

Mensch-Maschine-Interaktion

Das Interesse im Bereich Mensch-Maschine-Interaktion (MMI) richtet sich auf die Entwicklung von Methoden und konkreten Technologien, die dem Menschen die Nutzung technischer Systeme und die Steuerung komplexer Anwendungen erleichtern. Zentrale Fragestellungen gelten der Benutzungsfreundlichkeit von Software und technischen Systemen ebenso wie dem verlässlichen Zusammenwirken von Menschen und Maschinen. Techniken und Ansätze zur Mensch-Maschine-Interaktion betreffen einen weiten Bereich von Anwendungen, die von der Benutzung computergestützter Werkzeuge bis zu Formen der ‚sozialen‘ Interaktion mit technischen Artefakten reichen.

Als engeres Teilgebiet der Mensch-Maschine-Interaktion befasst sich die Mensch-Computer-Interaktion (*human-computer-interaction*) mit der benutzergerechten Gestaltung von interaktiven Systemen und Mensch-Maschine-Schnittstellen sowie mit der Gebrauchstauglichkeit der eingesetzten Soft- und Hardware. Der vornehmlich auf Computersoftware und interaktive Medien wie z.B. das Internet bezogene Begriff der Benutzerschnittstelle (*user interface*) umfasst nach ISO 9241 „alle Bestandteile eines interaktiven Systems, die Informationen und Steuerelemente zur Verfügung stellen, die für den Benutzer notwendig sind, um eine bestimmte Arbeitsaufgabe mit dem interaktiven System zu erledigen“ (*Handbuch Usability*). Die Gebrauchstauglichkeit von Benutzerschnittstellen wird danach bewertet, wie genau und vollständig Benutzer ein bestimmtes Ziel erreichen können (Effektivität), welchen Aufwand Benutzer im Verhältnis zur Genauigkeit und Vollständigkeit des erzielten Effekts erbringen müssen (Effizienz) und wie positiv ihre Einstellung gegenüber der Nutzung und wie frei diese von Beeinträchtigungen ist (Zufriedenheit).

Eine weitergehende Frage der Mensch-Maschine-Interaktion zielt auf die Übertragbarkeit von Aspekten der natürlichen Interaktion zwischen Menschen auf Maschinen ab (Mensch-Maschine-Kommunikation). Bezugspunkte sind hier kognitive Leistungen wie das Verstehen und Produzieren von Sprache, Gestik oder Mimik in dialogischer Interaktion bis hin zur Mensch-Technik-Kooperation und ihrer Realisierung durch künstliche Gesprächspartner sowie ‚soziale‘ Roboter. Die Mensch-Maschine-Interaktion ist somit ein interdisziplinäres Gebiet, für das neben der Informatik und der Künstliche-Intelligenz-Forschung auch Erkenntnisse aus der Medienpsychologie und dem Design, aus der Arbeitswissenschaft, der

Softwareergonomie, der Kognitionswissenschaft und der Techniksoziologie relevant sind.

Direkte Manipulation versus indirektes Management

Die Entwicklung von Benutzerschnittstellen für interaktive Systeme begann in der Ära der Mehrbenutzersysteme in den 1960er und 1970er Jahren, in der die Interaktion mittels Kommandozeileneingaben und alphanumerischer Bildschirme erfolgte, und führte in den 1980er Jahren zunächst zum Menü-orientierten Arbeitsplatzrechner. Das Vorhaben, leistungsfähigere und zugleich einfachere Schnittstellen zu entwickeln, resultierte wenig später in der bis heute dominanten grafischen Benutzerschnittstelle (*graphical user interface*). Basierend auf Fenstern, Ikonen, Menüs und Maus-unterstützter Interaktion ist der Leitgedanke dabei die Handhabung von sichtbaren Objekten auf einer Bildschirm-Schreibtischoberfläche. Diese sog. WIMP-Schnittstellen (*Windows Icons Menus Pointing Device*) werden heute – eingeschlossen neuere Entwicklungen wie *multi-touch* – unter der von Shneidermann (2002) geprägten Interaktionsmetapher der *direkten Manipulation* geführt.

Die Interaktionsform der direkten Manipulation ermöglicht zwar eine vordergründig komfortable Bedienung, verlagert jedoch sämtliche Aktionen wie Initialisierung, Durchführung und Beobachtung von Aufgaben auf die Seite des Benutzers. Eine weitere Beschränkung ist das unveränderte Verhalten gegenüber wiederkehrenden Aufgaben. Je mehr Leistungsmerkmale eine Anwendung bietet, desto komplizierter gestalten sich klassische WIMP-Schnittstellen. Als Herausforderung sah man deshalb die Schaffung intuitiv nutzbarer Schnittstellen, die Menschen von technischen Details entlasten. Dies führte zu der Entwicklung von natürlichsprachlichen oder auch gestischen bis hin zu multimodalen Schnittstellen (Hedicke 2002), bei denen etwa nebenläufige Sprach- und Gesteneingaben auf getrennten Kanälen registriert und für die Steuerung von Anwendungen integriert und interpretiert werden.

Parallel dazu entstand seit Mitte der 1980er Jahre – überschrieben mit dem Schlagwort ‚*virtual reality*‘ – eine neue Form interaktiver Mensch-Maschine-Schnittstellen. *Virtual reality* zielt auf eine möglichst unmittelbare Verbindung der sensorischen und aktorischen Fähigkeiten des Menschen mit der synthetischen Welt des Computers – generiert mit Hilfe der modernen Grafiktechnik als dreidimensionale virtuelle Umgebung (das sog. interaktive *walkthrough*) – oder vermittelt Informationen

aus entfernten Räumen mit Hilfe von Stereo-Videokameras bzw. -Sensoren (die sog. Telepräsenz). Im Vordergrund steht dabei die Echtzeitinteraktion mit rechnergenerierten Modellen oder mit Simulationsergebnissen, aber auch mit Ereignissen und Personen an anderen Orten und zu anderen Zeiten, zuweilen überlagert mit realen Darstellungen (*augmented* oder *mixed reality*).

Einer anderen Strömung gehören Unterstützungssysteme an, die mit Hilfe intelligenter Techniken Mittler- und Assistenzaufgaben in der Mensch-Maschine-Interaktion übernehmen (Sullivan/Tyler 1991). Der Grundgedanke besteht darin, die Systemschnittstelle selbst intelligenter zu machen, d.h. sie in die Lage zu versetzen, zwischen Benutzer und technischem System zu vermitteln, indem sie Anteile der Lösungsfindung eigenständig erbringt. Für diese Interaktionsform wurde von Kay (1990) die Metapher des *indirekten Management* geprägt. Hierbei handelt es sich um einen kooperativen, zum Menschen komplementär ausgerichteten Ansatz (*human complementary approach*, Terveen 1995), der sowohl das künstliche System als auch den Menschen als einzigartiges System mit besonderen Fähigkeiten und Stärken ansieht. Dies führt zu einer Asymmetrie, bei welcher dem künstlichen System und dem Menschen unterschiedliche Rollen übertragen werden, in denen die Verantwortlichkeiten derart aufgeteilt werden, dass ihre jeweiligen Stärken möglichst gut ausgeschöpft werden.

Interface-Agenten

Eine besondere Ausprägung des indirekten Management sind ‚Interface-Agenten‘ (Laurel 1990; Shneiderman/Maes 1997), die auf der Technik der ‚Software-Agenten‘ basieren, d.h. Computerprogrammen, die als autonome Einheiten agieren und Aktivitäten in ihrer Umgebung beobachten und initiieren können. Schnittstellen werden dabei als digitale Assistenten realisiert, denen Aufträge so übertragen werden wie menschlichen Bearbeitern. Die möglichen Anwendungen reichen vom automatischen Bildschirm-Layout über intelligente Filtersysteme für E-Mail bis hin zur Informationsbeschaffung im Internet. Als weiterer Schritt wird die Simulation von Schreibtischoberflächen durch animierte Umgebungen abgelöst, in denen Agenten als anthropomorphe Figuren in virtuellen Räumen agieren und delegierte Aufträge ausführen.

Interface-Agenten vereinfachen die Handhabung komplexer Anwendungen

insofern, als anstelle der direkten Bedienung über eine Kommandoschnittstelle oder Maus und Menüs die Kommunikation mit dem Anwendungssystem tritt, wobei Anteile seiner Funktion unabhängig von der direkten Steuerung durch Benutzer erbracht werden. Dem Interface-Agenten wird ein Auftrag des Benutzers etwa durch textuelle oder sprachliche Eingabe mitgeteilt, und er nutzt Wissen über die Anwendungsdomäne, um den Auftrag eigenständig auszuführen oder dem Benutzer Lösungsvorschläge zu präsentieren. Hinzu können auch Techniken des maschinellen Lernens kommen, z.B. zur automatischen Anpassung an den individuellen Benutzer.

Mit der Idee von Interface-Agenten, die ihren Benutzer in einen kooperativen Prozess einführen, in der sowohl der Mensch als auch der Agent Kommunikation initialisieren, Aufgaben ausführen und Aktivitäten beobachten können, wird ein komplementärer Interaktionsstil möglich. Als wesentliche Anforderungen an die Gestaltung von Interface-Agenten nennt Maes (1994) Kompetenz und Vertrauen. Kompetenz betrifft das Wissen, das erforderlich ist, um zu entscheiden, wann und wie Benutzer bei bestimmten Aufgaben unterstützt werden. Vertrauen bezieht sich auf die Bereitschaft des Benutzers, die Ausführung von Aufgaben an einen Agenten zu delegieren.

Künstliche Gesprächspartner und soziale Roboter

Die mit Interface-Agenten aufgekommene Idee eines kommunikativen Umgangs mit technischen Artefakten setzt sich fort in dem Ansatz künstlicher Gesprächspartner in Form synthetischer, computeranimierter Figuren (sog. virtueller Agenten), in breitem Umfang populär gemacht durch sog. verkörperte konversationale Agenten (*embodied conversational agents*; Cassell et al. 2000). Dabei handelt es sich um technische Systeme, die in Erscheinung und Verhalten menschenähnlich konzipiert sind und durch ein komplexes Zusammenspiel sensorischer, kognitiver und aktorischer Fähigkeiten eine dialogische, multimodale Kommunikation mit der Maschine realisieren. Um autonomes und flexibles Interaktionsverhalten zu erzielen, werden modellbasierte Techniken eingesetzt, die wesentliche Anteile des Verhaltens – etwa sprachsynchrone Gestik, Mimik oder Lippenbewegungen – automatisch erzeugen (z.B. Kopp/Wachsmuth 2004).

Zusätzlich können solche Agenten über ‚mentale‘ Eigenschaften verfügen, die mit Begriffen wie ‚Wissen‘, ‚Überzeugung‘, ‚Motivation‘, ‚Wunsch‘, ‚Intention‘ oder

‚Verpflichtung‘ charakterisiert werden können. Ein Beispiel ist die *belief-desire-intention*-Architektur (Rao/Georgeff 1991), die es Agenten ermöglicht, durch kognitive Verarbeitung – z.B. durch Bewertung der Situation oder durch planvolle Zielverfolgung – Entscheidungen über die Auswahl von Aktionen zu treffen. Ein weitergehendes Ziel ist eine kognitive Ausstattung künstlicher Agenten derart, dass eine Mensch-Maschine-Kooperation ähnlich wie die Kooperation zwischen Menschen durch gemeinsames Wissen über Ziele und gemeinsames Verfolgen von Plänen realisiert wird (*human emulation approach*; Terveen 1995).

Für die glaubwürdige Verhaltensausrüstung verkörperter Agenten in der Interaktion mit Menschen spielt darüber hinaus die Modellierung von Persönlichkeit sowie die Integration von Emotionsmodellen eine wichtige Rolle. Kontrovers diskutiert wird die Frage, ob es für einen positiven Eindruck auf menschliche Interaktanten tatsächlich notwendig ist, interne Emotionszustände zu simulieren, oder ob dafür der situationsbezogene Einsatz vorberechneter Emotionsausdrücke, etwa eines synthetischen Gesichts, ausreichend ist.

Während *embodied conversational agents* (ECA) als synthetische Charaktere Bestandteil einer simulierten, computergrafischen Umgebung sind, sind autonome Roboter typischerweise reale, in der physikalischen Welt befindliche Agenten. Von Belang für die Mensch-Maschine-Interaktion sind speziell Ansätze zu kooperativen oder ‚sozialen‘ Robotern, d.h. zu autonomen Robotern, die zielgerichtet mit Menschen interagieren und etwa selbst erkennen können, wann und wie Hilfe anzubieten ist. Solche Roboter weisen eine humanoide äußere Erscheinung auf, imitieren die Erscheinung von Tieren (z.B. Hund, Katze, Dinosaurier, Seehund) oder haben eine cartoonhafte Erscheinung (Fong et al. 2003). Um dem menschlichen Interaktanten glaubwürdiges, gewolltes Verhalten zu suggerieren, verfügen sie über annähernd lebensechtes expressives Verhalten, bis hin zur koordinierten Ein- und Ausgabe von Sprache und nichtverbalen Signalen, etwa Gesten und Emotionsausdrücke des Gesichts sowie Kopfbewegungen und Blickrichtung, mit denen z.B. ein Aufmerksamkeitsfokus übermittelt werden kann. Die Forschung zu kommunikationsfähigen Robotern ähnelt in ihrer Zielsetzung in vielerlei Hinsicht der zu künstlichen Gesprächspartnern in Form von *embodied conversational agents*, speziell was die Kombination verbaler und nichtverbaler Kommunikation angeht, so dass beide Ansätze wechselseitig voneinander profitieren können (Wachsmuth/Knoblich 2008).

Abschließende Bemerkungen

Bereits mit dem Aufkommen der Idee von Interface-Agenten – und mehr noch mit der des künstlichen Gesprächspartners oder sozialen Roboters – verändert sich die Sicht des maschinellen Systems als computergestütztes Werkzeug für die Interaktion, d.h. als *Kommunikationsmedium*, zu einer Sicht, die den künstlichen Agenten als *Teilnehmer* im Kommunikationsprozess betrachtet. In der Techniksoziologie hat die Verwendung des Begriffs der Interaktion für technische Artefakte zu einer kontroversen Diskussion über die Handlungsträgerschaft interaktiver Technologien geführt (Krummheuer 2008): So wird in den klassischen Techniktheorien der Terminus ‚Interaktion‘ zunächst für natürliche Personen reserviert und Technik jeweils als Werkzeug, als Objekt menschlichen Handelns gesehen. In den Bezügen zwischen Menschen und künstlichen Agenten lassen sich jedoch Verhalten feststellen, die eine Orientierung an sozialen Mustern zwischenmenschlicher Interaktionen aufweisen, indem der künstliche Agent behandelt wird *als ob* er eine Person, eine Art soziales Gegenüber sei.

Wenngleich der Charakter interaktiver Technologien nicht völlig der Interaktion zwischen Menschen gleichkommt, gibt es Evidenzen dafür, dass technische Artefakte zunehmend als soziale Akteure wahrgenommen werden. So wurde beobachtet, dass Menschen die Interaktion mit solchen Systemen vorziehen, die eine erkennbare Persönlichkeit aufweisen und in der Lage sind, ihren kommunikativen Stil an den Benutzer anzupassen (Reeves/Nass 1996). Untersuchungen im Bereich des *affective computing* haben gezeigt, dass maschinelle Systeme – selbst bei rein textlicher Kommunikation – effektiver mit dem Menschen kommunizieren, wenn sie Emotionen wahrnehmen und zum Ausdruck bringen (Picard 1997). Dies ist insbesondere für verkörperte Agenten mit konversationalem oder kooperativem Verhalten der Fall, die nichtverbale Signale in den sozialen Dialog einbeziehen (Bickmore/Cassell 2005).

Damit verschiebt sich die Sicht, dass Menschen ‚Benutzer‘ einer Anwendung sind, zu der einer Sicht von ‚Partnerschaft‘ mit künstlichen Agenten. Voraussetzung dafür ist jedoch, dass diese künstlichen Agenten als fähig betrachtet werden können, in Mensch-Technik-Kooperationen wie autonom handelnde Entitäten die Initiative zu ergreifen (Negrotti 2005). Zwar wird in der Programmatik zukünftiger Mensch-Technik-Kooperation *„technischen Artefakten an keiner Stelle der Status von handelnden Subjekten mit eigenem Bewusstsein zugeschrieben“*; es wird lediglich konstatiert, dass Artefakte zunehmend Funktionen übernehmen, die zuvor nur dem

Menschen zugemessen wurden und daher immer menschenähnlicher zu werden *scheinen*“ (Cuhls et al. 2009, 15). Inwieweit durch Fortschritte in der Künstliche-Intelligenz-Forschung künstlichen Agenten in der MMI ein Status von handelnden Subjekten (auch rechtlich) zugemessen werden kann, wie es in Science-Fiction-Szenarien vielfach diskutiert wird, wird jedoch die Zukunft zeigen.

Literatur

- Bickmore, Timothy/Cassell, Justine (2005): Social dialogue with embodied conversational agents. In: Jan van Kuppevelt/Laila Dybkjaer/Niels Ole Bernsen (Hg.): *Advances in Natural, Multimodal Dialogue Systems*. New York, 23–54.
- Cassell, Justine/Sullivan, Joseph/Prevost, Scott/Churchill, Elizabeth (Hg.) (2000): *Embodied Conversational Agents*. Cambridge (Mass.).
- Cuhls, Kerstin/Ganz, Walter/Warnke, Philine (Hg.) (2009): *Foresight-Prozess im Auftrag des BMBF, Zukunftsfelder neuen Zuschnitts*. Karlsruhe. http://www.bmbf.de/pubRD/Foresight-Prozess_BMBF_Zukunftsfelder_neuen_Zuschnitts.pdf.
- Fong, Terrence/Nourbakhsh, Illah/Dautenhahn, Kerstin (2003): A survey of socially interactive robots. In: *Robotics and Autonomous Systems* 42, 143–166.
- Handbuch Usability: ISO 9241. <http://www.handbuch-usability.de/iso-9241.html>.
- Hedicke, Volkmar (2002): Multimodalität in Mensch-Maschine-Schnittstellen. In: Klaus-Peter Timpe /Thomas Jürgensohn/Harald Kolrep (Hg.): *Mensch-Maschine-Systemtechnik*. Düsseldorf [2000], 203–230.
- Kay, Alan (1990): User interface. In: Brenda Laurel (Hg.): *The Art of Human-Computer Interface Design*. Reading, 191–207.
- Kopp, Stefan/Wachsmuth, Ipke (2004): Synthesizing multimodal utterances for conversational agents. In: *Computer Animation and Virtual Worlds* 15, 39–52.
- Krummheuer, Antonia (2008): Die Herausforderung künstlicher Handlungsträgerschaft. In: Hajo Greif/Oana Mitrea/Matthias Werner (Hg.): *Information und Gesellschaft*. Wiesbaden, 73–95.
- Laurel, Brenda (1990). Interface agents. In: Brenda Laurel (Hg.): *The Art of Human-Computer Interface Design*. Reading, 155–165.
- Maes, Pattie (1994): Agents that reduce work and information overload. In: *Communications of the ACM* 37, 31–40.

- Negrotti, Massimo (2005): Humans and naturoids. In: Massimo Negrotti (Hg.): *Yearbook of the Artificial*, Bd. 3. Bern, 9–15.
- Picard, Rosalind (1997): *Affective Computing*. Cambridge (Mass.).
- Rao, Anand/Georgeff, Michael (1991): Modeling rational agents within a BDI-Architecture. In: *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning*. San Fransisco, 473–484.
- Reeves, Byron/Nass, Clifford (1996): *The Media Equation*. New York.
- Shneiderman, Ben (³2002): *User Interface Design – Effektive Interaktion zwischen Mensch und Maschine [Designing the User Unterface – Strategies for Effective Human-Computer Interaction, 1987]*. Bonn [2001].
- Shneiderman, Ben/Maes, Pattie (1997): Direct manipulation vs. interface agents. In: *Interactions* 4, 42–61.
- Sullivan, Joseph/Tyler, Sherman (Hg.) (1991): *Intelligent User Interfaces*. New York.
- Terveen, Lauren (1995): Overview of human-computer collaboration. In: *Knowledge-Based Systems* 8, 67–81.
- Wachsmuth, Ipke/Knoblich, Günther (Hg.) (2008): *Modeling Communication with Robots and Virtual Humans*. Berlin.

Ipke Wachsmuth

Dieser Text ist – ergänzt um zahlreiche Querverweise – erschienen in: A. Stephan & S. Walter (Hrsg.), <i>Handbuch Kognitionswissenschaft</i> (pp. 361–364). Stuttgart Weimar: J.B. Metzler.
