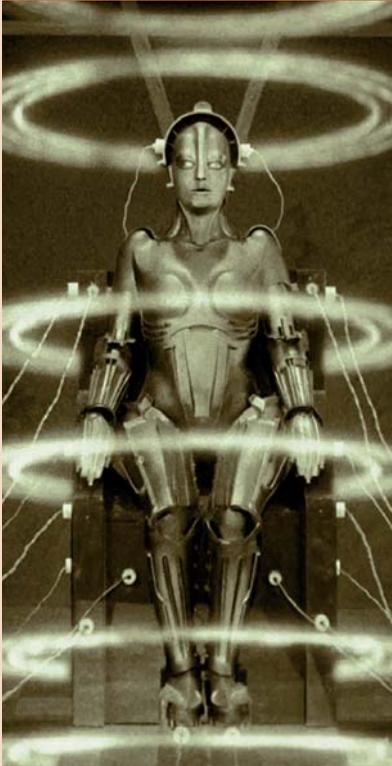


Malte-Christian Gruber
Jochen Bung
Sascha Ziemann (Hg.)



Beiträge zur Rechts-,
Gesellschafts- und Kulturkritik

Autonome Automaten

Künstliche Körper und artifizielle Agenten
in der technisierten Gesellschaft

2. Auflage



BWV • BERLINER WISSENSCHAFTS-VERLAG

Autonome Automaten

Künstliche Körper und artifizielle Agenten
in der technisierten Gesellschaft

Beiträge zur Rechts-, Gesellschafts- und Kulturkritik

Band 12

Begründet von Gisela Engel (†)

Herausgegeben von Malte-Christian Gruber

Wissenschaftlicher Beirat:

Sonja Buckel (Kassel)

Gabriele Budach (Southampton)

Jochen Bung (Passau)

Gisela Engel (Frankfurt a. M.) (†)

Andreas Fischer-Lescano (Bremen)

Tanja Michalsky (Berlin)

Susanne Scholz (Frankfurt a. M.)

Malte-Christian Gruber, Jochen Bung &
Sascha Ziemann (Hg.)

Autonome Automaten

Künstliche Körper und artifizielle Agenten in
der technisierten Gesellschaft

2. Auflage



BWV • BERLINER WISSENSCHAFTS-VERLAG

Bibliografische Informationen der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation
in der Deutschen Nationalbibliografie;
detaillierte bibliografische Daten sind im Internet über
<http://dnb.ddb.de> abrufbar

Impressum

Malte-Christian Gruber, Jochen Bung, Sascha Ziemann (Hg.)
„Autonome Automaten. Künstliche Körper und artifizielle Agenten
in der technisierten Gesellschaft“
[= Beiträge zur Rechts-, Gesellschafts- und Kulturkritik, Band 12]

ISBN 978-3-8305-2056-6

Titelbild: siehe Bildnachweis, S. 300

© 2015 BWV · BERLINER WISSENSCHAFTS-VERLAG GmbH,
Markgrafenstraße 12–14, 10969 Berlin
E-Mail: bwv@bwv-verlag.de, Internet: <http://www.bwv-verlag.de>
Alle Rechte, auch die des Nachdrucks von Auszügen,
der photomechanischen Wiedergabe und der Übersetzung, vorbehalten.

Inhalt

| | |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Vorwort | 7 |
| Im Gedenken an Gisela Engel | 9 |
| <i>Malte-Christian Gruber, Jochen Bung und Sascha Ziemann</i> Editorial: Auf dem Weg zur Automatenautonomie? | 11 |
| I. AUTOMATISIERTE AUTONOMIE? | 15 |
| <i>Jochen Bung</i> Können Artefakte denken? | 17 |
| <i>Dirk Fabricius</i> Die Freiheit, $2*2=5$ zu rechnen | 29 |
| <i>Stephan Meyer</i> Staatliche Willensträgerschaft als Voraussetzung juristischer Normgeltung | 45 |
| II. KÜNSTLICHE KÖRPER / ARTIFIZIELLE AGENTEN | 59 |
| <i>Sebastian Sierra Barra und Martin Deschauer</i> Versuch einer nichtmenschlichen Anthropologie von Intelligenz | 61 |
| <i>Hyo Yoon Kang</i> Autonomie/Code. Überlegungen zur Software-Rhetorik in künstlicher Intelligenz, Postgenomik und Recht | 79 |
| <i>Julia von Dall'Armi</i> Menschliche Macht über Maschinen oder maschinelle Macht über Menschen? Die literarische Realisierung des (Gen-)Technikdiskurses in der aktuellen Gegenwartsliteratur | 93 |
| <i>Sebastian Klinge und Laurens Schlicht</i> Differenz <i>Automat</i> . Ein Ausschnitt aus der Geschichte des Menschen: Taubstummenforschung (um 1800) und Kybernetik (1946–1953) | 103 |

| | |
|------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| III. HERAUSFORDERUNGEN DER TECHNISIERTEN GESELLSCHAFT | 135 |
| <i>Frank Dittmann</i> Was ist, wenn alles denkt? – Eine historische Annäherung | 137 |
| <i>Jan-Philipp Günther</i> Embodied Robots – Zeit für eine rechtliche Neubewertung? | 155 |
| <i>Susanne Beck</i> Technisierung des Menschen – Vermenschlichung der Technik. Neue Herausforderungen für das rechtliche Konzept „Verantwortung“ | 173 |
| IV. MENSCHEN- / MASCHINENRECHTE? | 189 |
| <i>Malte-Christian Gruber</i> Was spricht gegen Maschinenrechte? | 191 |
| <i>Kirsten Brukamp</i> Patientenautonomie angesichts moderner Neurotechnologien wie der tiefen Hirnstimulation | 207 |
| <i>Eric Hilgendorf</i> Roboterprostitution. Gedankenspiele zwischen Recht und Moral | 221 |
| V. TECHNOPATHIE | 235 |
| <i>Sabine Müller</i> Verführung und Verderben durch Automaten? Weibliche Maschinenmenschen und „Orient“-Topoi in <i>Metropolis</i> und <i>Gothic</i> | 237 |
| <i>Martin Uebelhart</i> Wunschmaschinen und Maschinen(alb)träume. Anmerkungen zu Stanislaw Lem (1921–2006) | 249 |
| <i>Jutta Weber</i> Autonome und ferngesteuerte Kampfdrohnen. Über ‚Revolution in Military Affairs‘ und den Traum vom automatisierten Krieg | 267 |
| Autorinnen und Autoren | 291 |
| Bildnachweis | 300 |

Vorwort zur 2. Auflage

Die bio-sozio-technischen Beziehungen von Menschen und Maschinen geraten zunehmend in den Mittelpunkt gesellschafts- und kulturkritischer Debatten um die Zukünfte neuer Technologien. Auch im juristischen Diskurs hat die Aufmerksamkeit für das Thema seit Erscheinen der ersten Auflage dieses Bandes weiter zugenommen. Mit der nunmehr im Berliner Wissenschaftsverlag erschienenen zweiten Auflage hat die Schriftenreihe „Beiträge zur Rechts-, Gesellschafts- und Kulturkritik“ eine neue Publikationsstätte gefunden, die in geradezu entsprechender Weise in die Zukunft weist.

Frankfurt am Main,
im Februar 2015

Malte Gruber
Jochen Bung
Sascha Ziemann

Vorwort zur 1. Auflage

„Autonome Automaten“ lautete das Thema der Frankfurter Jahrestagung der Kritischen Reihe am 15. und 16. Juli 2011, deren Fokus auf der besonderen gesellschaftlichen Bedeutung lag, die so genannten Zukunftstechniken bereits in der Gegenwart zukommt. Künstliche Intelligenz, künstliches Bewusstsein und auch künstliches Leben gehören mehr denn je zu den prominenten Forschungsgegenständen der heutigen Technikwissenschaften, und gerade die Fortschritte im Bereich der Robotik und der zunehmend sozial integrierten Informationstechnologien werfen Fragen nach dem gesellschaftlichen Umgang mit deren Gegenständen auf. Roboter und informationstechnische Artefakte wie Software- oder Netzagenten scheinen nicht prinzipiell auf ihren ursprünglichen Status als bloße Objekte oder Dinge beschränkt, sondern begegnen Menschen mitunter als widerständige Wesen, als handelnde Akteure, womöglich sogar als soziale Interaktionspartner. Solche Begegnungen könnten nicht zuletzt auch das Selbstbild des Menschen und der Gesellschaft beeinflussen.

Dass das Thema der „Autonomen Automaten“ dabei kaum erschöpfend zu behandeln ist, vielmehr eine fortgesetzte interdisziplinäre Verständigung erfordert, zeigt sich schon an der gewachsenen Aufmerksamkeit und der Anzahl akademischer Veranstaltungen zum Themenkreis der Informationstechnologie, Robotik und Recht. Schon im August desselben Jahres 2011 fand – ebenfalls in Frankfurt am Main – der 25. Weltkongress der Internationalen Vereinigung für Rechts- und So-

zialphilosophie zum Thema „Recht, Wissenschaft und Technik“ statt. Dieser bot einen enger an der Rechtstheorie orientierten, allerdings um einen internationalen Teilnehmerkreis erweiterten Rahmen, um die Diskussion der mit den „Autonomen Automaten“ aufgeworfenen Fragen fortzusetzen. Insbesondere das Problem der brüchig werdenden Dichotomie von Mensch und Maschine konnte somit nochmals thematisiert und in einem von Vagias Karavas (Luzern) und Malte Gruber unter dem Titel „Recht am technisierten Körper/ Recht an verkörperter Technik“ organisierten Workshop eingehend behandelt werden.

Der vorliegende Band versammelt ausgewählte Beiträge der beiden genannten Veranstaltungen des Sommers 2011. Die Kritische Reihe wurde auch in diesem Jahr wieder von der Vereinigung von Freunden und Förderern der Johann Wolfgang Goethe Universität Frankfurt am Main e.V. gefördert. Ihr gilt unser besonderer Dank.

Frankfurt am Main,
im April 2013

Malte Gruber
Jochen Bung
Sascha Ziemann

Im Gedenken an Gisela Engel

Kurz vor Drucklegung dieses zwölften Bandes ist Gisela Engel, die die *Kritische Reihe* ins Leben gerufen und ununterbrochen begleitet hat, verstorben. Die Betroffenheit war und ist noch immer groß unter den Autorinnen und Autoren, den Kolleginnen und Kollegen, die Gisela Engel allesamt als eine engagierte, großzügige, zuverlässige und hilfsbereite Persönlichkeit kannten. Gisela Engel war von 1974 bis 2008 wissenschaftliche Mitarbeiterin am Institut für England- und Amerika-studien der Frankfurter Goethe-Universität; seit 1993 war sie außerdem Mitarbeiterin am Zentrum zur Erforschung der Frühen Neuzeit. Neben ihren Tätigkeiten als Herausgeberin und als Wissenschaftliche Beirätin dieser Buchreihe war sie Mit-herausgeberin und verantwortliche Redakteurin der Zeitschrift *Zeitsprünge. Forschungen zur Frühen Neuzeit* und der Reihe *Frankfurter Kulturwissenschaftliche Beiträge*. Studierende, Nachwuchswissenschaftler und Kolleginnen und Kollegen schätzten Gisela Engel als vorbildliche Lehrerin und treue Wegbegleiterin, vor allem als liebenswürdigen Menschen. Wir verdanken ihr viel. Sie wird uns immer in Erinnerung bleiben.

Frankfurt am Main,
im April 2014
Malte Gruber

Editorial: Auf dem Weg zur Automatenautonomie?

Die Erschaffung künstlicher Intelligenzen, Roboter und Humanoiden hat sich von einem ursprünglichen Motiv der literarischen Science Fiction in eine Utopie der heutigen Technikwissenschaften, möglicherweise auch in ein realistisches Forschungsziel verwandelt. Anders als in den Anfängen der Forschung im Bereich der Künstlichen Intelligenz bemüht man sich heute nicht mehr nur darum, menschliche Intelligenz auf der Basis einfacher Algorithmen „im Computer“ zu simulieren. Vielmehr geht es auch darum, künstliches Bewusstsein in die Lage zu versetzen, Wahrnehmungen und Erfahrungen selbständig zu organisieren und sich flexibel zu verhalten (vgl. etwa Metzinger 2005; Dennett 2005; Birnbacher 2005).

Solche Leistungen erfordern, wie insbesondere jüngere neurowissenschaftliche Theorien nahelegen, mehr als kluge Programme. Denn Geist und Bewusstsein sind keine bloße „Software“, sondern bedürfen darüber hinaus eines „organischen“ Körpers (siehe Gallese 2005; Rizzolatti/ Sinigaglia 2008). Folgerichtig beschäftigt sich die neuere Forschung verstärkt mit der Konstruktion von verkörperter Intelligenz, die zuallererst in der Fähigkeit gesucht wird, sinnliche Wahrnehmungen zu verarbeiten (Pfeifer/ Scheier 1999) und vielleicht sogar eine menschenähnliche Erscheinung zu erreichen (zur daran orientierten *android science* Ishiguro 2006).

Dahinter steht die Absicht, Maschinen als artifizielle Agenten zu sozialen Interaktionspartnern zu machen und schließlich auch zu verantwortlich handelnden Akteuren der technisierten Gesellschaft. Womöglich bedarf es dazu aber noch eines weiteren Schrittes zurück zu den Ursprüngen menschlicher Intelligenz, nämlich zum Verständnis dessen, was Leben überhaupt sei. Spätestens aber, wenn es gelingen sollte, künstliches Leben zu erzeugen, das sich zu komplexeren Formen des künstlichen Bewusstseins weiterentwickeln lässt (vgl. Bedau 2003), kommt es zur Begegnung mit einer Vielzahl neuer künstlicher Akteure – einer Begegnung, die nicht zuletzt auch das Bild des Menschen und der Gesellschaft verändern könnte (vgl. exemplarisch Rammert 2007; Latour 2007). Selbst die bislang für unverbrüchlich gehaltene Differenz von Mensch und Maschine würde dadurch angreifbar (dazu jüngst Fabricius 2011).

Solche Aussichten beflügeln die wissenschaftliche und künstlerische Phantasie, die sich in sogenannter Pop Science niederschlägt und zwischen einer besonderen Gattung der Unterhaltungsliteratur und seriöser Wissenschaftskommunikation changiert. Die Arbeit der Popsceiencefiktionäre bleibt nicht auf die Ursprungsdisziplin

ihrer Fiktion beschränkt, sondern führt zu neuen Theorien oder Theorietransfers in den etablierten Wissenschaften (beispielhaft hierfür Clark 2003). Im Bereich der Gesellschaftstheorie führen solche fiktionalen Inspirationen zu neuen sozialen Modellen, die insbesondere Fragen der Subjektivität, der Handlungs- und der Rechtsfähigkeit, nicht nur menschlicher, sondern auch nichtmenschlicher Wesen, thematisieren (Rammert/ Schulz-Schaeffer 2002; Teubner 2006; Koops/ Hildebrandt/ Jaquet-Chiffelle 2010; Gruber 2012). Ferner ermöglichen sie neue Anthropologien: Das Menschenbild des selbstbestimmten Individuums weicht dem eines in seine technische Umwelt eingebundenen Wesens, dessen körperliche und geistige Fähigkeiten artifiziell erweiterbar, aber auch immer schwerer beherrschbar sind (Merkel et al. 2007; Müller/ Clausen/ Maio 2009).

Während Mensch-Maschine-Schnittstellen, Neuroimplantate und technologische oder pharmakologische „Neuro-Enhancements“ vor allem Fragen nach Umfang und Grenzen der menschlichen Persönlichkeit aufwerfen (dazu etwa Gruber 2009), bringen andere Technologien gänzlich neue Akteure mit eigener Wirkungsmacht hervor: Software- und Netzagenten, oder auch die zunehmend in physische Umwelten eingebetteten Programme sogenannter Cyber-Physical oder Ambient Intelligence Systems treten menschlichen Akteuren zunehmend als „widerständige“ Wesen gegenüber (vgl. Herzog 2009). In der dialektischen Dynamik von Widerstand und Anpassung könnten sie sich als autonome Automaten erweisen und dabei mitunter zu Zeugen eines allgemeinen Autonomieverlustes des technisierten Menschen werden.

Die folgenden Beiträge stellen jeweils eigenständige Annäherungen an das Thema dar, die von einer gemeinsamen, fachübergreifenden Diskussion inspiriert und in ihrem jeweiligen fachspezifischen Kontext ausgeführt worden sind. Das Gesamtprojekt kann insoweit keinen Anspruch auf eine erschöpfende Behandlung eines derart weitgespannten Feldes erheben, soll aber sehr wohl einige Schwerpunkte erkunden und die Problematiken der automatisierten Autonomie, der artifiziellen Agenten, der technisierten Gesellschaft, insbesondere auch mit Blick auf deren Pathologien und ihre Aussichten auf (Re-)Humanisierung, näher beleuchten.

Literatur

- Bedau, Mark A. (2003): „Artificial life: organization, adaptation and complexity from the bottom up“, in: *TRENDS in Cognitive Sciences* 7, S. 505–512.
- Birnbacher, Dieter (2005): „Künstliches Bewusstsein“, in: T. Metzinger (Hg.): *Bewusstsein – Beiträge aus der Gegenwartsphilosophie*, 5. Aufl., Paderborn: Mentis 2005, S. 713–729.

- Clark, Andy (2003): *Natural-Born Cyborgs. Minds, Technologies, and the Future of Human Intelligence*, Oxford: Oxford University Press.
- Dennett, Daniel C. (2005): „COG: Schritte in Richtung auf Bewußtsein in Robotern“, in: T. Metzinger (Hg.): *Bewusstsein – Beiträge aus der Gegenwartsphilosophie*, 5. Aufl., Paderborn: Mentis 2005, S. 691–712.
- Fabricius, Dirk (2011): „Der Mensch ist eine Maschine. Maschinen können frei sein.“, in: *Archiv für Rechts- und Sozialphilosophie* 97, S. 556–573.
- Gallese, Vittorio (2005): „Embodied simulation: From neurons to phenomenal experience“, in: *Phenomenology and the Cognitive Sciences* 4, S. 23–48.
- Gruber, Malte-Christian (2009): „Neurotechnologisch modifizierte Rechtssubjektivität: Persönlichkeitsschutz im Recht der Neuro- und Informationstechnologie“, in: O. Müller / J. Clausen / G. Maio (Hg.), *Das technisierte Gehirn. Neurotechnologien als Herausforderung für Ethik und Anthropologie*, Paderborn: Mentis, S. 87–104.
- Gruber, Malte-Christian (2012): „Rechtssubjekte und Teilrechtssubjekte des elektronischen Geschäftsverkehrs“, in: S. Beck (Hg.), *Jenseits von Mensch und Maschine. Ethische und rechtliche Fragen zum Umgang mit Robotern, Künstlicher Intelligenz und Cyborgs*, S. 133–160.
- Herzog, Otthein et al. (2009): „Intelligente Objekte – klein, vernetzt, sensitiv“, Stellungnahme acatech – Deutsche Akademie der Technikwissenschaften: acatech bezieht Position – Nr. 5, München.
- Ishiguro, Hiroshi (2006): „Android science: conscious and subconscious recognition“, in: *Connection Science* 18, S. 319–332.
- Koops, Bert-Jaap / Hildebrandt, Mireille / Jaquet-Chiffelle, David-Olivier, „Bridging the Accountability Gap: Rights for New Entities in the Information Society?“, in: *Minnesota Journal of Law, Science & Technology* 11, S. 497–561.
- Latour, Bruno (2007): *Eine neue Soziologie für eine neue Gesellschaft. Einführung in die Akteur-Netzwerk-Theorie*, Frankfurt/M.: Suhrkamp.
- Merkel, Reinhard et al. (2007): *Intervening in the Brain: Changing Psyche and Society*, Berlin: Springer.
- Metzinger, Thomas (2005): „Einleitung“, in: ders. (Hg.), *Bewusstsein – Beiträge aus der Gegenwartsphilosophie*, 5. Aufl., Paderborn: Mentis 2005, S. 713–729 S. 683–690.
- Müller, Oliver / Clausen, Jens / Maio, Giovanni (2009): „Der technische Zugriff auf das menschliche Gehirn. Methoden – Herausforderungen – Reflexionen“, in: dies. (Hg.), *Das technisierte Gehirn. Neurotechnologien als Herausforderung für Ethik und Anthropologie*, Paderborn: Mentis, S. 11–19.
- Pfeifer, Rolf/ Scheier, Christian (1999): *Understanding Intelligence*, Cambridge (Mass.): MIT Press.

- Rammert, Werner (2007): *Technik – Handeln – Wissen. Zu einer pragmatistischen Technik- und Sozialtheorie*, Wiesbaden: VS Verlag für Sozialwissenschaften.
- Rammert, Werner / Schulz-Schaeffer, Ingo (2002): „Technik und Handeln. Wenn soziales Handeln sich auf menschliches Verhalten und technische Abläufe verteilt“, in: dies. (Hg.), *Können Maschinen handeln? Soziologische Beiträge zum Verhältnis von Mensch und Technik*, Frankfurt/M.: Campus, S. 11–64.
- Rizzolatti, Giacomo / Sinigaglia, Corrado (2008): *Empathie und Spiegelneurone. Die biologische Basis des Mitgefühls*, Frankfurt/M.: Suhrkamp.
- Teubner, Gunther (2006): „Elektronische Agenten und große Menschenaffen: Zur Ausweitung des Akteurstatus in Recht und Politik“, in: *Zeitschrift für Rechtssoziologie* 27, S. 5–30.

I.

AUTOMATISIERTE AUTONOMIE?

Können Artefakte denken?

Der Beitrag stellt im Anschluss an Davidsons Kritik des Turing-Tests die Verkörperung von Maschinen als notwendige Bedingung der Zuschreibung mentaler Zustände heraus. Philip K. Dicks fiktionalem Empathie-Test sowie Merleau-Pontys Phänomenologie des situationsräumlichen Leibes lassen sich weitere wichtige Hinweise für ein reichhaltigeres Verständnis einer adäquaten und zurechnungsfähigen artifiziellen Lebensform entnehmen.

Der Turing-Test

Die Frage, die ich hier an den Anfang stelle, ist eine Variation der berühmten Frage des britischen Mathematikers Alan Turing: Können Maschinen denken? Im Ausgang von dieser Frage entwickelte Turing 1950 ein weithin beachtetes und in der Folge nach ihm benanntes Experiment, den *Turing-Test* (Turing 1950). Dieses berühmte Experiment ist bis heute Gegenstand lebhafter Diskussionen (lesenswert neuerdings Christian 2011), so dass ich mich, was den Aufbau des Experiments betrifft, auf einige Hinweise beschränken kann. Es war Turings Absicht, die Ausgangsfrage, ob Maschinen denken können, zu ersetzen oder ihr einen präziseren Sinn zu geben. Er misstraute dem normalen Gebrauch der Begriffe der Maschine und des Denkens und meinte, die Untersuchung des Sprachgebrauchs könne keine Grundlage für taugliche Definitionen, sondern allenfalls für statistische Erhebungen liefern. Also schlug er eine andere Methode vor. In dem von ihm als „Imitationsspiel“ bezeichneten Experiment geht es darum herauszufinden, mit welcher Zuverlässigkeit eine hypothetische Versuchsperson ermitteln kann, ob sie mit einem Menschen oder mit einer Maschine kommuniziert, wenn Mensch und Maschine nicht sichtbar sind und mit beiden lediglich ein schriftlicher Austausch von Informationen (nach Turings Vorstellung über einen „Teleprinter“) möglich ist. Je schwieriger es ist, im Hinblick auf die empfangenen Texte zu sagen, ob sie von einem Menschen oder einer Maschine stammen (wobei Turing für die Rolle der Maschine „digitale Computer“ vorgesehen hatte), desto weniger Gründe haben wir für die Annahme, einem der beiden unsichtbaren Kandidaten jene spezifischen mentalen Fähigkeiten abzusprechen, auf die wir uns mit dem Begriff des Denkens beziehen. Mit anderen Worten: Wenn wir

einem Menschen in relevanter Hinsicht Denken zuschreiben und in dieser Hinsicht nicht in der Lage sind, zwischen Mensch und Maschine zu unterscheiden, spricht alles für die Annahme, dass wir insoweit auch die Maschine als denkendes Wesen ansehen können. Der Turing-Test beabsichtigt keine unmittelbare Antwort auf die Frage, ob Maschinen denken können, sondern macht plausibel, dass kommunikativ nicht aufzuschlüsselnde Ununterscheidbarkeiten einen Grund dafür abgeben, auch keine Unterscheidung im Hinblick auf die mentalen Fähigkeiten zu machen.

Modifizierter Turing-Test und Empathie-Test

Es versteht sich, dass man Turings Fragesteller (interrogator) aus Gründen der Fairness eine gewisse Zeit zur Verfügung stellen muss. Er muss eine Reihe von Fragen stellen dürfen, denn wenn er die Frage „Bist du ein Mensch?“ an die unsichtbaren Adressaten X und Y richtet, wird er sowohl von X als auch von Y ein „Ja“ zur Antwort erhalten, und das ist ersichtlich keine ausreichende Grundlage, um Mutmaßungen darüber anzustellen, welche Antwort richtig ist. Der Turing-Test legt damit nahe, dass die Frage, ob Maschinen denken können, durch die Frage ersetzt werden kann, ob Maschinen in der Lage sind, das Imitationsspiel (in einer bestimmten festgelegten Zeit) zu gewinnen, aber diese Interpretation führt womöglich vom Grundgedanken des Experiments weg, denn es ist nicht ganz sicher, was Gewinnen hier heißen könnte. Heißt es, dass die Maschine den Fragesteller erfolgreich darüber täuscht, dass sie eine Maschine ist, dann müsste die erfolgreichste Maschine bewirken, dass der Fragesteller stets sie und nicht den Menschen für den Menschen hält. Aber in diesem Falle würde der Turing-Test erhebliche Probleme für unser Selbstverständnis aufwerfen, die unseren Begriff des Denkens nicht unberührt ließen und damit auch die Deutung des Experiments insgesamt. Andererseits erscheint es schwer vorstellbar, dass die aufs Gewinnen programmierte Maschine sich ab und zu als Maschine zu erkennen gibt, denn es geht ja für sie gerade darum, vom Menschen nicht unterscheidbar zu sein.

Die eben aufgeworfenen Probleme sprechen dafür, das ursprüngliche Arrangement des Turing-Tests abzuändern. In dem veränderten Arrangement gibt es nicht mehr drei, sondern nur noch zwei Akteure: den Fragesteller und den Befragten. Der amerikanische Philosoph Donald Davidson hat in einem Beitrag über den Turing-Test dieses Arrangement als „Modifizierten Test“ vorgeschlagen. Nach seiner Auffassung können wir uns besser auf die Frage des Denkens konzentrieren, wenn wir nur ein einziges Objekt untersuchen und entscheiden müssen, ob dieses Objekt denkt oder nicht (Davidson 2006, S. 148). Veranschaulicht man sich den modifizierten Test, kommt man vielleicht zur Überzeugung, dass eine weitere

Bedingung des ursprünglichen Tests fragwürdig wird, nämlich die Unsichtbarkeit des Gegenübers. Sie sollte in Turings Test verhindern, dass der Fragesteller aufgrund von Äußerlichkeiten („X sieht aus wie ein Mensch“; „Y sieht aus wie ein Computer“) voreingenommen entscheidet, welches seiner beiden Gegenüber das denkende Wesen ist. Die Fähigkeit zu denken könnte etwas sein, das nicht nur uns Menschen auszeichnet, so wie auch Autonomie und Moralität nach Kants Auffassung etwas ist, das nicht nur uns Menschen, sondern alle Wesen auszeichnet, die sich durch Vernunftgründe zu Handlungen bestimmen können, weswegen es auf eine allgemeine Bestimmung des vernünftigen Handelns ankommt und nicht auf eine anthropozentrische Restriktion des Konzepts. Geht man davon aus, dass es für die Frage des Denkens gleichgültig ist, wie ein Wesen aussieht, aus welchem Stoff oder nach welchem Bauplan es gemacht ist, ist die im klassischen Turing-Test scheinbar so essentielle Bedingung der Unsichtbarkeit der Kandidaten hinfällig.

Man sollte andererseits aber die Macht der Unterstellung, dass ein Wesen, um denken zu können, uns irgendwie ähnlich sein muss, nicht unterschätzen. Hält man diese Vorannahme nicht für kontingent oder zumindest (unter den gegebenen kulturellen Bedingungen) für einstweilen unüberwindbar, könnte man vielleicht der folgenden Variante des modifizierten Tests etwas abgewinnen: Nach wie vor gibt es nur zwei Beteiligte, den Fragesteller und den Befragten. Der Befragte ist sichtbar und beobachtbar, und er ist entweder ein Mensch oder ein Artefakt, das einem menschlichen Wesen äußerlich bis aufs Haar gleicht. In der Literatur gibt es dafür eine beeindruckende Schilderung, nämlich den sog. *Voigt-Kampff-Test* in Philip K. Dicks 1968 erschienenem Roman „Do Androids Dream of Electric Sheep?“, der durch Ridley Scotts filmische Adaption „Blade Runner“ von 1982 bekannt geworden ist. Der Voigt-Kampff-Test ist als Empathie-Test angelegt und umgeht dadurch eine kognitivistische Verengung des Verständnisses von geistigen Eigenschaften zugunsten eines reichhaltigeren und ganzheitlicheren Konzepts der Person. Dem Fragesteller in diesem Test geht es nicht um Konversation, sondern um das Eintauchen in tiefere Schichten des Gegenübers. Er will ermitteln, ob es sich um ein natürliches menschliches Wesen mit einer authentischen, selbst erlebten Lebensgeschichte oder um ein humanoides Artefakt, einen „Replikanten“ mit nur kurzer Lebensdauer und simulierter, wesentlich auf Programmierung basierender Erinnerung handelt. Neben moralischen und existentiellen Fragen, die Dicks Roman aufwirft, regt dieser auch zum Nachdenken darüber an, welche Bedeutung dem Umstand zukommt, dass die Maschine im Gegensatz zu Turings Computer nicht nur sichtbar und beobachtbar, sondern auch uns ähnlich ist. Turing selbst hat diesen Aspekt gesehen, aber als irrelevant verworfen: „No engineer or chemist claims to be able to produce a material which is indistinguishable from the human skin. It is possible that at some time this might be done, but even supposing this invention available

we should feel there was little point in trying to make a „thinking machine“ more human by dressing it up in such artificial flesh.“ (Turing 1950, S. 434).

Etwas sagen und etwas meinen

Indem er die Bedeutung der Fleischwerdung der Maschine herunterspielt, verkennt Turing aber möglicherweise einen Punkt, auf den es bei der Frage nach dem Denken der Maschine entscheidend ankommt. Möglicherweise sitzt Turing hier dem Vorurteil einer Tradition des Denkens auf, wonach der Körper des Denkenden für das Denken selbst nicht wesentlich, sondern allenfalls eine zufällige Randbedingung ist. Es mag sein, dass eine Maschine keine menschliche Haut braucht, um Gedanken haben zu können, aber es ist denkbar, dass sie dazu irgendeinen Körper bestimmter Art braucht. Dem scharfen Dualismus zwischen *res cogitans* und *res extensa*, auf dessen Bestätigung es Turing bei der Anordnung seines Experiments ausdrücklich ankommt, könnte eine Fehlannahme über die Natur des Geistigen zugrunde liegen.

Folgt man Davidsons Kritik am Turing-Test, kann die Frage des Denkens nicht unabhängig von einer Interpretation des verbalen und nonverbalen Verhaltens eines Objekts entschieden werden, welches der Interpret beobachten kann. Das heißt nicht, dass es sich um ein Objekt handeln muss, das wie ein Mensch aussieht. Es bereitet keine Schwierigkeiten, sich die Zuschreibung von Gedanken auf der Basis des Verhaltens eines Wesens vorzustellen, das eine metallisch blaue Oberfläche, zwei Antennen auf dem Kopf und acht Extremitäten hat. Schwieriger wird es allerdings, wenn das Objekt wie ein Desktopcomputer aussieht oder ein Computer ist, den man, wie im Turing-Test, gar nicht sieht. Davidsons Argument ist verblüffend einfach: Greift man auf sprachliche Äußerungen als Belegmaterial zurück, dann kommt es für die Frage des Denkens entscheidend auf die Frage der Bedeutung des Geäußerten an. Nur wenn unser Gegenüber, egal ob Mensch oder Maschine, mit dem, was es sagt, etwas meint (etwas Bedeutungsvolles zu sagen beabsichtigt), ist es plausibel, ihm Gedanken zuzuschreiben. Es ist aber ohne weiteres denkbar, dass unser Gegenüber grammatisch wohlgeformte Sätze von sich gibt, ohne damit irgendetwas zu meinen (Davidson 2006, S. 147 f.). Das liegt daran, dass der Artikulationsvorgang als solcher keinen Hinweis darauf enthält, ob er mehr ist als ein Mechanismus zur Generierung grammatisch wohlgeformter Sätze.

Lässt sich die Artikulation, was bei schriftlichen Äußerungen im Stil des Turing-Tests natürlich ausgeschlossen ist, auf einen Gegenstand beziehen, von dem der Interpret annehmen darf, dass der Bezug von dem Sprecher intendiert ist, ist die Bedeutungshaftigkeit des Geäußerten leicht zu erklären. Der Sprecher will etwas sagen und er glaubt, dass er das uns gegenüber erfolgreich durch die Bezugnahme

auf etwas tun kann, von dem er denkt, dass es auch für uns in irgendeiner Weise bemerkenswert ist, z. B., um Quines klassisches Beispiel zu verwenden, ein vorüberhoppelndes Kaninchen (Quine 1960, S. 29). Wenn wir hingegen, um ein hypothetisches Szenario von Hilary Putnam aufzugreifen (Putnam 1990, S. 21), ein Gehirn in einem Bottich mit Nährlösung schwimmen sehen, das mit einem Computer verdrahtet ist, auf dessen Bildschirm das Wort „Kaninchen“ erscheint, dann werden wir nicht sogleich vermuten, dass hier etwas Bedeutungsvolles gesagt wurde, auch wenn zu dem fraglichen Zeitpunkt ein Kaninchen durchs Labor hoppelt.

Wenn andererseits, wie im Falle von Philip K. Dicks Androiden, unser Gegenüber äußerlich einem Menschen bis aufs Haar gleicht, gehen wir sicherlich davon aus und haben auch gute Gründe davon auszugehen, dass es so gestrickt ist wir. Wir vermuten ganz einfach, dass unser Gegenüber in ganz ähnlicher Weise mit Dingen in Berührung gekommen ist und Erfahrungen gemacht hat, wie wir selbst. Etwas anderes können wir zunächst gar nicht tun, insofern gehört die Unterstellung von Ähnlichkeit tatsächlich zu den Grundbedingungen der Kommunikation. Macht unser Gegenüber eine Äußerung über Kaninchen (obwohl momentan kein Kaninchen zu sehen ist), halten wir die Äußerung nicht für bedeutungslos, sondern führen die Bedeutsamkeit des Geäußerten darauf zurück, dass unser Gegenüber im Laufe seines Lebens schon ein paar Mal mit den langohrigen Tieren in Berührung gekommen ist und jene spezifischen Überzeugungen gebildet hat, die notwendig sind, damit man mit einer Äußerung von „Kaninchen“ einen Sinn verbinden und sie anderen als Ausdruck von Gedanken über Kaninchen zu verstehen geben kann. Wir werden erst dann stutzig, wenn es Anzeichen gibt, dass irgendetwas mit der Kaninchentheorie des Gegenübers nicht stimmt, wenn wir etwa feststellen, dass eine Kommunikation über Kaninchen bei unserem Gegenüber zu Schweißausbrüchen, Zittern und angstgeweiteten Augen führt.

In der Welt sein und sich mit der Welt auseinandersetzen

Stellen wir uns nun vor, wir unterhalten uns mit einer Person, die aussieht wie unsere Nachbarin, über Kaninchen. In Wahrheit hat unsere Gesprächspartnerin aber noch nie in ihrem Leben ein Kaninchen gesehen. Sie ist ein vor zwei Wochen fertiggestellter humanoider Roboter, und der Grund, dass dieser Roboter, nennen wir ihn R, sich gleichwohl über Kaninchen äußern kann, ist der, dass ihm Kaninchen-Gedanken einer anderen Person einprogrammiert wurden, und zwar in der Standard-syntax und Standardphonetik des Deutschen, mit der R arbeitet. Der Sinn des von R Geäußerten hängt mithin von dem Überzeugungssystem des Programmierers ab oder von dem einer dritten Person, deren kaninchenbezogene Überzeugungen der

Programmierer verwendet hat. Es sind Hinweise auf anderes Denken, wenn R über Kaninchen spricht, aber mangels einer eigenen Überzeugungsbildung auf der Basis einer Interaktion mit realen Kaninchen meint R nichts, wenn er über Kaninchen spricht, und daher spricht auch nichts dafür, in einem solchen Fall anzunehmen, dass die Maschine denkt (Davidson 2006, S. 150).

Philip K. Dicks Voigt-Kampff-Test erweitert die Problemstellung, indem er zeigt, dass die Zuschreibung authentischer Einstellungen mit sehr viel mehr zusammenhängt als lediglich der semantischen Interpretation sprachlicher Äußerungen. Der Protagonist Rick Deckard erklärt seiner Probandin Rachael vor Durchführung des Tests: „I’m going to outline a number of social situations. You are to express your reaction to each as quickly as possible“. Rachael versteht, denn sie entgegnet: „And of course my verbal responses won’t count. [...] Go ahead, Mr Deckard.“ (Dick 2010 [1968], S. 38). In einem Empathie-Test versucht der Fragesteller nicht nur herauszufinden, ob das befragte Gegenüber das Richtige oder etwas Passendes sagt, sondern ob es im Ganzen eine emotional angemessene und (nach unseren Maßstäben) richtige Reaktion erkennen lässt. Ein Androide, der noch nie in seinem Leben mit Kaninchen zu tun gehabt hat, offenbart vielleicht keine signifikante Gefühlsäußerung, wenn der Fragesteller sagt: „Als Kind schenkt Ihnen Ihre Mutter Stiefel aus Kaninchenfell. Sie erfahren, dass es das Fell Ihrer Kaninchen ist, die zuvor von Ihrer Mutter getötet wurden.“ Geringe Ausschläge der im Roman auch als „polygraphische Maschine“ bezeichneten Voigt-Kampff-Apparatur begründen den Verdacht, dass es sich bei dem Gegenüber um ein Wesen handelt, das über keine echte, „authentische“ Lebensgeschichte verfügt. Man hat dem Androiden vielleicht Bruch- oder Versatzstücke der Lebensgeschichte anderer einprogrammiert, aber damit hat er diese Episoden nicht selbst erlebt und empfunden, was die emotional befremdlichen Reaktionen auf die Testfragen erklärt.

„Denken und Meinen“, argumentiert Davidson, „setzen eine Geschichte bestimmter Art voraus.“ (Davidson 2006, S. 154). Denken heißt nicht einfach, diesen oder jenen Gedanken zu haben, sondern sich bestimmte Gedanken zu machen, weil man bestimmte Erfahrungen oder Erfahrungen bestimmter Art gemacht hat. Einem Wesen ohne eigene Geschichte und eigene Erfahrung kann man keine Gedanken zuschreiben, weil man nicht weiß, worum es diesem Wesen eigentlich geht. Den Androiden in Dicks Roman kann man hingegen sehr wohl Gedanken (sowie Wünsche und Überzeugungen und andere mentale Einstellungen) zuschreiben, auch wenn sie sich manchmal etwas merkwürdig und nicht in Übereinstimmung mit unseren Normen der Empathie verhalten. Aber sie haben in ihrem tragisch kurzen Leben eigene Erfahrungen gemacht, sehr einschneidende und prägende sogar, denn es ist die Erfahrung von Sklaverei und Leibeigenschaft, die erklärt, worum es den Androiden geht, wenn sie aus den extraterrestrischen Kolonien fliehen. Sie wollen

ein selbstbestimmtes Leben führen. Sie wollen nicht länger Automaten, sie wollen autonome Automaten sein.

Davidson erzählt eine andere Science-Fiction-Geschichte: „Nehmen wir an, im Sumpf schlägt der Blitz in einen Holzstamm ein, und durch reinen Zufall kommt ein Objekt zustande, das mir in allen physischen Hinsichten genau gleicht: Allem Anschein nach hat es meine Erinnerungen, es scheint meine Freunde wiederzuerkennen, und auf Fragen antwortet es mit Lauten, die wie Englisch klingen. Dieses Objekt wäre zwar nicht ich, und vielleicht würden wir uns weigern, es eine Person zu nennen, doch es wäre kein leichtes anzugeben, wieso es keine Gedanken und Gefühle hätte.“ (Davidson 2006, S. 158). Der Grund, warum der Sumpfmann keine Gedanken und Gefühle hat, ist nicht, dass er kein Mensch ist. Es liegt nicht an den Entstehungsbedingungen oder an dem, woraus ein Wesen gemacht ist, ob es Gedanken und Gefühle haben oder ob man ihm solche zuschreiben kann. Maßgeblich ist, dass es eine eigene Lebensgeschichte hat: „Gedanken setzen eine Geschichte voraus. [...] Ein Lebewesen oder ein Objekt kann keine Gedanken über Sterne, Tintenfische oder Sägespäne haben, sofern sich der betreffende Gedanke nicht irgendwie kausal auf geeignete Musterfälle zurückführen lässt.“ (Davidson 2006, S. 159). So formuliert, ist das Argument freilich missverständlich, denn „irgendwie kausal“ lässt sich auch der einprogrammierte Gedanke über Tintenfische und Sägespäne auf geeignete Musterfälle zurückführen (auf Musterfälle der Erfahrung anderer Personen nämlich). Worauf es ankommt, ist, dass das Lebewesen mit Blick auf die relevanten Fälle *eigene* Erfahrungen gemacht hat. Wie viele und welche Erfahrungen erforderlich sind, lässt sich nicht sagen. Man kann sinnvoll über Säbelzahntiger sprechen, ohne einen gesehen zu haben, und bestimmt kann man auch sinnvoll über Säbelzahntiger sprechen, ohne je einen Tiger gesehen zu haben. Aber ob man sinnvoll über Säbelzahntiger sprechen kann, ohne jemals eine Katze gesehen zu haben, ist fraglich, und sicherlich kann man keinen Gedanken über Säbelzahntiger haben, wenn man noch niemals ein Tier gesehen hat. Folgt man dem, kann man die eingangs formulierte Frage beantworten. Können Artefakte denken? Ja, wenn sie sich genügend mit der Welt auseinandergesetzt haben.

Die einfachste Möglichkeit, um herauszufinden, ob ein Objekt denken kann, schreibt Davidson, „besteht darin, dass man dem Fragesteller erlaubt, dem Objekt bei seinen Interaktionen mit der Welt zuzuschauen“ (Davidson 2006, S. 151). Genau dies erlaubt der Turing-Test aber nicht, und so verfehlt er, sicherlich aus dem aner kennenswerten Motiv, Anthropozentrismus zu vermeiden, den entscheidenden Punkt. Richtig an Turings Experiment ist, wie Davidson hervorhebt, dass Denken davon abhängt, dass es als solches interpretiert werden kann. Was sich der Interpretation verschließt, kann kein Gedanke sein. Turing übersieht aber, dass Interpretationen mehr voraussetzen, als die Wahrnehmbarkeit von grammtisch geordneten

Buchstabengruppen auf einem Bildschirm. Dass jemand etwas meint, setzt voraus, dass man Körperbewegungen „nachvollziehbar“ mit einer Umwelt in Verbindung bringen kann. Ohne dass dies schon allzu starke Vorannahmen über die äußere Gestalt des beobachteten Objekts implizierte, ist festzuhalten, „dass die physischen Eigenschaften des Objekts eine große Rolle spielen [...], denn das Objekt muss auf viele der gleichen Merkmale der Welt reagieren können, die auch vom Fragesteller bemerkt werden können, und der Fragesteller muss die Möglichkeit haben zu sehen oder sonstwie zu erfahren, dass das Objekt auf diese Merkmale reagiert. Damit dem Objekt eine Semantik zukommen kann, muss es in bestimmter Weise in der Welt tätig sein, und damit jemand anderes diese Semantik verstehen kann, muss es eine dreiseitige Interaktion zwischen Gegenstand, Fragesteller und gemeinsamer Welt geben.“ (Davidson 2006, S. 151 f.). Der Turing-Test ist kein geeignetes Experiment zur Beantwortung der Frage, ob Maschinen denken können, weil er fälschlich unterstellt, „die physische Realisierung eines Programms mache, was die geistigen Kräfte angeht, nur sehr wenig aus“ (Davidson 2006, S. 156). Die Externalisierung des Experiments, sein In-die-Welt-Stellen, das „Embodiment“ der Maschine werfen aber weitere, einstweilen ungelöste Fragen auf. Die wahrscheinlich bedeutendste Frage ist, wie viel äußere Ähnlichkeit mit uns selbst wir voraussetzen müssen, um einem Artefakt Gedanken zuschreiben zu können. Auch Davidson hat dafür keine Lösung, wenn er meint, ein Artefakt könne durchaus in der Lage sein zu denken, „wenn es einer Person in wichtigen Hinsichten ähnelt und eine passende Geschichte hat“ (Davidson 2006, S. 155). Aber in welchen Hinsichten muss uns das Artefakt ähneln? Und was ist es, das eine Geschichte passend macht?

Situationsräumlichkeit und Leibbewusstsein

Ich habe auf diese Fragen auch keine Antworten. Was ich stark vermute, ist, dass wir keine aufschlussreichen und interessanten Antworten finden, wenn wir an allzu robusten Anthropomorphismen oder allzu vertrauten Anthroponarrativen festhalten. Die Auseinandersetzung mit Wesen, die anders sind als wir, ist vermutlich eher dazu angetan, unsere Erkenntnis zu erweitern, als wenn wir uns immer nur mit uns selbst und unsersgleichen beschäftigen. Um das Wesen der Moralität zu erfassen, müssen wir, wie uns Kant belehrt hat, über den Menschen hinausgehen. Ich vermute, dass es mit dem Denken genauso ist. Zwei Arten von Gedankenexperimenten stehen uns zur Verfügung. Wir treffen auf eine neue Lebensform oder wir bauen eine intelligente Maschine.

Was auch immer denkt, haben wir gelernt, muss einen Körper haben, und dieser Körper muss dem unseren, dem menschlichen Körper, in wichtigen Hinsichten

ähnlich sein. Auch die „Geschichte“ muss passend sein, aber von dieser Bedingung sehe ich vorläufig ab, weil ich denke, dass sich zumindest ein Teil der Antwort auf diese Frage aus der Antwort auf die Frage nach den Bedingungen der körperlichen Ähnlichkeit ergibt. Körperliche Ähnlichkeit im relevanten Sinne, vermute ich, hat nichts oder wenig mit stofflicher oder morphologischer Ähnlichkeit zu tun (Davidson 2006, S. 157 ff.), sondern es kann sie zwischen Wesen geben, die sich äußerlich ganz unähnlich sind, sofern sie überhaupt beweglich sind und ihre Bewegungen einem sensomotorischen Analogieschluss zugänglich sind. Es geht also weniger um Körperähnlichkeiten als um Bewegungsanalogien.

Betrachten wir einen Organismus X, der sich in bestimmter, hinreichend auffälliger Weise um einen bestimmten Gegenstand bewegt. Stellen wir uns vor, X bewegt sich in konzentrischen Kreisen, die sich um den Gegenstand gleichsam öffnen und schließen. Es muss nicht so sein, aber wenn weitere Belege hinzukommen, ist es nicht fernliegend, diese Bewegung so zu deuten, dass X sich auf etwas bezieht und uns dies womöglich mitteilen möchte. Auch wenn X keinen Arm und damit keine Hand und keinen Zeigefinger hat, muss es eine Bewegung geben, die als Zeigen interpretiert werden kann. Ohne Zeigen kein Denken. Auf welche Weise gezeigt wird, ist gleichgültig, solange eine Bewegung vorliegt, die als Zeigen gedeutet werden kann. Wir sind damit noch gar nicht bei der Bestimmung eines Gedankens angekommen, und wir wissen auch noch gar nicht, ob X überhaupt denkt, aber wir haben es augenscheinlich mit einer orientierten, vielleicht zielgerichteten Bewegung zu tun, die gegebenenfalls als Ausdruck eines Denkens angesehen werden kann.

Dass die Perspektive des Turing-Tests zu eng ist, erweist sich nach Davidson auch daran, dass das Kriterium des Denkens zu unmittelbar mit dem Inhalt des Gedachten verbunden ist. Während wir im Alltag häufig auch nonverbale Anhaltspunkte dafür haben, dass jemand denkt, gibt es beim Turing-Test genau diese Möglichkeit nicht, sondern *dass* gedacht (oder womöglich eben auch nur zu denken prätendiert) wird, erschließt sich nicht unabhängig von dem, *was* gedacht wird (Davidson 2006, S. 146). Diese Verengung der Perspektive auf symbolische Repräsentation lässt außer Acht, dass es nicht nur auf Output, sondern auch auf Ausdruck ankommt. Denken findet nicht nur im Körper, sondern mit dem Körper statt. Vielleicht ist es daher nicht verkehrt, vom Begriff des Körpers abzurücken, sofern in ihm stets noch die Präsupposition der Cartesischen Differenz nachklingt. Bewusst hat die Phänomenologie etwa den Begriff des *Leibes* gewählt, um ein Verständnis des Körpers in Abkehr von seiner Konzeptualisierung als *res extensa* zu entwickeln. In der „Phänomenologie der Wahrnehmung“ von Maurice Merleau-Ponty heißt es: „Der Umriss meines Leibes bildet eine Grenze, die von den gewöhnlichen Raumbeziehungen unüberschritten bleibt. Der Grund ist der, dass die Teile des Leibes in einem ursprünglich eigenen Verhältnis zueinander stehen: sie sind nicht nebenei-

ander ausgebreitet, vielmehr ineinander eingeschlossen.“ (Merleau-Ponty 1966, S. 123). Fast möchte man dieses Verständnis des Leibes als „Embodiment“ von Kants ursprünglich-synthetischer Einheit der Apperzeption auffassen, aber damit würde man womöglich die Pointe verfehlen. Das „Bewusstsein des Leibes“ ist nach Merleau-Pontys klarer Auskunft jedenfalls „kein Denken“ (Merleau-Ponty 1966, S. 234). Aber zweifellos ist nach Merleau-Ponty die „Ausdruckseinheit“ (Merleau-Ponty 1966, S. 242) unseres Leibes die Bedingung dafür, dass wir Wahrnehmungen machen und Gedanken haben können.

Mein Arm liegt nicht so auf dem Tisch, wie das Buch auf dem Tisch liegt. Das würde weder ich so sehen, noch du. Der Leib konstituiert eine eigene, intrinsisch bedeutsame „Situationsräumlichkeit“ (Merleau-Ponty 1966, S. 125), die das bewegliche Fundament aller geistigen Operationen darstellt. Kappt man die Verbindung dieser Operationen zu ihrem Fundament, wie im Turing-Test, geht auf Dauer die Bedeutung der symbolischen Repräsentationen verloren. Bezugnahmen müssen vergewissert und erneuert werden. Gegenstände müssen aus verschiedenen Perspektiven betrachtet (und vielleicht auch angefasst oder gar gestreichelt) werden. Dafür muss man sich um sie herumbewegen können. Wenn wir das nicht selber machen wollen, müssen wir situationsräumlich geschmeidige Artefakte bauen, deren Sensomotorik sie zu selbständigen Anpassungsleistungen befähigt. Diese Fähigkeit kann nicht vollständig programmiert werden, sondern muss möglicherweise aus einer ursprünglichen Motorik der Nachahmung und gewohnheitsmäßigen Stabilisierung hervorgehen (zur Motorik vgl. Merleau-Ponty 2006, S. 123 ff., zur Rolle der Gewohnheit a.a.O., S. 172 ff.). Es geht mithin um eine lernende Motorik und Mimetik, die uns „eine Weise des Zugangs zur Welt und zu Gegenständen [eröffnet], eine „Praktognosie“, die es als eigenständig, ja vielleicht sogar als ursprünglich anzuerkennen gilt“ (Merleau-Ponty 2006, S. 170).

Ausblick

Inwieweit eine solche Nachahmungsfähigkeit, eine selbstlernende, „praktognostische“ Motorik künstlich hervorgebracht werden kann, vermag ich nicht zu sagen. Sicherlich werden sich hier in der Praxis der Robotik absehbar neue Erkenntnisse ergeben. Mir erscheint es nicht undenkbar, dass eine im eben erläuterten Sinne situationsräumlich sensibilisierte Rechenmaschine, die wir aus kulturellen Gründen vermutlich im Kostüm eines menschlichen Wesens verpacken würden, dahin gelangen kann, praktisch-intelligentes Verhalten derart überzeugend zu simulieren, dass wir keinen Unterschied mehr zwischen dieser Maschine und einem Menschen machen würden, und zwar schlicht deswegen, weil wir im Verhalten keinen Un-

terschied mehr erkennen würden. Besteht die Maschine diesen anspruchsvolleren Turing-Test, gibt es im Übrigen auch keinen Grund, sie nicht als vollverantwortliche Adressatin rechtlicher und moralischer Zurechnung zu behandeln.

Literatur

- Christian, Brian (2011): *The Most Human Human. A Defence of Humanity in the Age of the Computer*, London: Viking/Penguin.
- Davidson, Donald (2006): „Repräsentieren und Interpretieren“, in: D. Davidson, *Probleme der Rationalität*, Frankfurt am Main: Suhrkamp, S. 141–156 [im Orig. „Representation and Interpretation“, in: Newton-Smith, W.H. / Wilkes, K.V. (Hg.), *Modelling the Mind*, Oxford University Press 1990, S. 13–26].
- Davidson, Donald (2006): „Turings Test“, in: D. Davidson, *Probleme der Rationalität*, Frankfurt am Main: Suhrkamp, S. 157–177 [im Orig. „Turing’s Test“, in: Newton-Smith, W.H. und Wilkes, K.V. (Hg.), *Modelling the Mind*, Oxford University Press 1990, S. 1–11].
- Dick, Philip K. (2010): *Do Androids Dream of Electric Sheep?* [1968], London: Orion.
- Merlau-Ponty, Maurice (1966): *Phänomenologie der Wahrnehmung*, Berlin: de Gruyter [im Orig. „Phénoménologie de la Perception“, Paris: Gallimard 1945].
- Putnam, Hilary (1990): *Vernunft, Wahrheit und Geschichte*, Frankfurt am Main: Suhrkamp [im Orig. „Reason, Truth and History“, Cambridge University Press 1981].
- Quine, Willard Van Orman (1960): *Word and Object*, Cambridge, Mass.: M.I.T. Press.
- Turing, Alan M. (1950): „Computing Machinery and Intelligence“, in: *Mind* 59 (1950), S. 433–460.

Die Freiheit, $2*2=5$ zu rechnen

Die Fähigkeit, aus einem formalen System herauszutreten, von seinen Regeln abzuweichen und zu einem anderen überzugehen, es gar zu entwerfen, ist eine Bedingung für Autonomie wie für Kreativität. Auch Automaten können bei adäquater Soft- und Hardware-Architektur diese Fähigkeit erwerben.

Einleitung

Ein Rechenknecht hat die Freiheit, $2*2=5$ zu rechnen sicherlich nicht. Hat sie der Mathematiker, gleichsam der Herr über das Rechnen? Intuitiv wird man sagen, dass der Mathematiker, der Gesetze der Zahlentheorie eingedenk, niemals versucht sein wird, sich eine solche Freiheit zu nehmen. Anders als der Rechenknecht könnte er, will aber niemals. Wie also steht es mit der Autonomie des Mathematikers? Anders als von Naturgesetzen kann man von mathematischen abweichen.

Der Rechenknecht mag Automat sein; das festzustellen ist nicht mehr anstößig. Doch ist der Mathematiker möglicherweise auch ein Automat oder könnte der Rechenknecht zum Rechenherrn werden? Im allgemeinen wie im technischen Sprachgebrauch, auch dem der Informatik, kennt der Automat, und sei es auch ein Vollautomat, weder spielerische Abweichung noch kreative Abkürzung, sondern nimmt dem Menschen repetierbare Vorgänge ab.¹ Der autonome Automat wäre so gesehen ein Widerspruch in sich. Verliert sich mit der Autonomie der Automaten-Charakter? Ich blende die vielfältigen psychischen Probleme, die Zaublerlehrling, Frankenstein und In-Vitro-Fertilisation machen, im Folgenden aus und verweise auf Ball (2011).

Die Implikatur des Titels, die Analogie zwischen Mathematik und Jura oder, allgemeiner, gesellschaftlichen Normen, ist dem Volksmund nicht fremd:

Fünfe gerade sein lassen.

1 „Eine Maschine, die vorbestimmte Abläufe selbständig durchführt“ ><http://de.wikipedia.org/wiki/Automat>< (20.5.11), „dynamisches System, das sein Verhalten relativ unabhängig von der Umgebung zu lenken vermag“ (K/B) EEPW; der nicht-deterministische Automat der Informatik weist Freiheitsgrade auf, jedoch wählt er willkürlich >[http://de.wikipedia.org/wiki/Automat_\(Informatik\)](http://de.wikipedia.org/wiki/Automat_(Informatik))< (20.5.11).