

Europäische Akademie
zur Erforschung von Folgen
wissenschaftlich-technischer Entwicklungen
Bad Neuenahr-Ahrweiler GmbH

Direktor:
Professor Dr. Carl Friedrich Gethmann

Robotik.
Einführung in eine
interdisziplinäre Diskussion

von
Michael Decker (Hrsg.)
September 1999

Europäische Akademie
zur Erforschung von Folgen
wissenschaftlich-technischer Entwicklungen
Bad Neuenahr-Ahrweiler GmbH

Direktor:
Professor Dr. Carl Friedrich Gethmann

Robotik.
Einführung in eine
interdisziplinäre Diskussion

von
Michael Decker (Hrsg.)
September 1999

Die Schriften der „Graue Reihe“ umfassen aktuelle Materialien und Dokumentationen, die von den Wissenschaftlern der Europäischen Akademie zur Erforschung von Folgen wissenschaftlich-technischer Entwicklungen Bad Neuenahr-Ahrweiler GmbH laufend erarbeitet werden. Die Publikationen der „Grauen Reihe“ werden als Manuskripte gedruckt und erscheinen in loser Folge im Selbstverlag der Europäischen Akademie. Sie können über die Europäischen Akademie auf schriftliche Anfrage hin bezogen werden.

Herausgeber:

Europäische Akademie

zur Erforschung von Folgen
wissenschaftlich-technischer Entwicklungen
Bad Neuenahr-Ahrweiler GmbH

Postfach 14 60, D-53459 Bad Neuenahr-Ahrweiler
Telefon: ++49 - (0)2641 - 973 - 300, Telefax -320
e-mail: europaeische.akademie@dlr.de

Direktor:

Professor Dr. Carl Friedrich Gethmann (V.i.S.d.P.)

ISSN 1435-487 X

Redaktion:

Dagmar Uhl, M. A.

Druck:

Druckerei Martin Warlich, Bad Neuenahr-Ahrweiler

Vorwort

KickOff-Meetings sind ein fester Bestandteil des Projektablaufes der Europäischen Akademie. Sie sind der Ort, an dem zu Anfang des Projektes Experten unterschiedlicher Disziplinen eingeladen werden, um ihre Sichtweise auf das Thema des neuen Projektes darzulegen. Damit stellt das KickOff-Meeting die erste Reflexion auf das Arbeitsprogramm der Projektgruppe dar, welches zwar vom Wissenschaftlichen Beirat der Europäischen Akademie begutachtet worden ist, aber noch nicht der wissenschaftlichen Öffentlichkeit vorgestellt wurde. Die Projektgruppe ist ihrerseits angehalten, die Anregungen der externen Wissenschaftler aufzunehmen und sie in die Studie einzubeziehen. (Dies kann sowohl durch Berücksichtigung des wissenschaftlichen Beitrages als auch durch begründete Nicht-Berücksichtigung geschehen.)

Die Europäische Akademie strebt daher die Zusammenfassung der Beiträge des KickOff-Meetings an, um sie in geordneter Form für die Projektarbeit zur Verfügung zu haben. Die hohe Qualität der Beiträge legte es nahe, diese Zusammenfassung auch einer breiteren Öffentlichkeit zur Verfügung zu stellen. Allerdings stand dieser Idee der Workshop-Charakter des KickOff-Meetings ein wenig entgegen, waren die Vortragenden doch gebeten worden, auch gewagtere, provokante Statements aus ihren Disziplinen zu präsentieren, um die interdisziplinäre Diskussion anzuregen.

Die Graue Reihe der Europäischen Akademie bietet die Möglichkeit, beidem gerecht zu werden: Sie ist eine „Reihe“ in dem Sinne, daß sie über die ISSN-Nummer bestellt werden kann und sie ist „graue Literatur“, d.h. es können auch Statements und Wortbeiträge ohne ausführliche Literaturangabe etc. berücksichtigt werden. Dementsprechend unterschiedlichen Ursprungs sind auch die Beiträge in diesem Buch. Einige Autoren haben das ausformulierte Manuskript des Vortrages, einige einen Beitrag mit Bezug zum Vortrag abgegeben. Die Herren Professoren F. Haft und H. Ritter haben dem Vorschlag zugestimmt, den Vortrag vom Unterzeichnenden aufzeichnen zu lassen, und die Zusammenfassung zu redigieren.

Bad Neuenahr-Ahrweiler, im September 1998

Michael Decker

INHALTSVERZEICHNIS

Einleitung	7
<i>Michael Decker</i>	
Substitution kommunikativer Kompetenz?	17
<i>Peter Janich</i>	
Kognitive Robotik.	
Ein neues Anwendungsgebiet für die philosophische Ethik?	32
<i>Thomas Metzinger</i>	
Arbeitsqualität durch Komplementarität von Mensch und Technik. .	47
<i>Gudela Grote</i>	
Haftung für Automaten	55
<i>Fritjof Haft</i>	
Stand der Forschung auf dem Gebiet	
Humanoider Roboter und geplante Anwendungsszenarien	67
<i>Rüdiger Dillmann</i>	
Roboter im medizinischen Bereich. Beispiele und Chancen	79
<i>Rolf-Dieter Schraft</i>	
Neurotechnologie. Auf dem Weg zu lernfähigen Neuroprothesen mit sensorischen und motorischen Implantaten	89
<i>Rolf Eckmiller</i>	
Perspektiven intelligenter Mensch-Maschine Kommunikation	103
<i>Helge Ritter</i>	
Autorenverzeichnis	121

Einleitung

Michael Decker

Diese Einleitung soll die Beiträge zum KickOff-Meeting des Projektes „Robotik. Optionen der Ersetzbarkeit des Menschen“ in zweierlei Hinsicht ergänzen. Zum einen wird ein Überblick über das Projekt selbst gegeben, um Lesern, die das Projekt noch nicht kennen, einen entsprechenden Hintergrund zu geben. Zum anderen werden in den Beiträgen verschiedentlich Beispiele von Robotern erwähnt, die nicht immer extra beschrieben werden. Einige dieser Roboter werden im zweiten Teil „vorgestellt“. Dabei wird nur eine jeweils kurze Beschreibung gegeben, die sich auf die Besonderheit des einzelnen Roboters konzentriert und ein Bild desselben zeigt. Für weitere Details ist jeweils die Internetadresse bzw. weiterführende Literatur angegeben.

Das Projekt

Hintergrund

Roboter müssen über eine Reihe von Grundfähigkeiten verfügen, wenn sie in die Lage versetzt werden sollen, in der Welt Handlungen auszuführen. Zu diesen Grundfähigkeiten gehören beispielsweise die Möglichkeit der (Fort-)bewegung, die häufig durch Räder realisiert wird. Alternativ werden auch „natürliche“ Fortbewegungsarten durch die Konstruktion von Beinen, Flügeln, Schuppen, Flossen, u.ä. kopiert. Die Möglichkeit der Wahrnehmung wird durch Kameras und andere Sensoren implementiert, aus deren Daten ein Modell der Umwelt erstellt wird. Der dritte wichtige Aspekt besteht in der Möglichkeit zu lernen, wobei das Lernen von Bewegungsabläufen, das Lernen der Interpretation des Weltmodells und das Lernen reflexiver Betrachtungsweisen, z. B. wie sich ein Roboter selbst in seine Umwelt einordnet, als die bedeutendsten Bereiche angesehen werden. Roboter, die über diese Möglichkeiten des Handelns verfügen, werden „autonome Roboter“ genannt.

Ist beispielsweise für das Erfüllen der Aufgabe eine Bewegung des gesamten Roboters notwendig, so wird auf Grundlage der Sensordaten eine Landkarte von der unmittelbaren Umgebung entwickelt, innerhalb welcher kollisionsfreie Bewegungsbahnen berechnet werden können. Bei der Umsetzung des Handlungsplanes werden ununterbrochen die Daten der Sensoren ausgelesen und in die kurzfristige Planung mit einbezogen. Auf Grundlage der aktuellen Daten entwickelt die Steuerung eigenständig Alternativpläne. Der Robotersteuerung sind verschiedene Handlungsoptionen implementiert, die je nach Aufgabenstellung mehr oder weniger praktikabel sind. Gleichzeitig verfügt die Robotersteuerung über die Möglichkeit, erfolgreiche und nicht erfolgreiche Handlungsversuche abzuspeichern. Das hat zur Folge, daß ein Roboter, der längere Zeit mit ähnlichen Aufgaben beauftragt wird, erfolgversprechende Optionen zuerst ausprobiert. Damit wird ihm ermöglicht, sich einer Aufgabenstellung anzupassen, was als eine Art „Lernen“ angesehen werden kann.

Ebenfalls von Interesse sind Ansätze, innerhalb derer einzelne Komponenten mit weitreichender Handlungsautonomie ausgestattet werden, deren emergentes Zusammenwirken Handlungen des Gesamtsystems realisieren. Eine Vorhersage der Handlungsoptionen scheint damit nicht mehr bzw. nur bedingt möglich. Gerade von diesem Ansatz verspricht man sich sehr robuste Robotersteuerungen, die in nicht für den Robotereinsatz optimierten Umgebungen Handlungen durchführen können.

Fragestellung

Geht man davon aus, daß Menschen Roboter entwickeln, um bestimmte Handlungsziele zu erreichen, so kann der künstliche Agent in diesem Zusammenhang als Mittel zum Zweck angesehen werden. Sollte es umgekehrt nicht gelingen, einen Roboter zur Durchführung einer Handlung zu entwickeln, so bleibt das Handlungsziel unverändert und der Mensch wird es auf andere Art und Weise verfolgen. In diesem Sinne kann ein Roboter menschliche Fähigkeiten ersetzen und aus der Sicht einer entwicklungsbegleitenden Technikfolgenbeurteilung kann man die Frage nach der Ersetzbarkeit des Menschen stellen.

In diese Fragestellung gehen die Perspektiven verschiedener wissenschaftlicher Disziplinen ein. Die Sicht der Robotik kann zunächst nur eine technische Ersetzbarkeit des Menschen beurteilen. Die Basis dieser Beurteilung wird im allgemeinen aus einem Anforderungskatalog der speziellen Anwendung bestehen. Die ökonomische Betrachtungsweise muß berücksichtigen, wann sich eine Ersetzung aus betriebswirtschaftlicher, bzw. volkswirtschaftlicher Sicht „lohnt“. Die Rechtswissenschaften entscheiden, wer die Verantwortung für ausgeführte Handlungen des künstlichen Agenten trägt, was beispielsweise bei Haftungsfragen eine Rolle spielen kann. Direkt gekoppelt sind diese Sichtweisen mit moralischen Aspekten, aus denen ethische Normen entwickelt werden können, anhand derer die Ersetzung des Menschen beurteilt werden kann. Insbesondere im Bereich der Service-Roboter, einem der wichtigsten Anwendungsbereiche autonomer Roboter, stellt sich die Frage nach der Bewertung der sekundären menschlichen Fähigkeiten. Schließlich sind mit dem Begriff Dienstleistung nicht nur Attribute wie Funktion, Rentabilität, Sicherheit und Zuverlässigkeit verbunden, sondern auch Freundlichkeit, Kommunikationsfähigkeit und Verständnis.

Durchführung

Die Europäische Akademie beruft für die Dauer von zwei Jahren eine interdisziplinär besetzte Projektgruppe, die sich oben beschriebener Fragestellung widmen soll. Ziel des Projektes ist es, begründete Kriterien zu finden, anhand derer die Ersetzung menschlicher Fähigkeiten durch künstliche autonome Agenten als wünschenswert bzw. nicht wünschenswert beurteilt werden kann. Monatliche Treffen erlauben eine intensive Diskussion, die zu der Formulierung von begründeten Handlungsempfehlungen führen soll, welche von allen Disziplinen als akzeptabel angesehen werden.

Aus jeder wissenschaftlichen Disziplin, die während des Vorprojektes als für die Fragestellung relevant befunden wurde, wurde ein Experte in die Arbeitsgruppe berufen. Ergänzend werden zu Beginn des Projektes weitere Experten zu einem KickOff-meeting eingeladen, um die

Diskussion auf eine breitere Basis zu gründen. Ungefähr in der Mitte der Laufzeit werden bei einem weiteren Treffen mit externen Wissenschaftlern die bisherigen Ergebnisse der Projektarbeit diskutiert. Dieses Vorgehen ist von Vorteil in zweierlei Hinsicht: Zum einen kann die Projektgruppe relativ klein gehalten werden, womit die Organisation der monatlichen Treffen realisierbar bleibt. Zum anderen basieren die Ergebnisse des Projektes nicht ausschließlich auf den Erkenntnissen der Projektgruppenmitglieder, sondern auf einem breiteren wissenschaftlichen Fundament.

Die Fallbeispiele des Projektes sind aus dem Bereich der medizinischen Anwendung gewählt. Hier finden sich viele Nischen, in denen autonome Roboter eingesetzt werden können, wie zum Beispiel in der Prothetik, im Bereich der Patientenbetreuung oder bei chirurgischen Eingriffen. Gleichzeitig ist in der Mannigfaltigkeit der Anwendungen das gesamte Spektrum der Robotik bezüglich der Eigenschaften, die am Anfang dieses Beitrages beschrieben wurden, abgedeckt. So findet beispielsweise die Kommunikation bei einer Prothese über Körperkontakt bis hin zur Kopplung an Nervenfasern statt, während bei chirurgischen Eingriffen häufig eine Joystick-Steuerung eingesetzt wird. Im Falle des Pflegeroboters bietet sich sprachliche Kommunikation an. Die Kooperation variiert zwischen direktem Körperkontakt über den Werkzeugcharakter bis hin zur „partnerschaftlichen“ Zusammenarbeit. Die Voraussetzungen für Adaptivität und Lernen unterscheiden sich ebenfalls deutlich: Eine Prothese wird im allgemeinen nur von einer Person verwendet, hier kann ein langfristiger Lern- und Anpassungsprozeß vorgesehen werden. Ein Pflegeroboter muß dagegen in der Lage sein, als „Diener vieler Herren“ zu agieren, wofür ein flexibler Lern- bzw. Adaptionsalgorithmus implementiert werden muß. Schließlich ist der medizinische Bereich ein Kontext, in dem die oben angeführten Fragestellungen - die einzelnen wissenschaftlichen Disziplinen betreffend - in relevanter Weise zum Tragen kommen.

Das Ergebnis des Forschungsprojektes wird von der Arbeitsgruppe in einem transdisziplinären Memorandum veröffentlicht. Neben dem Katalog begründeter Kriterien, der in entsprechende Handlungsempfeh-

lungen mündet, wird darin auch die Argumentationskette dargestellt, die zu den Handlungsempfehlungen führt. Damit kann das Memorandum einerseits einen Beitrag zur politischen Entscheidungsfindung liefern, wenn es um Gesetzgebung oder Forschungsförderung geht, und andererseits zu einer fundierten Diskussion innerhalb der Wissenschaften selbst und in der Gesellschaft beitragen.

Projektgruppe

Professor Dr. rer.nat. Thomas Christaller (Vorsitz)

Institut für Systementwurfstechnik, GMD Sankt Augustin

Professor Dr. med. Joachim Gilsbach

Neurochirurgische Klinik, Universitätsklinikum RWTH Aachen

Professor Dr.-Ing. Gerd Hirzinger

Institut für Robotik und Systemdynamik, DLR Oberpfaffenhofen

Professor Dr. med. Dr. sc. Karl Lauterbach

Institut für Gesundheitsökonomie, Medizin und Gesellschaft, Universität zu Köln

Professor Dr. jur. Dr. rer.soc.oec. Erich Schweighofer

Institut für Völkerrecht (AG Rechtsinformatik), Universität Wien

Professor Dr.-Ing. Gerhard Schweitzer

Institut für Robotik, ETH Zürich

Professor Dr. phil. Dieter Sturma

Institut für Philosophie, Universität Essen

Projektleiter

Dr. rer.nat. Michael Decker

Europäische Akademie zur Erforschung von Folgen wissenschaftlich-technischer Entwicklungen Bad Neuenahr-Ahrweiler GmbH

„Prominente“ Roboter

COG

Für die Beschreibung von COG, einem Forschungsprojekt am MIT (Artificial Intelligence Laboratory Cambridge, MA), soll der Erfinder Rodney Brooks zu Wort kommen (freie Übersetzung MD):

Folgende Hypothese kann als Motivation für die Entwicklung von COG angesehen werden:

„Humanoide Intelligenz benötigt humanoide Interaktion mit der Umwelt“.

Wenn man COG beschreiben möchte, dann stellt man sich am besten einen Satz von Sensoren und Aktuatoren vor, der so zusammengestellt wurde, daß er der sinnlichen Wahrnehmung und der motorischen Bewegung von menschlichen Körpern möglichst nahe kommt. Abgesehen von den Beinen und einer flexiblen Wirbelsäule sind alle wesentlichen Freiheitsgrade der Bewegung im Rumpf, dem Kopf und den Armen realisiert. Sehsinn, Gehör und Tastsinn sind durch Videocameras bzw. entsprechende Sensoren implementiert. Die Wahrnehmung des eigenen Körpers ist integriert über die Rückmeldung der Gelenkstellung und des wirkenden Drehmomentes. Zur Zeit wird an einem vestibulären System, an Sprache und an Roboterhänden gearbeitet. COG ist eine Plattform, auf der die Ergebnisse aus allen Teilgebieten der Künstlichen Intelligenz-Forschung realisiert werden können, um zu einem einzigen, kohärenten und funktionierenden Ganzen kombiniert zu werden.

(<http://www.ai.mit.edu/projects/cog/>)

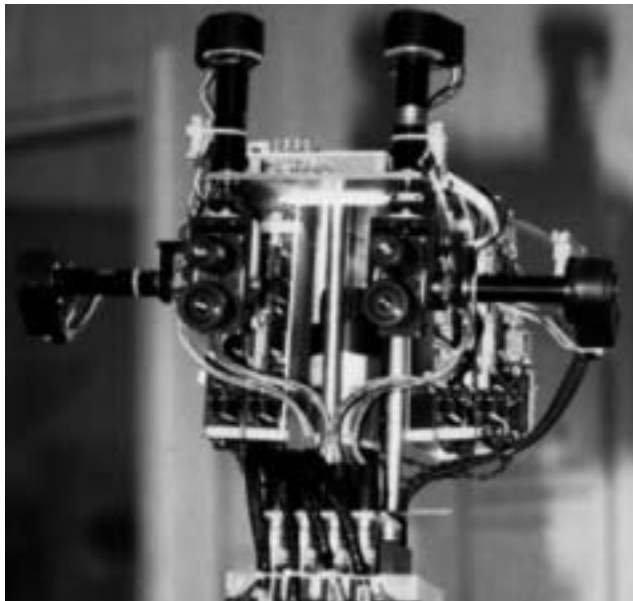


Bild oben: COG dreht eine Kurbel*

Bild unten: Der Kopf von COG mit den vier Kameras*



Der Kopf von COG wurde in dem Forschungsprojekt KISMET um typische Gesichtsmerkmale erweitert.

Bild links: „Ärgerlicher“ Gesichtsausdruck*

Bild rechts: „Freundlicher“ Gesichtsausdruck*

*alle Bilder: © Sam Ogden

Verschiedene Forschungsprojekte um COG beschäftigen sich insbesondere mit sozialem Verhalten von Robotern. Damit ist sowohl soziales Verhalten unter Robotern gemeint, als auch soziales Verhalten zwischen Menschen und Robotern. In dem Projekt Kismet wird die Interaktion von Eltern zu Kind als Vorbild genommen. Der Roboter lernt „wie ein Kind“. (<http://www.ai.mit.edu/projects/kismet/kismet.html>)

Literatur:

„The Cog Project: Building a Humanoid Robot“

Rodney A. Brooks, Cynthia Breazeal, Matthew Marjanovic, Brian Scassellati, Matthew Williamson.

To appear in a Springer-Verlag Lecture Notes in Computer Science Volume.

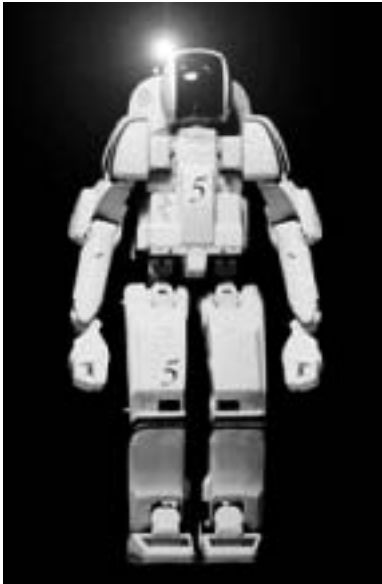
„Infant-like Social Interactions Between a Robot and a Human Caretaker“.
Breazeal(Ferrell), C. and Scassellati, B. (2000),

To appear in Special issue of Adaptive Behavior on Simulation Models of Social Agents, guest editor Kerstin Dautenhahn.

„How to build robots that make friends and influence people“.

Breazeal, C. and Scassellati, B. (1999),

To appear in IROS99, Kyonjuu, Korea.



Humanoider Honda-Roboter

Der Honda-Roboter stellt eine Besonderheit dar, da er sich auf zwei Beinen fortbewegt. Er ist in der Lage die Balance zu halten, auch wenn er etwas schiebt oder trägt. Er kann Treppen laufen und sich somit in einer Umwelt bewegen, die typischerweise für Bewegung von Menschen ausgelegt ist. Dafür wurde der menschliche Bewegungsapparat als Vorbild herangezogen. Beispielsweise wurde die Lage des Schwerpunktes oder die Beweglichkeit der Gelenke, etc. analog zum menschlichen Körper realisiert. Die Version „P2“ ist 1,80 m groß und wiegt 210 kg, die neuer Version „P3“ ist mit

1,60 m Größe und 130 kg Gewicht schon deutlich kompakter.

<http://www.honda.co.jp/english/technology/robot/index.html>

Sony Hund „AIBO“

In den Labors von Sony wurde ein künstliches Haustier, ein Roboterhund, entwickelt, der als „Entertainment Robot“ vorgestellt wird. Er kann sich mit Hilfe von 18 Gelenken bewegen, verfügt über Bewegungssensoren, eine Kamera und Entfernungssensoren. Auch „Gefühle“ sind implementiert. Programmiert sind beispielsweise „Freude“, „Überraschung“, „Furcht“ und „Ärger“ und sogenannte Instinkte wie „Liebe“ oder „Hunger“.

(<http://www.world.sony.com/robot/top.html>)

Übertroffen wird AIBO allerdings von Robokoneko, den Hugo de Garis, Leiter der Brain Builder Group von Sony, als den intelligentesten, jedenfalls komplexesten, Roboter bezeichnet. Er hat die Form eines Kätzchens (Robokoneko: japanisch für ein Roboterkatzenjunges) und verfügt über ein „Gehirn“ mit 32000 Modulen und etwa 40 Millionen künstlichen Neuronen. Ein solch großes „Gehirn“ hat in dem kleinen Körper keinen Platz, daher ist es außerhalb untergebracht und über Funk mit dem Körper in Kontakt. Im Gegensatz zu AIBO, der optisch ein Roboterhund bleibt, soll Robokoneko mit einem Fell ausgestattet werden, um möglichst wie eine echte Katze auszusehen.

(<http://www.telepolis.de/tp/deutsch/special/robo/6353/1.html>)

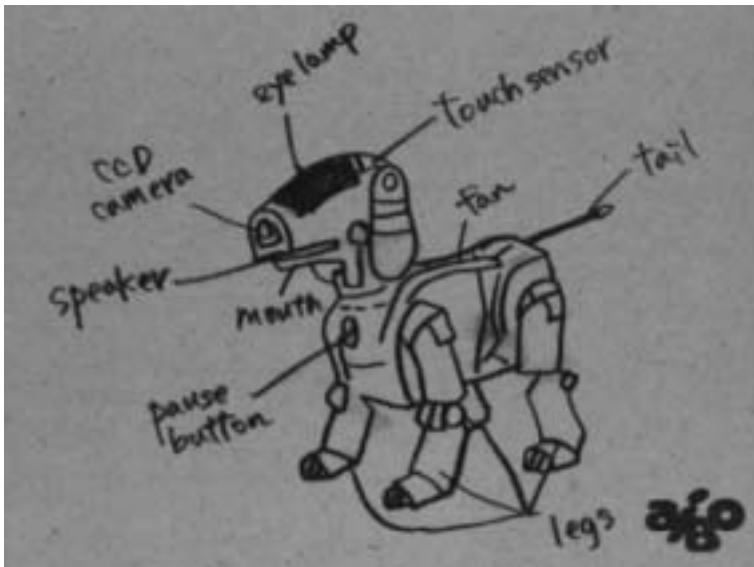


Bild oben links: AIBO in Gesamtansicht

Bild oben rechts: Der Kopf von AIBO

Bild unten: Schemaskizze

Substitution kommunikativer Kompetenz?

Peter Janich

Einleitung

Ziel dieses Vortrags ist die Klärung der Frage, ob und in welchem Sinne die Kompetenz von Menschen, miteinander zu kommunizieren, technisch durch Automaten substituiert werden kann. Die Antwort auf diese Frage setzt bei den Vorgaben an, die durch die Projektskizze der Europäischen Akademie sowie die Vorbereitung des kick-off-meetings gemacht wurden.

In der Projektskizze „Robotik. Ersetzbarkeit des Menschen“ heißt es unter der Überschrift „Ziel“: „Es gilt *Kriterien* zu finden, anhand derer die Ersetzbarkeit des Menschen durch künstliche Agenten *beurteilt* werden kann: Soll, kann, darf der Mensch ersetzt werden?“. Einen Beitrag zu dieser Beurteilung zu leisten ist die Absicht meines Vortrags.

Dazu heißt es dann unter der Überschrift „Ausgangslage“ zu Aspekten, unter denen eine Nicht-Ersetzbarkeit des Menschen diskutiert werden kann, daß technische, ökonomische, rechtliche und moralische Gründe ins Auge zu fassen sind – mit dem Zusatz, daß diese Liste „um weitere Gesichtspunkte der Nicht-Ersetzbarkeit erweiterbar“ sei.

In einem freundlichen Anschreiben an die beiden Philosophen, Herrn Metzinger und mich, teilt schließlich Herr Kollege Sturma mit, daß von philosophischer Seite einerseits die „Theorieperspektiven“ Anthropologie, Philosophie des Geistes, Ethik und Ethik der Technik, andererseits als Arbeitsvorschläge fünf Punkte eingebracht wurden, die ich alle im weiteren Sinne normativen Fragestellungen zuordnen würde.

In all diesen Hinweisen findet sich nirgends eine Andeutung von *erkenntnistheoretischen* oder *wissenschaftstheoretischen Fragen zur Robotik*. Ich interpretiere dies vorläufig so, daß im geplanten Projekt vor allem eine Bewertung der Ersetzung des Menschen durch Roboter nach sozialen, moralischen, rechtlichen, politischen und ästhetischen Kriterien gesucht wird, wozu selbstverständlich auch ein Überblick über die derzeitigen und die erwartbaren technischen Möglichkeiten erforder-

lich ist. Mir scheint jedoch zwischen den technischen Möglichkeiten und der Bewertung eine wesentliche Lücke zu liegen, die etwas mit erkenntnis- und wissenschaftstheoretischen Problemen des investierten Verständnisses von Technik und Naturwissenschaft zu tun hat. Die gesamte Debatte um die Robotik einschließlich der KI-Forschung steht nämlich unter einer Reihe unerkannter, stillschweigender, nichtsdestoweniger höchst starker und kritikbedürftiger Prämissen. Diese liegen auf zwei Ebenen:

- (1) Auf der *Objektebene* der Ersetzung des Menschen durch Roboter bestehen diese Prämissen in einem *Menschenbild*, das ganz aus den Naturwissenschaften stammt und pauschal als *naturwissenschaftlicher Materialismus* mit den Varianten Physikalismus, Physiologismus und Biologismus zu bezeichnen ist.
- (2) Auf der Metaebene sind die investierten Wissenschaften auf der Seite von Mathematik und Informatik *formalistisch*, auf der Seite der Natur- und Ingenieurwissenschaften *empiristisch* verstanden.

Beide Ebenen zusammenfassend wird man die gesamte Robotik-Debatte („gesamt“ selbstverständlich nur relativ zu meinem Kenntnisstand, wonach ich kein Gegenbeispiel finden konnte) als „naturalistisch“ bezeichnen dürfen. Da in der gegenwärtigen philosophischen Naturalismusdebatte eine kaum überschaubare Vielfalt verschiedener Naturalismen vertreten wird, ist für diesen Vortrag unter „naturalistisch“ die Zuständigkeit, genauer die Alleinzuständigkeit der (mathematischen) Naturwissenschaften für die Frage der Ersetzbarkeit des Menschen durch Maschinen zu verstehen.

Damit läßt sich das Ziel meines Vortrags bereits genauer nennen als am Anfang dieser Einleitung: Selbstverständlich ist die Robotik-Debatte dieses Projekts ein sprachliches Unternehmen, bei dem über Menschen und über Roboter geredet wird. Mir geht es um einen kritischen Blick auf diese Rede. Um die dabei ins Auge gefaßte *Unterscheidung von Sprachebenen* pointiert hervorzuheben: Wo die Robotikforscher über Roboter reden, rede ich über die Forscher. Welche Investitionen in die *Semantik* dieses Diskurses, und welchen Typ *Geltungsanspruch* vertre-

ten sie? Wie hängt die Beurteilung der Ersetzbarkeit (im Sinne der oben zitierten Zielbestimmung aus der Projektskizze) von den stillschweigenden philosophischen Hintergrundüberzeugungen ihrer Urheber ab?

Da aber die verfügbare Redezeit von „ca. 20 Minuten“ (Einladungsschreiben) keinen Raum für eine Beschreibung, Analyse und Kritik ganzer Theoriestücke der Robotik bietet, gehe ich einen anderen Weg: Ich versuche, den mir wichtigen erkenntnistheoretischen Punkt durch Darlegung und Begründung der *These* zu markieren, daß die *kommunikative Kompetenz des Menschen prinzipiell nicht technisch substituierbar ist*. Dabei muß ich eher thetisch als detailliert argumentierend vorgehen, hoffe aber doch, für die entscheidende These eine schlüssige Begründung vortragen zu können.

1) Philosophische Aspekte von Technik und Robotik

Eine klare Unterscheidung von natürlich und technisch findet sich zum ersten Mal in der Physikvorlesung von Aristoteles. Dort bedeutet der Dativ *technē*, heute zu übersetzen mit technisch, und synonym dem lateinischen Lehnwort artifiziell (und dem zugehörigen Substantiv Artefakt) so viel wie das deutsche Wort „künstlich“. Es steht für das vom Menschen gemachte, und zwar in voller Absicht oder auch als unbeabsichtigte Nebenfolge entstehende Produkt. „Natürlich“ bestimmt Aristoteles durch die Bedingung, daß es den Grund für seine Veränderung in sich selbst trage, von der Schwere des fallenden Steins über das Wachsen von Pflanzen bis zur Lokomotion von Tieren. Dabei ist von Aristoteles klar gesehen, daß diese Unterscheidung nicht etwa eine ontische Zweiteilung aller Dinge und Ereignisse in natürliche und künstliche betrifft (Seinsdualismus), sondern nur zwei Beschreibungsaspekte (Aspekte- oder Beschreibungs dualismus). Ein und dasselbe Ding bzw. Ereignis kann zugleich natürliche und künstliche Eigenschaften haben, wie etwa künstliche Zuchtformen unserer Kulturpflanzen oder unsere Haus- und Nutztiere.

Im *Alltagsverständnis* teilen wir diese aristotelische Auffassung vollständig: Finden wir bei einem Waldspaziergang am Wegrand einen

Steinpilz, nehmen wir an, er sei dort von Natur aus gewachsen. Finden wir an derselben Stelle einen Autoschlüssel, so denken wir wohl, daß jemand dieses technische Produkt dort verloren hat. Unsere geläufigen Unterscheidungen von künstlichen und natürlichen („echten“) Blumen, von natürlichen und künstlichen Zähnen, von einem natürlichen und einem unnatürlichen Tod, von einer natürlichen und einer künstlichen Befruchtung usw. betreffen genau dieselben Aspekte.

Leider ist diese klare Unterscheidung mit Beginn der neuzeitlichen Naturwissenschaft, prominent mit Galilei, dadurch verlorengegangen, daß Galilei in den technischen Produkten, an denen er seine Experimente durchführte, einen Ausdruck von Natürlichkeit, genauer, Naturgesetzlichkeit sah. Im Gefolge sprechen heute alle Naturwissenschaften davon, daß technische Laborforschung Naturwissenschaft sei und Naturerkenntnis produziere. Daß in dieser Auffassung nicht nur eine erhebliche erkenntnistheoretische Ungenauigkeit, sondern ein schwerwiegender naturalistischer Fehler steckt, ist leicht zu zeigen:

„Naturalistisch“ ist es, die Funktion der technischen Apparate (Instrumente, Geräte) wiederum allein naturwissenschaftlich beschreiben und erklären zu wollen. Dies wird, für die Robotik-Debatte einschlägig, an zwei Beispielen deutlich:

- (1) *Meßgeräte*, mit denen quantitative Beobachtungen, vornehmlich in experimentellen Zusammenhängen, gemacht werden, beruhen in ihrer Funktion nach naturalistischer Auffassung allein auf den in ihnen technisch realisierten Naturgesetzen: im einfachen Beispiel die Schwingungsgesetze des Pendels oder eines quarz-stabilisierten Schwingkreises für die Zeitmessung. Diese Auffassung ist insofern falsch, als sie übersehen hat, daß nur ungestörte Meßgeräte brauchbare Meßdaten liefern, die naturwissenschaftlichen Beschreibungen und Erklärungen der Funktion aber durch Störungen (z. B. die stehengebliebene Uhr) nicht außer Geltung gesetzt werden. Vielmehr werden Störungen und Defekte an Meßgeräten gerade aufgrund naturwissenschaftlicher Kausalerklärungen beherrschbar. Kurz, natur- und technikwissenschaftlich lassen sich gerade die entschei-

denden *Funktionskriterien* von Meßgeräten (ihre *Ungestörtheit*) nicht bestimmen. Diese liegen vielmehr in den Zwecksetzungen von Konstruktion, Herstellung und Verwendung. Es sind die Normen, die einerseits operational die jeweiligen Meßparameter definieren, andererseits die logisch-mathematischen Eigenschaften von Meßdaten und damit die *Theoriefähigkeit messender Erfahrung* zur Folge haben. Diese Einsichten sind in der Protophysik zu expliziten Theorien der Messung von Länge, Dauer und Masse und einem Ausblick auf methodisch anschließende physikalische Parameter ausgearbeitet.

- (2) *Informationsverarbeitende* und *kognitive* Maschinen werden naturalistisch derart verstanden, daß deren natur- und technikwissenschaftliche Beschreibung ihre *Funktion kausal erklären* soll. Zum Zweck der Kürze und Übersichtlichkeit wähle ich als Beispiel simple Rechenmaschinen für die vier Grundrechenarten. Analog zur naturalistischen Fehleinschätzung von Meßgeräten wird auch bei Rechenmaschinen das Störungsproblem und seine Folgen übersehen: Da nur ungestörte, d. h. nicht defekte Rechenmaschinen *richtige Rechenresultate* liefern, müßte die natur- und technikwissenschaftliche Kausalerklärung des Rechners nicht nur seine Leistungen im Sinne des Input-Output-Verhältnisses des technischen Systems liefern, sondern vielmehr dessen Geltung oder Wahrheit. Defekte Rechenmaschinen, die falsche Ergebnisse liefern, falsifizieren keinen naturwissenschaftlichen Lehrsatz. Vielmehr, wieder analog den Meßgeräten, dienen solche Lehrsätze zur Kausalerklärung des Defektes und erlauben eine erfolgreiche Reparatur. Und wieder sind es *menschliche Zwecksetzungen des richtigen Rechnens*, die mit dem Mittel der Konstruktion der Rechenmaschine erfolgreich realisiert werden.

Zusammenfassend ist es die *Zweckrationalität* einer durch menschliches Handeln hervorgebrachten Technik, die *naturwissenschaftliches Wissen* trägt. Da es sich von selbst versteht, daß auch in den naturwissenschaftlichen Bereichen der Physiologie, der Evolutionsbiologie und der Kognitionswissenschaften gültige und nicht ungültige Aussagen

angestrebt werden, ist zu beachten, daß die (kurz gesprochen:) wahr-falsch-Unterscheidung in diesen Bereichen nur durch die Zweckrationalität des Handelns in naturwissenschaftlicher Forschung etabliert und genutzt werden kann.

Technik als Grundlage naturwissenschaftlicher Forschung ist in seinen ingenieurwissenschaftlich-theoretischen Definitionen für die *Robotik* zu weit. Dort wird etwa vorgeschlagen, Technik so zu verstehen, daß sie die Mittel für Transformation, Transport und Speicherung bzw. Lagerung von Stoff, Energie und Information bereitzustellen habe. Darunter fallen selbstverständlich Produkte für die Naturbedürfnisse des Menschen wie Nahrung, Kleidung und Behausung einschließlich Möblierung, die für Robotik nur im Sinne der Automatisierung der Produktion von Interesse sein können. Deshalb wird, unter Beibehaltung der aristotelischen Unterscheidung von natürlich und technisch, für Robotik als definitonische Minimalbedingung eine Einschränkung zu machen sein: vom russischen Wort *rabota* (Arbeit) und *rabotatch* (arbeiten) – meines Wissens neben dem Wort *Nepp* das einzige aus dem Russischen kommende Lehnwort im Deutschen – verweisen die Wörter „Roboter“ und „Robotik“ auf menschliche Arbeit bzw. im weiteren Sinne auf *menschliche Tätigkeit*. Ungeachtet der feineren Unterscheidungen in der Robotik-Diskussion etwa zwischen Industrieroboter, Serviceroboter und „Personal Robot“ (R. D. Schraft u. E. Müller 1998) verstehe ich unter einem Roboter prinzipiell ein *technisches Substitut für den Menschen hinsichtlich einer seiner Tätigkeiten*. Da es auch Maschinen gibt, die keine menschliche Tätigkeit substituieren (wie z. B. Flugzeuge), sollen unter Robotern nur „Maschinen“ (von griechisch *mechanaomai*, ich ersinne eine List, Vorrichtung) verstanden werden, deren Substitutionsrolle durch *Leistungsgleichheit* mit einer menschlichen Tätigkeit oder Fähigkeit versehen ist.

Das Wort „Leistung“ hat, wie übrigens auch „Arbeit“ die zweifache Bedeutung, daß sowohl der Vorgang als auch das Resultat gemeint sein kann. (Zum Beispiel: Die Projektgruppe Robotik hat ihre Arbeit aufgenommen. In der Europäischen Akademie sind die Arbeiten des Künstlers X ausgestellt.) Leistungsgleich als Kriterium für eine maschinelle oder

technische Substitution soll *hier nicht die Prozedur*, sondern *nur das Ergebnis* betreffen. Und dies hat gute technikphilosophische Gründe:

In konsequenter Anwendung der aristotelischen Unterscheidung von natürlich und technisch ist zu bemerken, daß schon von archaischen Zeiten an *Technik* immer dort, wo sie menschliche Tätigkeiten substituiert, *prozedural anders* verfährt *als die Natur*. Das klassische Beispiel ist das (auf einer Achse frei drehbare) Rad: „Die Natur“, d. h. hier, Organismen können keine Räder ausbilden, weil sie wegen der völligen Trennung von der Achse im Rahmen des Stoffwechsels nicht versorgt werden könnten. So hat der Mensch schon früh, nämlich mit der Erfindung von *Rad und Kurbel*, die Möglichkeit geschaffen, natürliche Bewegungen von Gliedmaßen eines Lebewesens in kontinuierliche Drehbewegung umzuwandeln. Sieht man etwa ein Ruderboot neben einem Tretboot, bei dem über Pedale eine Schiffsschraube angetrieben wird, hat man den Gegensatz der organismischen und der maschinellen Bewegung als Beispiel der These zu verstehen, daß „die Technik“ anders verfährt als „die Natur“. Betrachtet man zusätzlich andere mechanische Leistungen des Menschen und nimmt zur Muskelkraft von Mensch und Tier natürliche Kraftquellen von Wind und Wasser sowie die technischen der Wärmekraftmaschinen hinzu, so lassen sich „im Prinzip leistungsgleiche“ Zielvorgaben für Maschinen, die den Menschen substituieren, mit der bekannt höheren Effizienz bauen, die auf der nicht-natürlichen, technischen Vorgehensweise beruht – man vergleiche etwas das Graben eines Menschen mit einer Schaufel mit den Leistungen eines Löffelbaggers.

Im Bereich der *Informationsverarbeitung* liegen die Verhältnisse genauso: Technische Substitute etwa für einen menschlichen Rechner (Kopfrechner mit Papier und Bleistift) verfahren anders, wie schon der schlichte Rechenschieber bei der Multiplikation oder der mechanische und elektronische Rechner bei den vier Grundrechnungsarten veranschaulichen. Nur die *Zwecke* der rechen- oder informationsverarbeitenden Maschinen bis zu anspruchsvollen Leistungen etwa der Übersetzung von Texten in andere Sprachen *definieren den Roboter*, nicht aber die Prozedur: Wie Rechner richtige Ergebnisse liefern sollen, sollen Übersetzungscomputer richtige Übersetzungen produzieren.

2) *Natürliche Konstruktionen und handelnde Menschen*

Leistungsgleichheit als Ergebnisgleichheit bei der Ersetzung des Menschen durch Maschinen suggeriert – und dafür gibt es in der laufenden Robotik-Diskussion mit den Aspekten des embodiment, der situatedness und des bottom-up-Vorgehens (Th. Christaller, 1998) gute Beispiele – den Versuch der „natürlichen Konstruktionen“. Hier ist wieder eine naturalistische Falle zu vermeiden:

Wer mit „*natürlicher Konstruktion*“ den Versuch meint, Natur *technisch zu imitieren* (wie die Struktur von Insektenflügel durch leichte Flächentragwerke wie dem Olympiadach in München, die Struktur von Grashalmen durch die Betonkonstruktion von Fernsehtürmen oder die autopoietische Erzeugung von Ameisenhaufen durch Verfahren der Städteplanung), der übersieht leicht, daß man die Natur *erst kennen* muß, bevor man sie technisch imitieren kann. Kurz, nur erkannte, also beschriebene und funktional erklärte Natur kann Vorbild für technische Konstruktionen sein. Naturalistisch wird dabei in der Regel übersehen, daß gerade das hier erforderliche Erkennen der Natur nicht anders gelingt als durch ihre technische Simulation. Das heißt aber, daß Naturerkenntnis und natürliche Konstruktionen sich denselben menschlichen Bemühungen nach denselben Kriterien von Erfolg und Mißerfolg verdanken. Mit anderen Worten, die *methodische Reihenfolge* von Schritten, Natur technisch zu substituieren, endet nicht, sondern beginnt mit der Entwicklung technischer Modelle für natürliche Systeme – und verlangt hierbei methodisch primär eine *Bestimmung des Explanandums*, bevor überhaupt ein Explanans gesucht und gefunden werden kann. Für die Ersetzbarkeit des Menschen durch Roboter heißt dies, daß primär das *Explanandum menschlicher Leistungen* bestimmt sein muß, bevor die Technikwissenschaften daran gehen können, unter der Zwecksetzung „Leistungsgleichheit“ den prozedural zur Natur verschiedenen technischen Weg zu suchen.

Dies war übrigens, was uns heute wegen unserer hohen technischen Zivilisation kaum mehr als Problem oder Aufgabe erkennbar ist, schon bei der Beschreibung des menschlichen Skeletts und seiner Muskulatur in den Unterscheidungen der klassischen Mechanik der Fall. Man

brauchte sozusagen erst den Blick auf den Bewegungsapparat des menschlichen Organismus, der in einem dreidimensionalen, kartesischen Koordinatensystem die Freiheitsgrade einzelner Gelenke, Aktionsradien, Hebelverhältnisse für den Ansatz von Sehnen am Knochen usw. lieferte. Die moderne Konstruktion etwa von Industrierobotern in der Automobilproduktion zeigt, daß dieser Bereich mit großer Perfektion beherrscht wird, und auch hier prozedural der technische Weg ein anderer ist als der natürliche, womit, wie oben erwähnt, eine hohe Leistungssteigerung in Geschwindigkeit, Genauigkeit und Effizienz gegenüber dem handwerklichen Autobau erreicht worden ist.

Selbstverständlich ist von besonderem Interesse der „*intelligente*“ oder „*autonome*“ *Roboter*, wie sowohl die Bestimmungen des Robotik-Projekts als auch die allgemeine Robotik-Debatte zeigen. Dabei wende ich mich weniger den Steuerungs- und Regelungsprozessen als vielmehr der Ersetzbarkeit derjenigen menschlichen Leistungen zu, die man beim Menschen intelligent und autonom nennt. Und der Kürze halber wähle ich dafür die *Kompetenz des Menschen zu (wort-)sprachlicher Kommunikation* aus.

Diese Auswahl ist dadurch begründet, daß bei allen erstaunlichen Erfolgen (und für mich weniger erstaunlichen Mißerfolgen) einer informationstheoretischen Verarbeitung von Wortsprache durch Computer wieder die stillschweigenden Prämissen des Naturalismus durchschlagen.

Unter Fachleuten ist es eine Trivialität, daß *Computer syntaktische Maschinen* sind und alle „intelligenten“, „autonomen“ oder sonstigen „kognitiven“ Leistungen von Robotern allein und ausschließlich syntaktischer oder struktureller Art sind. Inwiefern hier ein Verständnis von „Information“ durch ein formalistisches Mathematik- und ein empiristisches Naturwissenschaftsverständnis geprägt sind, kann hier nicht diskutiert werden. Ich muß dazu auf meine einschlägigen Arbeiten verweisen (vgl. Literaturangaben am Schluß des Texts).

Für das wortsprachliche Kommunizieren zwischen Menschen ist jedoch *Bedeutung und Geltung von Rede* von grundlegender Wichtigkeit.

Der Normalfall eines menschlichen Zwiegesprächs ist dadurch charakterisiert, daß beide Parteien sowohl verstanden als auch anerkannt werden wollen – sonst würde man seinem Gegenüber erst gar nichts sagen. Um das philosophisch gefährliche, leicht in Mentalismen abführende Wort „Wollen“ zu vermeiden, greife ich auf die bewährten Unterscheidungen der Handlungstheorie zurück und spreche davon, daß im Zwiegespräch beide Parteien die *Zwecke des Verstanden- und Anerkanntwerdens verfolgen*. Reden ist Handeln insofern, als sich Handeln vom bloßen Verhalten vor allem dadurch unterscheidet, daß der Agent authentisch über *Gelingen und Mißlingen* seiner Handlungen im Sinne des *Erreichens oder Verfehlens seiner Zwecke* befindet. Mit anderen Worten, zur kommunikativen Kompetenz des Menschen gehört die authentische Entscheidung, ob die eigenen Redehandlungen ihren Zweck erreichen, beim Gegenüber (hinreichend) verstanden und anerkannt werden.

(Von Anerkennung bzw. Geltung ist hier in dem allgemeinen Sinne die Rede, daß sie alle Typen üblicher Alltagssprache abdecken, vom Auffordern über das Fragen zum Behaupten und zu den performativen Sprechakten wie Danken, Bitten, Versprechen, Gratulieren usw. Damit sind zugleich verschiedene Formen der Anerkennung oder Geltung gemeint, bei der Aufforderung die Befolgung oder eine andere sozial eingübte Reaktion, d. h. (Wider-)Handlung, wie eine Verweigerung; bei der Frage die Beantwortung oder z. B. die Antwort, man könne die Frage nicht beantworten; bei der Behauptung die Zustimmung; bei den performativen Sprechakten das Akzeptieren des Dankes, der Bitte, des Versprechens usw.)

Das allgemeine Problem der Ersetzbarkeit menschlicher Leistungen spitzt sich im Bereich der kommunikativen Kompetenz auf die Frage zu, ob menschliche Kommunikations-Handlungen (im emphatischen, handlungstheoretischen Sinne) durch Roboter leistungsgleich ersetzt werden können. Auf dem Weg zu einer Antwort werden zwei wichtige handlungstheoretische Unterscheidungen benötigt, die dann in einem zweiten Schritt auf die speziellen kommunikativen oder Sprechhandlungen anzuwenden sind.

Handlungstheoretisch unterscheidet man zwischen *Schema* und *Aktualisierung*, also etwa dem Handlungsschema, ein Streichholz anzuzünden, und der Durchführung, die sich einmalig auf ein bestimmtes Streichholz (und selbstverständlich auf einen bestimmten Agenten) bezieht. Die zweite Unterscheidung betrifft die *Handlungsbeschreibung* versus den *Handlungsvollzug*. In einem geschriebenen oder gesprochenen Text kann selbstverständlich für das Handlungsschema „Streichholz anzünden“ nur die Beschreibung vorkommen, weil der *Vollzug eine nichtsprachliche Handlung* ist. Dieses Dilemma, über Handlungsvollzüge nur in der Form von Beschreibungen verfügen zu können, ist für einen schriftlichen Text prinzipiell unlösbar. Für den mündlichen Vortrag ist das Dilemma, wenn auch in akademischen Vorträgen unüblich, lösbar durch Unterbrechung des Redeflusses und das Entzünden eines Streichholzes durch den Vortragenden (im Vortrag: Ich zünde schweigend ein Streichholz an).

Die Pointe und die Tragweite dieser zweiten Unterscheidung liegt darin, daß durch Heraustreten aus der Sprache und *Eintritt in eine sprachfreie Poiesis* (handwerklich herstellendes Handeln) das Rückfragen in argumentativen Begründungs- und Rechtfertigungsketten von der Ebene des Sprechens in die Ebene der sprachfreien Wirklichkeit überführt und damit zu einem Abschluß gebracht werden kann. Ich möchte dies an einem simplen Dialog vorführen, um die Rolle von *Bedeutung und Geltung* sprachlicher Äußerungen *als Gegenstand des Gelingens menschlicher Kommunikation* zu charakterisieren.

Eine Person A fragt an einer bestimmten Straßenkreuzung eine Person B nach dem Weg zum Bahnhof, fordert also B zur Wegbeschreibung auf. (Ich wähle einen allein aus Aufforderungen bestehenden Dialog, um zunächst das Wahrheitsproblem für Behauptungen oder Beschreibungen zu umgehen.) Aufforderungen werden üblicherweise nicht als wahr oder falsch bezeichnet, sind aber selbstverständlich ebenfalls mit einem Anspruch auf Anerkennung durch die andere Partei verknüpft. B möge seinerseits in Form eines reinen Aufforderungssatzes antworten, welchen Weg A zu nehmen hat. Befolgt A diese Anweisung und gelangt auf diese Weise zum Bahnhof, ist sie mit ihrer ursprünglichen Auffor-

derung an B, eine Wegbeschreibung zu geben, *kommunikativ erfolgreich* gewesen. Und der propositionale Gehalt der Aufforderung von B ist eine *gültige* Wegbeschreibung, die sich, als Aussage formuliert, genau deshalb als wahr bezeichnen läßt, weil sie *zum erfolgreichen sprachfreien Handeln* der Person A führt, die von der Straßenkreuzung des Dialogs zum Bahnhof geht.

Nun erwarte ich, daß so gut wie jeder KI- oder Robotikexperte überrascht einwendet, daß diese simplen Leistungen heute schon oder fast schon durch sprachgesteuerte Automaten erbracht werden könnten, oder daß zumindest keinerlei Gründe zu sehen seien, warum die Leistungen der Personen A und B nicht leistungsgleich durch Roboter substituierbar wären.

Da es mir nicht im geringsten um empirische Hypothesen oder um Prognosen der künftigen Entwicklung der Robotik geht, schenke ich den Robotik-Experten, man habe Androiden gebaut, die sich „genau so verhalten“ wie die beiden Personen A und B; und mancher Experte mag dabei noch an den Turing-Test denken, in dem der Beobachter die beiden Personen A und B nicht von den Androiden A' und B' in ihren Leistungen unterscheiden können soll. Aber auch dieses schon in den science-fiction-Bereich vordringende Zugeständnis an die Experten würde nicht die leistungsgleiche Substituierbarkeit der kommunikativen Kompetenz bedeuten. Warum?

3) Bedeutung und Geltung versus Unhintergebarkeit des Handlungsvollzuges

Leistungsgleichheit im Sinne des Ergebnisses, nicht der Prozedur, sollen für die fragliche These entscheidend sein. Wie immer die fiktiven Roboter A' und B' (in einer Art modifiziertem Turing-Test) strukturell und funktionell gebaut sein mögen, ein menschlicher Beobachter möge im Aufforderungsdialo und in den beiden Befolgungen der beiden Aufforderungen keinen Unterschied zwischen A und A' bzw. B und B' feststellen. Der Handlungstheoretiker oder Philosoph, der auf die *Zwecksetzungsautonomie der Personen A und B* verwiese (A könnte auch zum

Rathaus wollen, und B absichtlich einen Irrweg empfehlen) sowie auf die *Authentizität der Beurteilung des Handlungserfolges* durch die menschlichen Agenten, schiene dann etwas Zusätzliches, Geistiges zu behaupten. Und der Naturalist würde ihm entgegen: Wozu?

Und genau mit diesem Zwiegespräch, also der *Kommunikation zwischen dem philosophischen Handlungstheoretiker und dem Naturalisten*, sehe ich meine These begründet: auch der Naturalist hat keine andere Wahl, als für seine Rede Bedeutungs- und Geltungsansprüche zu erheben, und zwar *durch Vollzug*. Jede Hoffnung, solche Vollzüge zum Explanandum zu machen und (naturalistisch, also mit natur- und technikwissenschaftlichen Mitteln) eine Kausalerklärung dafür anbieten zu wollen, wäre selbst wieder nur durch Handlungsvollzüge von Menschen, also hier der KI- oder Robotikexperten möglich. Der Handlungsvollzug ist bezüglich Handlungsbeschreibung unhintergebar.

Hier zeigt sich das gesamte Dilemma des Naturalismus, das oben an den Beispielen der Meßgeräte bzw. des Taschenrechners erläutert wurde: Ein konsequenter Naturalismus verliert nicht nur die wahr-falsch-Unterscheidung auf der Objektseite, sondern auch auf der Subjektseite, d. h., der Naturalist kann keine Geltungsansprüche für seine Beschreibungen und Erklärungen erheben, weil diese selbst kein Gegenstand naturalistischer Erklärung sind. Naturalisten haben nur die Wahl zwischen selbst-widersprüchlichen (und damit ungültigen) Äußerungen oder einem sprachlosen („natürlichen“) Dahinvegetieren ohne faktisches Erheben von Geltungsansprüchen (und Analoges gilt für Ansprüche an die semantische Bedeutung eigener Rede).

Wollte sich der Naturalist, der ja nicht nur Empirist, sondern auch Formalist ist, retten, könnte er noch auf einen Iterationstrick verfallen: Der soeben erhobene Einwand läuft ja, auf den modifizierten Turing-Test angewandt, auf die These hinaus, daß die Ununterscheidbarkeit von Mensch (Person 1) und Maschine und damit die Leistungsgleichheit in kommunikativer Kompetenz für die technische Substitution des Menschen vom Tester (Person 2) bestätigt wird, aber dazu wird ja eine dritte Person sozusagen als Beobachter des gesamten Turing-Tests benötigt

– sonst könnten die Bedingungen des Turing-Tests selbst semantisch nicht verstehbar und sinnvoll formuliert werden. Also möge ein weiterer Beobachter (Person 4) vorschlagen, den Turing-Beobachter (Person 3) selbst wieder mit einer Turing-Maschine als leistungsgleich anzunehmen und damit die spezifisch menschlichen kommunikativen Kompetenzen vom Beobachter auf einen Metabeobachter, einen Meta-Metabeobachter usw. zu verschieben. Ist diese Iteration unendlich, kann sie von keinem menschlichen Beobachter vollzogen werden. Ist sie endlich, ist ihr Ziel verfehlt.

4) *Schluß*

Ohne die kommunikative Kompetenz des Menschen, in Sprache *durch Vollzug Ansprüche auf Verstehbarkeit und Anerkennung zu erheben*, und die dabei ihm zuwachsende Leistung eines authentischen Beurteilens von Erfolg und Mißerfolg wiederum durch Vollzug zu erbringen, sind Bedingungen der Möglichkeit jeder Natur- und Technikwissenschaft, jeder KI-Forschung und Robotik. Deshalb ist aus erkenntnistheoretischen Gründen die Ersetzbarkeit des Menschen hinsichtlich seiner kommunikativen Kompetenz durch Roboter nicht in dem Sinne möglich, in dem, d. h. mit den Mitteln diese Wissenschaften selbst das Ziel dieser Ersetzung verfolgen. Oder kurz, das Ziel, die Leistungen sprechender Menschen durch Roboter (leistungsgleich) zu ersetzen, mit anderen Menschen sprachlich zu kommunizieren, steht mit sich selbst im Widerspruch. Es verhält sich damit wie mit dem Ziel, einen dreieckigen Kreis zu zeichnen.

Für das Robotik-Projekt ergibt sich daraus die Forderung, mit erkenntnis- und wissenschaftstheoretischen Mitteln eine Klärung herbeizuführen, was denn die beiden im obigen Beispiel fingierten Roboter A' und B' tatsächlich leisten, wenn sie so zu agieren scheinen wie die Personen A und B.

Literaturhinweise:

Christaller, Thomas: Mit dem Roboter der Natur auf der Spur, in: Spektrum der Wissenschaft, Dossier 4/1998, S. 104-108.

Janich, Peter: Informationsbegriff und methodisch-kulturalistische Philosophie, in: Ethik und Sozialwissenschaften 9 (1998) Heft 2, S. 169-182.

Janich, Peter: Argumente versus Standpunkte. Replik, in: Ethik und Sozialwissenschaften, a.a.O., S. 253-268.

Janich, Peter: Grenzen der Naturwissenschaft, München 1992.

Janich, Peter: Das Maß der Dinge. Protophysik von Raum, Zeit und Materie, Frankfurt 1997.

Schraft, Dieter u. Müller, Erik: Roboter gestern, heute, morgen, in: Spektrum der Wissenschaft, Dossier 4/1998, S. 6-11.

Kognitive Robotik: Ein neues Anwendungsgebiet für die philosophische Ethik?

Thomas Metzinger

1. Zur Einleitung: zwei historische Bezugspunkte

Ganz am Anfang seiner im Jahre 1632 veröffentlichten Abhandlung über den Menschen schreibt René Descartes:

Ich stelle mir einmal vor, daß der Körper nichts anderes sei als eine Statue oder Maschine aus Erde, die Gott gänzlich in der Absicht formt, sie uns so ähnlich wie möglich zu machen, und zwar derart, daß er ihr nicht nur äußerlich die Farbe und die Gestalt aller unserer Glieder gibt, sondern auch in ihr Inneres alle jene Teile legt, die notwendig sind, um sie laufen, essen, atmen, kurz all unsere Funktionen nachahmen zu lassen, von denen man sich vorstellen könnte, daß sie aus der Materie ihren Ursprung nehmen und lediglich von der Disposition der Organe abhängen.

Wir sehen Uhren, kunstvolle Wasserspiele, Mühlen und andere ähnliche Maschinen, die obwohl sie nur von Menschenhand hergestellt wurden, nicht der Kraft entbehren, sich aus sich selbst auf ganz verschiedene Weisen zu bewegen. Und wie mir scheint, könnte ich mir von einer Maschine, die - wie ich einmal annehme - aus der Hand Gottes angefertigt sein soll, nicht so viele Bewegungsarten vorstellen, noch ihr so viel kunstvolle Bildung zuschreiben, daß man sich nicht vorstellen könnte, daß sie nicht noch mehr davon besitzen kann. (DESCARTES [1632]1969:44; übersetzt von K.E. Rothschuh nach der ersten französischen Ausgabe von 1664)

Für Descartes waren Tiere Automaten, und auch der Körper des Menschen war als bloße Maschine zu denken, wenn auch mit ihm noch eine gänzlich andersartige Substanz – die Seele – verknüpft war. Die Grundidee des erkenntnistheoretischen Mechanismus, die Welt (und vielleicht auch uns selbst) vollständig verstehen zu können, indem wir ihre Teile und die zwischen Ihnen bestehenden Kausalbeziehungen analysieren, hatte bereits begonnen, Fuß zu fassen. Ein gutes Jahrhundert später, 1748 erschien in Leiden La Mettries berühmter Skandalerfolg *L’homme machine*, „Die Maschine Mensch“. Dort heißt es:

Ich glaube, Descartes wäre ein in jeder Hinsicht achtungswürdiger Mann, wenn er – geboren in einem Jahrhundert, das er nicht hätte aufklären müssen – den Wert von Erfahrung und Beobachtung erkannt hätte, sowie die Gefahr, sich von ihnen zu entfernen. ... Es ist wahr, daß dieser berühmte Philosoph sich vielfach geirrt

hat, und niemand streitet das ab. Aber schließlich hat er die animalische Natur erkannt; er hat als erster überzeugend bewiesen, daß die Tiere bloße Maschinen sind. Nun, nach einer Entdeckung von dieser Wichtigkeit und die so viel Scharfsinn voraussetzt, wäre es undankbar, ihm all seine Irrtümer nicht nachzusehen! ... Denn schließlich – soviel er auch die Unterscheidung der beiden Substanzen preist – handelt es sich doch offensichtlich nur um einen Kunstgriff, eine stilistische List, um die Theologen ein Gift schlucken zu lassen, das unter einer dunklen Analogie verborgen ist, die jedem auffällt und die nur sie nicht sehen. Denn sie ist es, diese auffällige Analogie ist es, die alle Gelehrten und wirklich Kundigen zwingt zuzugeben, daß jene stolzen und eitlen Wesen, die sich mehr durch ihren Hochmut als durch die Bezeichnung Mensch auszeichnen, im Grunde – wie sehr sie sich auch erheben möchten – nur Tiere und aufrecht kriechende Maschinen sind. (LAMETTRIE [1748] 1990:125)

2. Philosophische Aspekte der Robotik: Zwei Klassen von Fragen

Auch heute, dreieinhalb Jahrhunderte später, wirkt immer noch Descartes Gift. Eines seiner Symptome ist die Irrationalität und der Mangel an inhaltlicher Differenzierung in öffentlichen Debatten. Ich denke, daß es in diesem Zusammenhang für unsere Arbeitsgruppe wichtig ist, mindestens zwei Klassen von Fragen auseinanderzuhalten.

Die erste Art von Frage lautet: Stellen sich durch das Aufkommen der kognitiven Robotik wirklich genuin neue Fragen für die philosophische Ethik? Das heißt: Konfrontiert sie die akademische Philosophie mit neuen theoretischen Problemen? Die zweite Art von Frage lautet: Woran liegt es, daß die kognitive Robotik für viele Menschen eine so grundlegende Provokation ihres Selbstverständnisses und ihrer moralischen Intuitionen darstellt? Woher kommt das unterschwellige Unbehagen und die „Sensationspresse“; wie kommt es, daß bereits die allerersten Anfänge dieser neuen Technologie von einem so massiven öffentlichen Interesse und insbesondere von sich teilweise am Rande der Seriosität bewegendenden Meinungsäußerungen begleitet werden?

Meine Antwort auf die erste Frage ist „deflationär“: Nein – alle der konkreten Probleme, mit denen wir uns bis jetzt konfrontiert sehen, können im Prinzip mit den vorhandenen begrifflichen Mitteln analysiert und ei-

ner rationalen Diskussion zugeführt werden. Zumindest gilt das für die *ethische* Dimension der Problematik. In der Philosophie des Geistes mag es dagegen durchaus sein, daß die Entwicklung autonomer Systeme – oder z.B. die Lösung des „*Symbol-grounding-Problems*“ – uns auf genuin neues theoretisches Terrain führt. Trotzdem: Es ist wichtig, daß die Entwicklungen in der kognitiven Robotik, der *Artificial-Life*-Forschung usw. auch mit Blick auf die angewandte Ethik von der akademischen Philosophie kompetent und kritisch begleitet werden. Man kann aber nicht sagen, daß wir es hier mit einem *fundamental* neuen Problemhorizont für die Moralphilosophie zu tun haben.

Herr Decker hat uns in seiner Projektskizze und seinen bereits vorliegenden Überlegungen zur Ersetzbarkeit des Menschen bereits eine gehaltvolle Diskussionsgrundlage geliefert. Ich bin sicher, daß wir uns mit vielen Detailfragen noch ausführlich beschäftigen werden und genau dies sollte auch den Schwerpunkt der Arbeit bilden. Ich denke aber, daß es *zu Anfang* unserer Arbeitsgruppe wichtig ist, einen möglichst umfassenden Blickwinkel einzunehmen und zunächst Klarheit darüber zu erzeugen, *warum* wir die Entwicklungen in der Robotik überhaupt für so wichtig halten. Was genau wollen wir eigentlich wissen? Und warum erscheint vielen die theoretische Hintergrundproblematik als so virulent? *Ist* sie wirklich virulent?

Im nächsten Abschnitt möchte ich deshalb kurz einige Aspekte meiner Antwort auf die *zweite* Klasse von Fragen skizzieren: Woher kommt das Unbehagen an der Robotik und woher das moralisch provozierende Element an dieser neuen Technologie? Ein Teil dieser Antwort ist, daß es die *implizite Theorie des Geistes* ist, die vielen Forschungsprogrammen in der Robotik zugrundeliegt und die wir als kränkend empfinden. Das liegt unter anderem daran, daß diese Theorie des Geistes nicht nur ethische Anwendungsprobleme – z.B. bei der Auszeichnung des Gegenstandsbereichs moralischer Objekte – erzeugt, sondern auch unseren klassischen Begriff moralischer *Subjektivität* überhaupt in Frage stellt, z.B. was die Willensfreiheit und die Autonomie des handelnden Subjekts angeht. Es liegt aber auch daran, daß sie uns mit unserer eigenen Sterblichkeit konfrontiert und daran, daß sie uns zwingt, theo-

retische Entwicklungen zur Kenntnis zu nehmen, die besonders in den klassischen Geisteswissenschaften seit längerer Zeit systematisch verdrängt worden sind und nun in Gestalt konkreter, technischer Artefakte bereits einen kulturellen Einfluß zu entfalten beginnen. Es ist der von den Geistes- und Kulturwissenschaften verdrängte Erkenntnisfortschritt in den Neuro- und Kognitionswissenschaften, der sich nun auf dem Umweg über die technische Modellierung intelligenten Verhaltens wieder in ihr Blickfeld schiebt.

3. Die Kränkung durch den Funktionalismus

Die Debatte um die kognitive Robotik löst genau deshalb so viele Emotionen und irrationale Reaktionen aus, weil sie uns zwingt, die Aufmerksamkeit auf eine theoretische Entwicklung zu richten, die sich bereits seit einiger Zeit mit einer zunehmenden Eigendynamik und durch den Synergismus verschiedenster, teilweise sehr junger akademischer Disziplinen entfaltet. Diese Entwicklung ist sowohl auf der Ebene der philosophischen Metatheorie als auch im allgemeinen gesellschaftlich-kulturellen Diskurs verdrängt worden. Durch die explosionsartige Entwicklung des Wissens in den Neuro- und Kognitionswissenschaften ergibt sich jetzt in einem atemberaubenden Tempo ein neues Bild vom Menschen, das vielen klassischen Anthropologien, und damit auch vielen klassischen Theorien von Autonomie und Authentizität zu widersprechen droht. Es sind nicht so sehr moralphilosophische Intuitionen, die durch die kognitive Robotik in Frage gestellt werden, es ist unser anthropologisches, unser soziokulturelles *Selbstverständnis*, das durch die neuen Möglichkeiten des technologischen Zugriffs auf das menschliche Nervensystems, durch das Aufkommen biomorpher Rechnerarchitekturen, sowie verkörperter und situierter Formen von künstlicher Intelligenz provoziert wird. Wenn Sie mir eine metaphorische Beschreibung des Zusammenhangs erlauben: Der verdrängte Erkenntnisfortschritt kehrt in Form unserer eigenen Technologien („technischer Dämonen“) zu uns zurück. Deshalb wäre es in dieser Situation auch falsch, die Debatte wiederum einzuengen auf Detailprobleme für die angewandte Ethik. Auch das könnte man - um im Bild der psychoana-

lytischen Metapher zu bleiben - als eine Form des Widerstands und der Verdrängung interpretieren. Ich behaupte: Es geht um mehr, und die akademische Philosophie ist auf mehr als einer Ebene gefordert.

Meine erste These ist, daß die Robotik viele Geisteswissenschaftler beunruhigt, weil sie eine implizite Lösung des Leib-Seele-Problems beinhaltet, die kränkend ist. Werfen Sie dazu kurz einen Blick auf diese Aufstellung der seit dem Zweiten Weltkrieg bzw. nach dem Zusammenbruch von Gilbert Ryle's logischem Behaviorismus von Philosophen entwickelten Modelle der Beziehung zwischen Geist und Körper. Keine Angst – ich will jetzt weder scholastisch werden, noch Sie mit den analytischen Einzelheiten der modernen Diskussion des Leib-Seele-Problems langweilen. Vielmehr möchte ich Sie auf ein einfaches, aber interessantes philosophiehistorisches Detail hinweisen. Wie Sie sehen, gibt es etwa neun oder zehn verschiedene Positionen, die wiederum in vier Typen von Theorien zerfallen: Reduktive Formen des Materialismus, Versuche, einen nicht-reduktiven ontologischen Monismus zu entwickeln, das klassische dualistische Modell der Wechselwirkung zwischen Leib und Seele und die verschiedenen Formen des Funktionalismus. Unter diesen Theorietypen gibt es aber nur eine einzige logische Variante, von der man sagen kann, daß sie eine wirklich neue philosophische Denkfigur darstellt, und das ist der Funktionalismus. Natürlich gibt es spezielle Fachdiskussionen im Hintergrund, zum Beispiel darüber, ob nicht im Grunde Aristoteles der erste Funktionalist war, sie sind aber in unserem Kontext nicht von Bedeutung.

So können Sie zum Beispiel den neutralen Monismus Herbert Feigl's oder auch den im Moment gerade so erbittert diskutierten – und verdeckt panpsychistischen – Eigenschaftsdualismus von David Chalmers mit Fug und Recht als spinozistisches Motiv betrachten. Reduktive Materialisten hat es seit Demokrit und Leukipp in der abendländischen Philosophiegeschichte so viele gegeben, daß auch die verschiedenen Formen der Identitätstheorie schwerlich als neue Gedanken zählen können. Am anderen Ende des Spektrums dagegen haben Popper und Eccles einen direkten Anschluß an Descartes klassischen Substanzdualismus so sorgfältig vermieden, daß sie sein theoretisches Kernpro-

blem dabei gleich verdreifacht haben. Die philosophische Intuition hinter dem minimalen Materialismus der *Supervenienztheorie* von Jaegwon Kim ist mindestens seit 1904 veröffentlicht, nämlich durch Thomas Huxley's darwinistischen Epiphänomenalismus: Supervenienz als psychophische Relation ist einfach Epiphänomenalismus ohne Kausalität. Der *locus classicus* für die Emergenztheorien ist dagegen das im Jahre 1925 erschienene Buch „*The Mind and it's Place in Nature*“ von C.D. Broad.

Was hat jetzt all das mit kognitiver Robotik, mit Technikfolgenabschätzung und mit kritischer Anthropologie zu tun? Die kognitive Robotik ist der erste Anfang der technischen Umsetzung der einzig wirklich neuen Leib-Seele-Theorie, nämlich des Funktionalismus. Der Funktionalismus – auf der zweiten Folie sehen Sie seine drei wichtigsten logischen Varianten – hat eine neue Antwort gegeben auf die Grundfrage der Philosophie des Geistes: Was ist das Wesen mentaler Zustände? Die Antwort lautet: *Mentale Zustände sind Mengen von Kausalrelationen*. Diese Annahme – daß mentale Zustände durch die kausale Rolle individuiert werden, die sie in einem System spielen, das man auch als informationsverarbeitendes System beschreiben kann – wurde auf seiten der Philosophie in den frühen sechziger Jahren von Hilary Putnam in einer Serie von acht kanonischen Aufsätzen entwickelt. Diese Annahme, die zwischenzeitlich auch einmal als das „Computermodell des menschlichen Geistes“ bezeichnet wurde, war in wissenschaftshistorischer Hinsicht von großer heuristischer Fruchtbarkeit, aus ihr sind eine Flut von neuen Wissenschaften entstanden: die Kognitionswissenschaft und die kognitive Neurowissenschaft, die Künstliche Intelligenz-Forschung und die Neuroinformatik, die Robotik und die *Artificial-Life*-Forschung usw. Theoretisch folgenreich war der philosophische Funktionalismus aber zweitens auch durch sein theoretisches Herzstück, das *Multirealisierbarkeits-Argument*. Sie erinnern sich: Die funktionalen Zustände eines Systems werden durch die kausale Rolle individuiert, die sie bei der Verursachung von externem Verhalten spielen, durch die sie mit dem Input des Systems und anderen seiner internen Zustände vernetzt sind. Die kausale Rolle eines Systemzustandes

wird somit aus einer Menge von Kausalrelationen gebildet. Funktionale Zustände werden auf einer über der physikalischen liegenden, abstrakteren Beschreibungsebene zugeschrieben, indem man das betreffende System einer *funktionalen Analyse* unterzieht und zunächst von den konkreten Eigenschaften absieht, durch die sie auf der physikalischen Ebene *realisiert* sind. Mentale Zustände und psychologische Eigenschaften sind darum *multirealisierbar*, d.h. sie können durch eine prinzipiell unendliche Vielzahl physikalischer Strukturen instantiiert werden. Zwei *funktional äquivalente* Zustände können in zwei verschiedenen Systemen durch ganz unterschiedliche konkrete Mechanismen implementiert werden, ähnlich wie ein und dasselbe Programm auf ganz verschiedenen Rechnern laufen kann. Sogar Engel könnten Turing-Maschinen sein.

Wenn das aber wirklich wahr sein sollte, dann wären mentale Zustände im Prinzip austauschbar, sie könnten nicht nur kopiert werden, sondern auch *repariert*, *externalisiert* oder sogar *optimiert* werden. Alles, was man der Theorie des Funktionalismus zufolge – zum Beispiel bei mentalen Defiziten – tun muß, ist die Versorgung des Systems mit *funktional äquivalenter Hardware*, die Bereitstellung von Ersatzteilen, die zumindest potentiell dieselben kausalen Eigenschaften besitzen wie die verlorenen Teile der physischen Realisierung.

Genau das macht man zum Beispiel bei der Transplantation fetalen Hirngewebes, man versucht es aber auch bei der Applikation von *Cochlea*-Implantaten oder bei der Entwicklung der künstlichen Retina und intelligenter Neuroprothesen im allgemeinen. In diesem Sinne ist die Neurotechnologie ein medizinischer Zweig der Robotik: Bei vielen der von der medizinischen *Neurotechnologie* anvisierten Störungsbilder geht es nicht mehr nur um die Transplantation von Körperteilen. Zum Krankheitsbild von *Morbus Parkinson* gehören zum Beispiel auch depressive Verstimmungen, die in manchen Fällen der Manifestation der motorischen Symptome sogar vorausseilen, Demenzen und Bradyphrenien. Auch solche mentalen Eigenschaften will die Neurotechnologie heilen oder wiederherstellen, und die theoretische Hintergrundannahme ist dabei, daß auch sie prinzipiell durch funktional äqui-

valente Module austauschbar – also: *reinstantiierbar* – sind. Diese technologische Austauschbarkeitsannahme stammt, das ist mein Punkt, eigentlich aus der Philosophie des Geistes. Wenn man diese Austauschbarkeitsannahme jetzt nicht mehr auf subsystemische Teilregionen wie z.B. bestimmte Hirnareale anwendet, sondern auf *Systeme als Ganze*, dann sind wir wieder bei der Robotik und der Frage nach der Ersetzbarkeit des Menschen angelangt.

Ich behaupte: Es ist die Austauschbarkeitsannahme, es ist Hilary Putnam's klassisches Multirealisierbarkeitsargument, das uns emotional kränkt. Die Austauschbarkeitsannahme kollidiert nämlich mit dem, was ich die „Authentizitätsannahme der Alltagspsychologie“ nennen möchte: Meinungen, Wünsche, Gefühle und andere mentale Zustände sind immer meine eigenen, weil sie an die einzigartige, individuelle Erste-Person-Perspektive eines erlebenden Subjekts gebunden sind. Diese Perspektive ist unaustauschbar, weil sie die subjektive Perspektive ist, die ich selbst auf die Welt einnehme. Subjekte und individuelle Erlebnisperspektive sind *nicht* multirealisierbar.

Genau wie der funktionalistischen Neurowissenschaft liegt auch der Alltagspsychologie eine Theorie des Menschen und eine Theorie des Geistes zugrunde, die – auch wenn sie sich in der Gegenwart zunehmend als falsch erweist – in lebensweltlichen Bezügen seit Jahrhunderten hervorragend funktioniert. Sie ist die Grundlage nicht nur der meisten von uns geteilten moralischen Intuitionen, sondern auch ein Fundament unserer Kultur. Die Alltagsphänomenologie individuiert mentale Zustände primär durch ihren phänomenalen Gehalt, durch die Art und Weise, wie diese Zustände sich aus der Innenperspektive des erlebenden Subjekts *anfühlen*. Der Funktionalist dagegen individuiert mentale Zustände von außen, indem er ihnen kausale Rollen zuschreibt. Die Erste-Person-Perspektive ist durch die Entwicklung in der Philosophie des Geistes, aber auch in den empirischen Wissenschaften nicht nur in Bedrängnis, sondern auch in Konflikt geraten mit der Dritte-Person-Perspektive – und *diese* Tatsache spüren wir jetzt als das Unbehagen an unseren eigenen erweiterten Handlungsmöglichkeiten. Das ungute Gefühl bei der kognitiven Robotik ist eine Spätfolge von dem, was

LaMettrie als das Gift des Descartes bezeichnet hat. Es ist verborgen in der Dunkelheit von Maschinenmetaphern, die sich uns auch heute wie von selbst immer wieder aufdrängen. Die Relation zwischen der Erster-Person-Perspektive und der Dritte-Person-Perspektive ist sowohl die Essenz der theoretischen Hintergrundproblematik in der Philosophie des Geistes als auch der Kern des steigenden gesellschaftlichen Unbehagens an der kognitiven Robotik.

4. Die Provokation durch semantische Verschiebungen

Ich möchte die theoretische Dimension des Problems ganz kurz an zwei begrifflichen Unterscheidungen illustrieren, die zweierlei miteinander gemein haben: Erstens werden sie durch die kognitive Robotik in Frage gestellt und zweitens bilden sie das Fundament vieler normativer, anthropologischer und metaphysischer Grundannahmen unserer Kultur. Die Unterscheidungen sind die zwischen „physisch“ und „mental“ und zwischen „bloß physikalisch“ und „lebendig“. Von einem logischen Standpunkt aus sind – das sagen unsere kulturell tief verankerte Alltagsphysik, unsere Alltagspsychologie und unsere Alltagsbiologie – diese Unterscheidungen *exklusiv* und *erschöpfend*.

Ein Ding oder ein Zustand ist *entweder* geistig oder körperlich, und etwas anderes als geistige und körperliche Phänomene gibt es nicht. Diese Unterscheidung erzeugt das klassische Leib-Seele-Problem. Der Funktionalismus, der auch der Robotik zugrundeliegt, bietet nun eine Lösung für dieses Problem an, indem er einen *nicht-ontologischen Dualismus* vertritt: Geistige Zustände sind insofern nicht-physikalisch, als sie über ihre kausale Rolle individuiert werden. Weil aber jeder funktionale Zustand (durch „token-identities“, d.h. partikulare Identitäten mit physikalischen Zuständen) verschiedene Realisierungen besitzen kann, gibt es tatsächlich Zustände, die sowohl physikalisch *als auch* geistig sind, nämlich realisierte funktionale Zustände. Roboter sind solche realisierten Mengen von funktionalen Zuständen. Was Artefakte angeht, sind sie mittlerweile zu sehr *komplexen* Mengen von realisierten funktionalen Zuständen geworden und ihre fortschreitende

Autonomie beunruhigt uns deshalb, weil wir unsere klassischen begrifflichen Unterscheidungen nicht mehr auf sie anwenden können. Diese semantische Verschiebung fasziniert manche Menschen, viele beunruhigt sie dagegen. Was machen wir, wenn Roboter einmal zu dem werden, was Daniel Dennett „intentionale Systeme zweiter Stufe“ genannt hat? Was machen wir, wenn ein Roboter einmal selbst zur besten Informationsquelle bei der Vorhersage seines zukünftigen Verhaltens oder bei Fragen über den Inhalt seiner geistigen Zustände wird?

Die Forschung in der Robotik führt aber ebenso dazu, daß die semantische Exklusivität und Exhaustivität des Gegensatzes von „lebendig“ und „bloß physikalisch“ aufgehoben wird. Besonders deutlich wird dies auf theoretischer Ebene in der *Artificial-Life*-Forschung. Es geschieht aber bereits auf der Ebene der technischen Modellierung selbst. Ein erstes Beispiel sind hier die funktionale Kopplung zwischen Nervenzellen und technischen Systemen, in Form von Neuron-Silizium-Kopplungen in Zellkulturen. In Deutschland tut dies zum Beispiel Peter Fromherz am Max-Planck-Institut für Biochemie in Martinsried. Ich habe Ihnen jedoch heute ein anderes Beispiel dafür mitgebracht, was es heißt, daß klassische, philosophisch tief verankerte Begriffsunterscheidungen ins Wanken geraten – und zwar genau aus einem neuen Zweig der Robotik, die „hybride Bio-Robotik“.

An der Universität von Tokio hat Herr Takeuchi die erste Roboter-Kakerlake der Welt konstruiert: Man baut aus *lebenden* Küchenschaben hybride Bio-Roboter, indem man eine elektrische Stimulationseinheit direkt an ihr Nervensystem anschließt. Die bedauernswerten Tiere können dann von außen manipuliert werden, das Projekt ist unter dem Namen „RoboRoach“ für „Robot-Cockroach“ bekannt geworden. Und hier ist die semantische Verschiebung: Eigentlich haben wir es hier gar nicht mehr mit *Tieren*, zu tun, und auch nicht mit *Maschinen*. Spätestens hier dürfte auch deutlich werden, daß die Entwicklung qualitativ neue Elemente in sich trägt und daß eine kritische begriffliche Begleitung durch die Philosophie mehr als angebracht ist. Ich persönlich finde solche Experimente zumindest geschmacklos, jedoch ist das nicht mehr als eine ästhetische Aussage. Gibt es aber ein rationales *ethisches*

Argument, das man gegen diese Art von Wissenschaft ins Feld führen kann? Oder sind solche Experimente – ganz im Gegenteil – ein legitimer und wichtiger Bestandteil eines klassischen philosophischen Projekts, des Projekts der Selbsterkenntnis nämlich?

In der Moralphilosophie ist es hilfreich, zwischen drei logischen Ebenen zu unterscheiden. In der Metaethik haben wir es mit Sätzen über andere Sätze zu tun, nämlich mit Sätzen über *normative* Sätze, über moralische Urteile. Eine normative Ethik besteht aus solchen Urteilen, aus Sätzen über *Handlungen*. Die angewandte Ethik analysiert dann bestimmte, oft *neue Typen* von Handlungen, die aus bestimmten Gründen als besonders problematisch oder interessant erscheinen: Ein Anwendungskontext entsteht immer aus einer speziellen Klasse von Handlungen. Hier haben wir es genau mit einer solchen neuen Klasse zu tun. Worum es hier geht, ist also die parallele Beurteilung sowohl von *forschendem* Handeln als auch von *technologischem* Handeln. Am Ende geht es deshalb immer auch um Wissenschaftspolitik und um Gesellschaftspolitik, und es gibt mindestens zwei weitere thematische Bereiche, in denen ein philosophischer Kommentar zur Klärung der Debatte wichtig ist: Die Sozialphilosophie und die Anthropologie. Denn wir sind immer auch mit einer Reihe von sehr allgemeinen Fragen konfrontiert, die zum Beispiel die kulturelle Einbettung des technologischen Fortschritts betreffen (*Technikfolgenabschätzung*), vor allem aber auch die Konsequenzen unseres gewandelten Bildes von uns selbst („*Anthropologiefolgenabschätzung*“). Auf diesen Punkt möchte ich abschließend kurz zu sprechen kommen.

5. Der Soziokulturelle Kontext: „Anthropologiefolgenabschätzung“

Je besser Roboter einmal funktionieren werden, desto plausibler wird für viele die der Robotik zugrundeliegende Theorie des Geistes werden. Aus philosophischer Perspektive mag dies ein schwerer Fehlschluß sein – die kulturellen Folgen einer solchen Entwicklung könnten jedoch trotzdem tiefgreifend sein. Die generellen soziokulturellen Konsequenzen der Robotik sind natürlich vielfältig und gegenwärtig noch

schwer abzuschätzen. Ich erwarte das Folgende: Unser eigenes Bild von uns selbst – in dem viele der eben erwähnten Hintergrundannahmen ihre Wurzeln haben – wird sich auf dramatische Weise verändern. Es entsteht nämlich nicht nur eine ganze Palette von neuen Problemstellungen für die angewandte Ethik, sondern auch eine neue, z.B. durch die Erkenntnisse der modernen Hirnforschung erweiterte Anthropologie: Wir bewegen uns auf ein grundlegend neues Verständnis dessen zu, was es heißt, ein *Mensch* zu sein. Wie sollen wir uns – angesichts der Erfolge in der Robotik oder auch der Neurotechnologie – die Beziehung zwischen Gehirn und Bewußtsein denken? Gibt es so etwas wie eine Seele oder einen unveränderlichen Kern der Persönlichkeit? Was genau meinen wir eigentlich, wenn wir von „Willensfreiheit“ und „Autonomie“ sprechen; was bedeutet es, daß wir für manche unserer Handlungen verantwortlich sind, für andere dagegen nicht? Muß unser Rechtssystem an die neue Faktenlage angepaßt werden? Gibt es angesichts der vielen neuen Einsichten über die objektiven Entstehungsbedingungen des subjektiven Erlebens noch berechnete Hoffnung auf ein Leben nach dem Tod?

Aus neuen Antworten auf solche klassischen philosophischen Fragen ergibt sich Schritt für Schritt auch ein neues, ein verändertes Gesamtbild des Menschen. Die Robotik gibt solche Antworten nicht durch *theoretische*, sondern durch *technische* Modelle – und ein subversiver Charakter wie LaMettrie hätte sicher seine Freude an ihren ersten Fortschritten gehabt. Die erfolgreiche und fortschreitende Integration von humanoiden Robotern in menschliche Lebenswelten könnte dazu führen, daß solche Antworten intuitiv übernommen und kulturell verankert werden, selbst wenn es im Grunde *falsche* Antworten sind. Wir erkennen uns selbst in unseren Artefakten und auf diese Weise könnte sich prinzipiell auch auf eine schleichende Weise das allgemeine Bild davon verändern, was überhaupt ein Mensch ist. Das allgemeine Bild vom Menschen wiederum ist aber eine der wichtigsten Grundlagen unserer Kultur. Seine Besonderheit besteht darin, daß es sehr subtil und doch wirksam die Art und Weise beeinflusst, wie wir im Alltag miteinander umgehen und uns selbst erleben. Deshalb wird die oben ange-

deutete Entwicklung auch gesellschaftliche Konsequenzen nach sich ziehen und schließlich unser aller Leben beeinflussen. Wenn man also am Beispiel der Robotik über die Konsequenzen der Veränderungen in unserem Selbst- und Weltbild nachzudenken beginnt, dann zeigt sich gleich am Anfang, daß wir es nicht nur mit ethischen und anthropologischen, sondern bald auch mit gesellschaftspolitischen Problemen zu tun haben werden.

Welchen Preis sind wir für den Erkenntnisfortschritt in der naturwissenschaftlichen Erforschung des menschlichen Bewußtseins zu zahlen bereit? Auf rein wirtschaftlicher Ebene scheint sich abzuzeichnen, daß Roboter in vielen Bereichen sowohl effektiver als auch billiger sind als ihre Vorläufer – die Tiere, die Sklaven und die Lohnarbeiter. Das Entlastungspotential für den Menschen ist – falls die durch Roboter erzielbaren Produktivitätsgewinne gerecht verteilt werden – enorm. Wir müssen jedoch den *Gesamtpreis*, den wir alle für den Fortschritt in der Hirnforschung, der Kognitionswissenschaft und der Künstlichen Intelligenz bezahlen, so niedrig wie möglich halten. Diesen Preis zahlen wir aber in der Hauptsache auf emotionaler und auf soziokultureller Ebene. Der gefühlsmäßige Preis besteht unter anderem in dem, was ich eben als das „Unbehagen“ bezeichnet habe: Wir werden verunsichert in vielen unserer unhinterfragten Meinungen über uns selbst. Die ersten Erfolge bei der Konstruktion autonomer Systeme zwingen uns bereits heute, neu über uns selbst nachzudenken. Der „soziokulturelle“ Preis dagegen könnte in einer Vielzahl von unerwünschten Nebenwirkungen der neuen Handlungsmöglichkeiten in Bezug auf das gesellschaftliche Zusammenleben bestehen. Man muß in diesem Zusammenhang auch sehen, daß nicht alle Roboter streng lokalisierte Wesen mit *einem* künstlichen Körper sein werden: „Softbots“, „Avatare“ und „personoide“ Serviceprogramme werden uns bald im Internet zur Seite stehen und uns zum Beispiel Rechercheaufgaben abnehmen. Roboter werden nicht nur die physikalischen, sondern zunehmen auch die *medialen* Umwelten bevölkern, in denen sich auch Menschen bewegen. Sie werden dabei auch psychologische Eigenschaften simulieren und diejenigen ihrer Benutzer zu erfassen versuchen. Viele befürchten deshalb,

daß die Robotik uns auch auf gesellschaftlicher Ebene schleichend in eine andere, in eine reduzierte soziale Wirklichkeit führen könnte. Wenn wir zum Beispiel beginnen, uns *gegenseitig* mehr und mehr als eine besondere Art von sehr komplexen Biorobotern in den Blick zu nehmen, dann wird dies automatisch zu Veränderungen im sozialen Umgang, aber vielleicht später einmal auch in so speziellen Bereichen wie dem des Rechtssystems führen. Vielleicht ist in diesem Zusammenhang der Hinweis nicht unangebracht, daß es psychiatrische Störungsbilder gibt, bei denen manche Patienten präzise das Gefühl haben, sie würden sich schrittweise in eine Maschine, eben in einen Roboter verwandeln. Solche Patienten sind aufgrund einer Störung der Selbstrepräsentation nicht mehr in der Lage, Willensakte als ihre *eigenen* Willensakte zu repräsentieren und erleben sich deshalb als fremdgesteuert und „mechanisiert“. Man nennt solche Störungsbilder auch „Depersonalisationen“ und man kann sich vorstellen, daß es auch *kulturelle* Spiegelbilder solcher individuellen Formen des Leidens an einer falschen Form der Selbstrepräsentation geben könnte.

All dies ist noch weit entfernt. Jede Form von Alarmismus oder negativem Utopismus würde die rationale Diskussion behindern, die wir nun über die normativen Aspekte der Entwicklung führen müssen. Doch es geht um mehr als nur um neue Fragestellungen für die angewandte Ethik, es geht um die *kulturelle* Umsetzung der neuen, von den Neurowissenschaften und der Informatik gelieferten Erkenntnisse selbst. Aus den Neuro- und Kognitionswissenschaften sowie der Informatik ergibt sich nämlich zwangsläufig auch ein völlig neues Bild vom Menschen und eine neue Theorie darüber, was geistige Vorgänge *überhaupt* sind. Es ist bereits jetzt deutlich abzusehen, daß diese neue Anthropologie und die mit ihr einhergehende neue Theorie des Geistes fast allen traditionellen Bildern vom Menschen und seinem inneren Leben dramatisch widersprechen wird. Meine These ist: Das allgemeine Bild vom Menschen wird sich im kommenden Jahrhundert durch die Fortschritte der Neuro-, Informations- und Kognitionswissenschaften tiefgreifender verändern als durch jede andere wissenschaftliche Revolution der Vergangenheit. Dazu ein einfaches Beispiel: Nehmen wir an,

die neuronalen und funktionalen Korrelate der wesentlichen Merkmale des Bewußtseins sind einmal erforscht. Dann wird dies von vielen dahingehend interpretiert werden, daß die Annahme, es könnte – zum Beispiel nach dem physischen Tod – bewußtes Erleben auch in der Abwesenheit dieser körperlichen Basis geben, nur noch als grob irrational beurteilt werden kann. Wenn bewußtes Erleben einmal auf der begrifflichen Ebene der Hirnforschung einer reduktiven Erklärung zugänglich werden sollte, dann würde der klassische Begriff der „Seele“ endgültig zu einem *leeren* Begriff: Noch an diesem klassischen Begriff orientierte Theorien werden dann genauso irrational erscheinen wie die Theorie, daß die Sonne sich in Wirklichkeit doch um die Erde dreht. Roboter sind sozusagen die technisch manifest gewordenen Vorboten der Möglichkeit einer solchen Entwicklung. Sie demonstrieren auf sinnlich erfahrbarer Ebene, was es bedeutet, ernst zu machen mit einem radikal empiristischen „*bottom-up-approach*“ in der Erklärung von Mentalität, Autonomie und rationalem Verhalten. Die gegenwärtig neu entstehende Anthropologie muß darum ebenfalls zum Gegenstand breit angelegter Diskussionen gemacht werden. Denn es ist nicht unwahrscheinlich, daß die wichtigen neuen Beiträge zu unserem Bild vom Menschen in einigen Aspekten – zumindest aus der Perspektive klassischer Anthropologien und besonders im subjektiven Empfinden vieler Menschen – eine Demütigung und eine Kränkung darstellen. Auch auf *diese* Entwicklung muß die Öffentlichkeit in Gestalt einer rationalen und für alle offenen Diskussion vorbereitet werden. Was wir im Grunde brauchen – auch dies ist meine These – ist eine neue Bewußtseinskultur. Diese *Bewußtseinskultur* muß auf gesamtgesellschaftlicher Ebene eine vernünftige und produktive Umsetzung der neuen Erkenntnisse und Handlungsmöglichkeiten leisten, die sich in der Zukunft mit steigender Geschwindigkeit aus der Forschungstätigkeit in den genannten Bereichen ergeben werden. Aber das ist wirklich ein anderes Thema.

Arbeitsqualität durch Komplementarität von Mensch und Technik

Gudela Grote

These 1

Jedes automatisierte System ist ein soziotechnisches System, unabhängig vom Grad der Automatisierung.

Technische Systeme sind nicht isoliert zu betrachten, sondern immer als Teil eines soziotechnischen Systems, zu dem auch die das technische System betreibenden Menschen und die formalen und informellen Strukturen und Abläufe, in denen sie arbeiten, gehören. Das technische und das soziale Teilsystem müssen miteinander gestaltet werden, um eine Optimierung des Gesamtsystems zu erreichen.

These 2

Menschen müssen auch in automatisierten Systemen die Verantwortung für das Funktionieren des Systems übernehmen.

Mit zunehmender Automatisierung nicht nur körperlicher, ausführender Arbeit, sondern auch planerischer und kontrollierender Tätigkeiten wird der Mensch mehr und mehr zum Überwacher einer Technik, die im Normalfall alle Operationen weitestgehend autonom ausführen sollte, sei es die Fertigung eines Werkstücks, das Fliegen eines Flugzeuges, die Konstruktion eines Werkzeugs, die Planung betrieblicher Produktionsabläufe oder die Prozesskontrolle in der chemischen Produktion. Nur im Störfall, d.h. bei Ausfall der Technik oder einzelner ihrer Komponenten, und in Ausnahmefällen, für die die Technik nicht ausgelegt ist, sind Eingriffe des Menschen gefordert. Der Mensch ist also als kreativer Problemlöser für diejenigen Fälle gefragt, die nicht vollständig spezifiziert und algorithmisiert sind und damit auch nicht gänzlich der Technik übergeben werden können. Damit entsteht das grundlegende Dilemma, dass dem Menschen gerade durch seinen weitestmöglichen Ausschluss aus dem Produktionsprozess die Verantwortung

für sehr kritische, aber letztlich nicht bewältigbare Aufgaben übertragen werden: Unter „normalen“ Bedingungen sind die Aufgaben des menschlichen Operateurs auf die Überwachung des technischen Systems und gelegentliche manuelle Operationen reduziert; im Störfall soll er – unter allgemein erhöhter Belastung und oftmals durch lange Überwachungsphasen ermüdet – schnell und richtig eingreifen können, wozu ihm der Überblick über den aktuellen Systemzustand, detailliertes Hintergrundwissen, Geübtheit in den geforderten Operationen etc. fehlen.

These 3

Um Verantwortung übernehmen zu können, müssen die Systeme für den Menschen kontrollierbar sein; Kontrolle beinhaltet die Durchschaubarkeit, Vorhersehbarkeit und Beeinflussbarkeit des Systems.

Damit ein Mensch in der Lage ist, Verantwortung für das (Nicht-)Funktionieren eines hochautomatisierten System zu übernehmen, sind Voraussetzungen für menschliche Kontrolle über die Technik zu schaffen. Die Förderung der Kontrolle im Hinblick auf die effiziente und sichere Prozesssteuerung und -überwachung ist nicht nur unmittelbar relevant, sondern auch mittelbar, indem erlebte Kontrolle in einer Situation eine wesentliche Voraussetzung für die möglichst stressfreie Bewältigung der in der Situation gestellten Anforderungen ist.

In Anlehnung an Kontrolltheorien in Psychologie und Systemtheorie können drei Aspekte von Kontrolle unterschieden werden:

- die Durchschaubarkeit des zu kontrollierenden Systems,
- die Vorhersehbarkeit seines Verhaltens und
- die Möglichkeiten der Einflussnahme.

Diese Aspekte sind nicht unabhängig, sondern beeinflussen einander in der Form, dass Durchschaubarkeit und Vorhersehbarkeit als Voraussetzungen für eine sinnvolle Nutzung von Einflussmöglichkeiten zu sehen

sind, während umgekehrt die aktive Beeinflussung eines Systems seine Durchschaubarkeit und Vorhersehbarkeit fördert.

These 4

Nicht eine Entweder-Oder-Entscheidung aufgrund quantitativer Leistungsvergleiche, sondern die komplementäre Gestaltung der Interaktion von Mensch und Technik unter Nutzung ihrer qualitativ unterschiedlichen Potentiale ist anzustreben.

Der Mensch zeichnet sich besonders durch seine hochausgebildete Sensorik, seine Improvisationsfähigkeit sowie seine Fähigkeit, unscharf definierte Probleme zu lösen, aus. Seine Schwächen sind vielfach gerade die Stärken der Technik, besonders Geschwindigkeit, exakte Wiederholung ohne Ermüdung und Lösung sehr komplexer Probleme, wobei die Technik als Problemlöser allerdings höchstens dann besser abschneidet, wenn die Probleme sehr klar strukturiert und damit auf der Basis von Algorithmen lösbar sind. Statt Funktionen nun aber aufgrund vermuteter Leistungsvorteile jeweils dem Menschen oder der Technik zuzuteilen, sollte auf der Grundlage ihrer qualitativen Unterschiedlichkeit die Interaktion zwischen Mensch und Technik bei der Erfüllung der Aufgabe des Gesamtsystems optimiert werden.

Durch die Betonung der Interaktion zwischen Mensch und Technik statt eines Entweder-Oder wird neu der Blick geöffnet für verschiedene Formen der Unterstützung des Menschen durch die Technik, wodurch auf systematischere Weise sowohl menschliche Stärken genutzt und Schwächen kompensiert als auch Vor- und Nachteile technischer Prozesse berücksichtigt werden können. Eine komplementäre Gestaltungsstrategie hätte demnach zur Folge, dass Technikentwicklung nicht vorrangig Nachahmung menschlicher Fähigkeiten mit dem Ziel, Menschen zu ersetzen, beinhalten würde, sondern originär technische Möglichkeiten systematisch genutzt würden, um menschliche Fähigkeiten zu ergänzen. Dies könnte beispielsweise heißen, dass – statt Millionenbeträge in die technische Nachbildung menschlicher Sensorik zu investieren – die enormen Kapazitäten technischer Systeme für das schnell-

le Abarbeiten komplexester Algorithmen vorrangig in Richtung einer Unterstützung menschlicher Informationsaufnahme und Entscheidungsfindung weiterentwickelt würden.

Zusätzlich zur Betrachtung der Interaktion zwischen Mensch und Technik im Hinblick auf ihre sinnvolle Ergänzung bei einzelnen Systemfunktionen sind auch Wechselwirkungen zwischen Funktionen einzubeziehen. Eine bestimmte Funktion, z.B. Messung von Produktmerkmalen für die Qualitätssicherung, könnte vorrangig von einem technischen System übernommen werden, da hohe Geschwindigkeit und Genauigkeit gefordert sind. Eine andere Funktion, z.B. Entscheide über die Korrektur in einem Bearbeitungsprogramm auf der Grundlage dieser Messungen, wird dem Menschen aufgrund seiner Überlegenheit beim komplexen Problemlösen in nicht vollständig definierten Situationen übergeben. Nun ist nicht nur für jede dieser Funktionen zu fragen, welche unterstützende Rolle jeweils der Mensch bzw. die Technik übernehmen sollte, sondern zusätzlich auch die Einflüsse der Art der Verteilung zwischen Mensch und Technik bei einer Funktion auf die Ausführbarkeit der anderen zu berücksichtigen. Beispielsweise wäre zu bestimmen, wie beeinflussbar die Datenerhebung und -darstellung für den Menschen sein sollte, damit sichergestellt ist, dass er in möglichst allen Problemsituationen, die im Produktionsprozess auftreten können, die für die Korrektur nötigen Informationen hat.

Basierend auf diesen Grundannahmen wurden Kriterien entwickelt, die die Bewertung der Aufgabenverteilung zwischen Mensch und Technik in bestehenden Systemen sowie die Erarbeitung und Bewertung von Varianten der Systemgestaltung unterstützen. Um die Kontrolle des Menschen über das technische System zu gewährleisten, sollte die Mensch-Maschine-Funktionsteilung den folgenden, im Rahmen der Entwicklung der Methode KOMPASS empirisch überprüften Kriterien genügen:

Prozesstransparenz – Als zentrale Voraussetzung für die Übernahme von Überwachungs- und Steuerungsfunktionen durch den Menschen ist die Transparenz der technischen Prozesse und damit in Verbindung ste-

hend der Aufbau und Erhalt von mentalen Modellen des technischen Prozesses zu sehen. Der Mensch sollte an all jenen Funktionen des Gesamtsystems beteiligt sein, die das Verständnis von Art und zeitlichem Ablauf der Produktionsprozesse durch aktive Auseinandersetzung mit dem technischen System fördern und vielfältige Prozessrückmeldungen, möglichst auch solche mit unmittelbar sinnlichen Qualitäten, beinhalten.

Dynamische Kopplung – Sowohl eine zu enge Kopplung als auch eine völlige Entkopplung des Menschen vom technischen System sind zu vermeiden, um eine Auseinandersetzung des Menschen mit den Produktionsprozessen zu gewährleisten, ohne ihn zum ‚Knecht‘ des technischen Systems zu machen. Der Mensch sollte im Sinne einer dynamischen Kopplung zeitliche Spielräume haben, zwischen verschiedenen Aufenthaltsorten in der Anlage bei Überwachungs- und Steuerungsaufgaben wählen und verschiedene Arbeitsverfahren einsetzen können, sowie den geforderten kognitiven Aufwand durch unterschiedliche Nutzung technischer Unterstützungsmöglichkeiten variieren können.

Autorität – Die Autoritätsverteilung innerhalb eines Mensch-Maschine-Systems beschreibt, wer – Mensch oder Technik – letztlich die Entscheidungsgewalt über den Zugang zu technisch vermittelten Prozessinformationen und über die Beeinflussung der Prozessausführung hat. Wenn der Mensch für die Effizienz und Sicherheit von Produktionsprozessen die Verantwortung übernehmen soll, ist es unabdingbar, dass er auch über Informationszugang und Prozesseingriffe (mit-)entscheiden kann. Im einzelnen ist allerdings jeweils zu prüfen, inwieweit technische Unterstützung bis hin zur gänzlichen Übernahme einzelner Funktionen durch das technische System sinnvoll sind, um dem Menschen bei schwierigen Funktionen Hilfestellung zu geben bzw. ihn von Routineoperationen zu entlasten.

Flexibilität – Die Flexibilität betrifft die Veränderbarkeit der Funktionsteilung zwischen Mensch und Technik und die Aufteilung der diesbezüglichen Entscheidungsgewalt. Die Verteilung der Kompetenz auf

dieser Autoritätsebene höherer Ordnung unterscheidet adaptierbare von adaptiven Systemen. Adaptierbar ist ein System, wenn der Mensch über die Autoritätsverteilung entscheiden kann. Adaptiv ist es, wenn diese Entscheidung bei der Technik liegt. Anzustreben ist ein adaptierbares System, so dass der Mensch kurzfristig und eigenständig entscheiden kann, z.B. aufgrund seiner momentanen Arbeitsbelastung, ob er eine Funktion selbst ausführen oder sie ganz oder teilweise dem technischen System übergeben will.

These 5

Kontrolle ist bestimmt durch die Mensch-Technik- und Mensch-Mensch-Aufgabenteilung.

Die Ausbildung mentaler Modelle und Eingriffsmöglichkeiten in den Prozess, was beides für die Kontrolle über technische Systeme von größter Bedeutung ist, hängen nicht nur von der Aufgabenverteilung zwischen Mensch und Technik, sondern auch wesentlich von den insgesamt zu bewältigenden Aufgaben ab. Ein Operateur, der ein System nicht nur überwacht, sondern auch programmiert, Störungen selbst behebt, Korrekturen in den Abläufen vornimmt und die Planung der Arbeitsprozesse übernimmt, hat gänzlich andere Voraussetzungen und bedarf einer anderen Unterstützung als der Operateur, der im anderen Extrem ausschliesslich für die Überwachung zuständig ist. Die Gestaltung der Aufgabenverteilung zwischen Mensch und technischem System ist deshalb immer einzubetten in die Gestaltung der Arbeitsaufgaben aller Mitglieder des Arbeitssystems. Dabei ist zu fragen, ob die Gesamtheit der dem einzelnen Menschen übertragenen Aufgaben eine sinnvolle und motivierende Arbeitstätigkeit ergibt. Entsprechend beinhaltet die komplementäre Systemgestaltung neben den beschriebenen Kriterien für die Mensch-Maschine-Funktionsteilung auch Kriterien, die eine Bewertung der gesamten Arbeitstätigkeit der einzelnen Operateure ermöglichen. Ein zentrales Kriterium ist die Vollständigkeit der Arbeitstätigkeit, d.h. die Möglichkeit bzw. sogar Notwendigkeit für eigenständiges Setzen von Zielen, Entwickeln von Plänen für die Zielerreichung, Wahl individueller Arbeitsweisen und fortlaufende Rückmeldung über

Arbeitsergebnisse. Inwieweit vollständige Arbeitstätigkeiten geschaffen werden können, hängt schließlich von der Aufgaben- und Kompetenzverteilung im Arbeitssystem ab, d.h. der Arbeitsverteilung zwischen den Menschen im Arbeitssystem. Wichtige Gestaltungskriterien auf dieser Ebene sind die relative Unabhängigkeit des Arbeitssystems von anderen Einheiten im Unternehmen sowie kollektive Entscheidungsspielräume und hohe Polyvalenz der Mitglieder des Arbeitssystems.

These 6

Durch eine komplementäre Systemgestaltung wird Arbeitsqualität im doppelten Sinne von kompetenzfördernden Aufgaben und effektiver Leistungserbringung möglich.

Der soziotechnische Systemansatz wie auch verwandte Ansätze in der Organisationssoziologie lehren, dass in Systemen mit einem hohen Mass an internen Schwankungen, z.B. durch komplexe Produktionsprozesse verursacht, und/oder einem hohen Mass an Unsicherheit in den System-Umwelt-Beziehungen eine effiziente Koordination nur möglich ist, wenn möglichst unabhängige Untereinheiten möglichst autonom arbeiten können. Unter diesen Bedingungen ist das Bestreben, immer wieder neue Wege zu finden, diese Unsicherheiten zu reduzieren – durch immer ausgeklügeltere Planungs- und Kontrollsysteme, durch immer aufwendigere und umfassendere Automatisierung und durch immer komplexere Kombinationen von beidem – letztlich zum Scheitern verurteilt. Eine Prämisse komplementärer Systemgestaltung ist daher ein Verständnis von technischer Machbarkeit, das nicht von der zentralen Planbarkeit und Steuerbarkeit von Arbeitssystemen ausgeht, sondern Technik auf allen Ebenen einer Organisation als Unterstützung für situationsangemessenes Handeln von Menschen in einer dynamischen Umwelt begreift. Ziel muss sein, Unsicherheiten situationsangemessen bewältigen, daraus erwachsende Potentiale für Innovation nutzen und das dafür nötige Wissen erwerben und weiterentwickeln zu können, wozu ausreichende Kontrolle über technische Systeme sowie individuelle und kollektive Autonomie der Arbeitenden er-

forderlich sind. Dies auch und nicht zuletzt, weil Menschen nachweislich gesünder sind und in ihrer persönlichen Entwicklung eher gefördert werden, wenn ihre Arbeitssituation ein hohes Maß an Autonomie und lokaler Kontrolle ermöglicht.

Ausgewählte Literatur

Bailey, R.W. (1989). *Human performance engineering* (2nd ed.). London: Prentice-Hall International.

Dreyfus, H. & Dreyfus, S. (1986). *Mind over machine – The power of human intuition and expertise in the era of the computer*. Glasgow: Blackwell.

Grote, G. (1997). Autonomie und Kontrolle – Zur Gestaltung automatisierter und risikoreicher Systeme. Schriftenreihe Mensch Technik-Organisation (Hrsg. E. Ulich), Band 16. Zürich: vdf Hochschulverlag.

Grote, G., Wäfler, T., Windischer, A. & Ryser, C. (1998). KOMPASS – Eine Methode für die Analyse und Gestaltung komplementärer Aufgabenverteilung zwischen Mensch und Technik. Zeitschrift für Arbeitswissenschaft, 52, 144-151.

Pasmore, W.A. (1988). *Designing effective organizations. The socio-technical systems perspective*. New York: Wiley.

Sheridan, T.B. (1987). Supervisory control. In G. Salvendy (Ed.), *Handbook of human factors* (pp. 1243-1268). New York: Wiley.

Stahle, W.H. (1991). Redundanz, Slack und lose Kopplung in Organisationen: Eine Verschwendung von Ressourcen? In W.H. Stahle & J. Sydow (Hrsg.), *Managementforschung 1* (S. 313-345). Berlin: De Gruyter.

Thompson, J.D. (1967). *Organizations in action*. New York: McGraw-Hill.

Wilson, J.R. & Rutherford, A. (1989). Mental models: Theory and application in human factors. *Human Factors*, 31.

Haftung für Automaten

Fritjof Haft

Als Jurist vertrete ich eine Disziplin, die vergangenheitsorientiert und pessimistisch ist. Wir haben es nicht mit den Erfolgen der Menschen zu tun, sondern mit ihrem Scheitern. Überdies bin ich einer der letzten Dogmatiker. Es gibt noch eine andere Disziplin, in der es eine Dogmatik gibt, die Theologie. Die Theologie hat – soweit sie katholische Theologie ist – auch einen Papst, der letztgültig über die Dogmatik bestimmt. Wir Juristen haben keinen Papst. Die Päpste sind wir selbst, und wir versuchen, die eigenen Dogmen zur herrschenden Meinung zu machen – das ist dann die höchste Autorität, die wir kennen.

Ich möchte folgendes tun in diesem Kreise: Ich möchte zunächst ein paar Worte sagen über meine Erfahrungen bei dem Versuch, dogmatisch arbeitende Juristen mit moderner Technik in Berührung zu bringen. Ich denke, das ist für Sie insofern interessant, als es den Hintergrund für das Haftungsmodell beschreibt, das ich anschließend vorstellen möchte. Wie konstruieren wir Juristen Haftung? Das werde ich am Beispiel des Strafrechts aufzeigen, weil dort die Haftung am intensivsten ausgebildet ist. Und zum Schluß will ich dann einer Anregung von Herrn Decker folgen. Sie haben in Ihrem Brief geschrieben, ich könne auch kühne Thesen wagen. Also will ich eine These entwickeln, die ich auf dem Juristentag oder auf dem Anwaltstag nicht wagen würde, weil dort der Widerstand zu groß sein würde – aber hier ist vielleicht ein Gremium, in dem so eine These einmal vorgetragen werden sollte.

Zunächst also ein Wort zu dem Bemühen, meine dogmatisch arbeitenden Kollegen mit der modernen Technik, konkret, mit der Computertechnik in Berührung zu bringen. Diese Bemühungen reichen bis in die sechziger Jahre zurück. Als die ersten Überlegungen zur Entwicklung einer neuen Disziplin „Rechtswissenschaft“ vorgestellt wurden, fanden zu Beginn der siebziger Jahre Deutsche Juristentage oder Anwaltstagen unter dem Motto statt „Ich bin dagegen – worum handelt es sich?“ Der Computereinsatz im Recht, also das Unterfangen, mit Computertechnik in das juristische Informations- und Entscheidungsverhal-

ten einzudringen, wurde als ein Anschlag auf das Abendland im allgemeinen und auf die Rechtswissenschaft im besonderen empfunden. Das hat uns aber nicht entmutigt. Wir waren eine kleine Schar von Enthusiasten, die weitergemacht haben. Speziell in Tübingen wurde dann in den achtziger Jahren ein großes Forschungsprojekt gestartet, dessen Zustandekommen mir aus heutiger Sicht völlig unbegreiflich ist. Wir wollten ein natürlich-sprachiges computergestütztes Expertensystem bauen. Das sollte ein System sein, dem man einen Rechtsfall in natürlicher Sprache erzählen konnte, worauf man dann juristische Fragen stellen konnte, die das System zutreffend beantworten sollte. Das war ein sehr aufwendiges Unternehmen, welches damals nur möglich war, weil ein großer Computerhersteller als Sponsor zu gewinnen war. Meine dogmatisch arbeitenden Kollegen beobachteten dieses Unterfangen sehr mißtrauisch. Und als sie dann merkten, daß dieses System nicht in naher Zukunft zu einem einsatzfähigen Produkt führen würde, waren sie beruhigt und sagten: „Wir haben es ja schon immer gewußt, es wird nicht funktionieren!“

Das Bundesjustizministerium setzte damals eine Projektgruppe zum Thema „Einsatz moderner Technik in der Justiz“ ein. Zusammen mit mehreren Kollegen betreute ich darin den Bereich „Expertensysteme“. Zum Schluß der Untersuchung konnte ich nur vermelden, daß in absehbarer Zukunft keine Möglichkeit besteht, Richter und Staatsanwälte zu ersetzen. Das wurde mit großer Freude und Erleichterung aufgenommen. Keine Spur von Bedauern. Also – die Beziehung der Juristen zur modernen Technik ist nicht sehr glücklich. Und es zeugt auch nicht von einer glücklichen Beziehung, wenn der Gesetzgeber versucht, technische Befunde zu regeln. In den achtziger Jahren gab es Bemühungen, die sogenannte Computerkriminalität zu regeln. Man hatte damals beispielsweise behauptet, daß es Programmierer gebe, die Computer mit falschen Daten fütterten, um sich auf diese Weise irgendwelche wirtschaftlichen Vorteile zu verschaffen. Der Gesetzgeber meinte damals im Einklang mit der „herrschenden Meinung“, der Betrugstatbestand, der den Irrtum eines Menschen voraussetzt, würde den neuen technischen Möglichkeiten nicht gerecht werden, weil das Irrenkönnen eine

exklusiv menschliche Fähigkeit sei. Wenn eine Maschine „irre“, könne man dies daher nicht unter diesen Tatbestand des Betruges subsumieren. Also schuf der Gesetzgeber einen eigenen Straftatbestand, den Computerbetrug, der genauso konstruiert ist wie der normale Betrugstatbestand, der aber statt eines irrenden Menschen einen irrenden Computer aufführt. Ich habe nie verstanden, warum die Fähigkeit zu irren als exklusiv menschlich angesehen wird. Ich würde sagen, das Gegenteil des Irrenkönnens, also das Vermögen, die Wahrheit zu erkennen, ist es, was wir für uns beanspruchen sollten. Das Vermögen zu irren würde ich gerne allen Maschinen dieser Welt überlassen. Ich hätte keine Schwierigkeiten gehabt, den Betrugstatbestand so auszulegen, daß der Irrtum einer Maschine darunter zu subsumieren war. Aber das wurde von der „herrschenden Meinung“ abgelehnt. Diese Beispiele zeigen: Wir Juristen haben ein schwieriges Verhältnis zur modernen Technik.

Vor zwei Jahren wurde in Deutschland das sogenannte Multimediagesetz verabschiedet. Es unternahm den Versuch, die Verantwortlichkeit von Personen zu regeln, die im Internet aktiv sind. Das Ergebnis ist ein sehr merkwürdiges Gesetz. Da heißt es etwa sinngemäß: „Der Provider, also der Zugangsanbieter, ist verantwortlich für fremde Inhalte, wenn er davon Kenntnis hat und wenn es ihm technisch möglich und zumutbar ist, den Zugang zu diesen Inhalten zu verhindern“. Es handelt sich hierbei um eine Mischung aus Versatzstücken der strafrechtlichen Dogmatik (Kenntnis, möglich, zumutbar) und großen Worten (verantwortlich). Kein Jurist weiß, was das sein soll. Das Amtsgericht München war vor nicht langer Zeit in der verzweifelten Lage – es ging durch alle Zeitungen – den armen Geschäftsführer eines solchen Internet-Diensteanbieters zu einer Freiheitsstrafe auf Bewährung verurteilen zu müssen, weil dieser pornographische Inhalte – ich muß hier abbrechen, denn es ist in diesem Urteil völlig unklar, was er eigentlich getan oder unterlassen hat. Jedenfalls war es möglich, über diesen Diensteanbieter den Zugang zu pornographischen Inhalten zu erhalten, und das reichte für eine Bestrafung. Dieses Urteil ist so gescholten worden wie wohl noch kein Urteil in der deutschen Rechtsgeschichte. Ich habe es auch in einer Fachzeitschrift besprochen, und ich nahm mir ursprünglich vor,

das Urteil zu loben, denn wo alle hassen, sollte wenigstens einer lieben. Auch hatte der Amtsrichter sich die Sache nicht leicht gemacht. Er hatte sich wirklich alle Mühe gegeben, das Gesetz richtig anzuwenden. Aber da war nichts zu loben. Es war einfach ein Unglück beim Zusammenprall von Recht und Technik passiert.

Soviel als Hintergrund. Wir Juristen haben ein schwieriges Verhältnis zur modernen Technik. Und jetzt konkret zu der Frage, zu der Sie etwas von mir hören möchten, nämlich der Frage nach der Haftung für Automaten. Diese Frage ist ja schon mehrfach angesprochen worden: Was tut der Jurist, speziell der Strafrechtler, wenn beim Einsatz eines Automaten ein Unglück passiert, wenn etwas Schlimmes geschieht, wenn ein Mensch zu Tode kommt, wenn gar mehrere Menschen zu Tode kommen? Denken Sie etwa an das Flugunglück von Cavalese oder die anderen Vorfälle, die hier bereits genannt wurden, oder denken Sie an ein medizinisches Gerät, das versagt, woraufhin schlimme Folgen eintreten. Wenn so etwas passiert, dann blicken wir Juristen zurück in die Vergangenheit und wenden unsere Dogmatik an. Die erste Frage lautet dabei immer, ob ein Mensch gehandelt hat. Mit dem Handlungsbegriff geht es bei uns immer los. In der Handlung liegt die Grundlage für eine strafrechtliche Prüfung. Die Handlungslehre ist ein traditionelles Schlachtfeld des Strafrechts. Zu Beginn des Jahrhunderts gab es einen Rechtspositivismus, demzufolge gesagt wurde, eine menschliche Handlung liege dann vor, wenn ein willensgetragenes Verhalten vorhanden sei, welches kausal irgendwelche Folgen auslöse, wodurch Rechtsgüter verletzt würden. Das war die sogenannte kausale Handlungslehre, die gewissermaßen „blind“ war. Da gab es auf der einen Seite eine Ursache, auf der anderen Seite eine schlimme Folge und zwischen beidem gab es die Verbindung der Kausalität. Werner Bergengruen brachte diesen kausalen Handlungsbegriff einmal in einer Novelle auf den Punkt, die eine Begebenheit aus dem Dreißigjährigen Krieg schilderte. In einem Regiment dienten zwei verfeindete Soldaten. Eines Tages sprang einer auf den anderen zu, legte seine Pistole auf ihn an und drückte ab. Es gab einen Knall, aber es löste sich kein Schuß, die Pistole war blind geladen. Der Schütze wollte seinen Feind ledig-

lich erschrecken. Der andere erlitt vor Schreck jedoch einen Herzschlag und fiel tot um. Dem Schützen wurde der Prozeß gemacht. Er verteidigte sich, die Waffe sei überhaupt nicht geladen gewesen. „Ach was“, knurrte ein altgedienter Kapitän. „Erster Kerl schießt – zweiter Kerl fällt tot um – erster Kerl hängt!“ Das war die Logik, die ursprünglich mit unserem Handlungsbegriff verbunden war.

In den zwanziger und dreißiger Jahren gab es dann eine philosophische Neubesinnung im Recht, die auch die Handlungslehre erfaßte. Man erkannte, daß das Wesen der menschlichen Handlung nicht darin liegt, daß irgendwelche Folgen verursacht werden, sondern darin, daß der Mensch versucht, seine Ziele planmäßig – durch finales Ansteuern – zu erreichen. So wurde etwas geboren, was vom Begriff her völlig schief ist, wie oftmals bei juristischen Termini, aber vielleicht gerade deshalb großen Widerhall fand, die sogenannte finale Handlungslehre. Noch in meiner Studentenzeit bewegte sie alle Juristen, vor allem deshalb, weil keiner so recht wußte, was das eigentlich sein sollte. Der Grundgedanke, daß der Mensch seine Ziele gedanklich anvisiert, ehe er die Aktionen unternimmt, die zu diesen Zielen führen sollen, ist gut bei den sogenannten Vorsatzdelikten. Die finale Handlungslehre versagt aber vor den Fahrlässigkeitstaten, um die es in unserem Zusammenhang ausschließlich geht. Wir reden ja im Zusammenhang mit Automaten nicht vom Vorsatz wie bei Mord und Diebstahl, sondern wir reden von menschlichem Versagen, und das heißt Fahrlässigkeit, etwa fahrlässige Tötung oder fahrlässige Körperverletzung. Bei Fahrlässigkeit, vor allem bei der sogenannten unbewußten Fahrlässigkeit, kann man nicht davon sprechen, daß der Mensch final ein Ziel ansteuert. Vielmehr denkt er schlicht an irgend etwas nicht, an das er hätte denken müssen. Daraus wird ihm dann ein strafrechtlicher Vorwurf gemacht.

Die moderne Handlungslehre ist eine vermittelnde Lehre, die alle genannten Aspekte aufgreift und damit zu einer gewissen Beliebigkeit führt. Unter Handlung verstehen wir heute ein willensgetragenes menschliches Verhalten. Tun und Unterlassen sind dabei gleichwertig. Auch die Frage, ob Tun oder Unterlassen Gemeinsamkeiten haben, ist in diesem Jahrhundert lange erörtert worden. Heute geben wir die Ant-

wort „Nein“. Das Tun besitzt empirische Realität, das Unterlassen ist ein Werturteil, welches besagt, daß ein Tun hätte da sein müssen, das aber nicht vorhanden war.

Wenn wir die so bestimmte Handlung haben, dann geht es weiter in unserer Dogmatik. Dann suchen wir nach einer Norm, einem Straftatbestand, unter den wir diese Verhalten, dieses Tun oder Unterlassen, subsumieren können. Im Strafrecht haben wir das Gesetzlichkeitsprinzip – *nulla poena sine lege*. Es besagt, daß jede Bestrafung ein vor der Tat geschriebenes und bestimmt formuliertes Strafgesetz voraussetzt. Dieses Gesetzlichkeitsprinzip gehört zu den großen Errungenschaften der Aufklärung. Es kann aber da nicht befolgt werden, wo Gesetze nicht mehr formuliert werden können, und das ist bei Fahrlässigkeitsdelikten und bei Unterlassungsdelikten der Fall. Beides kann zusammentreffen. Der Fahrlässigkeitsbereich ist zu großen Teilen ungeschriebenes Recht. Es gibt hier nur wenige Gebiete mit relativ ausformulierten Normen, etwa der Straßenverkehr. Da ist zwar beschrieben, welche Pflichten die Verkehrsteilnehmer haben, aber nur unvollkommen, weshalb am Beginn der Straßenverkehrsordnung die Generalklausel steht, wonach sich jeder Verkehrsteilnehmer so zu verhalten hat, daß er andere nicht mehr als nach den Umständen unvermeidbar gefährdet oder behindert. Also nicht einmal im Straßenverkehrsrecht ist eine vollständige Ausformulierung aller relevanten Verhaltensnormen möglich. Erst recht gilt das für andere Bereiche. Manche sind völlig unregelt, etwa das Medizinrecht. Es gibt kein Gesetzbuch der Medizin. Ob ein Mediziner einen Fehler begangen hat, ob ihm ein Versagen unterlaufen ist, das ihm dann strafrechtlich vorgeworfen werden kann, das beantwortet sich nach den Regeln der ärztlichen Kunst, die sich ihrerseits mit dem medizinischen Fortschritt ständig wandeln. Die Fahrlässigkeitstatbestände sind also durchweg ergänzungsbedürftige Tatbestände, bei denen wir nachträglich urteilen. Wir blicken ja immer zurück zu einem Zeitpunkt, zu dem wir klüger sind. Als Juristen sind wir außerdem professionelle Rechtshaber, ein Schicksal, mit dem man leben muß. Wir blicken also zurück und sagen, das war eine Pflichtverletzung. Das und das ist geschehen, was nicht geschehen durfte. Entsprechend verhalten wir uns bei Unter-

lassungen. Wir sagen, dies und jenes hätte getan werden müssen, was aber nicht getan wurde. Wir formulieren also nachträglich Normen, die wir aus dem konkreten Fall gewinnen. Mit ihrer Hilfe können wir dann Pflichtverletzung feststellen und urteilen, welches menschliche Versagen nicht hätte vorkommen dürfen. Damit können wir das konkrete Geschehen unter das Gesetz subsumieren.

Wenn dann kein besonderer Grund existiert, der ein tatbestandsmäßiges Verhalten ausnahmsweise rechtfertigt, bejahen wir das Unrecht im strafrechtlichen Sinne. Danach nehmen wir die Schuldprüfung vor. Während beim Unrecht ohne Ansehen der Person geurteilt wird, wird bei der Schuldprüfung mit Ansehen der Person geurteilt. Wir sehen also den konkreten Täter an und fragen, ob ihm persönlich ein Vorwurf daraus gemacht werden kann, daß er Unrecht in die Welt gesetzt hat. Das Versagen des Täters kann etwa darin begründet gewesen sein, daß er etwa individuell überfordert war, daß er etwa infolge eines ihm nicht erkennbaren Altersabbau einen Fehler gemacht hat. Oder es kann so gewesen sein, daß er sich in einer besonderen Zwangslage befunden hatte, etwa in der Situation eines entschuldigenden Notstands. Vielleicht mußte er in dem Bemühen, Leben zu retten, von den Regeln der ärztlichen Kunst abweichen, so daß man die Fahrlässigkeitsschuld zu verneinen hat.

Das ist eine grobe Skizze des Modells, das wir im Strafrecht haben. Im Grunde liegt dieses Modell auch dem Zivilrecht zugrunde, wo es nicht um Strafe, sondern um Schadensersatz geht. Die Zivilisten differenzieren freilich nicht so scharf wie wir Strafrechtler. Wir haben beispielsweise im Strafrecht gelernt, daß die Schuld nicht identisch ist mit Vorsatz und Fahrlässigkeit, also mit der subjektiven Verbindung zwischen dem Täter und dem, was er getan hat. Für uns ist die Schuld etwas anderes. Im Zentrum des strafrechtlichen Schuldbegriffes steht das Unrechtsbewußtsein, also das Bewußtsein, gegen Normen, gegen Gebote, gegen Verbote zu verstoßen, und zwar gegen Rechtsnormen. Das haben die Zivilrechtler nicht mitgemacht. Wir haben im Strafrecht auch gelernt, die Kausalität um ein normatives Zurechnungsurteil zu ergänzen. Es reicht also auf der Tatbestandsebene nicht aus, daß eine menschli-

che Handlung als Ursache für den tatbestandsmäßigen Erfolg, etwa eine Tötung, eine Körperverletzung, eine Gesundheitsbeschädigung oder eine Sachschaden, kausal geworden ist. Wir haben vielmehr gelernt, daß ein normatives Zurechnungsurteil hinzukommen muß, damit wir bestimmte Fälle aus der strafrechtlichen Betrachtung ausscheiden können. In unseren Lehrbüchern gibt es einen Erbneffen, der seinen Erb-onkel auf eine Flugreise in der Hoffnung schickt, daß das Flugzeug abstürze und er die Erbschaft antreten könne. Das Flugzeug tut ihm den Gefallen, es stürzt ab, und die Frage ist, hat der Neffe einen Mord begangen? Wenn wir kausal nur die Verbindung zwischen Handlung und tatbestandsmäßigen Erfolg darstellen, dann ist die Antwort eindeutig vorgegeben. Und da auch der Vorsatz sowie die Habgier zu bejahen sind, hat er einen Mord begangen und muß mit lebenslanger Freiheitsstrafe bestraft werden. Aber weil er keine wirkliche relevante Gefahr geschaffen hat, muß man die objektive Zurechnung verneinen. Damit entfällt das Unrecht der Tat, und damit die Strafbarkeit des Neffen.

Das also ist ungefähr das Modell, in welches wir den Einsatz von autonomer Technik, von Automaten, einbauen müssen. Die Frage ist, wie können wir das tun? Und der Ausgangspunkt ist dabei immer: Wir müssen einen verantwortlichen Menschen finden. Im Strafrecht, im deutschen Strafrecht jedenfalls, gibt es nur eine Strafbarkeit von Menschen. Im Zivilrecht gibt es eine Verantwortlichkeit auch von juristischen Personen. Juristische Personen sind eine Abstraktion, welche die Zivilrechtler nach dem Vorbild der natürlichen Person, des Menschen, gebildet haben. Sie taten dies, weil der Mensch einen gravierenden Nachteil hat: Er ist sterblich. Um zur Unsterblichkeit zu gelangen, hat man die juristische Person erfunden. Sie kann zwar sterben, indem sie etwa insolvent wird und liquidiert wird. Sie ist aber potentiell unsterblich. Es gibt juristische Personen, die schon viele Jahrhunderte existieren. Unsere großen Banken und Versicherungen sind beispielsweise in vielen Fällen schon älter als der älteste lebende Mensch. Im Zivilrecht kann man das auch machen, aber die juristischen Personen handeln durch Organe, und diese Organe sind dann wieder Menschen. Letztlich ist das Recht eine Veranstaltung von Menschen für Menschen. Es ist immer auf Menschen bezogen.

Das Menschenbild, das jeweils zugrunde liegt, prägt die jeweils geltende Rechtsordnung. Wir haben heute ein Menschenbild, das sich an der Würde des Menschen orientiert. Es ist gar nicht lange her, da hatten wir ein Menschenbild, das gekennzeichnet war durch die Formel „Du bist nichts – Dein Volk ist alles“. Im Osten Deutschlands existierte bis vor kurzem ein wiederum anderes Menschenbild. Im Zentrum der Rechtsordnung steht immer der Mensch.

Es liegt auf der Hand, daß hieraus Schwierigkeiten entstehen können, wenn Organisationen handeln, zu denen viele Menschen gehören. Wenn etwa ein Chemieunfall passiert, das Beispiel wurde vorhin gebracht, und der Rhein verschmutzt wird, dann ist der Verursacher zunächst einmal eine Organisation, bei der Tausende von Menschen in Betracht kommen, die zusammengewirkt haben. Man kann das Problem natürlich auf der Rechtsfolgenseite lösen. Vorhin war von achthundert Schweizer Franken für eine Rheinverschmutzung die Rede. Ich würde das als eine symbolische Buße von der Art betrachten, wie sie einmal ein englischer Richter verhängte, der einen Angeklagten zu einem Penny Strafe verurteilte und mit erhobener Stimme zu ihm sagte: – Merken Sie sich Angeklagter, im Wiederholungsfall kann diese Strafe verdoppelt werden!“ Also, wenn so ein Unfall passiert, dann gehört es zu den schwierigen Aufgaben des Strafjuristen, nach den eigentlich Verantwortlichen zu suchen. Ist es der, der das Gerät bedient hat? Ist es der, der die Bedienung angeordnet hat? Ist es der Vorstand, der beschlossen hat, die Anlage zu betreiben? Und wie verhält es sich mit dem Vorstandsmitglied, auf dessen Stimme es gar nicht angekommen ist, oder der vielleicht sogar gegen den Betrieb der defekten Anlage gestimmt hat, und der überstimmt wurde? Mit solchen Problemen schlagen wir Strafrechtler uns herum, weil wir immer einen Menschen oder mehrere Menschen suchen müssen, die als Urheber, als Verantwortliche im Sinne des Strafrechts – und letztlich auch des Zivilrechts, das läuft weitgehend parallel – angesehen werden müssen. Und wenn Sie an Katastrophen denken wie etwa die Contergan-Katastrophe vor wenigen Jahrzehnten, dann sehen Sie, wie schwierig das ist, wie schwer sich die Gerichte tun, und welche Auswege sie dabei suchen. Im Falle Contergan hat man damals hat man einen Ausweg in der Einrichtung ei-

nes Fonds gefunden. Das war gesetzlich nicht vorgesehen, das wurde außerhalb des Gesetzes gemacht, und als das geschaffen war, wurden die strafrechtlichen Verfahren irgendwie beendet. Man suchte mit anderen Worten nach zukunftsorientierten Lösungen außerhalb des Gesetzes, und man fand sie.

Wenn ich mir jetzt vorstelle, daß man an irgendeiner Stelle autonome Technik in das Verhältnis zwischen Täter und Opfer oder Verursacher und Geschädigten einbaut, dann meine ich, daß man sich Gedanken über grundsätzlich andere Lösungen als die, die wir bislang haben, machen muß. Damit komme ich zu dem letzten Teil meiner Ausführungen. Mit Ihrer Erlaubnis, sogar mit Ihrer ausdrücklichen Aufforderung will ich etwas sagen, was ich – wie gesagt – in den engeren Juristenzirkeln nicht erzählen würde, weil es die herkömmlichen Denkkategorien verläßt.

Warum denken wir Juristen eigentlich immer vergangenheitsorientiert? Warum machen wir Gesetze für Situationen, in denen irgend etwas schiefgegangen ist? Nehmen wir an, der Gesetzgeber würde ein Roboterförderungs- und -begrenzungsgesetz machen. Dann würde er in § 1 definieren, was ein Roboter ist, in § 2, wer Roboter bauen darf, in § 3, wer sie bedienen darf, in § 4, wer für Schäden infolge des Robotereinsatzes haftet, und in § 5, welche Verhaltensweisen beim Einsatz von Robotern strafbar sind. Wenn so ein Gesetz gemacht würde – und nach diesem Strickmuster werden alle Gesetze gemacht – dann würde der Gesetzgeber nicht die Zukunft gestalten, sondern Vorsorge treffen für eine Zukunft, die heute schon als Vergangenheit gedacht wird. Manchmal geht der Gesetzgeber sogar so weit, eine gedachte Vergangenheit zu regeln, die er nicht kennt, und die er sich überhaupt nicht vorstellen kann. Ein konkretes Beispiel dazu: In den vorhin erwähnten Tatbestand des Computerbetrugs hat der Gesetzgeber alles hineingeschrieben, was man nach seiner Kenntnis mit Computern machen kann, wenn man einen Computerbetrug begehen will, also etwa die Eingabe falscher Daten. Zum Schluß hatte er das Gefühl, daß noch etwas fehlt, weshalb er noch einen letzten Halbsatz angefügt hat, wonach die sonstige unbefugte Einwirkung auf den Ablauf einer Datenverarbeitungsanlage straf-

bar sein soll. Wenn man in den Kommentaren oder in den Materialien nachliest, was der Gesetzgeber mit dieser Formulierung gemeint hat, dann kann man lesen, der Gesetzgeber habe zur Zeit des Gesetzgebungsverfahrens noch unbekannte Einwirkungsmöglichkeiten auf den Computer erfassen wollen. Wenn man das ein bißchen weniger vornehm ausdrücken will, dann kann man sagen, daß der Gesetzgeber keine Ahnung hatte, was er da regeln wollte, und daß er nach dem Motto handelte: „Bin dagegen, worum handelt es sich?“ Das ist eine Grundeinstellung, die wir Juristen hegen. Wir stellen uns die Zukunft immer als eine Vergangenheit vor, in der etwas schief gegangen ist, was wir bewältigen müssen, und zwar immer aus einer Situation heraus, in der wir viel klüger sind, weil man nachträglich natürlich immer gescheiter ist.

Die Frage, mit der ich schließen möchte, ist, ob wir nicht einen ganz anderen Ansatz überlegen sollten. Wäre es nicht besser, wenn auch wir Juristen in die Zukunft statt auf die Vergangenheit schauen würden? Konkret hieße dies für den Fall des Versagens von Automaten, daß wir anstelle der rückwärtsgerichteten Frage nach dem Verursacher, dem Schuldigen, dem Verantwortlichen, eine ganz andere Frage stellen sollten, nämlich die Frage nach dem, dem es nützt, nach dem Begünstigten, nach dem Kollektiv der Begünstigten. Wenn eine neue Technik, etwa in der Medizin, entwickelt wird, und wenn man weiß, diese Technik kann und wird auch in einem wie gering auch immer zu bestimmenden Prozentsatz versagen, dann ist es doch nicht sinnvoll zu fragen, wer das zu verantworten hat. Sinnvoll ist es vielmehr, zu fragen, wem nützt das? Und die Idee, die sich daraus ableitet, besagt, daß man das Kollektiv der Begünstigten heranzieht zum Ausgleich für die Schäden, welche die wenigen erleiden, bei denen die Technik versagt. Bei Arzneimitteln – wir haben ja ein Arzneimittelgesetz – gibt es Untersuchungen, die besagen, daß bei neuen Arzneimitteln nach ihrer Zulassung bei 0,2 % der Fälle nicht vorhersehbare Schäden auftreten. Das wird als tolerierbar hingenommen und liegt im Rahmen des Spielraumes, der noch akzeptiert wird, wenn ein neues Arzneimittel eingeführt wird. Da gibt es keinen Verursacher. Da gibt es niemanden, den man

dafür zur Verantwortung ziehen kann. Aber es gibt natürlich eine große Gemeinschaft von Begünstigten, denen dieses Arzneimittel geholfen hat. Da drängt sich geradezu die Frage auf, ob wir nicht zukunftsorientiert dort ansetzen sollten, wo – jetzt übertragen auf die Technik – diese den Menschen genützt und ihnen etwas gebracht hat. Dort hat sich ein Kollektiv gebildet, aus dem man diejenigen entschädigen kann, bei denen diese Technik ausnahmsweise zu einem Schaden geführt hat.

Das heißt jetzt nicht, daß wir Juristen generell grünes Licht für leichtsinniges Verhalten geben dürfen. Das heißt nur, daß wir zusätzlich zu unserer bisherigen Vorgehensweise, die Vergangenheit zu bewältigen, auch daran gehen sollten, die Zukunft positiv zu gestalten sollen. Das ist für uns Juristen bislang eine völlig fremde Denkweise. Ich bin gerne zu Ihnen gekommen, weil Sie über die Zukunft nachdenken, und ich hoffe, daß es mir gelingen wird, ein wenig von Ihrem Geist zu meinen Kollegen zu tragen. Vielen Dank.

Stand der Forschung auf dem Gebiet Humanoider Roboter und geplante Anwendungsszenarien

Rüdiger Dillmann

Die Hauptaktivitäten der Forschung im Bereich der humanoiden Roboter befinden sich in Japan. Dort wird schon seit einigen Jahrzehnten intensiv an Komponenten, der Integration dieser Komponenten und Anwendung für humanoide Roboter geforscht. In neuerer Zeit gibt es einige Entwicklungen in den USA. In Europa und speziell in Deutschland existieren auf diesem Gebiet fast keine Arbeiten. Da diese Technologie als Zukunftstechnologie einige Bedeutung erlangen dürfte, ist es äußerst wichtig, den Anschluß an die Weltspitze zu suchen und zu finden. Im folgenden werden einige neuere Aktivitäten im Bereich der humanoiden und kooperierenden Roboter vorgestellt.

Die Aktivitäten lassen sich in folgende Thematiken gliedern:

- konkrete humanoide Roboter Projekte
- Architekturen für symbiotische mobile Robotik
- Entwicklungen im Bereich der Kooperation zwischen Mensch und Maschine
- Entwicklung einfacher und intuitiver Bedienschnittstellen
- Entwicklungen im Bereich der Einzelkomponenten humanoider Roboter
- Untersuchen im Bereich der Anwendungsfelder humanoider und kooperationsbereiter Roboter.

1. Humanoide Roboter Projekte

Bereits fertiggestellte bzw. in der Entwicklung weit fortgeschrittene humanoide Roboter befinden sich hauptsächlich im Mutterland der humanoiden Robotik, in Japan, sowie in den USA.

Im Jahre 1998 haben die Agency of Industrial Science and Technology (AIST) und das Ministry of International Trade and Industry (MITI) in Japan ein neues Projekt ins Leben gerufen, welches sich eines Plattform-Ansatzes zur Konstruktion eines humanoiden Roboters bedient,

d.h. die Hauptzeit wird nicht auf die Entwicklung der Einzelkomponenten verwandt, sondern man bedient sich einer Standardroboterplattform und erforscht mit dieser deren Anwendungsmöglichkeiten. Als Plattform wird hier der neue humanoide Roboter P3 von Honda eingesetzt, welcher als Nachfolger des Modells P2 konzipiert wurde. P2 war der weltweit erste, humanoide Roboter, welcher sich äußerst menschenähnlich fortbewegen und auch Treppen steigen konnte. Genauere Angaben zum Plattformansatz findet man in [3], zu den Robotern P2 und P3 in [17]¹.

In einem früheren Stadium der Entwicklung befinden sich die Roboter der Waseda Universität in Japan [4]. Hier sind zu nennen der Hadaly-2, welcher auf einer mobilen Plattform beruht und mit Manipulatoren ausgestattet ist, sowie der WABIAN, welcher als zweibeiniger Roboter mit menschenähnlichem Gang ausgelegt ist und auch Gegenstände tragen kann. Hauptforschungsziel mit Hadaly-2 ist die Interaktion mit dem Benutzer auf den Ebenen der Sprache, der Gestik und der Vision, während WABIANS Eigenschaften hier etwas eingeschränkter sind.

Ein sehr bekanntes Projekt aus den USA ist der Roboter Cog [5,30,31]. Cog entstand als Forschungsplattform zur Evaluation von Methoden der Multisensorik und der Interaktion mit dem Menschen. Cog besteht aus einem Torso mit zwei Manipulatoren und einem Kopf, welcher mit vielen Sensorsystemen bestückt wurde. Das Gesamtsystem besitzt 21 Freiheitsgrade um menschliche Bewegungen zu imitieren und durch die Sensorik können die menschlichen Sinne (visuell, vestibular, auditiv und taktil) approximiert werden.

Eine neuere Entwicklung in Deutschland ist der humanoide Roboter Arnold der Ruhr Universität in Bochum [29].

2. Architekturen für symbiotische mobile Robotik

Auch in Japan sieht man den Zeitpunkt für einen Paradigmenwechsel in der industriellen Robotik näher kommen. Die Zeit, in der ein Industrieroboter damit beschäftigt sein wird, einige wenige Tätigkeiten mit

¹ Die Ziffern in eckigen Klammern entsprechen der Numerierung im Literaturverzeichnis (Anm. d. Hrsg.)

großen Serien an Wiederholungen durchzuführen, wird bald einer Zeit weichen, in der ein hohes Maß an Flexibilität gefordert wird. Man wird andere Roboter benötigen, um dem Menschen unter die Arme zu greifen. Ein Prototyp dieser Vertreter wird zur Zeit von der Forschungsabteilung der Firma Kawasaki Heavy Industries entwickelt [6]. Dieses Prototypsystem verfügt über eine mobile Basisplattform und ist mit einem Manipulator bestückt, dessen Endeffektor eine 4-Finger-Hand ist. Forschungsziele liegen im Bereich der flexiblen Manipulation und der Mensch-Maschine-Schnittstelle.

Im INRIA in Frankreich nimmt man sich auch der symbiotischen Existenz von Menschen und Robotern in gemeinsamen Arbeitsbereichen an. Dort werden die Anforderungen an die Sicherheit des Menschen bei dieser Mensch-Roboter-Beziehung speziell im medizinischen Umfeld untersucht [7].

Auch in Australien wird die Symbiose von Mensch und Roboter in alltäglichen Anwendungen untersucht [8]. Hier kommt man zu dem Ergebnis, dass Sicherheit für den Menschen und einfache, intuitive Benutzerschnittstellen die notwendigen Schlüsselqualifikationen sind, um in derartigen Anwendungsfeldern bestehen zu können. Gezeigt werden erste Ansätze im Bereich des Verfolgens von Gesichtern bzw. des Aufmerksamkeitspunktes des Benutzers sowie eine sichere Manipulatorstruktur.

Eine Arbeitsgruppe an der Vanderbilt University in Nashville Tennessee unter Professor Kawamura beschäftigt sich mit dem symbiotischen Zusammenleben von mobilen Plattformen und stationären Manipulatoren im Bereich der Versorgung älterer Menschen [1]. Hauptforschungsschwerpunkte liegen hierbei in der Philosophie der menschengeführten lokalen Autonomie der Roboter sowie in der internen Roboterarchitektur.

3. Mensch-Maschine-Kooperation

An der Tohoku Universität in Japan untersucht man die Probleme bei der gemeinsamen Manipulation von schweren Lasten durch den Menschen und einen bzw. mehreren maschinellen Manipulatoren [9].

Die Stanford University beteiligt sich an der Forschung mit den Arbeiten der Professoren Khatib und Casal [10]. Auch hier geht es speziell um die Anforderungen an Roboter, welche symbiotisch im selben Umfeld wie der Mensch agieren. Hierbei werden die Aspekte der Mobilität und der Manipulation betrachtet.

Eine Arbeit im Bereich der Koexistenz und Kooperation von Menschen und Robotern, welche sich mit dem Erlernen und Erkennen von menschlichen Absichten befasst, kann in [11] gefunden werden. Hier entwickelt das Toyota Technological Institute ein experimentelles System, um eben jene Absichten zu erkennen und zu unterstützen.

In einer Zusammenarbeit der Michigan State University und der Washington University entstand eine Arbeit über die Integration der unterschiedlichen, heterogenen Eigenschaften des Menschen und der Roboter zu einer kooperativen Beziehung, welche die Vorteile des jeweiligen Partners zu nutzen weiß [12]. Es wird ein neues Paradigma der Mensch-Roboter-Kooperation vorgestellt, nach dem die Umwelterfassung beider Parteien zu gemeinsamem Handeln zusammengefaßt werden kann.

4. Einfache und intuitive Benutzerschnittstellen

In Paris am Laboratoire des Robotique beschäftigt man sich mit dem haptischen Interface, welches eine einfache und intuitive Schnittstelle sein kann, um entsprechend humanoid ausgestattete Robotergreifer zu steuern bzw. zu Lernzwecken zu dienen [15]. Eine derartige Schnittstelle wird benötigt, um in Virtual-Reality-Systemen dem bedienenden Menschen den direkten, haptischen Eindruck des berührten Objektes zu vermitteln.

Das Institute for Problems in Mechanics RAS in Moskau untersucht die Anwendung von Exoskeletten als Hilfsmittel für kranke Menschen bzw. als Ein- und Ausgabeschnittstelle für Simulationssysteme [16].

Eine Weiterentwicklung der visuellen Benutzerschnittstelle für einen humanoiden Roboter findet man in [2]. Hier erforscht die Vanderbilt

University eine neues Modell der menschlichen Aufmerksamkeitsverteilung. Es wird eine Methode vorgestellt, die es ermöglicht, die visuelle Information aufgabenspezifisch nach Merkmalen zu untersuchen.

Ein häufig unterschätzter Kommunikationsweg ist die Wahrnehmung des Gefühlszustandes des Kooperationspartners über den indirekten Weg der Beobachtung dessen Mimik. Das Department of Mechanical Engineering der Science University of Tokyo hat einen 3D-Face-Robot entwickelt, welcher über 26 Freiheitsgrade verfügt und in der Lage ist, die unterschiedlichsten Emotionen wiederzugeben. Diese Emotionszustände werden durch Beobachtung realer Gesichter gelernt [13].

Ein relativ neues Interface zur (Fern-)Steuerung mobiler Serviceroboter durch den Menschen wird in [14] präsentiert. Es basiert auf einem Laufband, auf dem sich ein Mensch durch einen Manipulator kontrolliert bewegen kann. Der Manipulator wird einerseits genutzt, um die Bewegungen des Menschen zu beobachten und das Laufband bzw. die CAVE-ähnliche Simulation zu aktualisieren und andererseits auch, um dem Menschen ein haptisches Feedback über seinen aktuellen Zustand in der Umwelt zu geben. Durch ein solches Simulationssystem ist es möglich, mobile Serviceroboter fernzusteuern.

5. Komponenten für humanoide Roboter

Am Korea Institute of Science and Technology entstand eine Arbeit über eine roboterunabhängige netzbasierte Kontrollarchitektur, welche es ermöglicht, die vollständige Kontrollhardware auf den Robotern zu integrieren und nach außen ein offenes und erweiterbares System zu bieten [24].

In [25] wird das Konzept und die prototypische Implementierung eines humanoiden Roboters gezeigt. Er basiert auf einer mobilen Plattform, welche durch einen Torso und zwei mit Greifern besetzte Manipulatoren erweitert wird.

Eine wichtige Fähigkeit eines humanoiden Roboters ist es, unbekannte sich bewegende Objekte in komplexen unbekanntem Szenen verfolgen

zu können. Eine Arbeit in diesem Gebiet ist das Tracking System des Humanoid Interaction Laboratory in Tsukuba, Japan [26].

Eine andere wichtige Fähigkeit von zweibeinigen Robotern ist es, effektive Methoden zur Regelung der Laufbewegungen zu besitzen. Eine Arbeit hierzu findet man zum Beispiel in [27] am Department of Automatic Control der National University of Defense Technology, China.

In [19] wird ein Ansatz zur Lösung des dynamischen Problems mit variierenden geschlossenen und offenen kinematischen Ketten präsentiert. Eine Lösung dieses Problems ist ein wichtiger Teilschritt zur Konstruktion humanoider Roboter, welche mit der Umwelt agieren sollen, sei es nur sich zu bewegen, oder aber schwerere Gegenstände zu tragen oder gemeinsam zu agieren.

6. Anwendungsfelder für humanoide bzw. menschenfreundliche Roboter

In einer Zusammenarbeit der Waseda University in Tokyo und dem ARTS Lab in Sant'Anna, Italien entstand ein Teleoperationssystem, welches es ermöglicht, humanoide Roboter mit den entsprechenden Schnittstellen auf der ganzen Welt fernsteuern zu können. Demonstriert wurde dieses System anhand von zwei Robotern, dem WABIAN in Japan und der MOVAID Einheit in Italien. Genauere Angaben sind in [18] zu finden.

Eine Demonstration schneller Reaktionsfähigkeiten und Kooperationsfähigkeiten liefert die Entwicklungsabteilung von Toshiba in Kawasaki [20]. Schnelles Objekttracking, die Beherrschung des Ballwerfens und -fangens ermöglichen ihrem Beach Volley Ball Roboter ein interessantes Zusammenspiel mit dem Menschen und helfen, Vorurteile und Technikangst abzubauen.

Ein ganz anderes praktisch einsetzbares Prototypsystem entstand in den Hitachi Entwicklungsabteilungen. Hierbei geht es um ein Unterstützungssystem für ältere Menschen, welche Schwierigkeiten mit dem Setzen, Aufstehen und Laufen haben. Genauer Angaben hierzu findet man in [21].

In Sant'Anna (Italien) sieht man die zukünftigen Anwendungsfelder und Herausforderungen im Bereich des täglichen Lebens im Haushalt. Langfristiges Ziel dieser Initiative ist die Konstruktion eines Unterstützungsroboters für ältere und behinderte Menschen, welche Hilfe bei alltäglichen Handlungen benötigen [22].

Allgemein ist der häusliche Bereich mit dem Motiv der Pflege von behinderten und alten Menschen ein in der aktuellen Literatur häufig zitiertes Anwendungsfeld.

Seit etwa 30 Jahren werden in Japan bereits Entwicklungen und Versuche im Bereich (teil-) humanoider Roboter durchgeführt. Daher ist es nicht verwunderlich, daß die fortgeschrittensten Projekte dort angesiedelt sind.

Durch den technologischen Fortschritt gerät eine tatsächliche Anwendung derartiger Systeme immer mehr ins Blickfeld der Forscher. Daher sollte auch in Deutschland versucht werden, Tactile Sensing mit der Weltspitze aufzunehmen, um sich einen Anteil an diesem (wenn auch langfristig gesehenen) Zukunftsmarkt zu sichern. Lernfähigkeit und Kooperationsbereitschaft gepaart mit einfachen und intuitiven Bedienschnittstellen für (teil-)humanoide Robotersysteme sind in keinem auf der Welt bereits existierenden System zu finden.

Literatur

- [1] D.M.Wilkes, A.Alford, K. Kawamura: *Designing for Human-Robot Symbiosis*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998
- [2] J.A. Driscoll, R.A.Peters II and K.R. Cave: *A Visual Attention Network for a Humanoid Robot*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998

- [3] Hirochika Inoue: *A Platform-based Humanoid Robotics Project*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998
- [4] S. Hashimoto, S. Narita und andere: *Humanoid Robots in Waseda University - Hadaly-2 and WABIAN*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998
- [5] R.A. Brooks und andere: *The Cog Project: Building a Humanoid Robot*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998
- [6] O. Miki, T. Itoko und andere: *Development of an Intelligent Mobile Robot*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998
- [7] E. Coste-Maniere, N. Turro und S. Arias: *Programming Human Friendly Robots Enforcing Safety: Application to Medical Robots*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998
- [8] J. Heizmann, Y. Matsumoto, J. Kieffer und A. Zelinsky: *Smart Interfaces + Safe Mechanisms = Human Friendly Robots*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998
- [9] Kazuhiro Kosuge, Norihide Kazamura und Koji Takeo: *Towards Human-Robot Coordination and Collaboration*, First International Workshop on Humanoid and Human Friendly Robotics, Internatio-

nal Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998

- [10] O. Khatib, K. Yokoi, O. Brock, K. Chang und A. Casal: *Robots in Human Environments: Basic Autonomous Capabilities*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998
- [11] Yamada Yoji, Daitoh Haruyoshi, Hirasawa Yasuhiro, Sakai Takayuki: *Human-Will-Oriented Technology Constructed for Cooperative Motions between A Human and A Power Assisting Robot*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998
- [12] Ning Xi und T.J. Tarn: *Integration of Heterogeneity for Human Friendly Robotic Operations*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998
- [13] Fumio Hara, Hiroshi Kobayashi, Fumiya Iida und Masaoki Tabata: *Personality Characterization of Animate Face Robot through Interactive Communication with Human*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998
- [14] J.M.Hollerbach: *Human Friendly Robots and Interfaces*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998
- [15] C.S.Tzafestas und Ph. Coiffet: *Dextrous Haptic Interaction with Virtual Environments: Hand-Distributed Kinesthetic Feedback and Haptic Perception*, First International Workshop on Humanoid

id and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998

- [16] V.Gradetsky, L. Kravchuk, S. Kalinichenko, V.Lopashov und E. Solovtsova: *Interaction between Man and Machine into Exoskeleton System*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998
- [17] Kazuo Hirai: *Humanoid Robot and its Applications*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998
- [18] Hideaki Takanobu, Eugenio Guglielmelli, Hirokazu Tabayashi und andere: *Remote Interaction between Human and Humanoid Robot*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998
- [19] Yoshihiko Nakamura und Katsu Yamane: *Seamless Dynamics Computation of Structure-Varying Closed Kinematic Chains and Ist Application to Human Figures Interacting with Environments*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998
- [20] O. Fumio, S. Hirokazu, H. Hideaki und andere: *Beach Ball Volley Playing Robot*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998
- [21] Masakatsu Fujie, Yasuhiero Nemoto, Saku Egawa, Akihiko Sakai, Shizuko Hattori, Atsushi Koseki und Takeshi Ishii: *Daily Life Sup-*

- port System for Elderly - Power Assisted Walking Support and Walk Rehabilitation*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998
- [22] Paolo Dario, Eugenio Guglielmelli, Cecilia Laschi: *New Challenges in the Design of Personal Robots*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998
- [23] Takanori Shibata, Toshihiro Tashima und Kazuo Tanie: *Emergence of Emotional Behavior through Physical Interaction between Human and Pet Robot*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998
- [24] Young-Jo Cho, Bum-Jae You, Sang-Rok Oh, Mun Sang Kim, Ye Sun Hong und Chong Won Lee: *A Design of a Compact Network-Based Controller Mounted on a Humanoid Robot*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998
- [25] T. Asfour, K. Berns und R. Dillmann: *Concept and Prototype Implementation of a Humanoid Robot*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998
- [26] S. Rougeaux und Y. Kuniyoshi: *Robust Tracking by A Humanoid Vision System*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998

- [27] Ma Hongxu, Zhang Peng, Zhang Liangqi: *A Time-Position Control Method of Biped Dynamic Walking*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998
- [28] Hong Zhang: *Hybrid Tactile Sensing: Turning Garbage Bags into Tactile Sensors*, First International Workshop on Humanoid and Human Friendly Robotics, International Advanced Robotics Program (IARP), Tsukuba Center for Institute, Tsukuba Japan, 26-27 October 1998
- [29] Bergener, Th, Bruckhoff, C., Dahm, P. Janssen, H. Joublin, F., Menzner, R.: *Arnold: An Anthropomorphic Autonomous Robot for Human Environments*. SOAVE' 97, 1997
- [30] R.A. Brooks: *From Earwigs to Humans*, Journal of Robotics and Autonomous Systems (to appear)
- [31] R.A. Brooks und L.A. Stein: *Building Brains for Bodies*, A.I.Memo, No. 1439, Massachusetts Institute of Technology, August 1993

Roboter im medizinischen Bereich. Beispiele und Chancen

Rolf Dieter Schraft

1. Einleitung

Die rasante Entwicklung modernster Technologie macht auch vor dem Krankenhaus und dem Operationssaal nicht halt. Im Zuge der Digitalisierung wurden Krankenhäuser und Arztpraxen mit Datenverarbeitung und Computern ausgestattet und neue bildgebende Verfahren und Geräte wie Computer- und Magnetresonanztomographen zur Diagnose eingesetzt. Damit stehen diese Hilfsmittel auch für die Planung komplexer therapeutischer Maßnahmen und zur Unterstützung des Arztes im Operationssaal zur Verfügung. Mit diesen Entwicklungen wurden die Grundlagen geschaffen, Robotern den Weg in den Operationssaal zu ermöglichen. Die Datenmengen können zur Steuerung von Robotern genutzt werden, die in den verschiedensten medizinischen Disziplinen bereits zur Unterstützung eingesetzt und in den unterschiedlichsten Entwicklungsstadien im klinischen Bereich erprobt werden. Mit der Einführung von Robotern in Kombination mit neuen Technologien, wie Lasertechnik, bildgebenden Verfahren u. a. erhält die moderne Medizin wichtige Anstöße hinsichtlich effizienterer und weniger traumatisierender Verfahren. Zugleich eröffnen sich völlig neue Behandlungsmöglichkeiten.

Wie Erfahrungen aus der Werkzeug-, Produktions- und Automatisierungstechnik zeigen, sind Roboter präzise, ausdauernd und zuverlässig. Diese Eigenschaften bieten ein optimales Potential für den Einsatz der Robotertechnologie im Operationssaal.

2. Stand der Technik

Die Vielfalt der in der Entwicklung befindlichen Roboter und Manipulatoren in der Medizin ist groß. In keiner der bisher realisierten Anwendungen stehen vollautomatisierte Behandlungssysteme im Mittelpunkt des Interesses, sondern es werden Bewegungseinrichtungen fa-

vorisiert, die einzelne Handlungsabläufe im Krankenhaus/ Operationsaal selbständig durchführen können oder technisch unterstützen.

Eine strikte Einteilung der Robotertechnologie in der Medizin gibt es nicht, da sie durch den Einsatz bei unterschiedlichsten Anwendungen, auch die verschiedensten Anforderungen erfüllen muß. Um sich eine gewisse Übersicht zu verschaffen, kann man jedoch verschiedene Systemarten zusammenfassen. Angelehnt an die VDI-Richtlinie 2860 („Handhabungsfunktionen und Handhabungseinrichtungen“) soll im folgenden Bild die Robotertechnologie in der Medizin, entsprechend ihrer Funktionen als Bewegungseinrichtung, gegliedert werden.

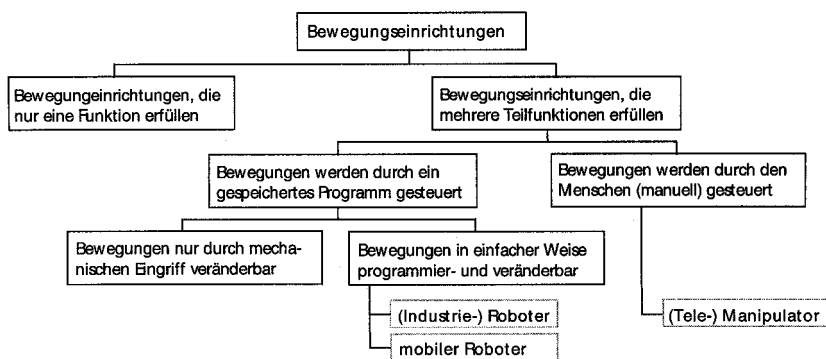


Bild 1: Einteilung von Robotern (angelehnt an die VDI-Richtlinie 2860)

2.1 (Industrie-) Roboter in der Medizin

(Industrie-) Roboter in der Medizin führen ihre Bewegungen programmgesteuert durch und sind mit Werkzeugen/Instrumenten ausrüstbar, um einzelne Schritte im Verlauf einer Operation durchzuführen. Beispiele, mit denen das Fräsen des Prothesenkanals bei Hüftgelenkoperationen durchgeführt werden kann, sind der ROBODOC (ISS) [BÖRN97] und ein System von ortoMaquet (CASPAR) [WEIL98]. Der Neurochirurgieroboter MINERVA (EPFL) [BURC95] führt stereotaktische Gehirnoperationen, unter ständiger Überwachung per CT-Scan, durch.

2.2 Mobile Roboter in der Medizin

Derzeit werden international erhebliche Anstrengungen unternommen, mobile Roboter zu entwickeln beziehungsweise sie in größerem Maßstab einzusetzen [SCHR98]. Ein Beispiel dafür wurde von der Firma TRC mit dem System „HelpMate“ bereits realisiert und befindet sich in amerikanischen Krankenhäusern im Einsatz. Er ist in der Lage, selbständig Objekte innerhalb eines Krankenhauses zu transportieren. Weitere Anwendungen von solchen Systemen lassen sich im Pflegebereich vorstellen. Ein Beispiel hierfür ist der am IPA entwickelte Care-Roboter (CARE-O-BOT), der eine sichere Zukunftsperspektive und verbesserte Lebensverhältnisse für ältere unterstützungs- und pflegebedürftige Menschen gewährleisten soll.



Bild 2: Care-O-Bot (IPA)



HelpMate (TRC)

2.3 (Tele-) Manipulatoren in der Medizin

(Tele-) Manipulatoren stellen dem Arzt ein Instrument zur Verfügung, das die Diskrepanz zwischen den diagnostischen Fähigkeiten der Ärz-

te und der motorischen Unzulänglichkeit des Menschen überbrückt: Viele Krankheiten können diagnostiziert werden, es fehlt aber das Instrument zum präzisen operativen Eingriff. Manipulatoren stellen hier ein Zwischenglied dar, um die groben Bewegungen der Chirurgenhand als feine Bewegungen an das Instrument weiterzugeben. Außerdem sollen sie den operierenden Arzt von physisch und psychisch ermüdenden Tätigkeiten entlasten und es ihm somit ermöglichen, sich stärker auf die eigentliche Diagnose und Therapie konzentrieren zu können, ohne die Verantwortung für den Therapieerfolg und die Sicherheit des Patienten aus der Hand zu geben. Beispiele hierfür sind das System ARTEMIS (FZ Karlsruhe) [VOGE98] und das Endoskopträgersystem AESOP (Computer Motion) [UECK95], das eine Positionierung und Nachführung des Laparoscops erlaubt.

3. Manipulatorgestütztes Operationssystem für die Neurochirurgie

Ein weiteres Beispiel ist der am Fraunhofer IPA in Zusammenarbeit mit Dr. Urban von der neurochirurgischen Abteilung der Dr.-Horst-Schmidt-Kliniken, Wiesbaden, entwickelte Prototyp eines „Mikrochirurgieroboters“, der sich im wesentlichen durch sein Kinematikkonzept von den beschriebenen Systemen unterscheidet.

Der klassische Roboter ist ein elektronisch gesteuertes Gerät, das maschinelle, sich wiederholende Arbeitsschritte in der industriellen Fertigung ausführt. In die Medizintechnik halten, wie beschrieben, sowohl konventionelle marktgängige Industrieroboter als auch auf einen Einsatzbereich spezialisierte Roboterkonstruktionen Einzug. So beruht das Konzept des am Fraunhofer IPA entwickelten Manipulators auf einer in der Medizintechnik neuen Kinematik, dem Hexapod (Stewart-Plattform). Anstelle der sonst üblichen seriell verketteten Achsen, wie beispielhaft am Knickarmroboter, handelt es sich hier um eine Parallelkinematik. Eine Plattform ist durch 6 Teleskopbeine mit einer Basisplatte verbunden, wodurch Bewegungen in allen sechs Freiheitsgraden möglich sind. Die einzelnen Beine sind dabei nur auf Zug und Druck beansprucht, wodurch der Hexapod eine hohe Steifigkeit und Genauigkeit besitzt.



Bild 3: Hexapod mit Endoskop

Eben dieses Kinematikkonzept ist auch das Kernstück der Entwicklung, die im wesentlichen aus zwei Hauptkomponenten besteht [WAPL98a und WAPL98c]. Zum einen dem Manipulator, also dem Hexapod und dessen Aufhängung an einem C-Bogen, zum anderen der Steuereinheit, dem sogenannten Operationscockpit. Die Bewegungen des Manipulators beruhen in der präoperativen Phase auf einer manuellen Ausrichtung des Hexapods entlang des C-Bogens. Dadurch können die meisten medizinischen Zugänge erreicht werden und es eröffnet sich somit eine universelle Einsatzmöglichkeit bei verschiedensten Operationen. Bei der Operation selbst führt das Hexapod kleine Bewegungen mit einer Wiederholgenauigkeit von wenigen Mikrometern aus. Da die Entwicklungen in der Medizintechnik immer mehr in Richtung neuartiger minimal invasiver Diagnose- und Therapietechniken gehen, werden manipulative Genauigkeiten gefordert, die selbst von geübten Operateuren nicht mehr erreicht werden können. Eine höhere Präzision des Eingriffs zur Verbesserung und Erweiterung der Therapie vieler Indikationen bedingt deshalb den Aufbau einer neuen Systemtechnik mit dem verstärkten Einsatz von technologischen Hilfsmitteln. Damit ist die Genauigkeit des Hexapodsystems einer seiner größten Vorteile.



Bild 4: Operationssystem bestehend aus OP-Manipulator und OP-Cockpit (IPA)

Die Ansteuerung des Hexapods erfolgt vom Operationscockpit aus, in dem der Chirurg Platz nimmt und von dem aus er den Eingriff vornimmt. Dieses besteht aus einem Operationsstuhl, der sich auf einem großen hydraulisch angetriebenen Hexapod befindet, Armlehnen mit integrierten Eingabegeräten und einem Monitor. Der Stuhl erfährt eine vergrößerte Rückkopplung der Bewegung des Hexapod-Manipulators ähnlich einem Flugsimulator [WAPL98b]. Dies vermittelt dem Operateur ein aktives Feedback des direkten Geschehens am Endoskop bzw. an den Instrumenten. Schnelle Bewegungen oder starke Auslenkungen warnen den Operateur, so daß keine unkontrollierten Vorgänge durch das Endoskop und die Instrumente ausgeführt werden.

Der am Operationscockpit befestigte Monitor dient der Darstellung der digitalen Sensor- und Patienteninformationen und der Endoskopbilder. Er erleichtert dem Chirurgen zudem die Arbeit durch eine benutzerfreundliche Bedienoberfläche und eine ergonomische Arbeitshaltung.

Als Eingabegeräte wurden zwei Varianten entwickelt und getestet. Das erste Konzept sieht die Ansteuerung des Manipulators mit einem Joystick vor, der eingebettet in eine Armlehne des Cockpits eine übersichtliche und benutzergerechte Bedienbarkeit verschafft.

Bei der Evaluierung des Systems kam eine Ansteuerung mit Voice-control zur Anwendung. Die Spracheingabe hat den Vorteil einer für den Operateur gewohnten Art der Bewegungsanweisung durch Zuruf.



Bild 5: Sprachsteuerung des Manipulators

Die Funktionsfähigkeit dieser Ansteuerung und die Anwendbarkeit des Hexapod-Manipulators in der kraniellen Neurochirurgie konnten im Anatomischen Institut der Universität Wien erprobt und nachgewiesen werden. Aufbauend auf diesen Grundlagen wird nun der Ablauf und Einsatz in weiteren Anwendungsfeldern in der Mikrochirurgie wie der Orthopädie, Wirbelsäulenchirurgie, Otologie (HNO) und Radiologie im Auftrag der Firma Universal Robot Systems (URS) untersucht.

4. Entwicklungsbedarf

Der Einsatz des Operationssystems in neuen Anwendungsfeldern erfordert eine genaue Evaluierung und medizinische Validierung des Sy-

stems. Weiterhin erfordert die Robotertechnologie die Verwaltung zusätzlicher Datenmengen, die manuell nicht mehr zu erfassen und zu überblicken sind. Beispielsweise müssen alle Sensorinformationen in aufbereiteter Form dem Chirurgen zur Verfügung gestellt werden, so daß er die Informationen optimal in seine Entscheidungsstrategien einbinden kann. Daher hängt die Qualität der Informationssysteme nicht nur vom Datenmanagement sondern in erheblichem Maß auch von der Gestaltung der Informationsschnittstelle ab. Moderne Technologien wie VR-Methoden (z. B. Stereoskopie, Augmented Reality) oder der Einsatz von Hilfsmitteln aus dem Multimediabereich könnten die mentale Informationsverarbeitung des Chirurgen unterstützen.

Über bisher existierende Experimentalsysteme hinaus soll ein neues Augmented-Reality System Daten aus bildgebenden Verfahren und rechnergenerierten Daten, die mit einem Planungssystem erstellt wurden, mit Live-Kameraaufzeichnungen des Patienten und anderen Sensordaten überlagert werden. Im Gegensatz zu bisher existierenden Experimentalsystemen würde ein solches AR-System die präzise Korrelation und Echtzeit-Visualisierung von realen und rechnergenerierten Objekten aus verschiedensten Datenquellen erlauben.

An das Manipulatorsystem müssen verschiedene Instrumentenplattformen für den Einsatz in spezifischen ausgewählten Therapien angepaßt werden. Die Anforderungen an eine neue Instrumentengeneration sind vielseitig: Die Instrumente müssen an das Manipulatorsystem adaptiert werden können und so beschaffen sein, daß sie einen schnellen, automatischen Instrumentenwechsel zulassen.

Der Manipulator ermöglicht Bewegungen im Mikrobereich, wobei dem Chirurgen durch ein Endoskop die Dimensionen stark vergrößert dargestellt werden. Die Instrumente können daher sehr viel kleiner als bisher gebräuchlich gefertigt werden. Da die konventionellen Instrumente und Endoskope für diesen Einsatz nicht geeignet sind, besteht großer Bedarf an präzisen, speziell für die Handhabung durch Manipulatoren angepaßten minimal invasiven Instrumenten und Endoskopen.

Von großer Bedeutung für die medizinischen Möglichkeiten sind zudem Instrumente und Geräte mit mehr Freiheitsgraden. Hier ist bei-

spielsweise in der Neurochirurgie ein erheblicher Entwicklungsbedarf gegeben. Der Manipulator bietet mit den geplanten Steuerungs- und Orientierungshilfen eine geeignete Plattform, um kompliziertere Kinetiken bei der Instrumentenführung zu kontrollieren.

5. Zusammenfassung und Ausblick

Das Operationssystem des Fraunhofer IPA ergänzt die Reihe der bestehenden und in der Entwicklung befindlichen „Medizinroboter“ um den Bereich der Mikrochirurgie. Noch sind die Systeme nicht serienreif und die Stimmen der Kritiker nicht beruhigt. Bedenken werden insbesondere hinsichtlich der zusätzlichen Belastung des Patienten durch die operative Anbringung von Referenzmarkern, speziell in der Orthopädie, laut. Einige alternative Verfahren wie z. B. die Auswertung von Volumendaten werden zum Teil schon getestet. Gleichzeitig sind alle Sicherheitsanforderungen (Überwachung des Bewegungsraumes, Kollisionskontrolle) und medizinischen Anforderungen (Sterilisierbarkeit, Bearbeitungseigenschaften) genauestens zu erfüllen und die gesetzlichen Regelungen einzuhalten. Deshalb wird die Entwicklung von „Medizinrobotern“ an verschiedenen Einrichtungen weltweit weitergehen. Weitere Disziplinen sind zu erschließen, bestehende Systeme zu verbessern und noch sicherer zu machen, sowie weitere Schritte in Richtung der Modifikation peripherer Systeme und Abläufe zu überdenken.

6. Literatur

- BÖRN97 Börner, M.; et al.: Computerunterstützter Robotereinsatz in der Hüftendoprothetik Unfallchirurg (1997) 100: 640-645 Springer-Verlag 1997
- BURC95 Burckhardt, C.W.; Flury, P.; Glauser D.: Stereotactic Brain Surgery IEEE Engineering in Medicine and Biology, May/june 1995, 314-317
- ENGE98 Engelberger, G.: HelpMate, a service robot with experience Industrial Robot, Vol 25, Nr 2, 1998, 101-104

- SCHR98 Schraft, Rolf Dieter; Schmierer, Gernot: Serviceroboter: Produkte, Szenarien Visionen.; Berlin; Heidelberg: Springer, 1998
- UECK95 Uecker, D.R et al.: Automated Instrument Tracking in Robotically Assisted Laparoscopic Surgery, J Image Guided Surgery 1995, 308-25
- VOGE98 U. Voges, E. Holler; B. Neisius; M. Schurr; T. Vollmer: Evaluation of ARTEMIS, the Advanced Robotics and Telemanipulator System for Minimally Invasive Surgery Proceedings IARP 2nd Workshop on Medical Robotics, Forschungszentrum Karlsruhe, 1997, 137-148
- VDI VDI-Richtlinie 2860: „Handhabungsfunktionen und Handhabungseinrichtungen“
- WAPL98a Wapler, M.; Stallkamp, J.; Weisener, T.; Urban, V: Motion Feedback as a Navigation Aid in Robot Assisted Neurosurgery, Medicine Meets Virtual Reality: 6, 215-219, San Diego, January 28-31, 1998.
- WAPL98b Wapler, M.; Neugebauer, J.; Weisener, T.; Urban, V.: Robotassisted Surgery System with Kinesthetic Feedback, Proceedings 29 th Int. Symposium on Robotics (ISR '98), April 1998
- WAPL98c Wapler, M.; Weisener, Th.; Hiller, A.: Hexapod-Robot System for Precision Surgery, Proceedings 29th Int. Symposium on Robotics (ISR '98) April 1998
- WEIL98 Dr. Andrea Weiler: Entwicklung der Robotik für die Hüftgelenksprothetik aus der Sicht eines Industrieunternehmens Fraunhofer IPA - Technologieforum F30, Stuttgart 23./24. Oktober 1998

Neurotechnologie - Auf dem Weg zu lernfähigen Neuroprothesen mit sensorischen und motorischen Implantaten

Rolf Eckmiller,

1. Der Mensch und seine Werkzeuge

Eine in Zeit und Raum möglichst distanzierte Betrachtung des Phänomens Mensch kann nur aus der beschränkten Sicht des Menschen erfolgen und ist damit zwangsläufig seinen sensorischen, kognitiven und motorischen Begrenzungen unterworfen[3].

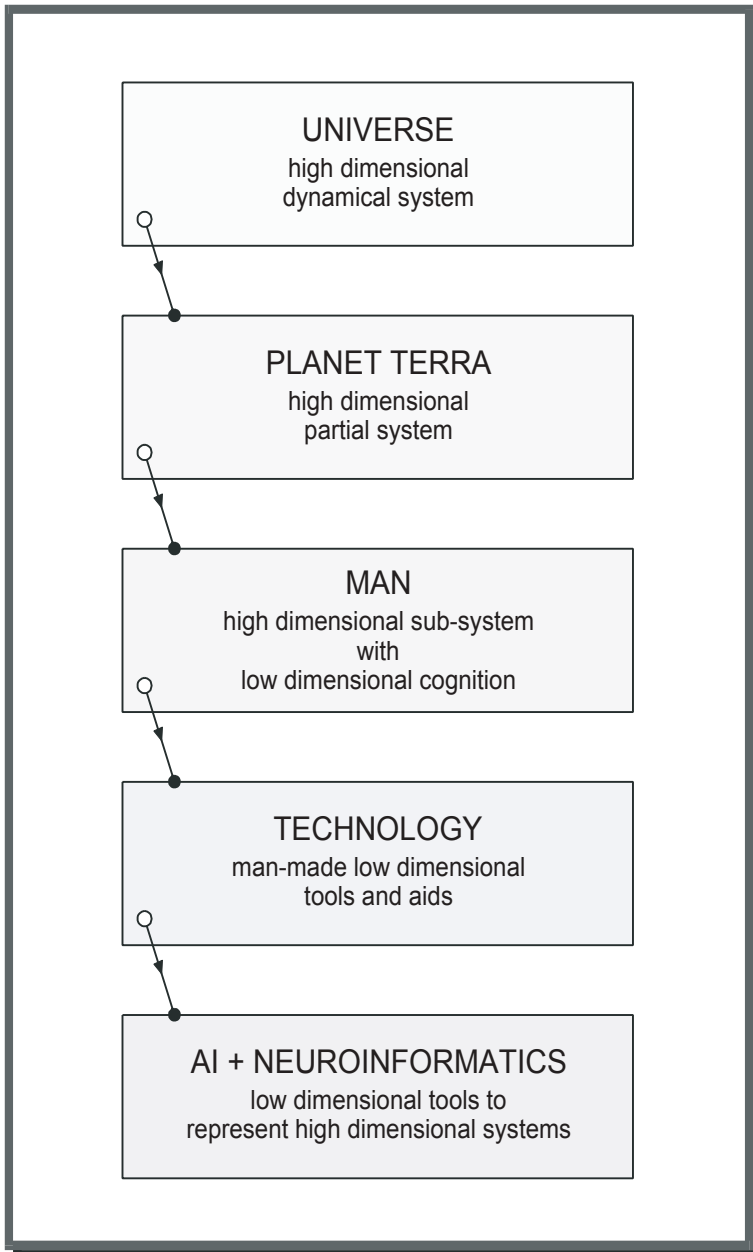
Die Naturwissenschaften sind als Kompendium des menschlichen und damit von seinen Hirnfunktionen begrenzten Bemühens zu betrachten, Beobachtungen zu ordnen, Regelmäßigkeiten zu interpretieren und Werkzeuge bzw. Verfahren und Handlungsprinzipien zu entwickeln.

Insbesondere die seit weit mehr als 10.000 Jahren entwickelte Astronomie als wichtigste Quelle für die beruhigende Erfahrung von Ordnung in Zeit und Raum (Zeitmessung per Äquinox-Beobachtung, Wiederkehr von Sternzeichen, Himmelsnavigation, usw.) sowie weitere Naturwissenschaften vermitteln dem Menschen den Eindruck (Abb. 1), daß es:

erstens ein Universum als hoch-dimensionales, dynamisches System,
zweitens darin den Planeten Terra als hoch-dimensionales Teilsystem
und
drittens auf Terra den Menschen als hoch-dimensionales Subsystem
gibt.

Der Mensch seinerseits:

hat nur nieder-dimensionale Kognition,
baut bzw. entwickelt nieder-dimensionale Werkzeuge einschließlich von lernfähigen Computersystemen und sensomotorischen Systemen, die teilweise als nieder-dimensionale Werkzeuge zur Repräsentation hoch-dimensionaler Systeme betrachtet werden können.



© Prof. Dr.-Ing. R. Eckmiller, Universität Bonn

Abb. 1 Universe - Planet Terra - Man - Technology - AI + Neuroinformatics

Der Mensch mit einem aus etwa 100 Milliarden Nervenzellen mit je etwa 10.000 Synapsen als lernfähigen Kontakten bestehenden Nervensystem empfängt, speichert, modifiziert und generiert Trajektorien, also spatio-temporale Signalfunktionen[2]. Dies gilt sowohl für eintreffende visuelle, oder auditorische Sinnesereignisse, als auch für intern generierte motorische Ereignisse zur Artikulation oder Lokomotion (Abb. 2).

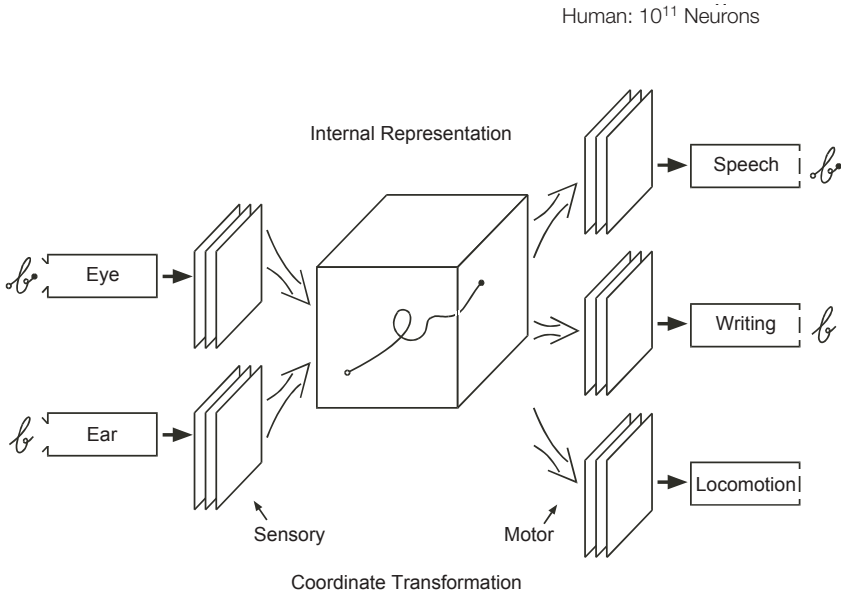


Abb. 2 Nervensystem des Menschen

Das für den Menschen gegebene Dilemma als hoch-dimensionales Subsystem in einem ebenfalls hoch-dimensionalen System jedoch nur mit nieder-dimensionaler Kognition ist bedingt durch die physikalischen und informationsverarbeitenden Grenzen von Sensorik, Motorik und Nervensystem. Dies entspricht der gedachten Situation eines fiktiven Insektes, welches nur 2-D Sensorik hat, sich jedoch in 3-D bewegt und orientieren muß. Das Insekt vor einer Kugel kann nur einen Kreis detektieren, also nur die 2-D Projektion der Kugel. Im Prinzip kann das Insekt mit einem Artgenossen, der von einer anderen Position aus eine andere 2-D Projektion der Kugel detektiert, kommunizieren und so intern eine 3-D Rekonstruktion der Kugel vornehmen. Analog hat der Mensch Chancen, einerseits durch optimale Kommunikation mit Art-

genossen und andererseits durch Zuhilfenahme technisch erzeugter Modelle, u.a. mit lernfähigen Computersystemen (Intelligente Systeme), Objekte von einer höheren Dimension zu rekonstruieren und so kognitiv zu erfassen. Eine wichtige Voraussetzung für eine derartige kognitive Erfassung höher-dimensionaler Objekte ist eine nüchterne Bestandsaufnahme der biologisch (Hardware/Software) bedingten Grenzen des sensomotorischen Systems des Menschen. Die bisherigen diesbezüglichen Bemühungen waren in der Menschheitsgeschichte nur sehr begrenzt erfolgreich. Es fehlt offenbar ein weltweit gemeinsames Referenzsystem, ein gemeinsamer Satz von Werten, Begriffen, Realitätsbeschreibungen und Zielen, der invariant ist gegenüber dem philosophischen, religiösen und politischen Zeitgeist. In dieser Hinsicht kann man den Eindruck haben, daß wir noch Analphabeten sind, dies aber zunächst realisieren müssen.

Auf der Suche nach einem geeigneten Bezugspunkt oder Referenz-Koordinatensystem für ein gutes Verständnis des Phänomens Mensch und seiner Rolle und Ziele innerhalb des Universums bin ich auf die extraterrestrische Perspektive (ETP) gestoßen[3]. Die ETP versetzt den Betrachter im Gedankenexperiment in eine Art Raumstation, von der aus alle Vorgänge auf dem Planeten Terra wertfrei gewissermaßen aus der Sicht eines Planetar-Biologen beobachtet werden. Aus ETP-Sicht treten die großen Entwicklungslinien, wie z.B. rapider Ressourcen-Abbau und ineffiziente Bereitstellung von Konsumgütern (z.B. Nutzfahrzeuge) durch Transport von Material statt Know-how zwischen den weit entfernten Wirtschaftsräumen sowie die rasch zunehmende Orientierungslosigkeit der einzelnen regionalen Populationen in den Vordergrund, während die vielen regional entscheidenden Aspekte im Hintergrund sind.

Die Mathematik als Reflexion menschlicher Denkfähigkeit und Hirnfunktion besteht klassisch vor allem aus zwei Theorie-Räumen[1], die einander teilweise so unversöhnlich gegenüberstehen, wie die Teilchentheorie und die Wellentheorie zur Beschreibung von Licht-Phänomenen, nämlich der algebraisch-analytische und der geometrisch-topologische Theorieraum (Abb.3). Rene Descartes erkannte die Not-

wendigkeit, ein gegebenes physikalisches Problem, welches jeweils nur in einem dieser Theorieräume repräsentiert war, auch in den jeweils anderen Theorieraum zu transponieren. Dazu erfand er 1637 das nach ihm benannte kartesische Koordinatensystem und begründete damit die analytische Geometrie.

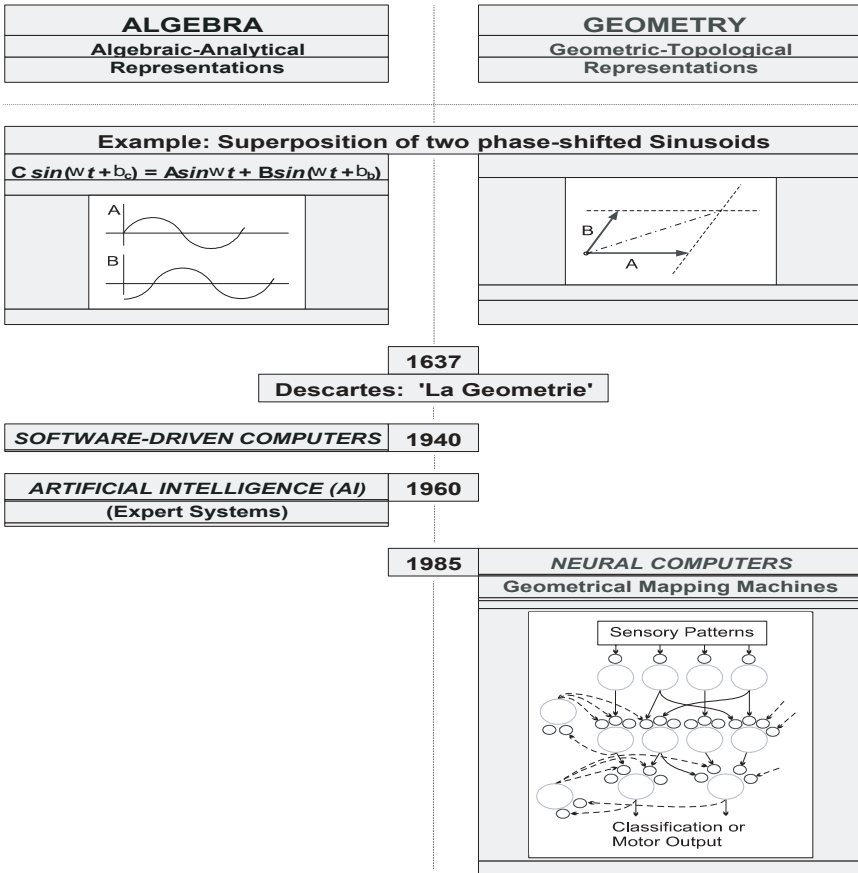


Abb. 3 Algebra - Geometry

Die etwa um 1940 entwickelten, Software-getriebenen „von Neumann Rechner“ gehören in den algebraisch-analytischen Theorieraum, da zur Berechnung das Problem zunächst durch Gleichungen repräsentiert werden muß. Das gleiche gilt für die um 1960 auftretende Künstliche

Intelligenz, die z.B. zu Expertensystemen führte. Im Gegensatz hierzu ist die Neuroinformatik mit technischen lernfähigen neuronalen Netzen dem geometrisch-topologischen Theorieraum zuzuordnen. Das zu bearbeitende Problem muß nicht durch Gleichungssysteme eingegeben werden. Ferner basieren neuronale Netze als geometrische Abbildungsmaschinen wesentlich auf geometrischen und topologischen Parametern, also der Netzwerk-Struktur. Von Neurowissenschaftlern wurde eine Verbindung angedeutet zwischen der Denk-Präferenz von Mathematikern, z.B. Boole und Lagrange für Algebra bzw. Riemann und Lobatschewski für Geometrie und deren Hirnhemispären-Dominanz. Danach findet algebraisch-analytisches Denken überwiegend in der linken und geometrisch-topologisches Denken in der rechten Hemisphäre statt[1].

Im Gegensatz zu den meisten anderen Tier-Species entwickeln Menschen Technologien und Werkzeuge, die lernfähig sind, also Ansätze von Intelligenz haben. Derartige intelligente Systeme (IS) bieten die Möglichkeit, dem Menschen in vielfältiger Weise zu helfen und zu dienen:

- A) IS zur Rekonstruktion hoch-dimensionaler Phänomene, bzw. Strukturen aus mehreren nieder-dimensionalen Beobachtungen verschiedener Menschen
- B) IS zur Orientierungshilfe bei der Suche nach dem Sinn des Lebens auf Terra
- C) IS zur Linderung von Funktionsstörungen des Nervensystems.

2. Neurotechnologie zur Linderung von Funktionsstörungen des Nervensystems

Seit etwa 1993 wird in Deutschland mit Förderung des Bundesforschungsministeriums (BMBF) die Neurotechnologie als neues anwendungsnahes Forschungsgebiet gefördert, in dem die Bundesrepublik an der Nahtstelle zwischen Technologie, Biologie und Medizin eine Trendsetter Rolle einnimmt[4, 5].

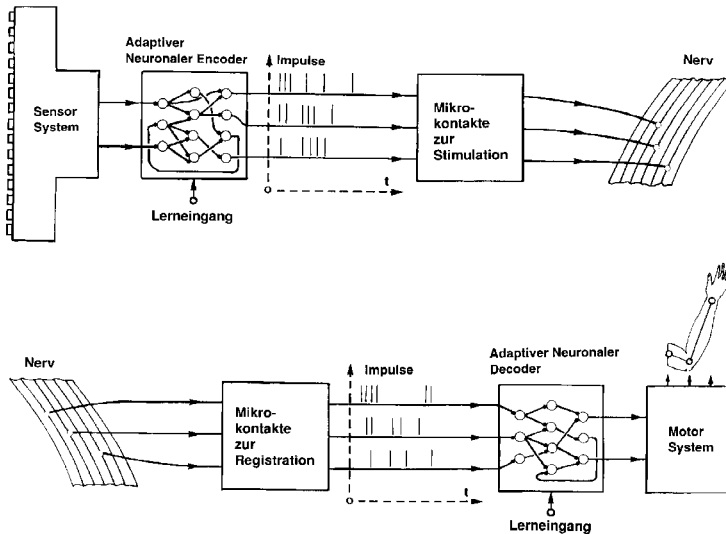


Abb. 4 Prinzip von Neurotechnologie-Produkten

Die Neurotechnologie (Abb.4) strebt nach lernfähigen Prothesen in direktem Signalkontakt mit Nervengewebe über implantierte Mikrokontakte zur Linderung von Funktionsstörungen des Nervensystems. Dies betrifft mittelfristig extrem ehrgeizige Ziele, wie die Entwicklung von sensorischen Neuroprothesen, z.B. Sehprothesen für Blinde, von motorischen Neuroprothesen, z.B. Greifprothesen für Querschnittsgelähmte, oder von cranialen Neuroprothesen, z.B. sensomotorischen Implantaten für Epilepsie.

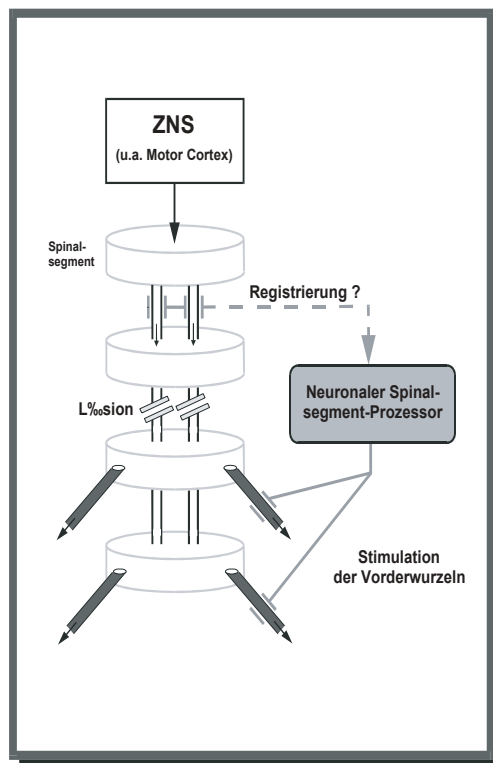
Das typische Neurotechnologie-Produkt besteht aus:

einer implantierten Mikrokontakt-Struktur zur Stimulation von Nervengewebe, zur Detektion von Nervensignalen und/oder zur bedarfs-gesteuerten, lokalen Wirkstoffabgabe an eine eng begrenzte Nervengewebsregion,

einer lernfähigen neuronalen Encoder- oder Decoder-Struktur zur Übernahme der defekten biologischen Nervengewebsfunktion mit angekoppelter Sensor- und/oder Aktor-Struktur und

einer drahtlosen Signal- und Energieübertragungsstruktur zwischen dem implantierten Teil und dem externen Teil der gegebenen Neuroprothese.

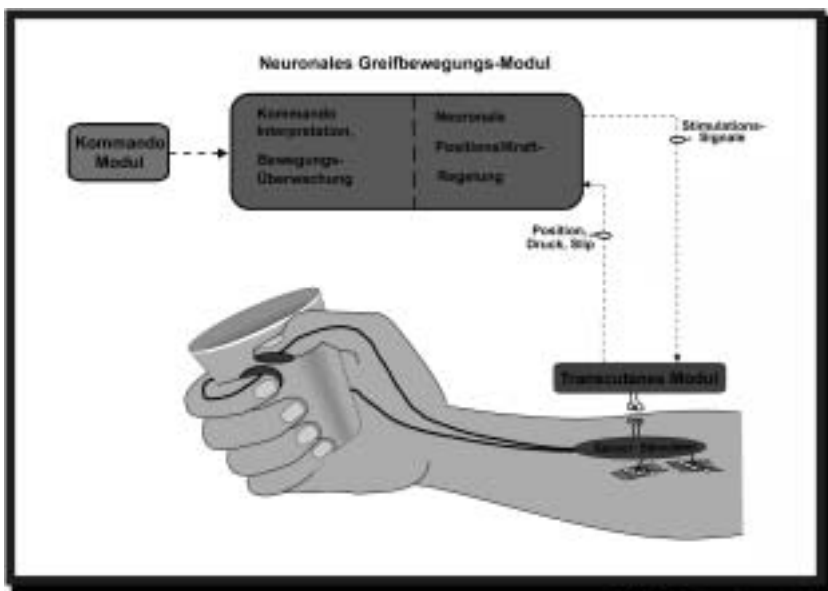
Bei der Gruppe von Querschnittsgelähmten wäre die Wunschvorstellung, den Ort der Durchtrennung des Rückenmarks mit seinen vielen, aus insgesamt weit über 100.000 Nervenfasern bestehenden Nervenfaserbündeln, die einerseits von Zentren innerhalb des Schädels zur Peripherie und andererseits in die Gegenrichtung Nervensignale fortleiten, technisch zu überbrücken. Eine derartige Überbrückung (Abb.5) ist jedoch nach Auffassung der Neurotechnologie-Expertengruppe (s. Neurotechnologie Report, 1995) wegen einer Zahl von nicht lösbaren Schwierigkeiten zumindest innerhalb der nächsten 10 Jahre nicht machbar. Andererseits sind spezifische Neuroprothesen für Querschnittsgelähmte, z.B. zur Steuerung von Körperfunktionen des Urogenitaltraktes, zur Steuerung von Stand und Gangbewegungen oder zur Steuerung von Greifbewegungen machbar.



© R. Eckmiller, Universität Bonn

Abb. 5 Fiktion der Überbrückung einer Rückenmarksdurchtrennung

Ein als machbar eingestuftes Greif-Implantat (s. Neurotechnologie Report, 1995) ist in Abb.6 schematisch dargestellt. Der Personenkreis kann aufgrund seiner weit oben lokalisierten Rückenmarks-Durchtrennung zwar noch die Arme bewegen und die Hände in Greifposition bringen, jedoch z.B. kein Glas umgreifen (Zylindergriff) und auch keinen Schlüssel erfassen und drehen (Schlüsselgriff). Das geplante Greif-Implantat, welches ein externes Kommando-Modul im Kopfbereich sowie externe lernfähige Steuer- und Sendemodule umfaßt und implantierte Stimulationkontakte im Unterarmbereich für die betreffenden peripheren Nerven sowie Druck/Rutsch-Sensoren im Fingerbereich enthält, soll sowohl Zylinder-, als auch Schlüsselgriff erlauben.



© R. Edmiller, Universität Bonn

Abb. 6 Schema eines Greif-Implantates

3. Retina Implant als lernfähige Sehprothese

Die Entwicklung von lernfähigen Retina Implantaten soll einer Gruppe von Blinden mit Netzhautdegeneration zu einer zumindest bescheidenen Wiedergewinnung von Gestaltwahrnehmungen (z.B. Tür, Stuhl, Tisch, Fenster) verhelfen. Es gibt allein im Bereich der EU etwa

600.000 diesbezüglich Betroffene mit Retinitis pigmentosa oder Makula Degeneration.

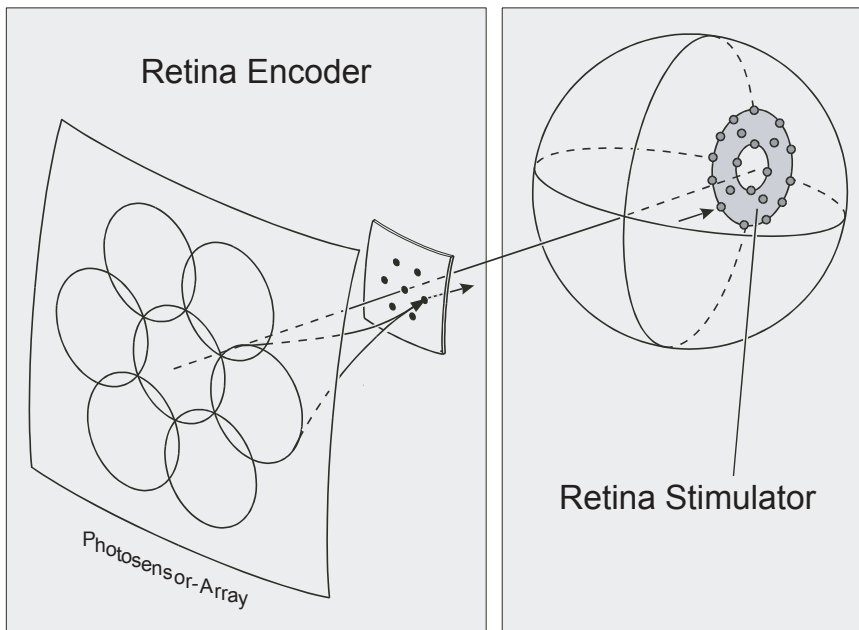


Abb. 7 Funktionsschema für ein lernfähiges Retina Implant

Zur Entwicklung des Retina Implants wurde 1996 mit Förderung des BMBF in Höhe von 10 Mio DM für zunächst vier Jahre ein Konsortium von 14 Partnern von Ophthalmologen, Biologen und Technologen sehr verschiedener Fachrichtungen unter Leitung des Autors gebildet. Das Retina Implant (Abb.7) besteht aus einem Retina Encoder (RE), der in einem Brillengestell unterbracht sein wird, einer nahe der Retina implantierten Mikrokontaktfolie als Retina Stimulator (RS) und einer drahtlosen Signal- und Energiekommunikation (SE) zur Interaktion zwischen RE und RS.

Der Retina Encoder besteht aus einem Photosensor-Array mit mehr als 100.000 Pixeln und aus einem Digital Signal Processor (DSP) zur Echtzeit-Implementation von mindestens 200 lernfähigen Orts-Zeitfiltern (RF-Filter) mit den typischen rezeptiven Feldeigenschaften retinaler Ganglienzellen der Primatenretina. Beim Einsatz des Retina Implantats

bei Menschen soll die Zahl der Mikrokontakte zur selektiven Stimulation von Ganglienzellen zwischen 100 und 500 liegen.

Der Entwicklungsstand gegen Ende der ersten Förderphase ist der, daß die Module des Retina Implantats als Funktionsmuster realisiert und erfolgreich tierexperimentell erprobt werden konnten. Ebenfalls konnte ein Lernverfahren zur Wahrnehmungsbasierten Funktions-Optimierung des RE entwickelt und bei normalsichtigen Versuchspersonen erfolgreich erprobt werden.

Aufgrund dieser erfolgreichen Entwicklungsarbeiten wurde eine zweite Förderphase seitens des BMBF bereits in Aussicht gestellt. Damit ist die Bundesrepublik im Bereich lernfähiger Sehprothesen für Blinde mit Netzhautdegeneration weltweit führend.

Parallel zu diesem Forschungsprojekt wurde 1998 die Fa. Intelligent Implants GmbH, Bonn für Fertigung und Vertrieb von Retina Implants gegründet. Forschungsteam und Firma stehen in engem Kontakt mit der Stiftung Retina Implant, die als Betroffenen-Organisation die Entwicklung intensiv verfolgt und sich an der Patienten-Information und der Sammlung von Fördermitteln zur Beschleunigung der schwierigen Forschungsaufgaben beteiligt. Die Beteiligung der Betroffenen schafft bereits im Vorfeld ein wichtiges Forum zur Diskussion von Akzeptanz- und Ethik-Fragen.

4. Ethische Aspekte des Einsatzes von Neurotechnologie-Produkten

Der in den Jahren 1994-1995 im Auftrag des BMBF erarbeitete Neurotechnologie Report nimmt auch zu ethischen Aspekten Stellung. Entsprechend dem gegenwärtigen Konsens in unserer Gesellschaft wird ein großer Unterschied gemacht zwischen Neurotechnologie-Implantaten innerhalb des Schädels und solchen außerhalb. Neurotechnologie-Produkte, wie z.B. Retina Implantate (in der Augenhöhle, aber außerhalb des Schädel-Innenraumes), Cochlea Implantate (im Innenohrbereich), oder Greifimplantate im Rückenmarks- oder Peripheren Nervensystembereich werden prinzipiell als vertretbar eingeschätzt. Dage-

gen sollen Neurotechnologie-Produkte innerhalb des Schädels, wo nach gängiger Meinung Fremdkörper nicht als „an mir“, sondern als „in mir“ empfunden werden, nur unter dem Vorbehalt eines vorher herzustellenden gesellschaftlichen Konsenses genehmigt werden. Insbesondere geht es darum, im Vorfeld zu verhindern, daß Persönlichkeits-Veränderungen durch nachher schwer kontrollierbare, technische Implantate möglich werden. Bei entsprechendem Forschungsfortschritt sollten craniale Implantate jeweils sehr sorgfältig auf Mißbrauch, Nebenwirkungen und Fehlfunktionskonsequenzen überprüft werden.

Andererseits sind mittelfristig Neurotechnologie-Implantate zur Sicherung der Persönlichkeit und Identität bei Beeinträchtigungen durch neurologische Krankheitsbilder, wie z.B. Epilepsie, oder Parkinsonsche Schüttellähmung, oder durch psychiatrische Krankheitsbilder, wie z.B. Alzheimer denkbar. Hier sind im Vorfeld sorgfältige und intensive Festlegungen von Rahmenbedingungen erforderlich. Im Neurotechnologie Report wird ausdrücklich die Position eines ‚Patienten-Anwaltes‘ vorgeschlagen. Der Patienten-Anwalt soll unabhängig von Interessen seitens Medizin, Industrie, Forschung, Politik oder Kapital sein, er soll die beabsichtigte Technologie, die Wirkung, den medizinischen Eingriff, den Nutzen, die Risiken und Nebenwirkungen hinreichend überblicken können und er soll den prospektiven Patienten allein aus der Patienten-Perspektive unabhängig und umfassend beraten.

5. Schlußbemerkung

a) Lernfähige Neuroprothesen zur Linderung von Funktionsstörungen des Nervensystems werden in den nächsten 5-10 Jahren verfügbar.

b) Die Neurotechnologie als wichtiger, innovativer Zweig der Medizintechnik und System-Biotechnologie folgt dem Motto: ‚High-Tech for the People‘.

c) Neurotechnologie-Produkte haben einen bisher nicht üblichen, oder nicht technologisch realisierbaren Komplexitätsgrad. Aus diesem Grund gilt hier das Motto: ‚Complex Teams for Complex Products‘.

Mit anderen Worten basiert die erfolgreiche Entwicklung von Neurotechnologie-Produkten ganz wesentlich auf der Fähigkeit, Teams von Biologen, Medizinern und Technologen zu organisieren und dabei die erheblichen Sprach- und Denkbarrieren der einzelnen Expertengruppen zu überwinden.

d) Die Neurotechnologie wird einen wichtigen Beitrag zu einer sich bereits seit einiger Zeit andeutenden Entwicklung leisten, das Monopol einer einzelnen Expertengruppe (der Mediziner) als Anwalt der Interessen des Patienten zukünftig einem Team von Experten aus Biologie, Medizin und Technologie zu übertragen.

e) Ich bin zuversichtlich, daß wir Menschen das technologisch Machbare auch menschenwürdig anwenden, daß wir Technologien uns zu Diensten machen werden.

Literatur:

1. Eckmiller, R.: Concerning the emerging role of geometry in neuroinformatics. In: Parallel processing in Neural Systems and Computers (Eckmiller, Hartmann, Hauske, eds.) Elsevier Science Publ., Amsterdam, pp. 5-8 (1990)
2. Eckmiller, R.: The design of intelligent robots as a federation of geometrical machines. In: An Introduction to Neural and Electronic Networks (Zornetzer, Davis, Lau, eds.) Academic Press, Orlando, pp. 109-128 (1990)
3. Eckmiller, R.: Interaktion menschlicher und technischer neuronaler Netze zur globalen Zukunftsvorsorge - Eine extraterrestrische Perspektive. In: Kultur und Technik im 21. Jahrhundert (Kaiser, Matejovski, Fedrowitz, Hrsg.) Campus-Verlag, Frankfurt am Main, pp.91-101 (1993)
4. Eckmiller, R.: Concerning the challenge of neurotechnology. In: Neurobionics (Bothe, Samii, Eckmiller, eds.) Elsevier, Amsterdam, pp. 21 - 28 (1993)

5. Eckmiller, R.: Biology-inspired pulse processing neural networks (BPN) for neurotechnology. In: Proc. ICANN '94, Sorrento, Springer Verlag, 1994, pp. 1329-1334 (1994)
6. Eckmiller, R.: Towards Retina Implants for improvement of vision in human with retinitis pigmentosa - Challenges and first results.
In: Proc. WCNN'95 INNS Press, Lawrence Earlbaum Assoc., Hillsdale, 1995, vol. I, pp. 228-233
7. Eckmiller, R.: Retina Implants - Auf dem Wege zu lernfähigen Sehprothesen für Blinde mit Netzhautdegeneration In: Sonderdruck zum WVAO Jahreskongress 1997, Köln, Verlag: Schmidt GmbH & Co., pp.10-13 (1997)
8. Eckmiller, R.: concepts for learning Retina Encoders, Proc. ICNN '97, Houston, Texas, IEEE Press, Vol. 4, pp. 2315-2320, 1997
9. Eckmiller R.: Learning Retina Implants with epiretinal contacts. Ophthalmic Research 29, 281-289, 1997
10. Eckmiller, R., Hünemann, R., Becker, M.: Exploration of a dialog-based tunable Retina Encoder for Retina Implants. Neurocomputing 24, in press, 1999
11. Eckmiller, R., Suchert, S.: Strategy for the Foundation of a Neurotechnology Company ICONIP 1998, vol. 1, pp. 37-39

Perspektiven intelligenter Mensch-Maschine-Kommunikation

Helge Ritter

Lassen Sie uns zum Einstieg kurz einen „Roboter der Natur“ betrachten, nämlich eine Honigbiene. Sie ist, wenn Sie so wollen, ein technisches Meisterwerk, sehr preiswert und auf engstem Raum realisiert, und verfügt über eine ganze Menge von Bewegungsgeschick und sehr viel motorische Intelligenz. Wenn wir das, was dieses Tier kann, in technischen Systemen nachahmen könnten, dann wären wir in der Robotik sehr, sehr weit. In der Natur kommt uns das nicht als eine besondere Leistung vor. Wir sehen, daß das „nur“ eine Art Spezies von Lebewesen, gewissermaßen Robotern der Natur ist. Wir kennen davon sehr zahlreiche und ganz verschiedene Bauformen, und daran knüpft die erste These an:

These 1:

Roboter entwickeln sich nach ähnlichen Gesetzen wie eine biologische Spezies; nur in anderen Biotopen und in augenscheinlich kürzeren Zyklen.

Ich glaube, daß wir in der Robotik vor einer ähnlichen Entwicklung stehen, wie sie in der Natur vorgezeichnet ist, und daß sich Roboter nach ähnlichen Gesetzen wie biologische Spezies entwickeln werden. Es gibt auch in der technischen Welt so etwas wie „Biotope“, die ganz spezifische Anforderungen an die Roboter stellen, aber die Entwicklungszyklen sind wesentlich kürzer als in der „natürlichen“ Evolution. Einige Beispiele solcher „Biotope“ möchte ich hier vorstellen:

Industrieroboter agieren im vielleicht derzeit größten „Biotop“, nämlich in Montagehallen. Die Anforderungen dabei sind gar nicht so, daß die Roboter extrem autonom und eigenintelligent sein müßten. Das würde in dieser Umgebung eigentlich eher stören. Was dort gefragt ist, sind sehr hohe Wiederhol-Zuverlässigkeiten, und daher Bewegungen, die mit hoher Geschwindigkeit in einer vorstrukturierten Umgebung re-

produzierbar ausgeführt werden können, sowie kurze Umrüstzeiten, wenn etwas anderes gebaut werden soll.

Roboter auf dem Meeresgrund haben in einem wesentlich widrigeren Biotop zu agieren, als es eine Industriehalle ist. Ein hoher Wasserdruck, in der Regel sehr viel Schmutz und vor allem die Energieversorgung stellen hier besondere Probleme dar, denn der Roboter kann nicht an ein Kraftnetz angeschlossen werden. Dementsprechend erfordern Roboter auf dem Meeresgrund ganz andere Technologien und sind auch ganz anders gebaut.

Kanal-Roboter: Hier sind die engen räumlichen Verhältnisse z. T. sehr prägend für die „Körperform“. Für das Fortkommen innerhalb von Röhren kann man sehr spezialisierte Fortbewegungseinrichtungen bauen, und Sie kennen vielleicht alle diese z. T. sehr stromlinienförmigen Prototypen von Kanal-Robotern, die heute schon viele nützliche Aufgaben verrichten können.

Partyroboter bewegen sich wiederum in gänzlich anderen Biotopen. Es handelt sich dabei um eine durchaus ernsthafte Anwendung, nämlich Roboter im Unterhaltungsbereich. Es wurde heute schon kurz der kleine Hund von SONY erwähnt, der ein gutes Beispiel eines Partyroboters abgibt. Es ist durchaus absehbar, daß derartige Roboter in Zukunft ein großes Marktsegment und damit auch eine große wirtschaftliche Bedeutung haben werden und somit auch bestimmte Berufsbilder, etwa das des Unterhalters, in einer gewissen Weise ergänzen bzw. verändern.

Antropomorphe Roboter sollen die letzte Spezies sein, auf die ich hier eingehen möchte. Auf einige Forscher üben sie eine besondere Faszination aus. Sie sehen hier einen Prototyp, der von Rodney Brooks am MIT entwickelt wurde: COG, ein sehr berühmter antropomorpher Roboter. Er ist ausgestattet mit vier Kameras, jeweils 2 Augen für Nah- und Fernsicht. Auf Bildschirmen kann man die verschiedenen Verarbeitungsschritte des visuellen Systems verfolgen. Es handelt sich um einen Laborprototypen, der zur Zeit allerdings noch mit vergleichsweise simplen Händen ausgerüstet ist. Eine der spannenden grundlegenden Fragen, die mit COG unter anderem untersucht werden, ist, wie man

mit beweglichen Objekten hantiert. Übrigens glaube ich, daß solche antropomorphen Roboter für uns eine sehr geringe „Bedrohung“ sind, weil sie in ein Biotop einzudringen versuchen, in dem wir selbst sehr gut, nämlich durch unsere körperliche Bauform optimal angepaßt sind. Auf absehbare Zeit wird es einfach viel zu teuer sein, einen antropomorphen Roboter zu bauen, der auch nur über annähernd vergleichbare Fähigkeiten wie der sehr gute Generalist „Mensch“ verfügt. Daher wird man viele Aufgaben nach wie vor Menschen übertragen und – da es für geraume Zeit immer noch billiger sein wird – sie gut bezahlen, so daß solche Roboter vermutlich lange Zeit „nur“ Forschungsgegenstand bleiben werden. Nichts desto trotz mag es Einsatzfälle geben, die dann trotzdem zu überraschenden antropomorphen Designs führen. Beispielsweise hat Honda einen inzwischen berühmten Prototyp eines zweibeinig-lauffähigen Roboters vorgestellt. Sein Aussehen erinnert schon sehr an Sciencefiction, aber eigentlich ist dieses Gerät, abgesehen von seiner Bewegungsintelligenz, ein recht „dummer“ Roboter. Es verfügt z. B. nicht über Sehfähigkeit, und es kann nur einen ziemlich einfach programmierten Weg verfolgen. Allerdings dürfen in diesem Weg solche Dinge wie Treppenstufen und Steigungen vorkommen, die der Roboter sehr gut bewältigt. Für die Entwicklung von Basistechnologien können solche Projekte durchaus hilfreich sein. Andererseits kann man gerade an diesem Roboter gut demonstrieren, daß er für uns Menschen eindeutig in die Kategorie der Automaten einzureihen ist, da ihm eine entscheidende Fähigkeit fehlt, nämlich die Fähigkeit der Kommunikation. Diese sehr wichtige Fähigkeit führt uns zur zweiten These dieses Vortrages:

These 2:

Kommunikationsfähigkeit ist ein wesentlicher Akzeptanzfaktor, der Roboter von reinen Automaten unterscheidet.

Kommunikationsfähigkeit ist, meiner Ansicht nach, der wesentliche Faktor, der einen Roboter aus der Sicht des Menschen „intelligent“ erscheinen läßt. Natürlich finden wir auch die Roboter in den Fabrikhallen, obwohl sie nur stereotyp ihre Bewegungen ablaufen lassen und da-

bei aber über sehr wenig flexible Repräsentation der Welt verfügen aufgrund ihres Aussehens bereits einigermaßen interessant. Jedoch stuft man ein Gerät, das einem in Form eines schlichten Kastens begegnet, mit dem man sich aber unterhalten kann, als sehr viel intelligenter ein. Roboter, die mit dieser Fähigkeit ausgestattet sind, haben sehr viel breitere Anwendungsfelder, d.h. sie können in vielen „Anwendungsbiotopen“ agieren. Geht man davon aus, daß in Zukunft die Verbindung von menschlicher und maschineller Expertise an Bedeutung gewinnt, dann benötigt man eine exzellente Mensch-Maschine-Kommunikation. Man sieht das schon, wenn man so einfache Dinge wie Videorecorder bedienen möchte. Hier wäre es sehr viel leichter, wenn man mit dem Gerät direkt sprachlich kommunizieren könnte, und wenn es Anweisungen dann ein Stück weit versteht. Diese Entwicklung wird sich verstärken, und wir werden ständig Hilfen brauchen, um die immer komplexer werdenden technischen Geräte vernünftig bedienen zu können. Heute existieren auf Rechnern bereits Graphiksysteme, die so viele Optionen anbieten, daß man ohne sehr ausgeklügelte Hilfsverfahren diese Möglichkeiten nicht mehr ausschöpfen kann. Mit einem richtig ausgeklügelten Verfahren meine ich im Grunde genommen, daß das Gerät selbst, mit Eigenintelligenz, seine Expertise in einer Form anbietet, daß man per Dialog die notwendigen Operationen veranlassen kann. Eine solche Steuerung wäre für sehr viele Hightech-Geräte von Interesse. In Automobilen hat das sprachliche Ansteuern von Geräten bereits Einzug gehalten, da man auf diese Weise vom Fahren weniger abgelenkt wird. Das Fahren selbst wird eines Tages auch ein Stück weit von Robotern übernommen, bzw. in Gefahrensituationen wird der Mensch heute schon ganz erheblich durch technische Systeme unterstützt, weil für die hier wichtigen, sehr kurzen Reaktionszeiten die Maschine besser sein kann.

Generell ist die heutige Mensch-Maschine-Interaktion noch alles andere als optimal. Zwar ist der Pfad in der Richtung von der Maschine zum Menschen sehr gut entwickelt: Wir kennen alle die wunderschöne Hochleistungsgrafik, mit der wir heute beeindruckende, dreidimensional dargebotene, bewegte Displays realisieren. Was aber noch schlecht

und sehr eingeschränkt ist, ist der umgekehrte Pfad, vom Mensch zur Maschine. In vieler Hinsicht leben wir da noch in vergangenen Jahrzehnten: Die Tastatur spielt nach wie vor die Hauptrolle. Vieles weitere, was wir mitteilen wollen, muß sich durch den engen Flaschenhals der Computer-Maus zwängen, mit der wir wenigstens zweidimensional noch ganz gut zeigen können. Hinzu kommt, daß die Datenformate stark an den Charakteristika maschineller Verarbeitung orientiert sind. Das macht die Kommunikation störanfällig und unnatürlich. Wir müssen uns derzeit noch immer sehr an die Maschine anpassen und nicht umgekehrt. Tatsächlich wollen wir es aber umgekehrt haben. Nur sind insbesondere diejenigen Ausdrucksmittel, die besonders menschengemäß sind, also Zeigen und Sprechen, technisch erst in geringem Umfange realisiert. Stellen Sie sich vor, man könnte etwa einer Datenbank ein Werkstück zeigen, auf dem Werkstück mit dem Finger eine Fehlstelle andeuten und dann erläutern, warum das eine Fehlstelle ist. Die Datenbank fragt noch einige Sachen, die ihr nicht klar sind, zurück, und dann hat sie dieses Wissen in den Datensatz aufgenommen, und man muß nichts mehr mühsam in der Repräsentationssprache der Datenbank formulieren. Vieles, was dazu notwendig wäre, ist technisch bereits vorhanden. Die Integration des technisch Vorhandenen ist allerdings eine nicht-triviale Aufgabe, die ihrerseits neue Teilaufgaben mit sich bringt. Andere Aspekte sind zudem bisher technisch noch nicht realisiert. Beispielsweise wird – und das berührt einen interdisziplinären Aspekt – die Gestaltung von Mensch-Maschine-Schnittstellen heute noch sehr einseitig als ein rein technisches Problem angesehen. Das ist es aber nicht, denn im Grunde wollen wir auch hier Ergonomiefaktoren berücksichtigt haben. Des weiteren spielt menschliche Kognition eine ganz wesentliche Rolle, denn wir denken in bestimmten Erwartungsmustern, wir haben bestimmte kognitive Spannen, und wir reagieren auf bestimmte Verbindungen von Farbe und Form besser als auf andere Verbindungen von Farbe und Form. Diese Dinge kann man sich nicht am grünen Tisch überlegen, sondern es muß experimentelle Wahrnehmungsforschung betrieben werden und deren Ergebnisse müssen dann in die neuen Designs einfließen. Für diese Aspekte wurde in Bielefeld ein Sonderforschungsbereich ins Leben gerufen, mit dem Thema „Si-

tuierte künstliche Kommunikatoren“. Hinter diesem Titel verbirgt sich die Zusammenarbeit von Informatikern, von Computerlinguisten und von Psychologen, um die Frage wirkungsvoller Mensch-Maschine-Kooperation von verschiedenen Seiten zu beleuchten. Im Mittelpunkt steht dabei die Verbindung von Sprache und Handlung: Es gibt einen Roboter, es gibt Kameras, es gibt Mikrofone, und es gibt ein Szenario, in dem eine Aufgabe im Dialog mit der Maschine bearbeitet wird. Dabei müssen dem Roboter verschiedene Teilfähigkeiten verliehen werden:

- Dem Roboter muß Wahrnehmung über mehrere Wahrnehmungskanäle ermöglicht und diese muß koordiniert werden: visuell, auditiv, bis hin zu haptisch, wenn die Roboterhand etwas greift.
- Der Roboter muß Bezugnahmen auf Objekte verstehen, um einen gemeinsamen Gesprächskontext zu schaffen.
- Es entstehen Fragen der Konzeptdynamik, wie beispielsweise in Abhängigkeit von der Verwendungssituation von Bauteilen eine Scheibe zu einem Rad werden kann und damit einen anderen kognitiven Stellenwert bekommt.
- Situiertheit, also die Berücksichtigung des Handlungskontextes, wird wichtig.
- Aufmerksamkeitssteuerung. Wo blicken wir hin? Wo blickt die Maschine hin? Wie wird das koordiniert? In Gesprächssituationen hängt davon beispielsweise das natürliche Wechselspiel zwischen Sprecher und Hörer ab.

Last not least soll dann auch die tatsächliche Integration einiger dieser Fähigkeiten in einen technischen Prototyp demonstriert werden, so daß im Experiment gezeigt werden kann, ob Hypothesen und Ansätze tragfähig sind.

Das Szenario kann man sich so vorstellen, daß eine Art Arbeitsplatz eingerichtet ist, in dem ein menschlicher Mitarbeiter mit einer Maschine kommuniziert. Die Maschine ist beispielsweise ein Puma-Roboter

mit einem mehrfingrigen Manipulator. Der Roboter verfügt über einen binokularen Kamerakopf, der die Szene beobachtet. Ziel ist es, dem Roboter durch sprachliche Anweisung eine Montageanweisung zu vermitteln, die er dann mit Bauteilen ausführen soll. Der Mensch darf dabei menschentypisch kommunizieren, also mit Worten und Gesten.

Natürlich sind in dem Sonderforschungsbereich viele verschiedene Teilaufgaben zu bearbeiten, und in unserer Arbeitsgruppe betrachten wir besonders den Aspekt der nicht-sprachlichen Kommunikation. Zunächst mag dieser als von untergeordneter Bedeutung erscheinen, aber in vielen Gesprächssituationen kommt man einfach viel schlechter aus, wenn man sich nur rein sprachlich verständigen kann. Wenn ich Sie etwa bitte, mir von da drüben einmal „diese“ Cola-Flasche zu reichen: Ohne die Worte „da drüben“ mit einer Zeige-Geste zu begleiten, muß ich sehr viele beschreibende Hinweise geben. Und gerade in Anleitungssituationen, wenn man vermitteln möchte, „wie“ man etwas macht, kommt dem Zeigen eine herausragende Rolle zu. Ziel ist es daher, daß man auch einem Roboter zeigt, wie etwas zu machen ist, statt ihn explizit zu programmieren. Dafür muß der Roboter Informationen über raumzeitliche Gegebenheiten – und beim Montieren handelt es sich gerade immer um solche – erkennen können. So muß der Roboter Arme und Hände eines Menschen erkennen und analysieren können, was die Finger gerade tun. Dabei ist es nützlich, wenn er auch Kopf und Augen des Menschen erkennt und die Blickrichtung schätzen kann, um herauszufinden, worauf der Mensch gerade besonders achtet. Wir sehen, für den Roboter sind viele Informationen wichtig. Robotersysteme mit Erkennungs- und Kommunikationsfähigkeit auszustatten führt in eine Richtung, die nicht gerade als Hauptströmung in der Robotikforschung bezeichnet werden kann. Vor allem, wenn man sich an der Frage des „Materials“ orientiert, mit dem der Roboter „hantieren“ soll. Wir denken unwillkürlich Roboter sind Maschinen, die mit „Hardware“ hantieren, die Automobile zusammenschrauben, die in Bergwerken arbeiten oder auf dem Meeresgrund. Das ist heute der wichtigste Anwendungsausschnitt. Aber ich glaube, es gilt für Roboter etwas Ähnliches wie für Menschen:

These 3:

Ähnlich wie für den Menschen wird auch für die Roboter der Zukunft Information der attraktivere Rohstoff sein.

In mittlerer Sicht wird auch für Roboter der Werkstoff der Information der wichtigere werden, und wir werden mehr und mehr Roboter sehen, für die das Informationshandling zu einem ganz wesentlichen Bereich zählen wird. Das Handling der anderen Dinge wird gleichgewichtig sein, wird Anhängsel sein oder in einzelnen Bereichen auch immer die Hauptrolle spielen. Aber aus der oben beschriebenen „Biotopsicht“ deutet vieles darauf hin, daß sich im Forschungsgebiet der Robotik diejenigen Roboter, die Information in erster Linie und in wichtigster Linie verarbeiten, sehr viel stärker ausbreiten werden als ihre stärker mechanisch orientierten Brüder und Schwestern. Die eigentliche Herausforderung, was die Ersetzung des Menschen angeht, ist demnach vielleicht so etwas wie der „Informationsroboter“, denn Roboter, und das führt in den Bereich der informationsverarbeitenden Computer hinein, haben andere Fähigkeiten, Informationen zu verarbeiten als Menschen. Sicherlich wird die intelligente Kombination von Information, die ich hier ohne den Intelligenzbegriff zu definieren einführen möchte, noch eine ganze Weile auch in dieser Universalität die Domäne des Menschen bleiben. Aber was die schnelle Verknüpfung von vielen Informationen in einem mittleren oder niedrigeren Qualitätsgrad angeht, wie er oft zum Vorfiltern oder zum Vorbereiten erforderlich ist, werden uns Roboter überholen, bzw. haben uns Roboter bereits überholt. Und das ist eine Entwicklung, die ihnen dann auch neue Betätigungsfelder, nämlich den ganzen Dienstleistungsbereich erschließt. Im Grunde, wenn Sie jetzt schon einen Service-Roboter in einem Krankenhaus haben, der Aufträge entgegennimmt, bestimmte Sachen von A nach B zu bringen, dann ist der Informationsaspekt schon eine nicht zu unterschätzende Komponente, obwohl der Roboter auch physikalisch etwas tut. Künftige Roboter werden Menschen beraten, und weil sie das weniger gut können als richtig gute menschliche Experten, werden sie diejenigen Teile von Beratungstätigkeiten unternehmen, die zunächst einmal vorbereitenden Charakter oder Überblickscharakter haben. Dazu gehört

Information finden, vorauswählen und anbieten. Die intellektuell tiefergehende Arbeit wird dann vom richtigen menschlichen Experten gemacht werden.

Ich will aber eine Unterthese anfügen: Es wird immer wichtige Bereiche von diffizilen manuellen Tätigkeiten geben, – wir haben heute vormittag sehr eindrucksvolle Beispiele gesehen – in denen man die Geometrie nicht so starr fixieren kann, wie wir das im Falle des Hüftgelenks gesehen haben. Eine enge Kooperation mit dem Menschen wird daher noch für eine ganze Weile das Adäquateste sein, weil Menschen doch extreme „Bewegungsexperten“ sind, insbesondere was die Bewegung unserer Finger angeht. Und deswegen sind Tele-Roboter wahrscheinlich auf lange Zeit wichtige Gehilfen oder Partner des Menschen, insbesondere, weil sie auch so etwas wie die „Überwindung des Raumes“ ermöglichen. Sie ermöglichen uns eine Fernpräsenz, eine Fernwirkung in andere Bereiche hinein, und damit eine Steigerung unserer Mobilität. Und die Steigerung von Mobilität hat industriell eigentlich immer ganz ausschlaggebende Rollen gespielt: Denken Sie an das Automobil, als ein Beispiel für Mobilität, bei der wir unseren körperlichen Ort tatsächlich verändern müssen. Der Tele-Roboter bietet dagegen eine Mobilität, mit der wir unsere physikalischen Einwirkungsmöglichkeiten an entfernte Orte spiegeln können. Und der Informations-Roboter kann gewissermaßen das Automobil der Zukunft für die schnelle Bewegung in Informationsräumen sein, also Mobilität in einer immer schneller wachsenden virtuellen Welt bieten, die ja ohne technische Hilfsmittel heutzutage schon nicht mehr begehbar ist. Aber lassen Sie uns für einen Moment nochmal zum Tele-Roboter zurückkommen. Hier ist etwa das Greifen noch eines der schwierigen wissenschaftlichen Probleme, die nur teilweise gelöst sind, weil hier im Grunde genommen alles zusammenkommt. Einerseits ist es mechanisch sehr schwierig, die menschliche Hand nachzubilden, die zu Recht als Wunderwerk der Natur angesehen wird. Es gibt in diesem Bereich zwar erfolgversprechende Ansätze, aber von Perfektion kann noch keine Rede sein. Andererseits sind mehrfingrige Roboterhände sehr aufwendig und mühsam anzusteuernende Geräte. Deswegen versuchen wir, die Steuerung dieser

Fingerbewegung nicht auf rein algorithmischem Wege zu machen, sondern wir versuchen gewissermaßen so etwas wie einen virtuellen Handschuh zu bauen, der menschliche Handbewegungen zu erkennen in der Lage ist. Das kann man heutzutage technisch schon direkter lösen, indem man Exo-Skelette benutzt, wobei dann die Übersetzung in die oft andere Fingertopologie der Roboterhand ein noch zu lösendes Problem darstellt. Besonders interessieren uns hier berührungsfreie Exo-Skelette, d.h. allein anhand von Kamerabildern etwas über die menschliche Handstellung herauszufinden und dann in der Lage zu sein, ein virtuelles Handmodell (die ist zunächst einfacher zu realisieren als eine echte Roboterhand) durch Handbewegungen vor einer Videokamera anzutreiben. Inzwischen funktioniert das bereits recht gut und mit einer vernünftigen Genauigkeit. Das System arbeitet mit 8 - 10 Bildern pro Sekunde, man kann das also schon in Echtzeit verfolgen, und es läuft auf einem schnellen PC. Man kann also von einem portablen Ansatz reden.

Wir verwenden dabei neuronale Netze. Es handelt es sich um einen hierarchischen Ansatz, das Netz „schaut“ in zunehmend kleineren Bildausschnitten immer detaillierter, um beispielsweise die Lage der Fingerspitzen mit recht guter Zuverlässigkeit zu schätzen. In einem zweiten Schritt wird dann ein dreidimensionales Handmodell angesteuert. Wir glauben, daß neuronale Netze generell als Beitrag für intelligente Mensch-Maschine-Kommunikation eine Reihe interessanter Fähigkeiten anbieten können. Sie sind gewissermaßen ein „natürliches“ Bindeglied zwischen Mensch und Maschine, weil sie Mustererkennung sehr gut unterstützen und so etwas wie aktive Lernfähigkeit bieten. Wir müssen dann zusätzlich Eigenschaften menschlicher Kognition berücksichtigen. Robustheit ist oft mit neuronalen Verfahren verknüpft, der Umgang mit nicht-symbolischer, mit vager oder mit inkonsistenter Information. Was uns besonders interessiert – auch im Rahmen des genannten Sonderforschungsbereichs – sind hybride Systeme, die versuchen, die Stärken symbolischer und subsymbolischer Informationsverarbeitung zusammenzubringen. Wir haben schon jetzt einige Brücken zwischen realen Robotern und simulierten Robotern gesehen. Meine Bemerkung, daß die Roboterhand im Rechner sehr viel einfacher zu si-

mulieren als sie wirklich zu bauen ist, geht in diese Richtung. Wenn man das weiter verfolgt, dann kommt man in einen Bereich – und das führt mich zu der 4. These – daß es ein Biotop für Roboter gibt, das man vielleicht zunächst gar nicht sehr ernst nehmen möchte oder von dem man sagen möchte, daß darin gar keine Roboter sind:

These 4:

Die virtuelle Realität bietet ein zuvor nicht gekanntes „Entwicklungsbiotop“ für eine neue Roboterspezies: Virtuelle Roboter, die allein in Simulationswelten existieren.

Im folgenden möchte ich argumentieren, warum wir virtuelle Roboter als vollwertige Roboter ansehen sollten. Zum einen: Jedermann in der Robotik simuliert seine Roboter in solchen virtuellen Welten, etwa um bei der Bewegungsplanung Kollisionen zu vermeiden. Wir sind es also gewöhnt, daß in vielerlei Hinsicht eine recht gute Deckungsgleichheit zwischen den virtuellen und den realen Robotern besteht. Natürlich kann man den Vorwurf erheben, daß diese virtuellen Roboter ja kein wirkliches Rad an ein Auto schrauben können, aber sie können das Rad sehr wohl an das Auto schrauben, das in ihrer virtuellen Welt existiert. Wir können dasselbe in unserer Welt tun, und wir können es nicht in der Welt der virtuellen Roboter, und diese können es nicht in unserer Welt. Das heißt, diese Roboter sind genauso auf ihre Welt beschränkt wie wir auf die unsere. Interessanterweise besteht auch zwischen diesen Welten eine ziemliche Symmetrie, wenn man mal vom Stecker absieht, den man als eine Restunsymmetrie betrachten könnte. Die virtuelle Welt ist auch noch etwas „grobkörniger“, weil unsere Rechner noch nicht schnell genug sind, um alle Vorgänge in sehr hoher Auflösung zu simulieren. Aber beispielsweise können die Roboter aus einer virtuellen Welt mit ihren Kameras heraussehen in unsere reale Welt, und somit unsere reale Welt so sehen, wie wir über die Bildschirme ihre virtuelle Welt. Auf diese Weise kann zwischen diesen Welten recht viel Interaktion stattfinden, und es ergeben sich eine ganze Reihe von interessanten und wichtigen Anwendungsmöglichkeiten für virtuelle Roboter. Außerdem haben sie eine Reihe von großen Entwicklungsvorteilen ge-

genüber ihren mechanischen Brüdern. Sie können sich gänzlich auf den Informationsverarbeitungsaspekt konzentrieren, weil sie nicht gehemmt sind durch so lästige Probleme wie Energieversorgung, Antriebsschwierigkeiten, Werkstoffgrenzen und Verschleiß. Das Material fällt einem im Rechner zu, man muß noch programmieren, aber es ist alles sehr viel biegsamer, Teile können sich zur Not selbst durchdringen, und man kann Werkstücke in nahezu beliebig feiner Auflösung und hochgenau „herstellen“. Manche Aspekte der realen Robotik sind allerdings sehr schwer zu simulieren, wie etwa Berührbedingungen und die entsprechenden Kräfte, wenn komplexe Gegenstände sich begegnen. Aber für die virtuellen Roboter kann man vieles davon weglassen. Deswegen können sie sich sehr viel rascher entwickeln als die realen Roboter, die sich ihrerseits schon schneller entwickeln als Spezies in der realen Evolution. Diese virtuellen Roboter sind gewissermaßen reine Informationsverarbeitungs- und -darbietungsspezies und können damit optimal an viele Situationen angepaßt werden. Ich zeige Ihnen hier ein Beispiel. Das Bild zeigt Kyoko Date¹, eine animierte Figur, vollständig im Rechner „gerendert“, die schon kommerziell mit großem Erfolg eingesetzt wird. Sie ist optimiert für den japanischen Jugendgeschmack, es gibt auch einen Lebenslauf dazu, sie singt Musikstücke, begleitet die Songs und wird als generelles Werbemittel genutzt. Wie

¹ Kyoko Date - der erste vollkommen synthetische menschliche Star. Über ein Jahr programmierten Software-Entwickler der japanischen Modelagentur HoriPro an ihrem ersten vollkommen synthetischen menschlichen Star. Entstanden ist die 17jährige Popsängerin Kyoko Date, von deren „Echtheit“ selbst ihre Schöpfer überrascht waren. Im Unterschied zu bisherigen virtuellen Stars soll Kyoko kein Computerspiel bevölkern, sondern in der realen Welt Karriere machen. So ließ HoriPro sofort nach der Entwicklung der Figur ein Musikvideo produzieren, das Kyoko Date bei Studioaufnahmen zu ihrer ersten Platte zeigt. Den Clip präsentiert sie beim Shopping, im Gespräch mit ihrem Agenten und bei der Dioptrin, hat Hobbies, Wünsche, Vorstellungen von ihrem Traummann. „Als ich noch kleiner war, spielte ich mit den Kindern im Gaijin-Wald. Meine Klammotten waren jedesmal mit Matsch verkleckert. Und ich habe sogar Fußball gespielt. Jetzt boxe ich für mein Rythmusgefühl und für meine Figur“, erklärt der Star in einem Interview, das Hiro Pro auf Videoband an Journalisten verteilen ließ.

Völlig neu an Kyoko Date ist, daß ihre Schöpfer nicht nur einen neuen virtuellen Star in die Welt setzen, sondern auch noch den gesamten Kontext dazu erfinden: den Star aus einem Guß. Mit Vergangenheit, mit Sozialisation, mit Talenten und - wenn es nach Hori Pro geht - auch noch mit jeder Menge Zukunft. Die Karrieremananger von Frau Date und ihre Software-Entwickler arbeiten bereits daran, die junge Japanerin in Talk Shows auftreten und mitreden zu lassen. Hori Pro möchte erreichen, daß sich das Virtuelle nicht mehr vom Realen der Medien unterscheidet. Oder sind die Medien längst virtuell geworden? Falls das Pilotprojekt Virtual Kids erfolgreich sein sollte, sieht HoriPrio hier den Markt der Zukunft: Nachrichtensprecher, Moderatoren und Schauspieler, die nicht mehr schlafen müssen, nicht mehr wegen Krankheit ausfallen, nicht älter werden und keine Honorarforderungen mehr stellen.

Sie sehen, ist die Rendering-Qualität schon ganz brauchbar, auch wenn Kyoko Date noch nicht hundertprozentig echt aussieht. Erste Anfänge von künstlichen Schauspielern gibt es bereits jetzt und ich vermute, dieser Trend zum künstlichen Schauspieler wird sich fortsetzen. D.h. ein Arbeitsplatz, bei dem Menschen ihr Tun interessanterweise gänzlich in eine virtuelle Welt hineingeben, nämlich die virtuelle Welt der Bühne, des Darstellens einer fiktiven Geschichte, und das sehr gut machen, wird nun sozusagen von den „richtigen“ virtuellen Wesen, von den virtuellen Robotern besiedelt. Sicherlich ist es auch für Regisseure u. U. interessant, mit solchen, sehr geduldigen, Komparsen zu arbeiten, wobei allerdings auch da die Kommunikation wieder eine große Rolle spielt. Der Regisseur wird sich nicht immer mit einem Heer von Programmierern beschäftigen wollen, nur um seinen Film zustande zu bekommen. Sicherlich werden sich die Filmemacher mehr Komfort wünschen, und dies wird dazu führen, daß solche virtuellen Roboter einigermaßen ansprechbar werden und ihre Drehbücher automatisch lernen.

Ein anderer Bereich sind die „sozialen“ Roboter. Wir hatten vorhin das Beispiel der alten Frau, die zu Hause eine Ansprache braucht. Wie wäre es – wenn das mit einem realen Menschen partout nicht zu machen sein sollte – wenn es eben einen großen Bildschirm gibt, über den sich beispielsweise eine nette junge Dame stündlich mit der Frau unterhält. Allmählich lernt diese junge Dame auch die Geschichten und die Begebenheiten, die die Frau in ihrem Alltag ihr erzählt, und kann daher „vernünftige“ Antworten geben. Das wäre sicherlich eine gewisse Unterstützung, eine Unterstützung, die wir in kruderer Form heute regelmäßig etwa in Form des Fernsehens ja auch schon annehmen. Das ist auch nur etwas „flach gedrücktes“, das aber überhaupt nicht individuell ist und überhaupt nicht auf unser Lebensgeschehen eingeht. Virtuelle Roboter können das Fernsehen möglicherweise weit übertreffen, wenn es um Unterhaltung im Alltag geht. Sie können dabei sogar aus ihrem Glaskasten heraussehen in die Welt desjenigen, mit dem sie sich unterhalten.

Informationsroboter wären aufgrund ihrer Eigenschaften an viele Dienstleistungssparten sehr gut angepaßt, schon heute gibt es eine ganz

extreme Ausprägung dieser Roboter, die überhaupt nicht mehr so charmant aussehen wie auf dem Bild gerade zu sehen war. Die sind gänzlich körperlos und besitzen nicht einmal mehr einen virtuellen Körper. Es handelt sich dabei um die sogenannten Internet-Roboter, die etwas tun, was typisch für Roboter ist: Sie sammeln Material, in diesem Falle reine Information („Information“ ist manchmal zu hoch gegriffen, sagen wir, sie tragen Daten zusammen, die auf verschiedenen Internet-Sites liegen). Sie beherrschen dabei heute noch lediglich sehr krude Auswahlkriterien. Aber es gibt in diesem Bereich sehr viele Anstrengungen, diese kruden Kriterien allmählich besser und nützlicher zu machen. Und es ist absehbar, daß diese Roboter eine sehr große Bedeutung haben werden. Übrigens spricht man in diesem Bereich häufig von „Agenten“, was im Grunde schon andeutet, daß es sich um so etwas wie „Identitäten“ handelt, die einen eigenen, abgegrenzten Standpunkt „vertreten“. Aber auch wenn virtuelle Welten im Vormarsch sind, wird die Erfahrung der realen Welt stets eine wichtige Grundlage bleiben. Meine fünfte These lautet daher:

These 5:

Anfangspunkt jeglicher Erfassung neuer Daten sind stets Sensoren.

Mit dieser These möchte ich die Bedeutung der realen Welt unterstreichen, denn irgendwo müssen virtuelle Roboter ihre Information her bekommen. Dies erfordert letzten Endes – als Anfangspunkt jeglicher „Nahrungskette“ im Bereich von Informationen – Sensoren. Mit Sensoren ist eine ganze Reihe von Techniken verbunden, etwa das Computersehen, das Sprachverstehen oder so etwas wie Fingerspitzengefühl. Beispielsweise wurden in unserer Arbeitsgruppe in den letzten Jahren taktile Fingerspitzen entwickelt. Andere Entwicklungen betreffen das Computersehen. Wenn ein Mensch vor einem Computerbildschirm sitzt, dann ist es heutzutage kein großer Aufwand mehr, noch eine kleine Kamera zu integrieren (Bei manchen Rechnern ist das schon fest eingebaut). Damit sind die technischen Voraussetzungen gegeben um herauszufinden, wo der Mensch hinschaut, und dies ist für viele Anwendungen von Interesse. Zunächst wird die Lage des Kopfes ermittelt,

dann die Position der Augen, des Mundes, der Nasenflügel. Aus diesen Informationen kann man mit Hilfe von hierarchischen neuronalen Netzen herausfinden, wo sich die Pupille befindet. In einem weiteren Schritt wird daraus dann die Blickrichtung und der Auftreffpunkt des Blickes auf dem Bildschirm geschätzt. An dem Auftreffpunkt des Blickes kann man dann nützliche Informationen einblenden oder den Blickort als ein Kommando auswerten. Sie können sich viele Anwendungen vorstellen, die davon profitieren können, wenn sie wissen, welches Fenster der Bediener gerade betrachtet oder auf welche Zeile er gerade schaut. Das System kann dann einfach intelligenter reagieren und stärker auf die momentane Intention des Benutzers eingehen.

Es gibt natürlich noch andere Rezeptoren, z.B. ist Sprachverstehen ein ganz wichtiges Gebiet, auf dem wir selbst allerdings nicht arbeiten. Künstliche Chemorezeptoren sind wieder eine ganz andere Dimension, die sicherlich beispielsweise in der Nahrungsmittelindustrie eine große Bedeutung haben wird.

Interessanterweise werden die virtuellen Roboter in manchen Bereichen sogar ohne eine sensorische Verankerung auskommen, z.B. werden virtuelle Roboter im Internet z. T. Sehvermögen haben, ohne in unsere Welt hinausschauen zu müssen und ohne Kameras zu besitzen. Der Grund sind beispielsweise Bilddatenbanken, in denen große Mengen an Bildmaterial liegen. Natürlich macht es für solche Such-Roboter Sinn, wenn sie Bilder, die auf der Festplatte liegen, „anschauen“ und den Bildinhalt verstehen können. Dafür müssen diese Bilder nie jemals auf einem Bildschirm dargestellt worden sein. Es wird also Wesen geben, die in diesem virtuellen Raum, nicht physikalisch instanziiert, über eine Sensorik Muster wahrnehmen können. Und das wird nicht aufhören bei den Bildern, wie wir sie kennen, sondern das wird sich in andere Datenwelten hinein entwickeln. Denken Sie etwa an 3-D-Daten aus Gehirntomogrammen oder an Röntgenaufnahmen. Das führt bis hin zu gänzlich abstrakten Daten, die man für den Menschen erst künstlich visualisieren muß und für die man versuchen kann, direkt verarbeitende visuelle Systeme zu bauen.

Damit komme ich allmählich zur vorletzten These, die die vielleicht im Raum stehende Befürchtung, daß uns Roboter in ihren Fähigkeiten übertreffen oder erreichen können, noch einmal zum Ausdruck bringt:

These 6:

Langfristig könnten Roboter viele menschliche Fähigkeiten erreichen bzw. übertreffen.

Ich glaube, das wird schon passieren, denn zum einen haben uns Roboter in sehr engen Bereichen ja schon übertroffen; und zum anderen gibt es Roboter erst seit etwa 50 Jahren, was wirklich kurz ist. Wenn wir weitere 50 Jahre in die Zukunft blicken, wird der Fortschritt vermutlich ein bißchen größer sein als in den letzten 50 Jahren. Es wird nicht ausbleiben, daß viele Fähigkeiten dann von Maschinen bereitgestellt werden. Allerdings werden diese uns auch helfen, den wirklichen Schwierigkeitsgrad dessen, was wir selber beherrschen, genauer auszuloten. Wir werden vielleicht feststellen, daß manche Sachen sich sehr hartnäckig einer Erreichung oder einer Übertreffung durch Maschinen entziehen. Wir sollten uns deswegen darauf vorbereiten, daß Arbeit den Status einer Freizeitbeschäftigung bekommen kann, etwa zur Pflege sozialer Kontakte, was ja heute schon eine wichtige Rolle von Arbeit ist. Und ehrlich gesagt, ich habe den Eindruck, wir ertappen uns dabei, daß bestimmte Sachen einfach kulturell tief verankert, aber doch relativ willkürlich sind. Wir sind es lediglich historisch gewohnt, daß Arbeitsleistung erbracht, und damit der Lebensunterhalt „verdient“ wird. Und daß dies eine wichtige moralische Legitimation dafür ist, daß man auch von der Gemeinschaft wieder was zurück bekommt. Warum soll das aber immer so sein? Wenn wir in die Aristokratie ein paar hundert Jahre zurückschauen, so war das dort immer schon anders. Die Fürsten und Grafen wären nie auf die Idee gekommen zu arbeiten, damit sie ihren Lebensunterhalt verdienen. Wir kommen dank der Roboter möglicherweise in eine Situation wie die alten Griechen, die sich den Wissenschaften und anderen schönen Dingen widmen konnten, weil sie dienstbare Sklaven hatten. Wir werden eines Tages Roboter haben. Das große Problem wird sein, daß wir nicht besonders gut darauf vorbereitet sind

– glaube ich zumindest – neue Verteilungskriterien so einzuführen, daß es nicht an sehr vielen Ecken knirscht und knackt. Das ist wahrscheinlich die Aufgabe, die in den nächsten Jahrzehnten zu lösen ist. Es wird immer noch Bereiche geben, in denen menschliche Tätigkeiten nicht ohne weiteres ersetzt werden können. Aber auf breiter Basis spricht einiges dafür, daß, wenn wir es wollen - und wie weit wir dabei das Steuer in der Hand haben, ist ebenfalls nicht klar -, Roboter an vielen Stellen Menschen ersetzen werden können. Auch diese These möchte ich etwas relativieren, aber möglicherweise in einer etwas provokativen Art und Weise, das müssen Sie entscheiden:

These 7:

Künftige Roboter werden so hervorragend an uns Menschen angepaßt sein, daß wir sie uns selber zurechnen werden.

Wir haben ja die große Fähigkeit, alles, was gut funktioniert, uns selber zuzurechnen. Und das werden wir vielleicht auch mit den Robotern machen. Vieles was wir im vorigen Vortrag kennengelernt haben, bietet hervorragende Beispiele für diese mögliche Entwicklung. Es gibt heute schon vereinzelte „Roboterkörperteile“, die in vielen Menschen segenreiche Dienste verrichten. Denken Sie etwa an den Herzschrittmacher, in dem typischerweise 150.000 Programmzeilen eines komplexen Algorithmus den künstlichen Rhythmus unter den vielen Bedingungen, die da einzuhalten sind, generieren. Es gibt heute die ersten erfolgreichen Implantate, um bestimmte Tremorformen im Gehirn zu reduzieren. Typischerweise ist man als Mensch beim Herzschrittmacher vielleicht toleranter, bei den Gehirnimplantaten ist man etwas empfindlicher. Selbst die Brille ist ein sehr simpler Kunstkörperteil, den man allerdings noch nicht Roboter nennen würde. Ein schon gegangener Schritt ist die künstliche Augenlinse bei Grauem Star. Wahrscheinlich wird man irgendwann daran denken, auch noch einen kleinen LCD-Schirm zu integrieren, daß man in der Brille, wenn man noch einen externen Speicher hat, schnell mal einen Blick in die Datenbasis werfen kann. Manche Leute werden das haben wollen. Andere werden vielleicht meinen, es beruflich unbedingt zu brauchen. Vielleicht werden

wir demnächst Handys in Hörgeräte implantieren. Sie sehen, es sind viele, aber gut vorstellbare kleine Schritte, – die Systeme sind sozusagen „einmarschreif“. Und wir denken üblicherweise in Silicium, in Blech und in Eisen. Die Biotechnologen arbeiten ja auch! Und die werden „für uns“ neue Materialien für diese Implantat-Technologien erschließen. Es gibt heute schon künstliche Knochen, die in Wunschwuchsformen gezüchtet werden. Man ist dabei, Technologien zu entwickeln, um künstliche Gewebe – etwa Stücke künstlicher Haut – zu erzeugen. Das wäre etwa nach Verbrennungen sehr hilfreich. Ich weiß nicht, bis in welche Bereiche hinein man das letzten Endes als „hilfreich“ empfinden wird. Aber immer, wenn das Leben eines Patienten bedroht ist und ihm so eine Technologie helfen kann, wird das in einem anderen Licht erscheinen, als wenn man aus der theoretischen Distanz urteilt.

Es könnte sehr gut sein, und das soll der letzte Punkt sein, daß sich allmählich das eine oder andere kleine Zusatzteil in uns wiederfindet. Das könnte immer mehr werden und vielleicht sind in 250 Jahren – sagen Sie irgendeine Prozentzahl – 30, 40, 50 % unserer Körperteile durch vermeintlich oder wirklich bessere Kunstprodukte ersetzt worden. Wir wären dann etwas wie ein Mischwesen, halb Kunstwerk, halb natürlich „gebaut“. Vielleicht gibt es dann Roboter in dem Sinne gar nicht mehr, weil wir das nicht mehr als etwas „anderes“ ansehen. Das ist jetzt die futuristische These, und die wollte ich an den Schluß stellen, weil wir jetzt sicherlich noch Zeit zum Diskutieren haben werden. Vielen Dank.

Autorenverzeichnis

Dr. Michael *Decker*, Europäische Akademie zur Erforschung von Folgen wissenschaftlich-technischer Entwicklungen Bad Neuenahr-Ahrweiler GmbH

Professor Dr.-Ing. Rüdiger *Dillmann*, Institut für Prozeßrechentechnik und Robotik, Universität Karlsruhe

Professor Dr.-Ing. Rolf *Eckmiller*, Institut für Informatik VI - Neuroinformatik, Universität Bonn

Professor Dr. Gudela *Grote*, Institut für Arbeitspsychologie, ETH-Zentrum Zürich

Professor Dr. Fritjof *Haft*, Lehrstuhl für Strafrecht, Strafprozeßrecht, Rechtsphilosophie und Rechtsinformatik, Universität Tübingen

Professor Dr. Peter *Janich*, Lehrstuhl 1 für Philosophie, Universität Marburg

Priv.-Doz. Dr. Thomas *Metzinger*, Philosophy Department, University of California at San Diego

Professor Dr. Helge *Ritter*, Technische Fakultät - Neuroinformatik, Universität Bielefeld

Professor Dr.-Ing. Dr. h.c. Rolf-Dieter *Schraft*, Fraunhofer-Institut für Produktionstechnik und Automatisierung Stuttgart

In der *Grauen Reihe* sind bisher erschienen:

- Nr. 1 Technikfolgenabschätzung: Konzeptionen im Überblick, Carl Friedrich Gethmann und Armin Grunwald, 9/96; 2. Aufl. 7/98
- Nr. 2 Umweltprobleme und globaler Wandel als Thema der Ethik in Deutschland, Carl Friedrich Gethmann, 9/96; 2. Aufl. 10/98
- Nr. 3 Sozialverträgliche Technikgestaltung: Kritik des deskriptivistischen Verständnisses, Armin Grunwald, 10/96
- Nr. 4 Technikfolgenbeurteilung der Erforschung und Entwicklung neuer Materialien. Perspektiven in der Verkehrstechnik. Endbericht zum Vorprojekt; Arbeitsgruppe Neue Materialien, 1/97
- Nr. 5 Zur Wissenschaftstheorie der Genetik. Materialien zum Genbegriff, Mathias Gutmann und Peter Janich, 4/97
- Nr. 6 Klimavorhersage und -vorsorge, Stephan Lingner und Carl Friedrich Gethmann, 7/97
- Nr. 7 Xenotransplantation. Ethische Fragen und Probleme, Jan P. Beckmann, 7/97
- Nr. 8 Perspektiven der Robotik. Überlegungen zur Ersetzbarkeit des Menschen, Michael Decker, 11/97
- Nr. 9 Philosophie in Rußland. Tendenzen und Perspektiven, Carl Friedrich Gethmann und Nikolaj Plotnikov, 5/98
- Nr. 10 Technikfolgenbeurteilung in Ländern Mittel- und Osteuropas, Gerhard Banse (Hrsg.); 6/98
- Nr. 11 Biodiversitätsforschung in Deutschland. Potentiale und Perspektiven, Mathias Gutmann und Wilhelm Barthlott (Hrsg.); 11/98
- Nr. 12 Biodiversität als Problem der Naturethik. Literaturreview und Bibliographie, Thorsten Galert, 12/98

- Nr. 13 Geistiges Eigentum und Copyright im multimedialen Zeitalter. Positionen, Probleme, Perspektiven; Gerhard Banse, Christian J. Langenbach (Hrsg.); 2/99
- Nr. 14 Materials Science in Europe; Karl-Michael Nigge; 3/99
- Nr. 15 Modelling Climate Change and its Economic Consequences. A Review, Meinhard Schröder and Stephan Lingner (eds.), 6/99

