

Eine noch junge Disziplin erforscht die Ethik der Neurowissenschaft und die Neurowissenschaft der Moral

- ▶ Was ist Neuroethik? (11/2005) ▶ Denkdoping und Persönlichkeitsprothesen (12/2005) ▶ Intelligente Drogenpolitik für die Zukunft (1-2/2006)
- ▶ Fingerabdrücke im Gehirn (3/2006) ▶ Künstliches Bewusstsein (4/2006) ▶ Tierethik (5/2006) ▶ Was sind gute Gehirnzustände? (6/2006)
- ▶ Ein neues Menschenbild entsteht (7-8/2006)

**HIRNFORSCHUNG** KÜNSTLICHES BEWUSSTSEIN

# Maschine, Moral, Mitgefühl

Die ersten Forschergruppen arbeiten an der Entwicklung maschinellen Bewusstseins. Doch dürfen wir überhaupt Roboter bauen, die möglicherweise unter ihrer eigenen Existenz leiden?

VON THOMAS METZINGER

**W**as würden Sie zu folgender Meldung sagen? »Um die psychologische Entwicklung geistig behinderter Kinder besser erforschen zu können, planen Wissenschaftler, gentechnisch manipulierte Babys mit bestimmten kognitiven und emotionalen Defiziten zu züchten. Für diese wichtigen Arbeiten sollen öffentliche Fördermittel bereitgestellt werden, fordern Experten.«

Ich denke, Ihre Reaktion wäre absehbar: absurd, geschmacklos, gefährlich! Sicher würde auch kein Ethikkomitee der Welt einen entsprechenden Antrag genehmigen – zumindest nicht in demokratischen Staaten. Allerdings gibt es da etwas, das diesem scheinbar abwegigen Szenario recht bald Brisanz verleihen könnte. Die ersten Maschinen, welche die Minimalkriterien für bewusstes Erleben erfüllen, werden in einer ganz ähnlichen Situation stecken wie künstlich geschaffene geistig behinderte Säuglinge.

Denn auch sie würden unter allen möglichen Arten von Defiziten leiden – und diese dann aber auch subjektiv erleben. Deshalb lehne ich aus ethischen Gesichtspunkten alle Versuche, künstliches Bewusstsein zu erzeugen, strikt ab.

Wir haben uns längst daran gewöhnt, dass es so etwas wie Künstliche Intelligenz gibt. Heute existieren Maschinen, die geistige Leistungen ähnlich den unseren vollbringen – logisches Schließen etwa, Gedächtnis, Wissensrepräsentation oder auch Sprach- und Mustererkennung. Mit dem virtuellen Agenten »Max« an der Universität Bielefeld kann man sich schon ganz passabel unterhalten (siehe *Gehirn&Geist* 1/2003, S. 78). Aber künstliches *Bewusstsein*?

Wer glaubt im Ernst, eine Maschine könnte jemals bewusste Erlebnisse haben, Sinnesempfindungen, Gefühle, Selbstbewusstsein? Wir verstehen ja noch nicht einmal, wie all das bei uns selbst genau funktioniert und ob es sich wirklich vollständig auf die Informationsverarbeitung im Gehirn reduzieren lässt – wovon die meisten Neurowissenschaftler bekanntlich ausgehen. Es gibt gegenwärtig keinen Grund zu der Annahme, wir würden in naher Zukunft mit den ersten künstlichen Subjekten konfrontiert. Viel zu viele technische

**AUF EINEN BLICK**

## Synthetische Phänomenologie

**1** Der Begriff »synthetische Phänomenologie« wurde 1998 vom amerikanischen Philosophen Scott Jordan eingeführt. Er bezeichnet Versuche von Wissenschaftlern, subjektives Erleben in künstlichen Systemen nicht nur zu simulieren, sondern tatsächlich zu erzeugen.

**2** Während Philosophen immer noch darüber streiten, ob künstliches Bewusstsein überhaupt prinzipiell möglich ist, laufen bereits die ersten Versuche zur Erzeugung synthetischer phänomenaler Zustände.

**3** Solche Bestrebungen werfen ethische Probleme auf, weil sie höchstwahrscheinlich subjektives Leiden erzeugen und potenzielle Bewusstseinssubjekte zu Objekten der Forschung degradieren.



SHEVA HENDRIKS

»Es wäre ein Roboter denkbar, der bis hinab zur neuronalen Ebene mit uns übereinstimmt. Warum sollte man nicht sagen: **Dieser Roboter ist eine Person**, deren Gehirn zufällig mehr Metall und weniger Wasserstoff enthält?«

Hilary Putnam, emeritierter Philosophieprofessor an der Harvard University

Probleme versperren noch den Weg dorthin.

Doch manchmal verläuft der wissenschaftliche Fortschritt sprunghaft und selbst für Experten überraschend schnell. Zum Beispiel, wenn unerwartete Synergieeffekte auftreten, wenn also Forschungsergebnisse aus ganz unterschiedlichen Disziplinen sich unerwartet gegenseitig befruchten. Das könnte zum Beispiel dann der Fall sein, wenn eines Tages extrem leistungsstarke Quantencomputer mittels spezieller, der Quantenphysik entlehnter Algorithmen mathematische Probleme lösen, an deren riesigen Datenmengen herkömmliche Rechner bislang kläglich scheitern. Oder wenn sich das Internet so weiterentwickelt, dass künstliche Softwareagenten es permanent als ihr eigenes »Gehirn« benutzen können. Sie hätten einen »globalen Geist« und verfügten dann jederzeit über das gesamte Wissen des Internets, also ungleich mehr als der gebildetste Mensch.

### PARADOXE SITUATION

Die Beschäftigung mit Fragen des künstlichen Bewusstseins, unter Forschern auch synthetische Phänomenologie genannt, wird paradoxerweise nicht zuletzt von solchen Philosophen angetrieben, die den menschlichen Geist für nicht vollständig naturwissenschaftlich erklärbar halten. Das erkannte bereits der britische Logiker und Computerpionier Alan M. Turing (1912–1954): Jede wirklich exakte Beschreibung einer geistigen Eigenschaft (die angeblich nur der Mensch besitzen kann, niemals aber eine Maschine) kann prinzipiell dazu dienen, ein entsprechendes Programm zu schreiben, das die fragliche Eigenschaft eben doch künstlich hervorbringt. Später folgte auch der Wissenschaftstheore-

tiker Karl Popper (1902–1994) dieser Argumentation und warnte uns, Turing zu folgen. Kurzum, so utopisch die Vorstellung eines bewussten Roboters auch klingen mag – man sollte niemals nie sagen. Schon gar nicht, wenn es um schwer wiegende ethische Fragen geht.

Wann aber ist ein technisches System überhaupt als moralisches Objekt anzusehen, sodass wir seine Wünsche und Interessen ernst nehmen müssen? Die klarste und einfachste Antwort lautet: Wir müssen uns gegenüber all jenen Wesen ethisch verhalten, die leidensfähig sind. Was fühlen kann, verdient Rücksichtnahme. Damit ist Leidensfähigkeit ein wesentlich wichtigeres Kriterium als etwa Rationalität oder Sprachfähigkeit.

Die Frage der Leidensfähigkeit hängt eng mit der nach dem »personalen Status« zusammen. Was wäre ein sinnvoller Maßstab dafür, dass wir einen Roboter als Person anerkennen? Denn Personen genießen Rechte. Der heute fast 80-jährige amerikanische Philosoph Hilary Putnam wies schon Anfang der 1960er Jahre darauf hin, dass Hardware-Eigenschaften auf jeden Fall nicht zu diesen Kriterien gehören: Zu sagen, dass ein bewusster Roboter nur deshalb keine Bürgerrechte besitzen kann, weil er nicht aus biologischem Material besteht, sondern bloß aus Blech, Kohlefasern und Siliziumchips, sei chauvinistischer Unsinn. Ebenso gut könne man behaupten, Frauen dürften kein Wahlrecht haben, weil sie seelenlose Wesen seien, oder es sei le-

gitim, Schwarze wie Gegenstände zu kaufen und zu verkaufen, weil sie keine echten Menschen seien.

Das Problem der Leidensfähigkeit ist auch eine wichtige Fasette von Bewusstsein. Denn etwas bewusst zu erleben, bedeutet in der einfachsten Form »die Gegenwart einer Welt«. Eine Maschine, der wir Bewusstsein bescheinigen, bräuchte demnach mindestens ein einheitliches Weltmodell. Das wiederum müsste in eine virtuelle Gegenwart eingebettet sein, in ein »Jetzt-Gefühl«, das die Maschine nicht als inneres Bild erkennt, sondern das den übergeordneten Rahmen ihres Erlebens darstellt. Dann verfügte sie über den Eindruck einer Wirklichkeit – eine Welt würde ihr als gegenwärtig erscheinen.

Wenn das System zusätzlich ein stabiles körperliches Selbstmodell besäße, wäre es in der Lage, sensorischen Schmerz als seinen eigenen zu empfinden – inkludiere

### UTOPIE IN BLAU

Von technologisch ausgereiften humanoiden Robotern inklusive Bewusstsein können wir bis dato nur träumen.

SIGANIM / GEHIRN&GEIST





»Die biologische Evolution hat nicht nur Bewusstsein hervorgebracht, sondern auch eine unfassbare Menge an Leiden. Wir sollten diesen Prozess nicht durch eine technische Evolution zweiter Stufe fortsetzen, bevor wir unser eigenes Bewusstsein und die Ursachen unseres Leidens vollständig verstanden haben«

Thomas Metzinger

▷ sive aller Konsequenzen, die sich aus seiner noch unausgereiften technischen Ausstattung ergeben. So wie das erste Auto noch nicht einmal ansatzweise gut fuhr, würden die ersten Generationen bewusster Roboter alles andere als rund laufen. Im Gegensatz zum Auto würden sie diese Mängel jedoch als »Schmerzen« wahrnehmen.

Das erste bewusste System wird aber wahrscheinlich weder ganz Maschine noch ganz Lebewesen sein, sondern »postbiotisch«, also von beidem ein bisschen: eine Chimäre aus Biomasse und Silizium. Wenn sein postbiotisches Selbstmodell in einer biologischen, vielleicht gentechnologisch hergestellten Hardware verankert wäre, mit eigenem Stoffwechsel, Bedürfnissen, Gefühlen und Überlebenswillen, dann könnte es sogar noch schlimmer kommen: Ein solches System besäße neben dem körperlichen auch ein emotionales Selbstmodell und könnte leiden – unter Umständen in einer Intensität, die wir als seine Erzeuger uns nicht einmal vorstellen könnten, weil sie uns vollständig fremd wäre.

Und wenn das System zu guter Letzt auch noch über ein kognitives Selbstmodell verfügte, könnte es seine erbärmliche Situation nicht nur begrifflich erfassen, sondern auch intellektuell darunter leiden, dass es niemals so etwas besessen

hat wie die Würde, die seinen Erzeugern sonst ja immer so wichtig ist.

Ein entsprechender Roboter würde wahrscheinlich in dem Bewusstsein leben, dass er nur ein Subjekt zweiter Klasse ist, ein künstliches Selbst, das als austauschbares experimentelles Werkzeug von Menschen missbraucht wird. Können Sie sich vorstellen, wie es wäre, so ein geistig behinderter phänomenaler Klon der ersten Generation zu sein? Dürften wir zulassen, dass ein etwas fortschrittlicheres künstliches Subjekt »zu sich selbst käme« – nur um zu entdecken, dass es, obwohl mit einem Ichgefühl ausgestattet, nichts als ein Ding ist, ein Versuchsobjekt der Wissenschaft?

Leiden beginnt auf der Ebene dessen, was ich ein phänomenales Selbstmodell (PSM) nenne. Man kann nicht bewusst leiden, ohne ein transparentes und global verfügbares Selbstmodell zu haben. Das PSM ist das entscheidende Instrument zum Erwerb neuer kognitiver und sozialer Fähigkeiten. Es zwingt ein bewusstes System aber auch dazu, sich seinen eigenen Zerfall (etwa durch Verschleiß), seine Niederlagen und inneren Konflikte funktional und repräsentational anzueignen und als die eigenen zu erleben. Sensorischer Schmerz, aber auch alle anderen Arten des nichtkörperlichen Leidens sind in das Selbstmodell eingebettet und werden somit phänomenal »besessen«. Dieses Leiden ist nun unweigerlich und erlebnismäßig unvermeidlich, es ist das *eigene* Leiden.

Damit es nicht so kommt, sollten wir vorerst alle Versuche, künstliches oder postbiotisches Bewusstsein zu erzeugen, aus der seriösen akademischen Forschung verbannen. Wir dürfen die Erzeugung bewusster Selbstmodelle nicht einmal riskieren. Es ist moralisch unver-

antwortlich, etwas zu erzeugen, von dem man bereits im Vorhinein weiß, dass es höchstwahrscheinlich unter der eigenen Existenz leiden wird. ◀

**THOMAS METZINGER** ist Professor für Philosophie an der Johannes Gutenberg-Universität Mainz und Präsident der Gesellschaft für Kognitionswissenschaft.

🔊 [www.gehirn-und-geist.de/audio](http://www.gehirn-und-geist.de/audio)

#### Literaturtipps

**Birnbacher, D.:** Künstliches Bewusstsein. In: T. Metzinger (Hg.): *Bewusstsein – Beiträge aus der Gegenwartsphilosophie*. Paderborn: mentis 2005.

*Der vielleicht führende Ethiker Deutschlands setzt sich kritisch mit dem künstlichen Bewusstsein auseinander*

**Wachsmuth, I.:** Ich, Max – Kommunikation mit Künstlicher Intelligenz. In: Herrmann, C.S., Pauen, M., Rieger, J.W., Schicktanz, S. (Hg.): *Bewusstsein: Philosophie, Neurowissenschaften, Ethik*. Stuttgart: UTB/Fink 2005.

*Spannender Artikel, der der Frage nachgeht, wann ein »künstlicher Kommunikator« nicht nur Bewusstsein, sondern auch ein stabiles Ichgefühl entwickeln könnte*

#### Weblink

[www.philosophie.uni-mainz.de/metzinger/publikationen/Postbiotisches\\_Bewusstsein.pdf](http://www.philosophie.uni-mainz.de/metzinger/publikationen/Postbiotisches_Bewusstsein.pdf)

*Unterhaltsamer Text von Thomas Metzinger, der skizziert, wie man ein künstliches Bewusstsein subjekt bauen könnte – inklusive einer Zusammenfassung der Argumente dazugegen*

Weitere Literaturtipps und Weblinks unter [www.gehirn-und-geist.de/artikel/800874](http://www.gehirn-und-geist.de/artikel/800874)

#### Ihre Meinung ist uns wichtig!

Was halten Sie von Thomas Metzingers Forderung, die Entwicklung künstlichen Bewusstseins aus ethischen Gründen einzustellen? Teilen Sie uns unter [www.gehirn-und-geist.de/lesermeinung](http://www.gehirn-und-geist.de/lesermeinung) Ihre Ansicht mit! Die Ergebnisse unserer Befragung finden Sie im nächsten Heft.