

NORBERT CAMPAGNA

DIE ROBOTER ALS GEGENSTAND DER ETHIK

Prof. Dr. Norbert Campagna, geb. 1963, Studium der Philosophie und Anglistik an den Universitäten Heidelberg (Staatsexamen für das Lehramt) und Cambridge, Promotion im Fach Philosophie an der Universität Trier und Habilitation im Fach Philosophie an der Université Paris Est. Professeur-associé an der Université du Luxembourg und zugleich Studienrat für Philosophie am Lycée de Garçons Esch. Mitglied des Institut Grand Ducal des Sciences Morales et Politiques und Vizepräsident des Ethikkomitees des Centre Hospitalier Kirchberg. Erhielt 2012 in Paris den *Trophée de l'Ethique – Catégorie recherche* für seine Arbeiten auf dem Gebiet der Sexualethik. Autor von 23 und Mitherausgeber von 2 Büchern. Hauptforschungsgebiete: Staats- und Rechtsphilosophie sowie Angewandte Ethik und dort besonders Sexualethik.

Neueste Veröffentlichungen: *La sexualité des handicapés: Faut-il seulement la tolérer ou aussi l'encourager?* (Genève, Labor et Fides 2012); *Staatsverständnisse im spanischen siglo de oro* (Baden Baden, Nomos 2013) ; *Torture et droits* (hrsg. zusammen mit L. Delia und B. Garnot, Paris, Imago 2014).

1. Eine kleine Geschichte zur Einführung

Sie betreten eine Bäckerei, begrüßen die Person hinter der Theke, welche die Begrüßung freundlich erwidert, und Sie geben ihre Bestellung auf. Während die Person Sie bedient, lächelt sie ganz freundlich Ihrem fünfjährigen Sohn, der Sie begleitet, und bietet ihm einen Bonbon an. Ihr kleiner Sohn bedankt sich bei der Dame, was Sie auch tun, nachdem sie Ihnen das gegeben hat, was sie gefragt hatten. Sie zahlen und beim Heraustreten wünschen Sie der Bäckerin einen schönen Tag, was sie auch ihnen gegenüber tut.

Bis dahin haben wir es mit einem Szenario zu tun, wie es sich tagtäglich überall auf der Welt abspielt. Ein Soziologe würde von einer Interaktion zwischen drei Menschen sprechen, die nach eingespielten sozialen Codes funktioniert – Codes, nach denen wir fast schon automatisch handeln und von denen einige Zyniker behaupten würden, dass es sich lediglich um leere Floskeln handelt.¹ Doch wie leer diese Floskeln auch immer sein mögen, sie drücken

¹ Es ist sicherlich nicht dasselbe, wenn ich einer mir vielleicht unbekanntem Bäckerin einen schönen Tag wünsche und wenn ich einem ganz engen Freund, von dem ich weiß, dass er einen Ausflug plant, einen schönen Tag wünsche.

etwas aus, das weit wichtiger als ihr bloßer Inhalt ist. Durch sie geben wir kund, dass wir, um uns der Sprache I. KANTS zu bedienen, unser Gegenüber nicht nur als ein bloßes Mittel betrachten, sondern auch immer als Zweck an sich selbst. Ihre Äußerung hat also eine sozusagen integrative Funktion: Wenn wir sie jemandem gegenüber äußern, dann integrieren wir dieses Wesen implizit in unser moralisches Universum und stellen es dort auf dieselbe Stufe wie uns selbst.

Aber nehmen wir jetzt an, dass man Sie, sobald Sie die Bäckerei verlassen haben, auf dem Bürgersteig anhält, und man teilt Ihnen mit, dass die „Frau“ hinter der Theke in Wirklichkeit ein hochgradig entwickelter Roboter war, der nicht nur genauso aussieht wie eine wirkliche Frau, sondern der auch so programmiert wurde, dass er nach den vorhin erwähnten sozialen Codes funktioniert, sich also, wie in unserem Beispiel, wie eine richtige Person verhält, zumindest auf der Ebene der alltäglichen sozialen Interaktionen. Wissenschaftler haben sich während Monaten angesehen, was in Bäckereien vor sich geht, und haben dann einen Roboter so programmiert, dass er auf die üblichen Situationen so reagiert, wie es normalerweise ein Mensch tut. Und Sie wurden als Versuchsperson gebraucht, mit dem Ergebnis, dass Sie sich haben täuschen lassen.

Nun stellt man Ihnen folgende Frage: „Wie werden Sie sich das nächste Mal verhalten, wenn Sie diese Bäckerei wieder betreten und der Roboter, den Sie ab jetzt nicht mehr mit einer wirklichen Frau verwechseln werden, erneut hinter der Theke stehen wird?“ Werden Sie den Roboter begrüßen, wie Sie es das erste Mal getan haben – bloß dass Sie bei diesem ersten Mal dachten, einen Menschen zu begrüßen? Werden Sie Ihren Sohn, wenn er es nicht von sich aus tut, darauf aufmerksam machen, dass er sich zu bedanken hat, wenn man ihm etwas gibt? Werden Sie dem Roboter beim Weggehen noch einmal einen schönen Tag wünschen? Oder werden Sie aber fortan den Roboter so behandeln, wie Sie einen normalen Automaten behandeln würden? Denn wenn Sie Geld auf einem Bankautomaten abheben, bedanken Sie sich nicht beim Automaten. Und Sie tun es auch nicht bei einem Getränkeautomaten, aus dem eine Dose herauskommt, nachdem Sie Geld eingesteckt und auf einen Knopf gedrückt haben. Und Sie hätten es auch nicht getan, wenn in der Bäckerei statt des hochgradig entwickelten Roboters ein banaler Brotautomat gestanden hätte. Und außer der Frage, wie Sie sich tatsächlich verhalten werden, muss auch die moralphilosophisch relevante Frage gestellt werden, wie Sie sich eigentlich gegenüber einem Roboter mit Menschenantlitz, der sich auf der Ebene der banalen sozialen Interaktionen wie ein Mensch verhält, verhal-

ten *sollen* oder *dürfen*. Macht das Aussehen und das Verhalten des Automaten einen moralisch relevanten Unterschied, und wenn ja, ab welchem Punkt und warum? Hat es primär etwas mit dem Roboter als solchem zu tun, so dass man unmittelbar *ihm* gegenüber respektlos ist, wenn man sich nicht auf eine bestimmte Art und Weise verhält, oder hat es etwas mit der Art und Weise zu tun, wie wir uns mit wirklichen Menschen verhalten werden, wenn wir uns respektlos gegenüber Robotern mit Menschenantlitz verhalten? Und gesetzt den Fall, dass Sie die Bäckerei betreten und der Roboter vom ersten Mal steht nicht mehr hinter der Theke. Wie können Sie wissen, ob das Wesen, das jetzt hinter der Theke ist, ein Mensch ist und kein neuer Roboter, der genauso menschenähnlich aussieht wie der erste?

Dass Roboter einen immer größeren Platz in unserem Leben einnehmen, ist eine Tatsache. Ob diese Tatsache allerdings begrüßenswert ist, ist eine für die Moralphilosophie wichtige Frage, und zwar nicht nur, weil die Roboter große Katastrophen auslösen können, sondern auch, weil die zunehmende Präsenz dieser Roboter, und vor allem ihre immer größere Ähnlichkeit mit uns Menschen, Fragen aufwirft, welche die Grundlagen der Moral betreffen. Mag es auch noch längere Zeit dauern, bis die Roboter tatsächlich so perfektioniert sind wie der vorhin beschriebene Roboter in der Bäckerei – Fakt ist, dass wir uns schon jetzt die Frage stellen müssen, welche Auswirkungen die Existenz solcher Roboter auf die Moral haben wird. Diese und ähnliche Fragen stehen im Zentrum der sogenannten Roboterethik.²

Ich werde in folgendem Beitrag zunächst ganz allgemein auf die Roboterethik als neuem Zweig der Angewandten Ethik eingehen und in diesem Zusammenhang zwei große Modelle der Roboterethik unterscheiden, und zwar das anthropozentrische und das, wie ich es nennen werde, robozentrische Modell. Daran anschließend werde ich auf die Roboter als Interaktionspartner eingehen. Hier liegt nämlich ein wesentlicher Unterschied zwischen Robotern und den üblichen Maschinen, deren sich der Mensch schon seit längerem bedient. Im dritten Teil soll auf die Roboter als potentielle moralische Entscheider eingegangen werden. Die Roboter werden nämlich immer häufiger in Situationen eingesetzt, in denen moralische Entscheidungen getroffen werden müssen, Entscheidungen, bei denen es um Leben oder Tod geht. Der letzte Teil befasst sich schließlich mit der Frage, ob und inwiefern wir eines Tages

² Die auch auf europäischer Ebene ernst genommen wird. Erwähnt sei nur Euron, ein europäisches Netzwerk, dem über 200 Forschungslabore und industrielle Einheiten angeschlossen sind, die auf dem Gebiet der Robotik tätig sind.

dazu kommen werden – und ein Teil des Weges dahin ist schon zurückgelegt –, den Robotern – wenn auch nicht allen, so doch vielleicht bestimmten – Rechte zuzusprechen.

Die folgenden Überlegungen verstehen sich in allererster Linie als bescheidener Versuch, wichtige Problemfelder der Roboterethik zu identifizieren und die sich dort jeweils stellenden Probleme auf den Begriff zu bringen.³ Es würde den Rahmen des Beitrags sprengen⁴, auch fertige Lösungen für diese Probleme zu geben. Einige dieser Probleme stellen sich schon auf eine sozusagen akute Art und Weise. Wir haben jetzt schon die Möglichkeit, Roboter zu bauen, die einem Menschen äußerlich ganz ähnlich sehen und relativ komplexe Interaktionsprozesse mit den Menschen eingehen. Wir haben auch schon die Möglichkeit, einen Roboter konsequentialistisch oder deontologisch zu programmieren, zumindest auf einer ganz elementaren Ebene. Bei anderen Problemen ist noch nicht sicher, ob sie sich überhaupt jemals akut stellen werden. Man denke hier etwa an die Frage, ob man einem Roboter ein moralisches Gewissen – oder zumindest ein funktionales Äquivalent dazu – einbauen sollte. Aber auch wenn sie sich noch nicht stellen – und vielleicht nie stellen werden –, so sollte uns dies nicht davon abhalten, sie schon jetzt zu thematisieren, damit wir nicht ganz ratlos dastehen, wenn die Wissenschaftler über die technische Möglichkeit verfügen.

2. Die Roboterethik

Aufgrund der rasanten Entwicklungen, die in den beiden letzten Jahrzehnten auf dem Gebiet der Robotik stattgefunden haben, hat sich, parallel dazu, auch ein neuer Zweig der praktischen Philosophie entwickelt: die sogenannte *Roboterethik*.⁵ Die Roboterethik kann dabei durchaus als Zweig der Technikethik betrachtet werden, da Roboter schließlich Produkte der Technik sind. Allerdings bringt sie eine neue Dimension innerhalb der Technikethik zum Tragen, eine Dimension, die man auch in der sogenannten Computerethik wiederfindet. Während die traditionellen Produkte der Technik weder wie Menschen aussehen noch sich in sozialer Hinsicht wie Menschen verhalten,

³ Für einen kurzen synthetischen Überblick sei verwiesen auf N. CAMPAGNA: *Roboterethik* (2012).

⁴ Ganz davon abgesehen, dass dadurch auch der Eindruck vermittelt werden könnte, dass die Philosophen über fertige und angemessene Lösungen für ein Problem verfügen, das zunächst einmal im Rahmen eines gesamtgesellschaftlichen Diskurses thematisiert werden sollte.

⁵ Im Englischen spricht man entweder von „roboethics“ oder „robot ethics“.

sehen Roboter immer mehr wie Menschen aus und verhalten sich in sozialer Hinsicht auch immer mehr wie Menschen. In einer vor 30 Jahren durchgeführten Studie hat die Soziologin SHERRY TUCKLE das Verhältnis des Menschen zum Computer untersucht und dabei die These aufgestellt, dass Geräte wie der Computer – mit seiner künstlichen Intelligenz – einen Einfluss auf unser Denken über uns selbst haben.⁶ Was TUCKLE von den Computern sagt, gilt in noch größerem Maß für die Roboter, denn diese besitzen nicht bloß eine künstliche Intelligenz, sondern oft auch eine künstliche Ähnlichkeit mit dem Menschen.⁷ Genauso wie andere Begriffe die mit dem Suffix „-ethik“ gebildet wurden und bereichsspezifische Ethiken bezeichnen, besitzt auch der Begriff der Roboterethik eine ihm inhärente Mehrdeutigkeit, so dass man zwischen radikalen und weniger radikalen Verständnissen des Begriffs unterscheiden kann. In diesem Zusammenhang lohnt sich ein kurzer Vergleich mit dem Begriff der Tierethik⁸, also jener Disziplin der angewandten Ethik, die sich mit unserem Verhalten zu den Tieren befasst, und die besonders seit dem Erscheinen von PETER SINGERS Buch *Animal Liberation* vor fast 40 Jahren einen zentralen Platz in der Angewandten Ethik einnimmt. Tierethik kann einerseits jenen Zweig der praktischen Philosophie bezeichnen, der, um mit KANT zu sprechen, lediglich unsere Pflichten *im Ansehen* der Tierwelt bezeichnet. Als Tierethik kann man aber auch jenen Zweig der praktischen Philosophie bezeichnen, der unsere Pflichten *gegenüber* der Tierwelt benennt. Während im ersten Modell keine moralischen Rechte für Tiere vorgesehen sind, ist im zweiten Modell durchaus Platz für solche Rechte, wenn sie nicht sogar als mögliches Fundament für die Zuschreibung von Pflichten gegenüber den Tieren vorausgesetzt werden müssen. Das erste Modell verbleibt im anthropozentrischen Rahmen und sieht in den Tieren primär Wesen, deren Wohlergehen für das Wohlergehen der Menschen relevant ist. Wenn ich den Hund nicht brutal behandeln darf, dann entweder weil es nicht mein Hund ist und sein Besitzer traurig darüber sein wird, dass ich seinem Hund Leid zufüge, oder weil ich dadurch ein brutales Verhalten an den Tag lege, das, wenn es zum Habitus werden sollte, negative Konsequenzen für mein Verhalten anderen Menschen gegenüber haben wird. Zentral ist in beiden Fällen der Bezug zum Leiden anderer Men-

⁶ SH. TURKLE: *The Second Self* (1984).

⁷ Manche Autoren warnen vor der Unterscheidung zwischen dem Künstlichen und dem Natürlichen und sehen hier nur eine linguistische Falle (siehe etwa TH. M. GEORGES: *Digital soul* (2003), S. 5).

⁸ Man hätte auch den Begriff der Umweltethik nehmen können. Hier unterscheidet man zwischen einer anthropozentrischen und einer physiozentrischen Umweltethik, wobei die Bewegung, die mit Letzterer assoziiert wird, im Englischen oft als „deep ecology“ bezeichnet wird.

schen. Der Hund als Hund soll nicht um seiner selbst willen geschont werden. Man kann insofern von einem anthropozentrischen Modell der Tierethik sprechen. Das zweite Modell ist hingegen ein zoozentrisches Modell, da es das Wohlergehen der Tiere um ihrer selbst willen betrachtet. Wenn ich den Hund nicht brutalisieren soll, dann soll dies primär deshalb sein, weil man im Namen des Hundes – also advokatorisch, wie man zu sagen pflegt – einen moralischen Anspruch geltend machen kann, dass er nicht um seiner selbst wegen brutalisiert werden darf.

Bei der Roboterethik könnte man ganz ähnlich zwischen einer anthropozentrischen und einer, falls der Ausdruck erlaubt ist, robozentrischen Roboterethik unterscheiden. Im Rahmen des ersten Modells geht es lediglich um das Wohlergehen der Menschen, während man sich beim zweiten auch, wenn nicht sogar primär, um das Wohlergehen der Roboter um ihrer selbst willen sorgt. Es ist etwas ganz anderes, ob man bei der Zerstörung eines Roboters nur den Schaden sieht, der seinem Besitzer dadurch entsteht – ein Schaden der durchaus auch emotionaler Natur sein kann⁹ –, oder ob man auch, und vielleicht in erster Linie, den Schaden oder gar das Leid in Erwägung zieht, die den Roboter als solchen betreffen. Dabei muss allerdings zunächst im Vorfeld geklärt werden, ob man einem Roboter überhaupt einen Schaden oder ein Leid zufügen kann. Dass man es bei vielen Tierarten kann, steht heute außer Zweifel.¹⁰ Aber kann man es auch bei Robotern? Wird der in der Einleitung erwähnte Roboter in der Bäckerei sich beleidigt fühlen, wenn wir ihn nicht begrüßen? Wird er in eine tiefe Depression fallen, wenn kein Kunde ihn mehr begrüßt und niemand mehr freundliche Worte zu ihm spricht – wie es bei Menschen geschehen würde, die sich in einer solchen Situation befinden?

Im Fall der Roboterethik kommt allerdings noch eine dritte Ebene ins Spiel. Während es nämlich ausgeschlossen scheint, aus den Tieren moralisch handelnde Wesen zu machen, scheint dies im Fall der Roboter zumindest nicht prinzipiell ausgeschlossen werden zu können, und dies in erster Linie deshalb, weil Roboter sich intelligent verhalten können. Einem Tier wird man

⁹ Ein ganz banaler Fall wäre hier derjenige eines Roboters, den man selbst gebaut hat, so dass man auf seine eigene Leistung stolz ist. Wer einen solchen Roboter zerstört, zerstört nicht bloß eine ersetzbare Maschine, sondern gleichzeitig auch den Gegenstand unseres Stolzes. Außerdem verlieren die vielen Arbeitsstunden, die wir in die Konstruktion gesteckt haben, ihren Sinn mit der Zerstörung des Gegenstandes, den wir konstruiert haben.

¹⁰ Während Descartes vor fast 400 Jahren die Tiere mit Maschinen verglich und ihnen jedes bewusste Leben absprach, stehen wir heute vor der Frage, ob wir Roboter mit Tieren vergleichen und ihnen zumindest dieselbe Art Schutz gewähren sollten.

nicht beibringen können, moralisch zu überlegen¹¹, bei den Robotern versucht man es mit Hilfe bestimmter Algorithmen. Es gibt einige Autoren, die behaupten, dass wir sogar gut daran täten, den Robotern so schnell wie möglich moralisches Überlegen und Verhalten beizubringen, sie also zu „moralischen Maschinen“¹² zu machen – genauso wie wir gut daran tun, unseren Kindern moralisches Verhalten beizubringen¹³. Je mehr die Roboter nämlich in soziale Interaktionsprozesse integriert werden, so das Argument, desto wichtiger ist es, ihnen auch jenes zentrale Element der sozialen Interaktionsprozesse beizubringen, das wir Moral nennen. Und dies wird auch noch deshalb wichtig, weil die Roboter immer autonomer werden. Aus den weitgehend stationären Robotern der Vergangenheit, die sich nur dann im Raum fortbewegten, wenn ein Mensch sie führte, sind inzwischen Roboter geworden, die sich von selbst fortbewegen können. Was nichts anderes bedeutet, als dass die Roboter, sieht man vielleicht vom Konstruktionsprozess ab, immer unabhängiger vom Menschen werden. Bestimmte batteriebetriebene Roboter – meistens Staubsauger – sind heute schon in der Lage, sich selbst an eine Steckdose anzuschließen, um sich wieder selbst aufzuladen. Der Roboter braucht also nicht mehr zu warten, bis in ihn ein Mensch wieder an die Steckdose anschließt, sondern er kann dies von selbst tun. Autonomie bedeutet aber noch mehr, und zwar bedeutet es auch, dass die Roboter immer mehr in der Lage sind, nicht im Vorfeld programmierte Reaktionen zu zeigen. Anstatt den Roboter so zu programmieren, dass er immer X tut, programmiert man ihn so, dass er selbst eine große Anzahl von Variablen bearbeitet, um dann auf eine bestimmte Weise zu handeln.

3. Roboter als Interaktionspartner

Wie allgemein bekannt, stammt das Wort „Roboter“ aus dem Tschechischen und bedeutet so viel wie „mühselige Arbeit“ oder „Zwangsarbeit“. Der Roboter kann insofern als eine Maschine betrachtet werden, deren primärer Zweck

¹¹ Was nicht ausschließt, dass es zumindest ein moralähnliches Verhalten bei bestimmten Tierarten, wie etwa den sogenannten Menschenaffen, geben kann.

¹² *Moral machines* lautet der Titel eines 2009 erschienenen Buches zum Thema. Die Autoren, WENDELL WALLACH und COLIN ALLEN, sprechen von *artificial moral agents* (abgekürzt AMA). Bisher kannten wir nur natürliche moralische Agenten, also moralische Agenten, die ihre, nicht unbedingt immer gleichen, moralischen Dispositionen von Natur aus hatten. Die Erziehung – ein an sich künstlicher Faktor – konnte der Aktualisierung dieser natürlichen Dispositionen förderlich sein – oder auch nicht.

¹³ Der „Lernprozess“ wird in beiden Fällen selbstverständlich nicht derselbe sein.

es ist¹⁴, mühselige Arbeiten zu verrichten und somit die Menschen von diesen Arbeiten zu befreien¹⁵. In dieser Hinsicht unterscheidet der Roboter sich nicht von vielen anderen Maschinen, die der Mensch im Laufe der Jahrhunderte gebaut hat. Was ihn aber von diesen Maschinen unterscheidet, ist die Tatsache, dass er anthropomorphe Züge hat.¹⁶ Und wo er keine anthropomorphen Züge aufweist, weist er zumindest zoomorphe Züge auf.¹⁷

Diese Ähnlichkeit zwischen einem Roboter und einem Menschen oder einem Tier betrifft zunächst einmal das bloß äußere Aussehen. In den letzten Jahren ist es den Roboterdesignern gelungen, Roboter zu bauen die, zumindest von einer bestimmten Distanz aus betrachtet, einem Menschen so ähnlich sehen, dass man nicht merkt, dass es sich nicht um einen wirklichen Menschen handelt. Große Fortschritte gibt es auch hinsichtlich der Ähnlichkeit des äußeren wahrnehmbaren Verhaltens mit dem eines Menschen oder Tieres.

¹⁴ Wobei man diesen Zweck allerdings nicht als ein konstitutives Element der Definition betrachten sollte.

¹⁵ Wir mögen es nicht, mühselige Arbeiten zu verrichten, und in der Vergangenheit wurde das Verrichten solcher Arbeiten manchmal als Strafe angesehen. Mögen die Roboter solche Arbeiten verrichten? Peterson schlägt vor, die Roboter so zu gestalten, dass sie jene Arbeiten mögen, die wir nicht mögen. Auf diese Weise, so Peterson weiter, würden wir ihnen nicht schaden, wenn sie diese Arbeiten verrichten (S. PETERSEN: *Designing people to serve* (2012), S. 284).

¹⁶ In Kapeks Roman, tragen sie solche Züge. In der Zwischenzeit hat sich der Gebrauch des Begriffs aber erweitert, so dass man etwa im Französischen den Ausdruck „robot de cuisine“ verwendet. Damit bezeichnet man jenes motorisierte Gerät, mit dem man eine Suppe mixen oder einen Kuchen- oder sonstigen Teig machen kann. Und auch den von selbst laufenden Rasenmäher oder Staubsauger bezeichnet man inzwischen oft als Roboter, auch wenn hier keine anthropomorphen Züge vorliegen. Was nicht verhindert, dass man trotzdem manchmal zur Anthropomorphisierung tendiert. Es sei mir hier erlaubt, ein Beispiel aus der eigenen Privatsphäre zu erwähnen: Zu Weihnachten bekam unser ältester Sohn eine kleine batteriebetriebene Staubbürste, die mit zwei „Augen“ versehen war.

Dieses Ziel einer größtmöglichen Ähnlichkeit zwischen der Maschine und dem Menschen oder anderen Lebewesen ist übrigens nicht ganz neu, denn schon JACQUES DE VAUCANSON (1709–1782) versuchte, mit seinem Flötenspieler (1737), seinem Tamburinspieler (1738) und seiner Ente (1738) Automaten zu schaffen, die den realen Lebewesen so ähnlich wie möglich waren. Aber während Vaucanson zu seiner Zeit nur auf die bloße Mechanik zurückgreifen konnte und auch bei den Materialien eine ziemlich begrenzte Wahl hatte, kann die heutige Automatenkonstruktion nicht nur auf ganz neuartige synthetische Materialien zurückgreifen, die eine größere Ähnlichkeit mit dem menschlichen Aussehen erlauben, sondern auch auf die Elektronik. Nicht zu sprechen von den Möglichkeiten, die durch die Nanotechnologie eröffnet werden.

Wichtige Elemente der Anthropomorphisierung sind die Augen und der Mund. NAO ist etwa ein Beispiel für einen humanoiden Roboter.

¹⁷ Bekey definiert den Roboter als „a machine, situated in the world, that senses, thinks, and acts“ (G.A. BEKEY: *Current trends in Robotics* (2012), S. 18). Dabei werden die drei Kompetenzen relativ minimalistisch aufgefasst: Der Roboter empfängt Signale von der Außenwelt („senses“), verarbeitet diese Informationen („thinks“) und übersetzt das Resultat dieser eigenen Verarbeitung in eine Reaktion („acts“).

Roboter können sich fortbewegen, sie können sprechen, usw. Bemerkenswerte Fortschritte sind ebenso auf dem Gebiet der Reaktionen auf äußere Reize zu vermerken. Roboter können gesprochene Befehle ausführen und sie können sogar sprachlich mit Menschen interagieren. Außerdem können Roboter „erkennen“, in welchem emotionalen Zustand sich ein Mensch befindet, um dann diesem Zustand entsprechend zu reagieren. Vor allem die Fortschritte auf den beiden letzten genannten Gebieten lassen die Frage auftauchen, ob man nicht eines Tages in der Lage sein wird, Roboter zu bauen, die auch selbst Gefühle haben¹⁸ und die auch autonom Entscheidungen bezüglich ihres Verhaltens treffen können, Entscheidungen, die nicht vorprogrammiert sind.

Doch warum bauen wir Roboter, die den Menschen immer ähnlicher sind? Hier muss man zwischen äußerer und innerer Ähnlichkeit unterscheiden und bei der äußeren Ähnlichkeit wiederum zwischen der Ähnlichkeit des Aussehens und der Ähnlichkeit des Verhaltens. Für beide großen Typen von Ähnlichkeit gilt allerdings, dass es sich nicht nur darum handelt, unser technisches Können unter Beweis zu stellen. Die menschen- oder tierähnlichen Roboter werden heute nicht nur jeweils einmalig konstruiert, sondern man strebt ihre massenweise Produktion an. Für die kommenden Jahrzehnte wird eine immer größere Präsenz solcher Roboter in unserem alltäglichen Leben vorausgesagt. Und wenn dem so sein sollte, dann heißt das, dass wir in unserem Alltagsleben nicht nur Wert darauf legen, dass Maschinen uns bestimmte Aufgaben abnehmen, sondern dass wir ebenso großen Wert darauf legen, dass man diese Maschinen so weit wie möglich nicht mehr als bloße Maschinen identifiziert.¹⁹

Auf dem Gebiet der Sexualität leuchtet das unmittelbar ein. Je menschenähnlicher der nicht-menschliche Sexualpartner ist, umso mehr glaubt man, sexuell mit einem Menschen zu verkehren und umso größer – so lässt sich zumindest vermuten – dürfte die sexuelle Befriedigung sein.²⁰ Schon bei den aufblasbaren Puppen konnte man eine ähnliche Entwicklung feststellen, aber diese Entwicklung hatte ihre Grenzen. Einige dieser Grenzen können durch perfektionierte Roboter aufgehoben werden, wie etwa die sexuelle Passivität

¹⁸ Hier wird sich die Frage stellen, wie wir wissen, ob ein Wesen Gefühle „hat“. Am MIT haben Wissenschaftler einen in seinen großen Zügen menschenähnlichen Kopf gebaut – KISMET genannt –, an dem bestimmte Reaktionen auf dem Gesicht provoziert werden konnten, die bei menschlichen Gefühlen präsent sind.

¹⁹ Es kommt noch ein weiteres Element hinzu. Während wir in einem ersten Schritt Maschinen gebaut haben, die für uns Arbeiten verrichteten, wobei wir aber diese Maschinen überwachen mussten, geht es jetzt in einem zweiten Schritt darum, Maschinen zu bauen, die sich um sich selbst kümmern (dazu etwa D. MCFARLAND: *Guilty robots, happy minds* (2009), S. 14.

²⁰ Replize Q1 des Japaners HIROSHI ISHIGURO sieht einem jungen japanischen Mädchen von weitem ganz ähnlich.

der Puppen. Ein Roboter kann so programmiert werden, dass er manchmal die Initiative für den Geschlechtsverkehr ergreift. Manche Autoren meinen, die Entwicklung solcher Roboter könne auf Dauer eine Lösung für das Problem der Prostitution bieten.²¹

Auch auf therapeutischem Gebiet lässt sich der Wunsch nach Ähnlichkeit relativ leicht erklären, wobei hier vor allem die Ähnlichkeit mit Tieren im Mittelpunkt steht. Eine Therapie mit einem tierähnlichen Roboter ergibt bessere Ergebnisse als eine Therapie mit einem rein „maschinenähnlichen“ Roboter. Der tierähnliche Roboter wird somit die bewegungslosen Plüschtiere oder die realen Tiere ersetzen, die man manchmal im Rahmen von Therapien einsetzt.²² Dasselbe gilt auch für menschenähnliche Roboter, die sich um kranke oder ältere Menschen kümmern können.²³

Sowohl auf sexuellem als auch auf therapeutischem Gebiet lassen sich beim Roboter jene Züge eliminieren, die für das angestrebte Ziel unerwünscht sind und die man aber oft bei natürlichen Lebewesen antrifft. So hat eine Roboterfrau niemals Migräne – es sei denn, man habe sie so programmiert, dass sie manchmal Migräne vortäuscht – und sie wird auch nicht schwanger. Und auch die Übertragung von Geschlechtskrankheiten ist ausgeschlossen. Und, weiterer Punkt, man hat keine Verpflichtungen gegenüber einem Roboter – zumindest noch nicht. Und sieht man sich die im Rahmen einer Therapie eingesetzte Roboterkatze an, so hat sie den Vorteil, dass sie nicht aus Versehen jemanden kratzt oder dass sie plötzlich wegläuft.²⁴ Und geht man noch einen Schritt weiter, so muss man auch die ökonomischen Aspekte erwähnen: Im Rahmen der Krankenpflege kostet der Roboter etwas bei der Anschaffung, danach ist er aber wesentlich billiger als menschliche Arbeitskräfte, so dass die ursprünglichen Kosten ziemlich schnell amortisiert sind.

Wenn Maschinen in der Vergangenheit vor allem für Arbeiten in einem ganz engen Sinn des Wortes eingesetzt wurden, so werden sie heute immer öfter für Dienstleistungen verwendet, und bei vielen dieser Dienstleistungen kann das Aussehen des „Dienstleistenden“ eine Rolle für den Erfolg der Dienstleis-

²¹ Siehe hierzu D. LEVY: *Love & sex with robots* (2008); ders.: *The ethics of robot prostitutes* (2012).

²² Ein Beispiel ist hier PARO, der einem kleinen Seehund ähnlich sieht und bei der therapeutischen Behandlung von Stress eingesetzt wird. RIBA, IWARD, ERNIE sind die „Namen“ von Robotern, die manche Aufgaben von Krankenschwestern übernehmen. Und ARES sowie Vinci Surgical System können, nachdem sie gestartet worden sind, selbständig medizinische Untersuchungen durchführen.

²³ RI-MAN, PaPeRo, PALRO, QRIO oder CareBot sind hier einige Beispiele.

²⁴ In beiden Fällen besteht allerdings das Risiko einer technischen Panne.

tung spielen. Wenn es uns lediglich darum geht, eine Last von einem Ort A zu einem Ort B zu befördern, ist es uns gleichgültig, ob die Maschine einem Zugtier gleicht oder nicht. Bei bestimmten Dienstleistungen ist dem aber nicht so. Wenn wir uns demnach dafür entscheiden, die betreffenden Dienstleistungen nicht mehr durch wirkliche Menschen ausführen zu lassen, sondern durch Roboter, dann scheint uns oft keine andere Wahl zu bleiben – falls wir an einer möglichst effizienten Dienstleistung interessiert sind²⁵, was man gewöhnlich voraussetzen darf –, als auf eine immer größere Ähnlichkeit der Roboter mit Menschen oder Tieren zu setzen – eine Ähnlichkeit, die sich aber, wie schon angeklungen, oft nur auf die für die betreffenden Dienstleistungen relevanten Eigenschaften beschränkt und damit all jene Eigenschaften ausklammert, die dem erfolgreichen Absolvieren der Dienstleistung im Wege stehen könnten. Wenn wir Roboter für bestimmte Dienstleistungen einsetzen, dann wollen wir nicht nur, dass sie uns von den betreffenden Dienstleistungen befreien, sondern wir wollen auch, dass sie sie die Leistung genauso gut, wenn nicht sogar noch besser – sprich: fehlerfreier – erbringen als ein normaler Mensch es tut. Wenn demnach von einer größtmöglichen Ähnlichkeit gesprochen wird, dann sollte gleich präzisiert werden, dass es sich nur um eine größtmögliche *selektive* Ähnlichkeit handelt.

4. Moralische Entscheider

Das Problem der selektiven Ähnlichkeit stellt sich auf eine akute Weise, wenn wir an Roboter denken, die moralisch relevante Entscheidungen treffen müssen bzw. Entscheidungen, die aus unserer Sicht moralisch relevant sind – ob der Roboter sie selbst auch als moralisch relevant erlebt, ist eine andere Frage. Vor allem Roboter, die man zu militärischen Zwecken einsetzt²⁶, können mit Situationen konfrontiert werden, die, wenn Menschen sie erleben würden, Anlass für moralische Dilemmas sein können. Ein bekannter Fall ist derje-

²⁵ Die Wirksamkeit lässt sich hinsichtlich unterschiedlicher Dimensionen messen, wobei in unserer Gesellschaft die ökonomische eine große Rolle spielt. Wenn ein Kassiererinnenähnlicher Roboter in einem Supermarkt in allen relevanten Hinsichten dieselbe Funktion erfüllen kann, wie eine menschliche Kassiererin, aber ohne dass man ihm einen Lohn zahlen muss, dann werden die Kassiererinnen nach und nach ihre Stellen verlieren. Hier stellt sich dann die Frage, ob Roboter die Menschen in bestimmten Dienstleistungen ersetzen sollten oder ob sie nur eine komplementäre Rolle haben sollten. Dazu J. BORENSTEIN/Y. PEARSON: Robot caregivers (2012).

²⁶ Es gibt eine große Anzahl solcher Roboter, die, wie es sich gehört, schreckenerregende Namen tragen: Predator, Reaper, Big Dog, Crusher, Harpy, BEAR, Global Hawk, Dragon Runner. Sie werden zu ganz unterschiedlichen Zwecken eingesetzt, von der Überwachung über die Entschärfung von Bomben bis hin zu Angriffsmissionen.

nige jener kleinen Gruppe amerikanischer Soldaten, die auf einer geheimen Spezialmission in den afghanischen Bergen von einem Schafhirten entdeckt wurden. Für diese Soldaten stellte sich folgende Frage: Den Hirten seinen Weg gehen lassen, mit dem Risiko, dass er die Taliban benachrichtigt, oder ihn umbringen? Ihn gefangen nehmen kam nicht in Frage. Die Soldaten entschieden sich dafür, ihn gehen zu lassen. Einige Stunden später wurden sie von den Taliban angegriffen.

Spätestens hier sind wir mit der Frage konfrontiert, was unsere moralischen Entscheidungen bestimmt bzw. was diese Entscheidungen bestimmen sollte. Grob gesehen lassen sich hier zwei Modelle unterscheiden, das logozentrische und das pathozentrische. Während das erste Modell die praktische Vernunft als einzig legitimen Ursprung des moralischen Handelns sieht, bringt das zweite Modell dieses Handeln unmittelbar mit Gefühlen oder allgemeiner mit sinnlichen Triebfedern in Verbindung, wie etwa die Sympathie bei Smith oder das Mitleid bei Schopenhauer.²⁷ Wenn wir von einem rein logozentrischen Modell des moralischen Entscheidungsfindungsprozesses ausgehen und demnach, wie KANT es tut, die Moralität einer Handlung davon abhängig machen, dass wir sie auch dann ausgeführt hätten, wenn wir keine Neigung zu ihr gehabt hätten, dann werden wir darauf bedacht sein, Roboter zu bauen, die keine Gefühle, Neigungen usw., also keine pathologischen Bestimmungsgründe kennen.²⁸ Auch wenn KANT nicht behauptet, pathologische Bestimmungsgründe dürften niemals bei der Entscheidung für eine Handlung eine Rolle spielen, so sagt er doch, dass sie auf keinen Fall notwendige Bestimmungsgründe sein dürfen. Ihre Präsenz beim Entscheidungsfindungsprozess kann Zweifel hinsichtlich der Moralität der Handlung aufwerfen, da wir niemals sicher sein können, ob wir die Handlung auch dann noch ausgeführt hätten, wenn wir sie nicht gehabt hätten.²⁹ Bei einem Wesen das keine pathologischen, sondern nur rationale Bestimmungsgründe kennt, kann ein solcher Zweifel nicht auftauchen.

²⁷ Hughes hält es für möglich, Maschinen zu bauen, die Empathie haben können (J. HUGHES: *Compassionate AI and selfless robots* (2012), S. 79).

²⁸ W. Wallach und C. Allen sehen es als einen Vorteil für moralische Entscheidungen an, dass Roboter sich nicht durch Gefühle u.a. beeinflussen lassen (W. WALLACH/C. ALLEN: *Moral machines* (2009), S. 71).

²⁹ KANT behauptet keineswegs, wie ihm manchmal unterstellt wird, man müsse alle sinnlichen Triebfedern in sich zum Schweigen bringen oder sie gar ausmerzen, sondern er sagt nur, dass diese sinnlichen Triebfedern nicht die determinierende Rolle bei moralischen Entscheidungen spielen sollen. Wenn also etwa A, B, C und D sinnliche Triebfedern und E eine rationale Triebfeder ist, dann muss E die determinierende Triebfeder sein, mögen A, B, C und D präsent sein oder nicht. So muss man sich um seine Kinder kümmern, ob man sie liebt oder nicht. Dass

Ein Wesen das sich nur durch rationale Gründe bestimmen lassen kann, nennt KANT ein autonomes Wesen. Und autonome Wesen sind für ihn Wesen, die bloß einen Wert und keinen Preis haben und die deshalb eine Würde besitzen und absoluten Respekt verdienen. Sind Roboter in diesem Sinne autonome Wesen bzw. lassen sich eines Tages Roboter bauen, die in diesem Sinne autonome Wesen sind? Was die neuen von den alten Robotern unterscheidet, ist die Tatsache, dass sich die neuen Roboter in Situationen befinden können, in denen ein Handeln von ihnen erwartet wird, das, wenn es in einer ähnlichen Situation von einem Menschen erwartet würde, durchaus als moralisches Handeln bezeichnet werden könnte.

Nehmen wir noch einmal die Situation der amerikanischen Soldaten und vergleichen wir die drei folgenden Szenarien:

Szenario 1: In den afghanischen Bergen befinden sich Menschen, die selbst darüber zu entscheiden haben, was sie mit dem Hirten tun werden.

Szenario 2: In den afghanischen Bergen befinden sich ferngesteuerte Roboter, und die Entscheidung darüber, was sie tun werden, wird von Menschen getroffen, die vielleicht mehrere hundert Kilometer von ihnen entfernt in einer Steuerzentrale sitzen.

Szenario 3: In den afghanischen Bergen befinden sich Roboter, die man so programmiert hat, dass sie auf Grund der ihnen von der Außenwelt zugeleiteten Informationen eine „Entscheidung“ treffen können. Und in ihrem Programm steht: Töte niemanden, der Dich nicht gefährdet!

In den zwei ersten Szenarien sind es Menschen, die unmittelbar die Entscheidung treffen. Allerdings gibt es einen wichtigen Unterschied. Wie mutig die Soldaten im ersten Szenario auch immer sein mögen, es kann kaum angenommen werden, dass sie gar keine Angst um ihr Leben haben werden. Ebenso wenig kann angenommen werden, dass sie gar kein Mitleid für den Hirten empfinden werden. Im zweiten Szenario fällt die Angst um das eigene Leben weg, da derjenige, der die Roboter steuert, sich nicht in einer Situation befindet, in der sein eigenes Leben von der Entscheidung betroffen wäre. Allerdings wird man bei ihm Mitleid mit dem Hirten vermuten dürfen.

man sich primär aus Pflicht und sekundär auch noch aus Liebe um sie kümmert, ist durchaus zulässig. Die Präsenz pathologischer Triebfedern kann allerdings die Erkenntnis erschweren, ob wir uns zur Handlung entschieden hätten, wenn nur rationale Triebfedern vorgelegen hätten. Diese Präsenz tangiert also nicht primär die *ratio essendi*, sondern die *ratio cognoscendi* des moralischen Handelns.

Wenden wir uns dann dem dritten Szenario zu. Wie hätten Roboter in der entsprechenden Situation gehandelt? Hätten sie den Hirten erschossen oder nicht? Sie hätten sicherlich die Möglichkeit in Betracht gezogen, dass der Hirte den Taliban ihre Präsenz mitteilt. Sie hätten dieser Möglichkeit auch eine gewisse Wahrscheinlichkeit zugeordnet. Aber welche Wahrscheinlichkeit und auf Grund wovon? Und ab welcher Wahrscheinlichkeit hätten sie auf eine mögliche Gefährdung geschlossen?³⁰ Und gesetzt den Fall, sie hätten auf eine solche Gefährdung geschlossen: Hätten sie den Hirten erschossen? Schließlich hat man sie dazu programmiert, niemanden zu töten, der sie nicht gefährdet, und nicht jeden zu töten, der sie gefährdet. Man hätte sie natürlich so programmieren können, genauso wie man sie auf beide Weisen hätte programmieren können. Aber wie hätte ein Roboter in einem solchen Fall entschieden und vor allem, wie hätte er eine Entscheidung getroffen? Was geschieht mit einem Roboter, der in sich zwei sich widersprechende Handlungsanweisungen hat, wobei keine Hierarchisierung vorgesehen ist?³¹

NOEL SHARKEY meint, man sollte den selbstgesteuerten militärischen Robotern *nicht* die Möglichkeit geben, Menschen zu töten, da Roboter nicht zwischen Kombattanten und Nicht-Kombattanten bzw. zwischen einem gefährlichen und einem nicht-gefährlichen Gegenüber unterscheiden können. Eine solche Unterscheidung setzt nämlich voraus, dass man die Intentionen des Gegenübers erkennen kann, was voraussetzt, dass man, so SHARKEY, über eine, wenn auch nur implizite, Theorie des Geistes verfügt.³² Wenn wir den von uns dargestellten Fall nehmen, müssten die Roboter in der Lage sein, den Menschen als einen afghanischen Hirten zu identifizieren; sie müssten des Weiteren in der Lage sein, Vermutungen anzustellen über das, was man von einem afghanischen Hirten erwarten kann, was wiederum voraussetzt, dass sie genügend Bescheid über die geopolitische Lage und über die Haltung der lo-

³⁰ Hier ein reeller Fall: 1983 wurde ein japanischer Angestellter, der in einer Fabrik arbeitete, von einem Roboter getötet, weil der Roboter ihn fälschlicherweise als eine Gefahr ansah. Der Roboter war so programmiert worden, dass er mögliche Hindernisse aus dem Weg räumen sollte, wobei aber nicht daran gedacht wurde, dass er einen Menschen als solches Hindernis identifizieren konnte. Nachdem er den Angestellten weggeschoben hatte – tragischerweise in eine funktionierende Maschine – fuhr der Roboter mit seiner Arbeit fort.

³¹ Asimovs drei Gesetze sind hierarchisch geordnet: Gesetz 1: Ein Roboter darf keinen Menschen verletzen oder durch Unterlassung zulassen, dass ein Mensch verletzt wird; Gesetz 2: Ein Roboter muss Befehlen gehorchen, die ihm ein Mensch gibt, es sei denn diese Befehle würden mit dem ersten Gesetz im Widerspruch stehen; Gesetz 3: Ein Roboter muss seine eigene Existenz bewahren, sofern diese Bewahrung nicht mit dem ersten und zweiten Gesetz im Widerspruch steht (I. ASIMOV: I, Robot (1968)).

³² N. SHARKEY: Killing made easy (2012), S. 116 ff.

kalen Bevölkerung zu den Taliban wissen.³³ Hier stellt sich dann die Frage, ob ein Roboter einerseits über die nötigen Informationen verfügen kann und ob er andererseits über die nötige Erfahrung verfügt, um mit dieser Information umzugehen. Welche Wichtigkeit man auch dem Besitz von Informationen bei der ethischen Entscheidungsfindung zuerkennen kann, so bleibt doch außer Frage, dass auch die Erfahrung eine Rolle spielt. Sie bildet den Hintergrund, vor dem wir die Informationen verarbeiten, um dann zu einer Entscheidung zu kommen.

Während für SHARKEY keine tötenden militärischen Roboter gebaut werden sollten, glauben G.J. LOKHORST und J. VAN DEN HOVEN, dass Roboter zuverlässiger sind als Menschen, und dass sie unter dem Strich weniger Menschen töten werden als es menschliche Soldaten tun würden. Für sie steht fest, dass wir eine größere Kontrolle über die Handlungen programmierter Roboter als über diejenigen ausgebildeter menschlicher Soldaten haben.³⁴ Programmiert man einen Roboter so, dass er niemanden tötet, der ihn nicht unmittelbar mit einer Waffe bedroht, dann wird er keinen solchen Menschen töten. Anders bei den menschlichen Soldaten, die unter permanentem Stress stehen und manchmal „durchdrehen“ können.³⁵

Bevor wir uns im nächsten Teil mit der Frage befassen, ob man Robotern Rechte zusprechen sollte, ist hier noch abschließend die Frage zu stellen, was wir tun können, um unsere eigenen Rechte gegenüber Robotern zu schützen. M. SCHEUTZ erwägt die Möglichkeit, dass uns rein rationale Roboter manipulieren können, indem sie von unseren gefühlsmäßigen Bindungen profitieren. Während wir uns nämlich gefühlsmäßig an Roboter binden können, gilt dies nicht für die Roboter uns gegenüber. Sein Vorschlag zur Vermeidung einer solchen Situation: „Hier könnte [ein] radikaler Schritt vonnöten sein, nämlich künftige Roboter mit menschenähnlichen Emotionen und Gefühlen auszustatten.“³⁶ SCHEUTZ geht von dem Gedanken aus, dass es letztendlich immer Gefühle sind, die uns davon abhalten, einem anderen Wesen Schaden zuzufügen, wenn eine Schadenszufügung in unserem Interesse läge. Will man demnach

³³ Dazu N. SHARKEY/A. SHARKEY: *The rights and wrongs of robot care* (2012), S. 272: “Robots are able to follow well-specified rules, but they are not good at understanding social context and predicting likely intentions.”

³⁴ G.-J. LOKHORST/J. VAN DEN HOVEN: *Responsibility for military robots* (2012), S. 148ff.

³⁵ Aber hier darf man nicht vergessen, dass auch die Technik pannenfällig ist, so dass auch ein Roboter durchaus, wenn auch in einem anderen Sinn, durchdrehen kann, wobei aber die Konsequenzen dieselben sein können.

³⁶ M. SCHEUTZ: *The inherent dangers of unidirectional emotional bonds* (2012), S. 218: “[A] radical step might be necessary to achieve this: to endow future robots with human-like emotions and feelings.”

vermeiden, dass Roboter uns manipulieren, muss man sie mit menschenähnlichen Gefühlen ausstatten bzw. mit einem funktionalen Äquivalent zu menschlichen Gefühlen. Sollen also Roboter eines Tages zu moralischen Entscheidern werden, dann genügt es nicht, ihre künstliche Intelligenz zu erweitern, so dass sie auch die moralische Intelligenz mit umfasst, sondern man muss ihnen auch künstliche Gefühle – oder zumindest ein funktionales Äquivalent – mit auf den Weg geben. Damit hätten wir eine Welt, in welcher Roboter leiden könnten, womit sich die Frage stellen würde, ob wir ihnen auch Rechte zusprechen müssten. Was in den Augen von SCHEUTZ notwendig erscheint, um uns gegenüber den Robotern zu schützen, wird somit zu einem Faktor, der die Notwendigkeit begründen könnte, die Roboter vor uns zu schützen.

5. Rechte für Roboter?

Für JEREMY BENTHAM war die ethisch relevante Frage nicht, ob ein Wesen denken oder sprechen konnte, sondern ob es fühlen und besonders ob es leiden konnte. Viele Tiere waren in seinen Augen leidensfähige Wesen, so dass man ihr Wohlergehen im utilitaristischen Kalkül mitberücksichtigen musste.³⁷

Können Roboter leiden?³⁸ Wir könnten den Leidensbegriff so fassen, dass er nur auf Wesen angewendet werden kann, die über eine organische *Hardware* verfügen, zu der u.a. ein Nervensystem gehört. Somit wäre ein metaphysisches Problem durch einen semantischen Trick gelöst – oder aufgelöst.

Nun ist es aber so, dass die Menschen nicht auf die Entdeckungen der Anatomie und Physiologie gewartet haben, bevor sie die Leidensfähigkeit anderer Wesen erkannten. Unser alltäglicher Leidensbegriff ist ein phänomenologischer, d.h. er fußt auf der Art und Weise, wie wir die Wesen in unserer Umwelt erfahren und inwiefern wir uns mit ihnen oder sie sich mit uns identifizieren können. Das ich ein mir noch völlig unbekanntes Tier als leidend erfahre, nicht aber eine mir noch völlig unbekannt Pflanze, hat nichts mit einer Kenntnis des inneren Organismus dieser Wesen zu tun, sondern mit dem, was ich beobachte.³⁹ Und ein weiteres Element das hier berücksichtigt werden muss, ist die

³⁷ Wobei Bentham allerdings keine moralischen Rechte, zumindest so wie sie traditionellerweise aufgefasst werden, anerkennen will.

³⁸ Einige Autoren verbinden die Beantwortung dieser Frage mit der Frage, ob man Roboter bestrafen kann. So schließt etwa Asaro eine mögliche Bestrafung von Robotern aus, da sie nicht leiden können – und Strafe immer eine Zufügung von Leiden ist (P.M. ASARO: *A body to kick, but still no soul to damn* (2012), S. 182).

³⁹ Und dass ich diese Wesen überhaupt als Tier oder Pflanze identifiziere, hat auch nicht unbedingt etwas mit einer Kenntnis ihrer internen Physiologie zu tun.

Interaktion, die sich zum einen aus der Identifikation ergibt, diese zum andern auch weiter verstärkt.

Anstatt, wie BENTHAM es tut, zu fragen, ob ein Wesen leiden kann, sollten wir vielmehr fragen, ob und inwiefern wir mit einem Wesen interagieren und uns mit ihm identifizieren können, wobei sowohl das äußere Aussehen⁴⁰ als auch das Verhalten eine Rolle spielen. Es ist erst auf Grund einer solchen Interaktion und Identifikation, dass wir einem Wesen die Fähigkeit zusprechen, leiden zu können. Die Frage, ob ein solches Wesen dann auch *tatsächlich* leidet, ist dann eigentlich irrelevant.⁴¹ Im Vokabular KANTS ausgedrückt: An einem Wesen ist uns nicht das Leiden an sich zugänglich, sondern immer nur das Leiden für uns, und dieses Leiden für uns offenbart sich im Rahmen möglicher Interaktionen mit dem betreffenden Wesen. Was auch immer im tiefsten Innern dieses Wesens vorgehen mag, an erster Stelle steht unser Verhältnis zu dem Wesen. Und dieses Verhältnis wird zum Teil durch angeborene, im Rahmen der Evolution entstandene Mechanismen bestimmt.⁴²

Was die Tiere anbelangt, können wir nur in ganz begrenztem Maße bestimmen, wie sie aussehen werden und wie sich unsere Interaktionen mit ihnen gestalten werden. Wir können zwar Tiere züchten und erziehen und neuerdings auch genetisch manipulieren, aber wir können bis auf weiteres kein Tier herstellen, das einem Menschen so ähnlich sieht wie manche heutige Roboter und das sich so menschenähnlich verhält wie diese Roboter. Wir müssen die Tiere so nehmen, wie sie sind. Bei den Robotern stehen wir aber vor einer fundamentalen Wahl: Es steht in unserer Macht, Roboter zu schaffen, denen wir früher oder später Rechte zuerkennen müssen, da sie zumindest für uns jene Bedingungen erfüllen werden, die uns als Grundlage für die Rechtezuschreibung dienen. Was auch immer der Roboter als noumenales Wesen sein mag und wie sehr er sich auch von meinem eigenen noumenalen Selbst unterscheiden mag – als phänomenales Wesen, das mir gegenübersteht und das mir allein zugänglich ist, kann er eines Tages in allen relevanten Hinsichten jenen phänomenalen Wesen gleichen, denen ich Rechte zuschreibe.

⁴⁰ Ich denke nicht, dass das *innere* Aussehen, etwa die neuronalen Zusammenhänge im Gehirn, relevant ist.

⁴¹ David Levy prophezeit, dass es in einem guten Jahrzehnt neben der Künstlichen Intelligenz auch Künstliche Gefühle geben wird (D. LEVY: *Love & sex with robots* (2008), S. 86).

⁴² So etwa M.R. CALO: *Robots and privacy* (2012), S. 196. Verwiesen sei in diesem Zusammenhang auch auf die sogenannten Spiegelneuronen. Einige Autoren erwägen die Möglichkeit, Roboter mit einem funktionalen Äquivalent dieser Spiegelneuronen auszurüsten. Dazu etwa J. HUGHES: *Compassionate AI and selfless robots* (2012), S. 79; M. GUARINI/P. BELLO: *Robotic warfare* (2012), S. 141.

Dem wird man sicherlich entgegenhalten, dass ein Roboter aus anorganischer Materie besteht, wohingegen die Wesen, denen wir bislang Rechte zugestanden haben, immer aus organischer Materie bestanden haben. Aber sollte die Natur der Materie, aus der ein Wesen besteht, moralisch relevant sein? Wenn wir eines Tages mit Wesen eines anderen Sonnensystems in Kontakt treten sollten und wenn sich herausstellen sollte, dass wir mit diesen Wesen die wichtigsten Interaktionen eingehen können, die wir tagtäglich mit unseren Mitmenschen eingehen, werden wir dann wirklich die Anerkennung von Rechten von einer biochemischen Analyse abhängig machen? Und selbst wenn die Natur der Materie, aus der ein Wesen besteht, moralisch relevant sein sollte, so kann darauf hingewiesen werden, dass schon heute versucht wird, Roboter zu bauen, deren zentrales Steuerungssystem – deren Gehirn, wenn man so will – aus organischer Materie besteht.⁴³ Spätestens wenn die ersten biologischen Roboter da sein werden, wird sich uns die Frage stellen, wie wir sie moralisch behandeln sollten und ob wir ihnen auch Rechte zuerkennen sollten.

Aus Kantischer Sicht wird man vielleicht einwenden, dass nur ein phänomenales Wesen das sich selbst als noumenales Wesen denken kann, es verdient, als moralisches Wesen behandelt zu werden. Aber wie wissen wir, ob ein phänomenales Wesen sich als noumenales Wesen denken kann?⁴⁴ Stammt dieses Wissen nicht aus unseren Interaktionen mit diesem Wesen? Aber die Interaktionen spielen sich immer in der Welt für uns ab, also auf phänomenaler Ebene. Die zentrale Frage ist somit nicht, ob das phänomenale Wesen sich tatsächlich und an sich als noumenales Wesen denken kann, sondern ob es sich als phänomenales Wesen so verhält, dass wir annehmen können, dass es sich als noumenales Wesen denken kann. Wir werden den Robotern dann ein noumenales Wesen zusprechen, wenn sie sich in unseren täglichen Interaktionen mit ihnen so verhalten werden, dass wir keinen triftigen Grund mehr haben, ihnen ein noumenales Selbst abzusprechen.

Aber, so ein weiterer möglicher Einwand, ist der Roboter letztendlich nicht immer nur eine bloße Maschine? Gemeint ist damit, dass der Roboter keine typisch menschlichen Züge hat. Sagen wir nicht auch manchmal von be-

⁴³ Dazu etwa K. WARWICK: *Robots with biological brains* (2012). Umgekehrt gilt aber auch für den biologischen Menschen, dass er in immer zunehmendem Maße durch nicht-biologische Materie ergänzt wird – künstliche innere oder externe Organe. Ab welchem Prozentsatz künstlicher Materie hört man auf, ein biologisches Wesen zu sein bzw. ab welchem Prozentsatz biologischer Materie hört man auf, ein künstliches Wesen zu sein?

⁴⁴ Wobei ich hier einfach einmal voraussetze, dass diese Kantsche Unterscheidung überhaupt brauchbar ist – was ich glaube.

stimmten Menschen, dass sie wie Maschinen wirken: Ihr Verhalten ist mechanisch, voraussagbar und lässt keinen Platz für Gefühle oder Spontaneität? Dass Roboter bis vor kurzem tatsächlich nur bloße Maschinen waren, steht außer Frage. Die Fortschritte auf dem Gebiet der Robotik stellen uns heute aber vor die Frage, ob wir nur Roboter haben wollen, die für uns bloße Maschinen sind, mit denen wir also so gut wie überhaupt nicht interagieren können, oder ob wir weiter auf dem Weg gehen wollen, den wir eingeschlagen haben und der uns dann zu einem Punkt führen wird – falls wir ihn nicht schon erreicht haben –, den R. SPARROW wie folgt beschreibt: „Maschinen werden dann Menschen sein, wenn wir sie nicht sterben lassen können, ohne vor dem gleichen moralischen Dilemma zu stehen, dem wir ausgesetzt sind, wenn wir darüber nachdenken, ob wir ein menschliches Wesen sterben lassen.“⁴⁵ Man wird sich hier allerdings die Frage stellen müssen, was es überhaupt bedeutet, einen Roboter sterben zu lassen. SPARROW hätte das Wort „die“ – sterben – zumindest zwischen Anführungsstriche setzen müssen.

Dass sich die hier erwähnten Probleme nur für Philosophen stellen, zeigen etwa die Existenz – bereits seit 1999 – der *American Society for the Prevention of Cruelty against Robots (ASPCR)* und die in Südkorea geführten Überlegungen, bis zum Ende des jetzigen Jahrzehnts den Schutz von Robotern vor Zerstörung und Misshandlung in das Gesetz zu integrieren.⁴⁶

Schluss

Wir stehen heute vor der grundsätzlichen Frage, ob wir weiter danach streben sollen, Roboter zu bauen, die sich für uns letzten Endes nicht mehr in allen, oder doch zumindest in den wesentlichen, moralischen Hinsichten von Menschen unterscheiden lassen, oder ob wir dem gegenwärtigen Trend in der Robotik ein Ende setzen sollten.⁴⁷ Gekoppelt an diese Frage ist eine andere: Wollen wir, dass Roboter immer mehr jene Arbeiten und Dienstleistungen übernehmen, die bislang nur Menschen ausüben konnten, weil ihre Ausfüh-

⁴⁵ R. SPARROW: *Can machines be people?* (2012), S. 307: “Machines will be people when we can’t let them die without facing the same moral dilemma that we would when thinking about letting a human being die.”

⁴⁶ R. IKONICOFF: *Faudra-t-il accorder des droits aux machines intelligentes?* (2013), S. 108.

⁴⁷ Für einige Autoren werden die Fakten die moralische Frage „beantworten“: “Autonomous systems are coming whether we like it or not” (W. WALLACH/C. ALLEN: *Moral machines* (2012), S. 16. Anstatt sich also in einem Nachhutgefecht zu üben, wo der Sieger sowieso schon feststeht, sollte man sich vielmehr mit der Frage befassen, wie man die Interaktionen zwischen Menschen und menschenähnlichen Robotern auf moralischer Ebene gestalten sollte.

rungen jene Kompetenzen voraussetzten, die bisher eben nur Menschen besaßen? Wir können die weniger attraktiven Aspekte der Krankenpflege weiterhin Menschen überlassen, aber wir können diese Dienstleistungen auch Robotern übertragen. Aber wenn wir sie Robotern übertragen, dann müssen wir uns die Frage stellen, ob wir diese Roboter so menschenähnlich wie nur möglich gestalten sollen, damit den Kranken ein in allen relevanten Hinsichten äquivalenter Dienst angeboten wird – wenn nicht sogar ein noch besserer Dienst. Ob der Roboter, der Autos am Fließband herstellt, einem Menschen ähnlich ist oder nicht, spielt keine Rolle. Es spielt aber eine große Rolle, ob ein Roboter der sich um Kranke kümmert einem Menschen ähnlich sieht oder nicht und sich so wie ein Krankenpfleger verhält oder nicht.

In einem Interview an die Tageszeitung *Le Monde*, stellt KATE DARLING fest, dass vor allem die Spielzeugindustrie den Bau von Robotern vorantreibt, zu denen wir gefühlsmäßige Bindungen aufbauen können.⁴⁸ Man denke hier zwei oder drei Jahrzehnte zurück, als die sogenannten *Tamagochis* Ströme von Tränen bei Kindern auslösten, die es vergessen hatten, sie zu „ernähren“, so dass der *Tamagochi* „starb“. Für DARLING hat die Spielzeugindustrie – aber, könnte man hinzufügen, die Robotikindustrie allgemein – eine derart große Macht, dass man die Entwicklung nicht mehr stoppen kann. Und je früher man Kinder emotional, im Rahmen bestimmter Interaktionen, an Maschinen bindet, umso mehr werden sie sich mit Maschinen identifizieren können.⁴⁹

Wie wichtig es auch sein mag, sich mit der Frage zu befassen, wie wir uns vor Robotern schützen können, so sollten wir diese Frage mit einer anderen verbinden, und zwar wie wir uns vor der Industrie schützen können, die diese Roboter hervorbringt, vor denen wir uns zu schützen haben werden. Die Roboterethik verweist somit über sich selbst hinaus auf die Frage nach der menschlichen Gestaltung der Zukunft. Vielleicht haben DARLING und andere Recht, wenn sie die Entwicklung als unaufhaltbar ansehen. Aber es gilt einzusehen, dass diese Unaufhaltbarkeit nicht den Charakter eines reinen Naturgeschehens hat. Wenn genügend Menschen es wollten, könnte sie aufgehalten werden.

⁴⁸ K. DARLING: „Donnons des droits aux robots“ (2013), S. 3. Vielen Dank an Hubert Hausemer, der mich auf dieses Interview aufmerksam gemacht hat.

⁴⁹ Dazu Levy: “[A]s robots become increasingly lifelike in their behavior and as these children influence the adults around them and grow into adults themselves, more and more people will treat robots as if they were mental, social, and moral beings – thus raising the perception of robotic creatures toward the level of biological creatures” (D. LEVY, *Love & sex with robots* (2008), S. 98).

Zusammenfassung

CAMPAGNA, NORBERT: **Die Roboter als Gegenstand der Ethik.** ETHICA 22 (2014) 2, 99–120

In den letzten Jahrzehnten hat die Robotik eine rasante Entwicklung erlebt, wobei besonders die immer größer werdende Ähnlichkeit zwischen Robotern einerseits und Menschen und Tieren andererseits hervorsticht. Diese Ähnlichkeit beschränkt sich nicht nur auf das äußere Aussehen, sondern tangiert auch das Verhalten und zunehmend bestimmte innere Prozesse. Die Entwicklung der Robotik hat dazu geführt, dass sich ein neuer Zweig der Angewandten Ethik herausgebildet hat, die Roboterethik. In diesem Beitrag werden einige der fundamentalen Fragen der Roboterethik angeschnitten, wie etwa die Frage, ob wir die Maschinen überhaupt immer menschenähnlicher gestalten sollten. Wir sollten uns nämlich bewusst werden, dass mit der zunehmenden Menschenähnlichkeit der Roboter sich früher oder später die Frage stellen wird, ob wir Roboter menschenähnlich behandeln sollten, ob wir ihnen also Rechte zuerkennen sollten.

Anthropomorphisierung
Maschine
Moral
Rechte
Roboter

Summary

CAMPAGNA, NORBERT: **The robots as subject of ethics.** ETHICA 22 (2014) 2, 99–120

In the last decades, robotics has experienced a rapid evolution, a central feature of this evolution being the growing similarity between robots on the one hand and humans and animals on the other hand. This similarity is not limited to external appearance but can also be found at the level of behaviour and, even more so, of certain inner processes. The evolution of robotics led to the emergence of a new branch of applied ethics, viz, roboethics. In this article some of the most fundamental questions of roboethics are introduced, e.g. the question whether we should continue to construct machines which are more and more like human beings. We should, in fact, realize that with the growing similarity between robots and human beings we will sooner or later be confronted with the question whether we should treat robots as we treat human beings, i.e. whether we should grant them rights.

Anthropomorphization
machine
morality
rights
robot

L i t e r a t u r

ASARO, PETER M.: A body to kick, but still no soul to damn. Legal perspectives on robotics. In: Patrick Lin/Keith Abney/George Bekey (eds.): Robot ethics. Cambridge (Ma): The MIT Press, 2012.

ASIMOV, ISAAC: I, Robot. London: Grafton Books, 1968 [Ersterscheinung: 1942].

BEKEY, GEORGE A.: Current trends in Robotics: Technology and Ethics. In: Patrick Lin/Keith Abney/George Bekey (eds.): Robot ethics. Cambridge (Ma): The MIT Press, 2012.

BORENSTEIN, JASON/PEARSON, YVETTE: Robot caregivers. Ethical issues across the human lifespan. In: Patrick Lin/Keith Abney/George Bekey (eds.): Robot ethics. Cambridge (Ma): The MIT Press, 2012.

CALO, M. RYAN: Robots and privacy. In: Patrick Lin/Keith Abney/George Bekey (eds.): Robot ethics. Cambridge (Ma): The MIT Press, 2012.

- CAMPAGNA, NORBERT: Roboterethik. *Information Philosophie* (2012) 5, 58–61.
- DARLING, KATE: „Donnons des droits aux robots“. Interview erschienen in *Le Monde* vom 16. Februar 2013, S. 3.
- GEORGES, THOMAS M.: *Digital soul. Intelligent machines and human values*. Oxford: Westview, 2004.
- GUARINI, MARCELLO/BELLO, PAUL: Robotic warfare: Some challenges in moving from non-civilian to civilian threats. In: Patrick Lin/Keith Abney/George Bekey (eds.): *Robot ethics*. Cambridge (Ma): The MIT Press, 2012.
- HUGHES, JAMES: Compassionate AI and selfless robots: A Buddhist approach. In: Patrick Lin/Keith Abney/George Bekey (eds.): *Robot ethics*. Cambridge (Ma): The MIT Press, 2012.
- IKONICOFF, ROMÁN: Faudra-t-il accorder des droits aux machines intelligentes? *Science & Vie* (avril 2013) 1167, S. 108.
- LEVY, DAVID: *Love & sex with robots*. London: Duckworth Overlook, 2008.
- The ethics of robot prostitutes. In: Patrick Lin/Keith Abney/George Bekey (eds.): *Robot ethics*. Cambridge (Ma): The MIT Press, 2012.
- LOKHORST, GERT-JAN/VAN DEN HOVEN, JEROEN: Responsibility for military robots. In: Patrick Lin/Keith Abney/George Bekey (eds.): *Robot ethics*. Cambridge (Ma): The MIT Press, 2012.
- McFARLAND, DAVID M.: *Guilty robots, happy dogs. The question of alien minds*. Oxford: Oxford University Press, 2009.
- PETERSEN, STEVE: Designing people to serve. In: Patrick Lin/Keith Abney/George Bekey (eds.): *Robot ethics*. Cambridge (Ma): The MIT Press, 2012.
- SHARKEY, NOEL: Killing made easy. From joysticks to politics. In: Patrick Lin/Keith Abney/George Bekey (eds.): *Robot ethics*. Cambridge (Ma): The MIT Press, 2012.
- SHARKEY, NOEL/SHARKEY, AMANDA: The rights and wrongs of robot care. In: Patrick Lin/Keith Abney/George Bekey (eds.): *Robot ethics*. Cambridge (Ma): The MIT Press, 2012.
- SCHEUTZ, MATTHIAS: The inherent dangers of unidirectional emotional bonds between humans and social robots. In: Patrick Lin/Keith Abney/George Bekey (eds.): *Robot ethics*. Cambridge (Ma): The MIT Press, 2012.
- SPARROW, ROB: Can machines be people? Reflections on the Turing Triage Test. In: Patrick Lin/Keith Abney/George Bekey (eds.): *Robot ethics*. Cambridge (Ma): The MIT Press, 2012.
- TURKLE, SHERRY: *The Second Self. Computers and the human spirit*. New York: Simon and Schuster, 1984.
- WALLACH, WENDELL & ALLEN, COLIN: *Moral machines. Teaching robots right from wrong*. Oxford: Oxford University Press, 2012.
- WARWICK, KEVIN: Robots with biological brains. In: Patrick Lin/Keith Abney/George Bekey (eds.): *Robot ethics*. Cambridge (Ma): The MIT Press, 2012.