

The Edge Question 2015 Einmal im Jahr stellt der Literaturagent John Brockman prominenten Wissenschaftlern und Künstlern eine Frage

Was, wenn sie leiden müssen?

Das menschliche Denken ist deshalb so effizient, weil Menschen Leid kennen. Erkenntnis auf höchster Ebene ist eine Sache, die sich selbst befeuernde Motivation eine andere. Künstliches Denken könnte schon bald effizienter werden, aber wird es dann zwangsläufig Leid erfahren? Wird das Leiden sogar notwendiger Teil einer post-biotischen Intelligenz sein müssen, die diesen Namen verdienen soll? Oder sind negative Gefühle von der Evolution nur für Menschen vorgesehen?

Menschen haben zerbrechliche Körper, sie werden in eine gefährliche Umwelt hineingeboren und sie befinden sich in einem beständigen Bemühen, ihre Sterblichkeit leugnen zu müssen. Unser Hirn arbeitet fortwährend daran, dass die Wahrscheinlichkeit, böse überrascht zu werden, minimal bleibt. Wir sind deshalb schlau, weil wir verletzen und Bedauern empfinden können und weil wir uns bemühen, für uns selbst eine überzeugende Form der Täuschung und der wenigstens symbolischen Unsterblichkeit zu finden. Die Frage also lautet: Benötigt auch künstliche Intelligenz den zerbrechlichen Körper, die unsichere Umwelt und ein Wissen um die eigene Sterblichkeit? Es wird irgendwann denkende Maschinen geben. Werden sie sich um ihre eigene Gedanken scheren? Warum sollten sie?

Das Wesen des Leidens liegt in der Tatsache, dass ein denkendes System nicht anders kann, als sich mit seiner Gefährdung zu identifizieren. Leiden hat natürlich verschiedene Aspekte. Doch es ist diese Identifikation damit, die zählt. Denn was dem

System solches Unbehagen bereitet, ist untrennbar mit ihm verbunden, ein Zustand, der dessen Autonomie bedroht. Wenn man das versteht, sieht man, warum die „Erfindung“ des bewussten Leidens durch die biologische Evolution so erfolgreich war. Sie ist ein zugleich innovatives wie absolut böses und grausames Konzept.

Klar, diese Phänomenologie der Zugehörigkeit erklärt Leiden nicht hinreichend. Wir alle können uns leicht selbstbewusste Lebewesen vorstellen, die nicht leiden. Leiden benötigt einen Zustand mit negativer Wertigkeit, der vollständig in das Selbstbild und Selbstverständnis eines Systems integriert ist. Durch diese Integration erst werden aus negativen Präferenzen negative subjektive Präferenzen – und das sind die bewussten und klaren Vorstellungen darüber, dass es eben meine eigenen Präferenzen, Wünsche, Vorstellungen sind, die scheitern, gescheitert sind oder in Zukunft scheitern werden.

Das bedeutet nicht, dass ein System mit künstlicher Intelligenz ebenfalls ein völliges Verständnis dessen haben muss, was diese Präferenzen genau sind. Es reicht, dass es nicht will, dass sich die Erfahrung eines negativ empfundenen Zustandes wiederholt, dass es diesen beenden will.

Ich bin oft gefragt worden, ob wir nicht Maschinen bauen könnten, die sehr intelligent sind und leidensunfähig zugleich. Doch kann es echte Intelligenz ohne existenzielle Sorge überhaupt geben?

THOMAS METZINGER, *Philosoph, Universität Mainz*



Was denken Sie über Maschinen, die denken?

Dieses Jahr geht es um künstliche Intelligenz. Hier schon mal eine Auswahl von Antworten. Mehr als 100 werden in voller Länge am Samstag auf der Webseite edge.org veröffentlicht



In der Debatte um die künstliche Intelligenz gibt es Pessimisten, die eine Machtübernahme der Maschinen fürchten, und Optimisten, die auf die *Singularity* hoffen, die Erweckung des synthetischen Geistes. Momente wie der Sieg des Computers Deep Blue über den Schachmeister Kasparov 1996 und des Programms Watson über die Champions der Quizes- und „Jeopardy“ 2011 gelten als erste Schritte. FOTOS: REUTERS, AFP

Man muss die Frage anders stellen

Edsger Dijkstra hatte recht, als er 1984 sagte, ob Maschinen denken können, sei „ungefähr so relevant wie die Frage, ob U-Boote schwimmen können“. Er zielte darauf ab, dass es bei beiden Fragen ganz darauf ankommt, was man unter „schwimmen“ oder eben „denken“ versteht. Auf Englisch würde man nicht sagen, dass ein U-Boot „schwimmt“, im Russischen aber schon. Für die Eigenschaften eines U-Bootes ist das unwichtig. Deshalb sollten wir lieber herausfinden, was Maschinen tun können und ob wir uns vor ihren Fähigkeiten fürchten sollten.

Pessimisten warnen davor, dass wir nicht wissen, wie wir sichere komplexe Systeme künstlicher Intelligenz erschaffen können. Da ist etwas dran. Wir wissen aber auch nicht, wie wir sichere Systeme erschaffen können, die nichts mit künstlicher Intelligenz zu tun haben. Jede komplexere Erfindung wird eine Mischung aus positiven Ergebnissen und nicht beabsichtigten Konsequenzen haben. Aber gibt es Besorgniserregendes, das nur die künstliche Intelligenz betreffen würde? Ich glaube, die interessantesten Themen in dieser Hinsicht sind Anpassungsfähigkeit, Autonomie und Universalität der Systeme.

Lernende Maschinen können sich anpassen. Sie verändern sich, je nachdem, was sie hinzulernen. Anpassungsfähigkeit ist nützlich. Wir möchten zum Beispiel, dass unsere Rechtschreibkorrektur zügig neue Wörter wie „Bitcoin“ lernt. Aber die Anpassung kann zu neuen Fehlern führen. Entwickler müssen deshalb lernen, mit Anpassungsfähigkeit umzugehen.

Die zweite Sorge gilt der Autonomie. Wenn künstliche Intelligenz allein agiert, können Fehler entstehen, die nicht passie-

ren würden, wenn ein Mensch beteiligt wäre. Die Sorge ist ebenfalls ernst zu nehmen, doch auch hier gilt: Das ist keine Besonderheit der künstlichen Intelligenz. Wir machen dauernd Kompromisse, wenn wir autonom handelnde Technik einsetzen.

Die dritte Besorgnis gilt der Universalität der intelligenten Maschinen. 1965 schreibt I. J. Good, „eine ultraintelligente Maschine könnte selbst immer bessere Maschinen entwerfen; das würde dann eindeutig zu einer ‚Intelligenzexplosion‘ führen, so dass die Intelligenz des Menschen weit abgeschlagen wäre. Deshalb wird die ultraintelligente Maschine die letzte Erfindung sein, die der Mensch machen muss.“

Ich finde, in dieser Aussage wird „Intelligenz“ zu einer überragenden Superfähigkeit stilisiert. Ich glaube, die Realität ist etwas differenzierter. Die klügste Person, die ich kenne, ist nicht zwangsläufig die erfolgreichste. Die klügste Politik ist nicht immer die, die umgesetzt wird. Und: Ich kenne viele Probleme, die sich mit Intelligenz überhaupt nicht bewältigen lassen, denn ganz egal, wie clever man ist, wird man niemals ausreichende Rechenkraft haben, um diese Probleme zu lösen.

Aber natürlich gibt es viele Probleme, bei denen Intelligenz sehr hilft. Computer sind Werkzeuge für eine komplexe Welt. Künstliche Intelligenz als Teil unseres Baukastens verändert die Dinge nicht grundlegend. Mein Vorschlag ist: Lasst uns vorsichtig sein, wenn wir Mechanismen entwerfen, und lasst uns so oder so die besten Werkzeuge verwenden. Ganz unabhängig davon, ob auf einem Gerät der Sticker „Enthält künstliche Intelligenz“ klebt.

PETER NORVIG, *Forschungschef, Google*

Unsere Geschöpfe beherrschen uns

Unsere Gehirne und unsere Fähigkeit zum Denken wurden nicht von einem großen, intelligenten Schöpfer im Himmel entworfen, der entschied, wie wir denken und was unsere Motivationen sein sollen. Unsere Intelligenz und unsere Motivationen entwickelten sich. Die allermeisten Forscher der künstlichen Intelligenz würden dem zustimmen. Dennoch glauben manche offenbar noch, dass wir Menschen intelligente Schöpfer sind, die Maschinen erfinden können, die genau unseren Wünschen gemäß denken werden und die auch die von uns gewünschten Absichten haben werden. Wenn ich die Technologieschichte richtig verstehe, haben sie unrecht.

Das Problem ist eine Art irreführender Anthropomorphismus: Wir stellen uns vor, dass eine denkende Maschine so arbeitet wie wir. Doch wir schätzen uns selbst so extrem falsch ein, dass wir dasselbe auch mit den Maschinen tun. Deshalb erkennen wir nicht, dass die riesigen Denkmaschinen sich nach denselben Prinzipien entwickeln, wie es unsere Gehirne früher taten. Evolution, nicht ein großer Schöpferwille, prägt ihre Art zu denken.

Wir dürfen uns nicht länger für clevere Schöpfer mit voller Kontrolle über unsere Kreationen halten. Wir sollten lieber über unsere zukünftige Rolle nachdenken. Steht auch uns vielleicht das Schicksal des kleinen Mitochondrion bevor, jener einfachen Zelle, die vor langer Zeit von einer größeren absorbiert wurde? Das Mitochondrion gab das unabhängige Leben auf, um ein Kraftwerk für seinen Wirt zu werden, während der Wirt seinerseits die Energie-

produktion zugunsten anderer Tätigkeiten aufgab. Beide gewannen in diesem Prozess der Endosymbiose.

Sind wir so ähnlich? Digitale Information entwickelt sich überall um uns, auf Milliarden Smartphones, Tablets, Computern, Servern und winzigen Chips in Kühlschränken, Autos und Kleidern, sie bewegen sich um den Globus, durchdringen unsere Städte, unsere Wohnungen, sogar unsere Körper. Und wir füttern sie willig. Pro Tag werden mehr Telefone hergestellt als Babys geboren. Jede Minute werden 100 Stunden Video ins Netz gespeist, Milliarden Fotos werden in die immer größer werdende Cloud geladen. Clevere Programmierer schreiben immer cleverere Software, darunter auch Programme, die andere Programme schreiben, die Menschen nicht mehr verstehen oder nachvollziehen können. Die denkenden Maschinen sind unter uns und folgen ihren eigenen evolutionären Pfaden, während sie immer weiter wachsen.

Werden wir diese Maschinen kontrollieren? Können wir sichergehen, dass sie sich um uns kümmern werden? Nein. Und selbst wenn wir das erkennen: Wir wollen das, was die Maschinen uns geben, viel zu sehr, um dafür nicht mit unserer Unabhängigkeit zu bezahlen. Was denke ich also über denkende Maschinen? Ich denke, dass ich mich aus einer kleinen, unabhängigen denkenden Maschine in einen winzigen Teil einer riesigen denkenden Maschine verwandle.

SUSAN BLACKMORE, *Psychologin, London*

Oberflächliches Lernen

Sie können einem schon leidtun, die Leute in der National Security Agency: Sie überwachen jedermann, und jedermann ist nervt von ihnen. Aber wenigstens überwachen sie uns, um uns vor Terroristen zu schützen. In diesem Moment, wo Sie das hier lesen, geht irgendwo auf der Welt wieder ein Pop-up-Fenster auf einem Computerbildschirm auf, das sagt: „Sie haben gerade zwei Tonnen Stickstoff-Dünger gekauft. Leute, die zwei Tonnen Stickstoff-Dünger gekauft haben, interessierten sich auf für diese Zünder.“ Amazon, Facebook, Google und Microsoft überwachen uns nämlich auch. Aber weil das Spionieren dieser E-Giganten uns viele Dinge ermöglicht, ist es angelegentlich in Ordnung.

Dabei wird viel Rechnerleistung darauf verwendet, dass Maschinen darüber nachdenken, was wir vorhaben. Die gesamte Computerpower, die Datensammelfirmen auf uns loslassen, hat die Größe eines Exaflops – eine Trillion Rechenoperationen pro Sekunde. Aber was machen diese Maschinen, wenn sie über unser Denken nachdenken? Sie knüpfen Verbindungen zwischen großen Mengen persönlicher Daten und sie erkennen Muster.

Der aktuelle Trend bei Denkmaschinen wird dabei „Deep Learning“ genannt. Als ich das erste Mal davon hörte, fand ich die Idee aufregend, dass Maschinen nun endlich tiefere Aspekte unsere Existenz enthüllen würden – Wahrheit, Schönheit, Liebe. Ich wurde schnell eines Besseren belehrt. Die Tiefe im „Deep Learning“ bezeichnet nur die innere Architektur der Lernmaschinen: Sie bestehen aus vielen Schichten verknüpfter logischer Elemente, in Analogie zu den „tiefen“ Schichten verknüpfter Neuronen im Gehirn. Wie sich zeigt, ist es schwierig, eine hingekritzelte 7 von einer hingekritzelten 5 zu unterscheiden. Tief in den Achtzigerjahren scheiterten die ersten neuronalen Netzwerk-Computer an dieser Aufgabe.

Gleichzeitig sagten uns die Forscher auf dem Feld der neuronalen Computer, wenn sie nur viel größere Rechner hätten und viel größere Trainings-Sets mit Millionen hingekritzelte Zahlen statt nur Tausenden – dann würde die künstliche Intelligenz das schon hinkriegen.

Die haben wir jetzt. „Deep Learning“ ist informationstechnisch breit aufgestellt – riesige Datenmengen werden analysiert. Aber konzeptionell ist es flach. Computer können uns nun sagen, was die neuronalen Netzwerke im Gehirn schon immer wussten. Aber wenn sie einen handbeschriebenen Briefumschlag zur richtigen Postleitzahl lotsen können, sage ich: Nur weiter so!

SETH LLOYD, *Quantenmechaniker, Massachusetts Institute of Technology*

Roboter essen Seele auf

Je mehr wir über das Denken lernen, umso deutlicher wird, dass das menschliche Denken als Knotenpunkt unterschiedlicher Faktoren verstanden werden muss. Als Konglomerat aus Aktionen und Reaktionen des menschlichen Körpers, der menschlichen Emotion, Kultur und speziellen Eigenschaften des gesamten Gehirns. Einer der größten Fehler der westlich geprägten Philosophie war es, dem cartesianischen Dualismus des berühmten Statements: „Ich denke, also bin ich“ Glauben zu schenken. Es ist nicht weniger wahr zu sagen: „Ich verbrenne Kalorien, also bin ich.“ Noch besser wäre es, zu sagen: „Ich bin das Ergebnis menschlicher Evolutionsgeschichte, deshalb kann ich darüber nachdenken, dass ich bin.“

Wenn das so ist, dann ist Intelligenz kaum verwandt mit einer „Intelligenz“, die mit all diesen Faktoren nichts zu tun hat. Ich glaube an „künstliche Intelligenz“, so lange wir dabei bleiben, dass sie künstlich ist. Vergleicht man die Problemlösungen, die Computer heute anbieten, ihr Schachspiel, ihre „Argumentationen“, mit den menschlichen Fähigkeiten, dann ist es, als vergleiche man den Flug eines Airbus 320 mit einem Adler.

Es gibt viele Gründe dafür, dass künstliche Intelligenz nicht dasselbe wie reale Intelligenz ist. Erstens: Das „semantische Problem“. Es besteht darin, dass ein Computer, der Englisch in Mandarin übersetzt, weder Englisch noch Mandarin spricht. Es gibt keinen Computer, der eine menschliche Sprache erlernen kann, es gibt nur Bits und Anweisungen.

Zweitens gibt es jenes Problem, das ich die „dunkle Materie“ nenne, manche Philosophen sprechen von „implizitem Wissen.“ Wir besitzen die Fähigkeit, unsere Wahrnehmungen nach Wahrscheinlichkeiten zu bewerten, wegen unserer Gefühle, unserer Erfahrungen, unsere Sensibilität und unserer starken sozialen Bindungen.

Computer sind in der Lage, sehr viele Probleme zu lösen. Aber sie können nicht lieben. Sie können nicht urinieren. Sie können keine sozialen Bindungen eingehen. Die Furcht, dass Maschinen uns steuern könnten, ist nichts anderes als die in wissenschaftliche Fachsprache übersetzte Idee einer „Seele“ in der Religion. Sie lenkt von wirklichem Verständnis ab.

Künstliche Intelligenz kann eines Tages durchaus weniger künstlich als heute sein, wenn es ums Erschaffen von Körpern, Emotionen, sozialen Rollen, Werten geht. Bis es so weit ist, werden Staubsauger, Taschenrechner und niedliche Roboter, die Triviales von sich geben, nützlich sein.

DANIEL L. EVERETT, *Linguist und Autor, Bentley University, Massachusetts*

Menschsein

Was bedeutet es, Mensch zu sein? Obwohl es mittlerweile sieben Milliarden von uns gibt, hat auf diese Frage merkwürdigerweise noch niemand eine zufriedenstellende Antwort gefunden. Unbestreitbar ist indes, dass wir Menschen beharrlich Dinge entwickeln, mit denen sich unser Menschsein immer neu und immer anders ausdrücken lässt. Das Radio bescherte uns Hitler und die *Beach Boys*. Stachelnd und Klimaanlagen bescherten uns das westliche Nordamerika. Das Internet bescherte uns einen verschwindenden nordamerikanischen Mittelstand und Katzenbilder.

Andauernd hört man nun, neue Technologien würden uns befremden. Dabei ist es ja nun mal so, dass nicht irgendwelche Außerirdischen im Ufo kamen und uns die mitbrachten – wir haben sie selbst geschaffen. Deswegen können sie uns im besten Sinne nur vertraut sein. Und hier kommt die künstliche Intelligenz ins Spiel.

Viele Menschen gehen davon aus, dass denkende Maschinen einmal über eine Intelligenzform verfügen werden, die der unseren gänzlich fremd gegenübersteht. Das allerdings ist überhaupt nicht möglich. In Ermangelung wohlwollender Aliens, werden am Ende allein wir Menschen jede Form von neu aufkommender künstlicher Intelligenz geschaffen haben. Insofern wird sie, in welcher Weise auch immer, unser Menschsein und die unverkennbaren Charakteristika unserer Spezies spiegeln.

Deswegen denke ich, dass Leute, die sich wegen fremder Intelligenz und Singularität sorgen, eigentlich Angst vor all unseren unschönen Facetten haben, welche momentan noch unausgesprochen bleiben, mit den Möglichkeiten künstlicher Intelligenz aber offen zutage treten werden.

DOUGLAS COUPLAND, *Schriftsteller, Vancouver*

Gewohnte Ahnungslosigkeit

Ich bin gerade in meinem Landhäuschen. Sobald die Zentralheizung angeht, werde ich aufstehen und mir einen Tee und einen Haferbrei machen. Ich schalte den World Service an, um Nachrichten zu hören und mache ein paar Anrufe, weil hier was abgedichtet werden muss. Dann pflanze ich wahrscheinlich ein paar Osterglockenzwiebeln für den Frühling (auf der Packung heißt es, dass das jetzt sein sollte). Ich glaube, danach gehe ich in den Supermarkt und hole mir was fürs Mittag- und Abendessen. Vielleicht nehme ich dann den Bus nach Norwich und schaue mir ein neues Bett an. Ich habe keinen Breitband-Anschluss in meinem Landhäuschen, deswegen werde ich in Norwich auch meine E-Mails checken, einen Zug nach London reservieren und eine Stromrechnung bezahlen.

Was ich dabei alles nicht verstehe: Ich weiß nicht, wie das Öl, das meine Zentralheizung befeuert, von weit entfernten Ölfeldern in mein Haus kommt. Ich weiß nicht, wie der Hafer oder der Tee zu mir geschafft wurden. Ich weiß nicht, wie mein Telefon oder mein Digitalradio funktionieren oder wie die Nachrichten, die mich erreichen, gesammelt und redigiert wurden. Ich verstehe nichts von der Komplexität, den Bus- und Zugverkehr zu organisieren. Ich werde keine Ahnung haben, wie man eine Supermarktkette führt oder wie Betten hergestellt werden oder was genau passiert, wenn ich bei einer E-Mail „senden“ drücke, oder wie Geld überwiesen wird. Ich verstehe auch nicht, wie man ein Haus isoliert oder warum ich die Osterglocken nicht schon im Dezember pflanzen konnte.

Der Witz ist aber, dass mich meine bodenlose Unwissenheit über so ziemlich alles, was sich an diesem Morgen tue, nicht stören wird. Ich bin daran gewöhnt. Vor langer Zeit haben die Menschen aufgehört, multi-kompetente Individuen zu sein, die ihr eigenes Essen jagen, ihr eigenes Feuer anzünden und ihre eigenen Werkzeuge herstellen, damit sie Spezialisten werden konnten, Teil einer größeren Gemeinde von Menschen, die untereinander all das tun konnten, was zu tun ist. Ist diese enorme Struktur aus Kompetenzen und Möglichkeiten wirklich von „künstlicher Intelligenz“ zu unterscheiden? Die Sorte, die digitale Computer produzieren, ist eigentlich nur ein neues Fraktal in einem großen Zusammenhang. Letztendlich haben wir seit Tausenden Jahren sehr glücklich mit einer künstlichen Intelligenz gelebt.

BRIAN ENO, *Musiker, Produzent, London*

Unberechenbar

Warum lässt sich das Sein nicht berechnen? Oder das Glück? Glück ist nicht berechenbar, weil es als Zustand eines physischen Objekts außerhalb der Welt der Berechnungen liegt. Computer und Software erschaffen und beeinflussen Physisches nicht. (Zwar können sie andere, gleichsam angehängte Maschinen dazu bewegen, dies zu tun. Doch was diese Maschinen verrichten, ist nicht die Leistung von Computern. Roboter können fliegen, Computer nicht. Zudem ist überhaupt nicht garantiert, dass irgendein computergesteuerter Apparat Menschen glücklich machen kann; aber das ist eine andere Geschichte.) Sein ist nicht berechenbar: eine wichtige Tatsache, die bislang übersehen wurde – nicht von ungefähr. Computer und der menschliche Geist existieren – wie etwa Kürbisse und Puccini – in unterschiedlichen Universen und sind daher kaum miteinander vergleichbar.

Kann es ohne das Sein trotzdem eine denkende Maschine geben? Nein. Unser Geist wird durch das Nachdenken und das Sein (oder Fühlen) definiert. Am oberen Ende des Spektrums – bei maximaler Wachsam-

keit oder Konzentration – wird der Geist im Zustand des Nachdenkens alle Emotionen abwehren, die ihn ablenken könnten.

Am unteren Ende der Skala rangieren jene Zustände und Aktivitäten, bei denen wir besonders wenig denken, etwa der Schlaf und das Träumen. Während wir haluzinieren, sind wir mit Eindrückern, häufig auch Gefühlen beschäftigt (Träume können äußerst emotional sein), in jedem Fall mit dem Fühlen oder, anders gewendet, mit dem Sein.

Warum braucht es so viel Platz, um solch eine simple Argumentation zu entfallen? (Und warum werden wahrscheinlich nur wenige Denker bereit sein, sie zu akzeptieren?) Vielleicht, weil sich die meisten Philosophen und Wissenschaftler nach einem Geist sehnen, der mit Fühlen und Sein nichts zu tun hat, weil er ausschließlich Denkleistungen erbringt. Dieser Wunsch ist so stark, dass seine Anhänger ihn am Ende mit der Wahrheit verwechselten. Philosophen sind nur Menschen.

DAVID GELERTER, *Computerswissenschaftler, Yale University*