nachgebildet und können vom Benutzer mit

natürlichen Gesten oder im Sprachdialog referenziert und in die Interaktion einbezogen werden.

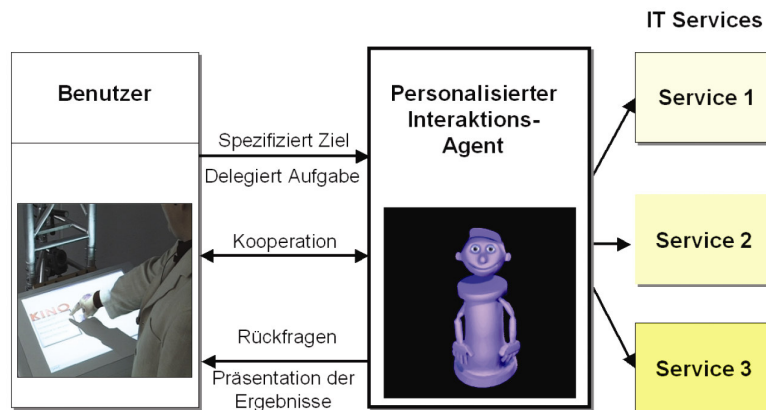


Bild 2: Das situierte, delegationszentrierte Dialogparadigma von SmartKom

SmartKom bietet anhand von drei ausgewählten Anwendungsszenarien - einem öffentlichen Informationskiosk (Public), einer Interaktionszentrale für das Wohnzimmer (Home), und einem mobilen Kommunikationsassistenten zur Wegsuche (Mobile) - die Grundfunktionalitäten für die nächste Generation intuitiver Benutzerschnittstellen, die bezüglich der Integration und Orchestrierung immer weiterer Interaktionsmodalitäten skalierbar sind.

Begleitend zu Entwurf und Entwicklung des technischen Systems wurden umfangreiche Datensammlungen (Wizard-of-Oz Experimente) durchgeführt. Nach Fertigstellung des ersten SmartKom-Prototypen wurde mit der Evaluation des Systems durch naive Benutzer begonnen. Zur Auswertung der Experimente wurde ein neues Standardverfahren für das Labeling multimodaler Daten entwickelt (Steininger, Schiel und Glesner 2002). Als Ergebnis steht die derzeit größte Sammlung multimodaler Dialoge auf über 100 DVDs zur Verfügung (BAS 2003). Die daraus gewonnenen Erkenntnisse fließen umgehend in die Entwicklungsarbeiten ein.

Das SmartKom-System setzt sich aus einem multimodalen Dialog-Kernsystem zusammen, das je nach Anwendungsszenario verschiedenste Funktionalitäten und Interaktionsmedien bedient. In Bild 3 sind die drei Hardwareausprägungen der Szenarien abgebildet. Im Informationskiosk-Szenario von SmartKom Public wird der Interaktionsinhalt auf eine Oberfläche projiziert. Der Benutzer kann mit natürlichen Gesten diese Inhalte manipulieren. Die Gestenerkennung verwendet hochentwickelte Bildverarbeitungsalgorithmen, so dass die Interaktion keine Instrumentierung der Hand erfordert. Als zusätzliche und wissenschaftlich sehr anspruchsvolle Modalität wird im Szenario Public Mimikerkennung zur Erkennung des Benutzerzustandes erforscht. Derzeit werden zwei Hauptzustände unterschieden, nämlich neutral und ärgerlich. Erkennt das System, dass der Benutzer verärgert auf eine Interaktion reagiert, kann mittels der Hilfefunktionalität versucht werden, dem Benutzer bessere und auf die persönlichen Vorlieben zugeschnittene Information anzubieten. Im Home-Szenario wird ein Touchpad verwendet, das Gesten mit dem Finger oder dem Stift aufnimmt und weiterverarbeitet. Eine Besonderheit dieses Szenarios ist es, dass die beim Benutzer verfügbaren Modalitäten bei der Interaktion berücksichtigt werden. Sieht er beispielsweise fern, kann keine graphische Ausgabe verwendet werden (siehe unten). Die Interaktion im Mobile-Szenario erfolgt über einen PDA mit Stifteingabe. Die unterschiedlich großen Ausgabegeräte

erfordern dabei eine szenarioangepasste Präsentation. Zum Beispiel wird in der Medienallokation berücksichtigt, welche Modalitäten und Medien gerade verfügbar sind, und auf welche der Benutzer derzeit achtet.

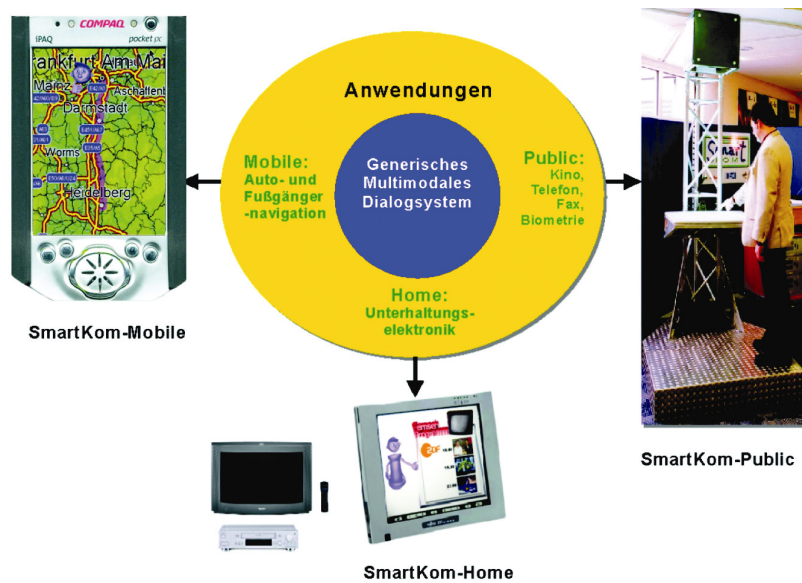


Bild 3: Ein multimodales Dialog-Kernsystem für alle Szenarien

Das multimodale Dialog-Kernsystem und die Anwendungsszenarien sowie besonders ihr Zusammenspiel bei unterschiedlichen Anforderungsprofilen – die sich beispielsweise aus einem variierenden Vorhandensein von Ein-/Ausgabe-Modalitäten ergeben - werden anhand einer breiten Palette von Funktionalitäten erprobt (siehe Tabelle 1).

Tabelle 1: Funktionalitäten in den drei Anwendungsszenarien

Szenario Home			
EPG (Electronic Programming Guide)	<ul style="list-style-type: none"> • Programm • Kanalauswahl • Senderanzeige • Genregesteuerte Auswahl 	<ul style="list-style-type: none"> • Info zu Sendung • Zeitangaben • Hilfe zu Genres 	7
TV Steuerung	<ul style="list-style-type: none"> • TV an/aus 	<ul style="list-style-type: none"> • Senderauswahl 	2
VCR Steuerung	<ul style="list-style-type: none"> • VCR an/aus • Aufnehmen • Abspielen • Pause 	<ul style="list-style-type: none"> • Vor- und Rücklauf • Programmierung in Kombination mit EPG und Kalender 	6
Lean-Forward/ Lean Backward	<ul style="list-style-type: none"> • Einschalten • Ausschalten 	<ul style="list-style-type: none"> • Modalitätsabhängige Präsentation 	3
Summe Home			18
Szenario Public			
Telefonieren	<ul style="list-style-type: none"> • Manipulative Tastenfunktionalität • Telefonie 	<ul style="list-style-type: none"> • Audiohandling • Adressbuch 	4
Handkontur- biometrie	<ul style="list-style-type: none"> • Anfrage nach Biometriotyp • Hand-Biometrie 	<ul style="list-style-type: none"> • Präsentation und Kamerasteuerung • Adressbuch (s.o.) 	3
Stimmbiometrie	<ul style="list-style-type: none"> • Präsentation und Audiosteuerung • Sprach-Biometrie 	<ul style="list-style-type: none"> • Adressbuch(s.o.) • Anfrage nach Biometriotyp (s.o) 	2
Fax-Versand	<ul style="list-style-type: none"> • Bereitstellung der Fax-Präsentation • Fax-Handling 	<ul style="list-style-type: none"> • Adressbuchsuche (s.o.) • Dokumentenkamera-steuerung (s.o) 	3
E-Mail-Versand	<ul style="list-style-type: none"> • Bereitstellung der Mail-Präsentation • E-Mail Handling 	<ul style="list-style-type: none"> • Adressbuchsuche (s.o.) • Dokumentenkamera-steuerung (s.o) 	3
Summe Public			15
Szenario Mobile			
Fahrzeug- navigation	<ul style="list-style-type: none"> • Auswahl Start und Zielstadt • Auswahl Routentyp 	<ul style="list-style-type: none"> • Auswahl der Parkhäuser • Wegberechnung 	4
Fussgänger- navigation	<ul style="list-style-type: none"> • Auswahl der Karte • Auswahl Start und Zielort • Wegberechnung 	<ul style="list-style-type: none"> • Auswahl der Points-of-Interest • Information zu Points-of-Interest 	5
Karten- manipulation	<ul style="list-style-type: none"> • Größe ändern 	<ul style="list-style-type: none"> • Ausschnitt ändern 	2
Summe Mobile			11
Gesamtsystem			44

Ein Vergleich mit derzeit im Einsatz befindlichen Sprachdialogsystemen zeigt, dass SmartKom in eine neue Dimension der dialogischen Interaktion vorstößt. Ein Zugauskunftssystem kann beispielsweise nach Start- und Zielort, Reisedatum und Uhrzeit, Gleis, und Preis befragt werden. Die Anzahl der Parameter in solchen Systemen ist dabei auf etwa 5 – 20 beschränkt. SmartKom dagegen interagiert in einem kohärenten Ansatz mit dem Benutzer über 44 Funktionalitäten aus verschiedensten Themenbereichen, die mit derzeit 126 Parameter kontrollierbar sind. Diese Funktionsmächtigkeit ist nur durch einen integrierten Verarbeitungsansatz möglich, der applikationsübergreifend die Dialogfähigkeiten modelliert.

3. Integrierte Modalitätenverarbeitung

Die Leitvorstellung von SmartKom ist, einen einheitlichen Verarbeitungsansatz für alle Anwendungen und alle Modalitäten zu verwenden. Der Ansatz basiert auf den folgenden Grundsätzen:

1. Die Verarbeitung basiert auf übergreifenden, wissensbasiert parametrisierten Ansätzen: Es gibt keine Speziallösungen für Anwendungen.
2. „No processing and presentation without representation“: Für alle Ein- und Ausgaben wird ein einheitlicher, allgemein im System verwendeter Repräsentationsformalismus verwendet. Nur so sind generische Interaktionsmodelle einsetzbar.
3. Das System basiert auf einer Multi-Blackboard-Systemarchitektur, die in einer weiterentwickelten Version der im Projekt Verbmobil verwendeten Integrationsmiddleware realisiert ist (Wahlster 2000). Bild 4 zeigt den Kontrollbildschirm des Testbeds mit allen wichtigen Modulen des Systems. Das Testbed dient zur Kontrolle und Visualisierung aller Module im System.

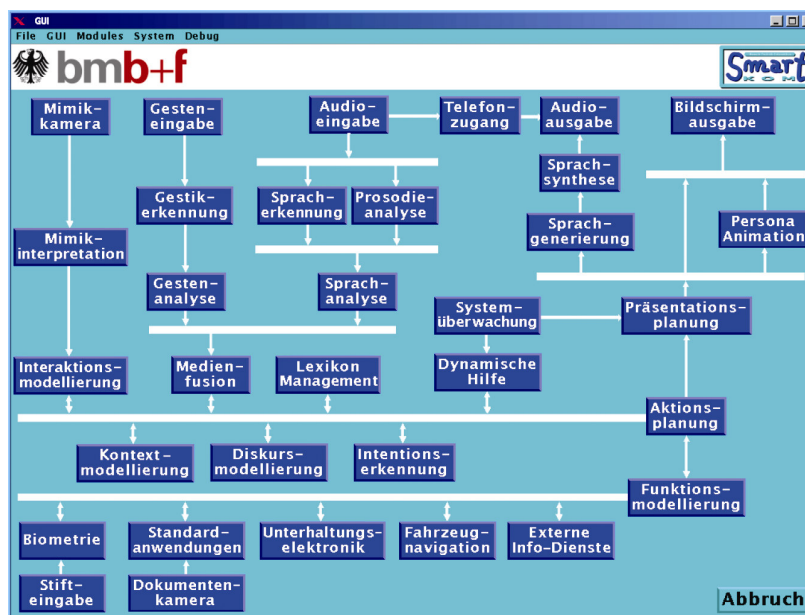


Bild 4: Der Kontrollbildschirm des Testbeds

Die Schnittstellen und die übergreifende Domänenrepräsentation sind in M3L (Multimodal Markup Language) definiert. M3L ist ein XML-Dialekt und definiert mittels XML-Schemata die Syntax, in der die Daten zum Austausch zwischen dem Modulen von SmartKom notiert

werden. Integriert in diese Sprache ist die einheitliche, hierarchische Domänenmodellierung, die alle Anwendungen umfasst. Die unterliegende Wissensquelle ist eine Ontologie, die in OIL (OIL 2003) definiert wird und in ein XML-Schema umgesetzt wird. In SmartKom werden alle Verarbeitungsschritte und Zwischenergebnisse mittels M3L repräsentiert. Das inkludierte ontologische Wissen wird in den verschiedenen Modulen zusätzlich für Inferenzprozesse verwendet.

Insbesondere können somit multimodale Interaktionsbeiträge von Benutzer und System einheitlich dargestellt werden. Sprache und Gesten auf der Eingabeseite, sowie Sprache und visuelle Darstellungen auf der Ausgabeseite, werden intern einheitlich auf einer semantisch-pragmatischen Ebene dargestellt.

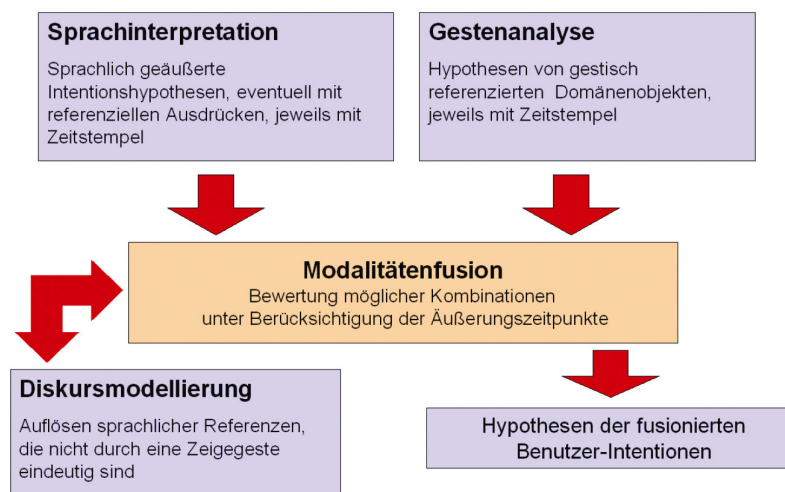


Bild 5: Modalitätenfusion im Kontext von SmartKom

In SmartKom kann der Benutzer mittels Sprache oder Gesten mit dem Benutzer interagieren (Wahlster 2002). Bild 5 zeigt den schematischen Ablauf der Modalitätenfusion. Sprachliche und gestische Eingaben werden in M3L-Intentionshypothesen bzw. Objektbeschreibungen abgebildet. Die robuste Sprachinterpretation verwendet dazu ein templatebasiertes Verfahren, das einen Chunk-Parser mit einschließt (Engel 2002). Die Gestenanalyse liefert Hypothesen von Domänenobjekten: Da zu jeder Visualisierung eine Beschreibung des Inhalts vorhanden sein muss, wird aus der Geometrie der Geste und deren Beziehung zu den dargestellten Objekten eine Liste von möglicherweise referenzierten Domänenobjekten geliefert. Die Modalitätenfusion kombiniert beide Informationen mit Hilfe des Dialoggedächtnisses und der zeitlichen Information und erzeugt vervollständigte Intentionshypothesen.

Das Dialoggedächtnis ist der zentrale Speicherort für dialogspezifische Kontextinformation. Im Rahmen von SmartKom wurden bekannte Ansätze aus der Computerlinguistik und Sprachverarbeitung erweitert zu einem multimodalen Dialoggedächtnis (Löckelt, Becker, Pflieger und Alexandersson 2002). Bild 6 zeigt einen schematischen Ausschnitt des Dialoggedächtnisses. Es baut auf einem dreistufigen Modell auf. Zu jedem Objekt in einer Ein- oder Ausgabemodalität wird ein modalitätsabhängiger Platzhalter eingefügt, der auf ein Diskursobjekt verweist. Dabei können mehrere Modalitätsobjekte zu einem Diskursobjekt gehören. Diskursobjekte wiederum sind mit der internen Beschreibung des Anwendungsobjekts verlinkt, speichern aber auch die Eigenschaften von sprachlichen und visuell dargestellten Objekten. Im Beispiel wurde visuell eine Liste von Titeln dem Benutzer

präsentiert. Der Benutzer kann jetzt sprachlich auf ein Element der Liste referieren, ohne eine Geste benutzen zu können. Über die Struktur des Diskursobjekts ist eine einfache Dereferenzierung möglich.

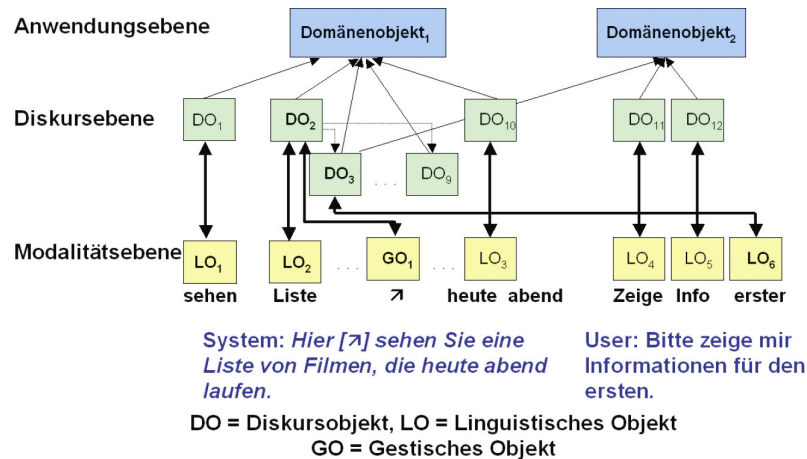


Bild 6: Ein Beispiel für das multimodale Dialoggedächtnis

Die Aktionsplanung ist der „Geist“ der Maschine. In Planoperatoren verfügt sie über das Wissen, wie auf Benutzerintentionen reagiert werden muss, welche Anfragen an welche Anwendungsfunktionalität gestellt werden können und welche Inhalte an den Benutzer übermittelt werden sollen. Die Anfragen an die Funktionalitäten sind dabei abstrahiert und werden von einem Modul zur Funktionsmodellierung in die realen Anfragen umgesetzt. Damit sind funktionspezifische Anfragedetails weitgehend vom Dialogsystem abgekoppelt.

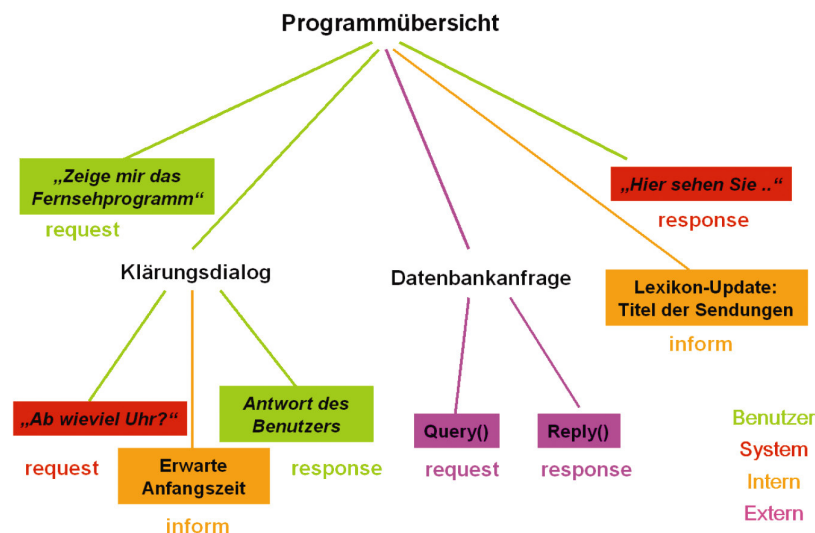


Bild 7: Communicative Games

Der theoretische Ansatz der Aktionsplanung sind „Communicative Games“ zur Modellierung von Dialogen. Hierbei wird eine einheitliche Behandlung aller Kommunikationskanäle, also der Benutzerinteraktionen, der internen und externen Funktionalität angestrebt. Die Planung und Durchführung einer Reaktion des Systems basiert auf einer Abfolge von paarweisen Kommunikationsakten z.B. *request*, *response*, *inform*, oder *greeting*. Die Aktionspaare

können ineinander verschachtelt sein, wobei eine synchrone oder asynchrone Abarbeitung möglich ist: der Benutzer kann also vorziehen, auf eine Anfrage nicht zu antworten und das Thema zu wechseln, ohne dass das System darauf nicht mehr reagieren kann. Die Aktionsplanung erstellt somit eine Abfolge von Games und führt diese aus, um die Ziele des Benutzers zu erreichen.



Bild 8: Eine kombiniert graphisch-sprachliche Präsentation

Ein Beispiel für einen Aktionsplan ist in Bild 7 dargestellt. Um auf eine Anfrage des Benutzers nach dem Fernsehprogramm angepasst reagieren zu können, wird zuerst mit einer Klärungsfrage der Zeitraum eingeschränkt. Danach ist eine Anfrage an die Applikation möglich. Damit der Benutzer die Titel und weitere Informationen zum aktuellen Programm in einer weiteren Anfrage sprachlich verwenden kann, muss als nächstes das Lexikon des Spracherkenners um die bisher nicht in der Wortliste enthaltenen Wörter erweitert werden. Erst danach wird die Präsentation des Ergebnisses initiiert.

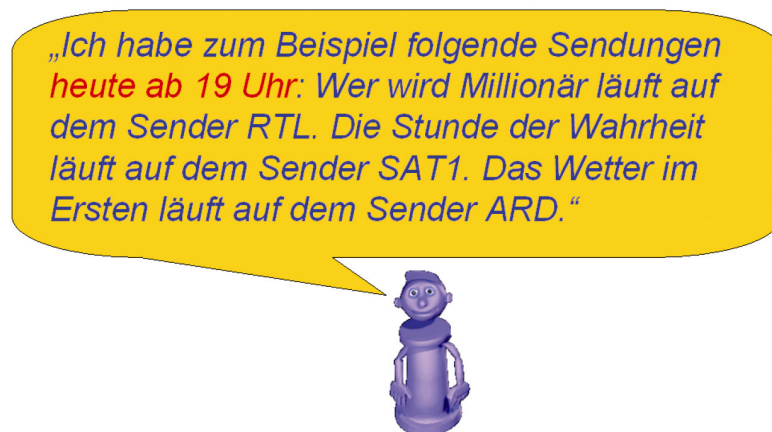


Bild 9: Eine rein sprachliche Präsentation

Analog zur Modalitätenfusion wird das Ergebnis der internen Verarbeitung in einer Modalitätenfission oder Präsentationsplanungskomponente benutzer- und situationsangepasst visualisiert (Poller und Müller 2002). Ein wichtiges Kriterium ist dabei, welche Modalität derzeit verfügbar ist, bzw. welche Modalität der Benutzer präferiert. Für das Szenario Home wurden zwei unterschiedliche Modi definiert:

- Lean-Forward: Der Benutzer sieht auf das zur Steuerung verwendete Touchpad, dessen Inhalte er wahrnimmt. In diesem Modus kann SmartKom alle Modalitäten verwenden. Bild 8 zeigt das Ergebnis einer Anfrage an den Fernsehprogramm-Service. Der Benutzer beachtet den visuellen Kontext und die Präsentation enthält Detail-Information in der Grafik. Smartakus abstrahiert vom Inhalt und äußert einen allgemeinen Hinweis.
- Lean-Backward: Der Benutzer wendet seine visuelle Aufmerksamkeit von der Benutzeroberfläche ab und sieht beispielsweise fern. Die visuelle Modalität ist also nicht frei verwendbar. Das System stellt seine Präsentationsstrategie um und verbalisiert nur die ersten Ergebnisse, wobei Information sinnvoll zusammengefasst wird (siehe Bild 9).

4. Zusammenfassung

SmartKom realisiert einen wichtigen Schritt auf dem Weg zur selbsterklärenden, multimodalen und benutzeradaptiven Schnittstellen für die Interaktion von Mensch und Technik. Es ist gelungen, unter einem einheitlichen Interaktionsparadigma verschiedenste Funktionalitäten und Szenarioausprägungen zusammenzuführen. Bis zum Abschluss des Projektes im dritten Quartal 2003 werden noch weitere Applikationsfunktionalitäten wie Kinoauskunft und inkrementelle Zielführung hinzugefügt werden, mit denen zusätzliche wissenschaftliche und technische Fragesellungen untersucht werden. Das implementierte und erfolgreich demonstrierte System zeigt die Stärke des gewählten übergreifenden, wissenschaftlich parametrisierten Ansatzes. Zudem hat sich die Stärke und Allgemeingültigkeit der Formel „No processing and presentation without representation“ beim Einsatz generischer Interaktionsmodelle gezeigt. Dieser Ansatz war auch die Grundlage der Erstellung einer englischsprachigen Version für Teilapplikationen des Szenarios Mobile. Da der Systemkern sprachunabhängig ist, mussten nur der Spracherkenner und die Sprachsynthese durch englische Module ersetzt werden. Einige Wissensquellen für Sprachanalyse und -generierung sowie die Präsentation für das Englische wurden an die geänderte Dialogsprache adaptiert.

Die Ergebnisse des BMBF-Leitprojekts Smartkom haben bereits zur Laufzeit Eingang in die Entwicklungen der Projektpartner gefunden und zu verbesserten Interaktionsmöglichkeiten im Mensch-Technik Bereich geführt. Sie beeinflussen bereits jetzt sichtbar zukünftige Forschungsthemen (Roadmaps). Die Zukunft multimodaler Interaktionssystemen wird in der weiteren Konvergenz weiterer Forschungsrichtungen liegen. Einen Schwerpunkt wird dabei die Verbindung zwischen kommunikativer Kompetenz allgegenwärtiger Hintergrundsysteme (ubiquitous computing) mit realitätsnahen, hochqualitativen virtuellen Interaktionsräumen in Echtzeit bilden. Erste Ansätze hierzu erforscht das BMBF-Verbundprojekt VirtualHuman (2002-2006).

5. Literatur

- BAS 2003. <http://www.phonetik.uni-muenchen.de/Bas/BasHomedeu.html> (Letzter Zugriff: 31.01.2003).
- DFKI 2003. <http://www.dfki.de> (Letzter Zugriff: 31.01.2003).
- Engel, R.: SPIN: Language Understanding for Spoken Dialogue Systems Using a Production System Approach", In: *Proceedings of 7th International Conference on Spoken Language Processing (ICSLP-2002)*. Denver, U.S.A.: 2002.
- Löckelt, M.; Becker, T.; Pflieger, N.; Alexandersson, J.: Making Sense of Partial. In: *Proceedings of the Sixth Workshop on the Semantics and Pragmatics of Dialogue (EDILOG 2002)*. (Hrsg. Bos, J.; Matheson, C.). Edinburgh, UK: 2002.

- Maybury, M.; Wahlster, W. (Hrsg.): *Readings in Intelligent User Interfaces*. San Francisco: Morgan Kaufmann, 1998
- MTI 2003. <http://www.dlr.de/IT/IV/MTI> (Letzter Zugriff: 31.01.2003).
- OIL 2003. Ontology Inference Layer. <http://www.ontoknowledge.org/index.shtml> (Letzter Zugriff: 31.01.2003).
- Poller, P.; Müller, J.: Distributed Audio-Visual Speech Synchronization. In: *Proceedings of the 7th International Conference on Spoken Language Processing, (ICSLP-2002)*. Denver, USA: 2002.
- SmartKom 2003. <http://www.smartkom.org> (Letzter Zugriff: 31.01.2003).
- Steininger, S.; Schiel, F.; Glesner, A.: Labeling Procedures for the Multi-modal Data Collection of SmartKom. *Proceedings of the 3rd Language Resources & Evaluation Conference (LREC)*. Las Palmas, Spain: 2002.
- VirtualHuman 2003. <http://www.virtual-human.org> (Letzter Zugriff: 31.01.2003).
- Wahlster, W. (Hrsg.): *Verbmobil: Foundations of Speech-to-Speech Translation*. Berlin: Springer, 2000.
- Wahlster, W., Reithinger, N.; Blocher, A.: SmartKom: Towards Multimodal Dialogues with Anthropomorphic Interface Agents. In: *Proceedings of International Status Conference "Human-Computer Interaction"*. (Hrsg. Wolf, G.; Klein, G.). Berlin: DLR, 2001a.
- Wahlster, W.; Reithinger N.; Blocher, A.: SmartKom: Multimodal Communication with a Life-Like Character. In: *Proceedings of Eurospeech 2001, 7th European Conference on Speech Communication and Technology*. Aalborg, Dänemark: 2001b.
- Wahlster, W.: SmartKom: Fusion and Fission of Speech, Gestures, and Facial Expressions In: *Proceedings of the 1st International Workshop on Man-Machine Symbiotic Systems*. Kyoto, Japan: 2002.

6. Kurzbiographien

Dr. Norbert Reithinger, geboren am 15.11.1959 in Altdorf bei Nürnberg



Nach dem Studium der Informatik an der Universität Erlangen – Nürnberg (1978-1984) und kurzer Industrietätigkeit von 1985-1993 wissenschaftlicher Mitarbeiter an der Universität des Saarlandes (Lehrstuhl für künstliche Intelligenz und deduktive Datenbanken, Leitung Prof. Dr. Dr. h.c. mult. Wolfgang Wahlster). Dort Promotion im Bereich der praktischen Informatik zum Thema Generierung multimodaler Dialogbeiträge. Seit 1993 als Projektleiter für den Themenbereich Dialogverarbeitung am Deutschen Forschungszentrum für Künstliche Intelligenz GmbH (DFKI) verantwortlich. Im Jahr 2000 Ernennung zum Research Fellow des DFKI.

E-Mail: reithinger@dfki.de

Dr. Anselm Blocher, geboren am 01.08.1962 in Karlsruhe



Nach dem Studium der Informatik an der Universität des Saarlandes (Abschluss 1994) bis 1999 wissenschaftlicher Mitarbeiter an der Universität des Saarlandes (Lehrstuhl für künstliche Intelligenz und deduktive Datenbanken, Leitung Prof. Dr. Dr. h.c. mult. Wolfgang Wahlster). Dort Promotion im Rahmen des Sonderforschungsbereichs 378 „Ressourcenadaptive kognitive Prozesse“. Seit 1999 als Projektmanager für das BMBF-Leitprojekt SmartKom am Deutschen Forschungszentrum für Künstliche Intelligenz GmbH (DFKI) verantwortlich.
E-Mail: blocher@dfki.de

7. Das Deutsche Forschungszentrum für Künstliche Intelligenz

Das Deutsche Forschungszentrum für Künstliche Intelligenz (DFKI GmbH) mit Sitz in Kaiserslautern und Saarbrücken ist auf dem Gebiet innovativer Softwaretechnologien die führende wirtschaftsnahe Forschungseinrichtung in Deutschland (DFKI 2003). In der internationalen Wissenschaftswelt zählt das DFKI zu den weltweit wichtigsten "Centers of Excellence", dem es gelungen ist, Spitzenforschung rasch in praxisrelevante Anwendungslösungen umzusetzen.

1988 von namhaften deutschen Unternehmen der Informationstechnik und zwei Großforschungseinrichtungen als gemeinnützige GmbH gegründet, hat sich die DFKI GmbH inzwischen durch ihre proaktive und bedarfsorientierte Projektarbeit national und international den Ruf eines kompetenten und zuverlässigen Partners für Innovationen in der Wirtschaft erworben.

Da durch zunehmend kürzere Innovationszyklen in der Informationstechnik Vorlaufforschung, anwendungsnahe Entwicklung und die Umsetzung in Produkte immer enger zusammenwachsen, wird in DFKI-Projekten das gesamte Spektrum von der anwendungsorientierten Grundlagenforschung bis zur markt- und kundenorientierten Entwicklung von Produktfunktionen abgedeckt.

Die Geschäftsführung der DFKI GmbH bilden Prof. Dr. Wolfgang Wahlster (Vorsitz, Technisch-wissenschaftlicher Geschäftsführer) und Dr. Walter G. Olthoff (Kaufmännischer Geschäftsführer).

Die Projekte am DFKI gliedern sich in fünf Forschungsbereiche:

- Wissensmanagement (Leitung: Prof. Dr. Andreas Dengel)
- Intelligente Visualisierungs- und Simulationssysteme (Leitung: Prof. Dr. Hans Hagen)
- Deduktions- und Multiagentensysteme (Leitung: Prof. Dr. Jörg Siekmann)
- Sprachtechnologie (Leitung: Prof. Dr. Hans Uszkoreit)
- Intelligente Benutzerschnittstellen (Leitung: Prof. Dr. Wolfgang Wahlster)

Seit Anfang 2002 ist außerdem das Institut für Wirtschaftsinformatik (IWi) (Leitung: Prof. August-Wilhelm Scheer) im DFKI eingebunden.

Der Auftrag der drei Transferzentren ist es, die Forschungsergebnisse des DFKI in kommerzielle Anwendungen zu transferieren:

- SISO - "Der Weg zur sicheren Software"
- TransLect - "Sprachtechnologie im Einsatz"
- AICommerce - "Intelligenz im eBusiness"

Das Ziel der drei im Augenblick existierenden Kompetenzzentren, in denen technologisches und fachliches Können des DFKIs konzentriert wurde, ist die Beantwortung wichtige Fragestellungen aus den Bereichen:

- e-Learning
- Sprachtechnologie
- Semantisches Web

Zur Zeit beschäftigt die DFKI GmbH mehr als 180 Angestellte, davon 145 wissenschaftliche Mitarbeiter. Hinzu kommen noch rund 140 studentische Hilfskräfte als Teilzeitkräfte. Im Jahr 2001 wurde ein Volumen von rund 15,4 Mio. € im Rahmen praxisorientierter Auftragsforschung erzielt.

Alle Arbeiten sind in der Form zeitlich befristeter und klar fokussierter Projekte organisiert, die u.a. zu patentierten Lösungen, Prototypen oder Produktfunktionen führen. Es werden derzeit über 59 Projekte bearbeitet. Der Projektfortschritt wird einmal im Jahr durch eine unabhängige Gutachtergruppe namhafter internationaler Experten überprüft.