

III Die Position der harten KI

Die Debatte um die künstliche Intelligenz hat weniger die real existierende KI als deren anthropologische Implikationen zum Gegenstand. Die Annahme, derzufolge es möglich ist, intelligente Maschinen zu bauen, beinhaltet zwar eine optimistische Auffassung von der Leistungsfähigkeit einer Technologie, mehr noch aber drückt sie eine spezifische Vorstellung von der Operationsweise des menschlichen neuronalen Systems aus: Die harte KI-These behauptet, daß das menschliche Gehirn Informationen verarbeitet, und daß prinzipiell informationsverarbeitende Maschinen konstruierbar sind, die die gleichen Operationen durchführen können. Anders formuliert: Die harte KI behauptet die Invarianz der informationsverarbeitenden Hardware. Die harte KI behauptet nicht, daß real existierende Maschinen über Intelligenz auf einem Niveau verfügen, das mit den kognitiven Leistungen von Menschen verglichen werden könnte.

Die Debatte um das starke KI-Postulat ist teilweise polemisch geführt worden. Sie erkennt in der real existierenden KI deren inhärente Annahme der harten KI-These und benutzt fehlgeschlagene oder mangelhaft funktionierende KI-Systeme als Argument für deren Falsifikation. Das bon-mot der KI-Debatte "if it works, it's not AI" charakterisiert den common sense der Diskussion. Die explizite Behauptung der harten KI-These setzt selbst renommierte Wissenschaftler in die Nähe zur Unseriosität, mit der Folge, daß die These nur selten explizit vertreten wird¹. Allzuleicht wird die Position der harten KI mit science fiction gleichgesetzt. Dennoch ist eine Grenzziehung möglich: Die literarische science fiction betrachtet Szenarien, die sich als Folge der Entwicklung intelligenter Maschinen ergeben. Die harte KI, dort wo sie in den Selbstbeschreibungen der künstlichen Intelligenz vorkommt, untersucht die Bedingungen der Möglichkeit zur Entwicklung informationsverarbeitender Hardware, die hinsichtlich der Rechengeschwindigkeit und Speicherkapazität dem menschlichen neuronalen System ein funktionales Äquivalent gegenüberstellt.

Die soziologische Rezeption der künstlichen Intelligenz kann die harte KI-These aus mehreren Gründen nicht ignorieren: 1. Aus methodologischer Sicht ist ihre Rezeption schlicht eine Frage der Vollständigkeit. 2. Das Postulat der harten KI birgt wichtige kommunikationstheoretische und wissenssoziologische Implikationen². Gelänge die Konstruktion von intelligenten, bewußtseinsfähigen Maschinen, wäre das soziale System um neue, quasipsychische Systeme ergänzt. 3. Die Rezeption der harten KI-These ist geeignet zu zeigen, daß die systemtheoretische Soziologie über ein theoretisches Arsenal verfügt, das die Implikationen der harten KI-These fassen kann. Das gilt insbesondere für das soziologische Konstrukt eines über Kommunikation generierten Begriffes von Technik. Folgt man den Ausführungen von Moravec oder Tipler, wird evident, daß hingegen eine anthropologische Konzipie-

¹ Eine Besprechung von Moravecs Buch "mind children" in der "ZEIT" beschreibt Moravec als einen der bestgehaßten Autoren der jüngeren Wissenschaftsgeschichte. "DIE ZEIT", 28.06.1996, S. 62.

² Vgl. Rammert 1995, S. 8 ff.

rung von Technik als Stoffwechsel mit der Natur oder als Entlastung von Organen wenig geeignet ist, die Rekonstruktion eines Bewußtseins angemessen zu beschreiben. (Die Konzipierung als Stoffwechsel mit der Natur wäre obsolet, weil die harte KI-These eine Technik anvisiert, für die "Stoffumwandlung" allenfalls einen marginalen Begleiteffekt darstellt. Technik als Entlastung könnte dann allenfalls paradox als Entlastung von sich selbst gefaßt werden.)

Obwohl bei der Rekonstruktion der harten KI-These auf deren interessante Varianten in der science fiction Literatur referiert werden könnte, beschränkt sich die Darstellung auf die Interpretation der These durch die Physik (Tipler) und die Robotik als Disziplin der künstlichen Intelligenz (Moravec). Die Darstellung Tiplers ist unter dem Gesichtspunkt interessant, daß sie aufgrund ihrer zeitlichen Aktualität die bekannte KI-Kritik (Searle, Dreyfus etc.) antizipieren konnte und um eine offensive Auseinandersetzung bemüht ist. Tipler diskutiert die informationstheoretischen Rahmenbedingungen, während Moravec aus der Innenperspektive der Robotik die technischen Extrapolationen thematisiert.

3.1 Menschen und Maschinen als endliche Systeme

Tipler vertritt dezidiert die harte KI-These in ihrer starken Variante. Danach ist es nicht nur prinzipiell möglich, eine intelligente Maschine zu konstruieren, sondern deren Realisation liegt im Bereich des technisch Machbaren: "Das Ziel des Programms für starke KI ist es, ein Computerprogramm zu entwickeln, das sich aus eigener Kraft Intelligenz aneignen kann, wenn es auf einer Computerhardware abläuft, die ihm mit der Welt zu kommunizieren ermöglicht... Ich behaupte, es wäre kurzsichtig und ein Auswuchs von Angst und Unsicherheit, nicht aber eine Folge zweckmäßiger Überlegung, wenn Männer und Frauen, die einen intelligenten Roboter herzustellen in der Lage sind, daran gehindert werden, es zu tun. Wir sind selber intelligente Maschinen."³ Zu beachten ist, daß Tipler nicht eine Maschine fokussiert, die intelligentes Verhalten simuliert, sondern die im ontologischen Sinne intelligent ist, die also über Persönlichkeit, Bewußtsein und Gewissen verfügt. Folgt man Tipler, lernt eine solche Maschine qua Kommunikation und ist bereits deshalb ein interessantes Thema einer Soziologie, die Technik qua Kommunikation definiert. Tipler geht es explizit nicht um die Konstruktion einer autopoietischen Maschine, die abgekoppelt von sozialen Kontexten existiert und so, streng genommen, mit einem über Kommunikation generierten Technikbegriff nicht mehr zu fassen wäre.

Als Minimalkonsens für seine Argumentation setzt Tipler Einverständnis über die Geltung physikalischer Naturgesetze in konventioneller Interpretation voraus. Als Kriterium für das Vorhandensein von Intelligenz schließt sich Tipler den bekannten, von Turing formulierten und weithin akzeptierten Definitionen an, wonach eine Maschine dann als intelligent zu gelten hat, wenn ein menschliches Gegen-

³ Tipler 1995, S. 72.

über nicht unterscheiden kann, ob es mit einer Maschine oder mit einem Menschen kommuniziert. Wie aber kann entschieden werden, ob es prinzipiell möglich ist, eine Maschine zu bauen, die den Turingtest versteht? Zunächst eine Vorüberlegung: Wenn ein System A alle Systemzustände des Systems B in sich abbilden kann, kann System A alle Operationen des Systems B ebenfalls ausführen. Dazu muß System A im Verhältnis zu B über einen Komplexitätsüberschuß verfügen. Übertragen auf den Vergleich zwischen Computer und psychischen Systemen bedeutet das, daß ein Computer die Leistungen eines psychischen Systems erbringen kann, wenn er mindestens über die gleiche Komplexität verfügt. Tipler setzt im folgenden mathematisch die Komplexität eines psychischen Systems in Relation zu der erforderlichen Mindestkomplexität eines funktional äquivalenten Computers. Die Speicherleistung eines Gehirns veranschlagt er:

"Die Speicherfähigkeit eines menschlichen Gehirns berechnet sich folgendermaßen: Das Gehirn hat ungefähr 10^{10} Neuronen, von denen jedes über 10^5 Verbindungen zu anderen Neuronen verfügt. Angenommen, jedes *Neuron* codiert ein Bit, dann ergibt das 10^{10} Bits. Angenommen, jede *Verbindung* codiert ein Bit, dann erhält man 10^{15} Bits.." Nach ausführlichen Referenzen auf die Neurobiologie veranschlagt Tipler die tatsächliche Speicherleistung des Gehirns zwischen 10^{13} und 10^{17} Bits. Um die Komplexität der neuronalen Informationsverarbeitung zu bestimmen, ergänzt Tipler die Speicherleistung mit der Informationsverarbeitungsgeschwindigkeit, angegeben in Flops⁴. Unter Bezugnahme auf neurobiologische Forschungen schätzt Tipler die Informationsverarbeitungsgeschwindigkeit des Gehirns auf etwa 10 Teraflops. Die Gesamtkomplexität ergibt sich aus den beiden Werten: "Wir wollen von 10^{15} Bits und zehn Teraflops als den derzeit bestmöglichen Schätzungen für die Speicherkapazität bzw. die Geschwindigkeit der Informationsverarbeitung im menschlichen Gehirn ausgehen."⁵ Die Informationsverarbeitungskapazität von Computern ist präzise bestimmbar. Die Firma Thinking Machines lieferte im Januar 1992 eine 100 Giga-Flop Computer an die Forschungslabors in Los Alamos aus und bietet derzeit einen 2 Teraflop Computer für 200 Mio. Dollar an (ebd.). Speicherkapazitäten in der Größenordnung von 10^{15} Bits sind bereits realisiert. Auf der Grundlage von linearen Extrapolationen der Rechenleistung von Supercomputern (Vertausendfachung der Rechenleistung in jeweils zwanzig Jahren) gelangt Tipler zu dem Schluß. "Gegen Ende des Jahrzehnts müßte es demnach Computer geben, die Informationen genauso schnell verarbeiten können, wie das menschliche Gehirn."⁶ Da die Leistung von Personalcomputern im Vergleich zu Großrechnern relativ langsam wächst, "können wir damit rechnen, daß wir im Jahre 2030 PCs mit einer dem menschlichen Gehirn entsprechenden Verarbeitungskapazität zu einem dem heutigen vergleichbaren Preis kaufen können."⁷ Tipler rechnet vor, daß ein Irrtum um den Faktor 100 bei den Ausgangsgrößen Informationsverarbeitungskapazität

⁴ Flops (Floating point operations per second) als Maß für die Rechengeschwindigkeit von Computern bezeichnen Gleitkommaoperationen pro Sekunde.

⁵ Tipler 1995, S. 49.

⁶ Tipler 1995, S. 49.

⁷ Ebd.

des menschlichen Gehirns und Zunahme der Rechenleistung von Computern die Angleichung von elektronischer und biologischer Informationsverarbeitungskapazität nur um sieben Jahre verzögern bzw. beschleunigen würde.

Die Reduktion aller intelligenten Phänomene auf Informationsverarbeitung ermöglicht es Tipler überzeugend, menschliche und maschinelle Informationsverarbeitungskapazitäten aufeinander zu beziehen. Maschinen und Menschen (inklusive ihrer Körper) erscheinen als endliche, miteinander vergleichbare Größen. Seitens der KI-Kritik immer wieder vorgetragene Einwände, Menschen und Maschinen seien wegen des Verhaftetseins der menschlichen Intelligenz in einem biologischen Körper nicht miteinander vergleichbar, treffen nur für die weiter oben skizzierte klassische KI zu, deren Konzeptualisierungen augenfällig von biologischen Systemen differieren. Computer, wie sie von Tipler prognostiziert werden, bewegen sich in Größenordnungen, die jede Facette biologischer Informationsverarbeitung nachbilden können. Penrose⁸ hat das biologistische Unvergleichbarkeitsargument quantentheoretisch reformuliert. Danach können für Phänomene wie Intentionalität nicht prognostizierbare Quanteneffekte ausschlaggebend sein, die den Menschen als biologisches System mit quasi-unendlich vielen Systemzuständen kennzeichnen. Computer wären demzufolge endliche, Menschen unendliche Maschinen. Nach einer ausführlichen, sich in der Hauptsache auf quantenphysikalische Berechnungen stützenden Argumentation, gelangt Tipler zu dem Fazit: "Mein Haupteinwand gegen Penroses Anti-KI-Behauptung ist jedoch die Bekenstein-Grenze, laut der es eine Obergrenze für die Anzahl distinkter Quantenzustände in einem Gebiet begrenzter Größe und mit begrenzter Energie gibt sowie eine Obergrenze für die Geschwindigkeit, mit der diese Zustandsänderung stattfinden kann.... Der Physiker Jakob Bekenstein hat gezeigt, daß Quantensysteme - und von der Physik her ist alles Sichtbare ein Quantensystem - nur eine endliche Anzahl von Zuständen haben. Im besonderen kann ein Mensch sich in einem von $10^{10^{45}}$ Zuständen befinden und höchstens 4×10^{53} Zustandsänderungen pro Sekunde durchmachen... Dennoch sind diese Grenzen endlich und beruhen auf den grundlegenden Gesetzen der Quantenmechanik. Sie beweisen also, daß der Mensch eine Maschine mit endlich vielen Zuständen ist: Nicht mehr und nicht weniger."⁹

Tiplers Argumentation für das starke KI-Postulat ist in eine großangelegte Kosmologie eingebettet, in der es in erster Linie um den Nachweis geht, daß das Universum durch die Transformation von Materie und Energie notwendig in einen Zustand wachsender Informationsdichte übergeht. Die Thematisierung der künstlichen Intelligenz steht im Kontext der Diskussion makrokosmischer Prozesse, in der Tipler zeigt, daß Information in kosmischen Maßstäben nicht verloren gehen kann. In terrestrischen Maßstäben gilt analog: Die Gesamtmenge der (künstlich generierten) Information nimmt zu und wird

⁸ Penrose 1991, S. 395 ff.

⁹ Ebd. S. 57.

sich notwendig in Größenordnungen bewegen, die die gesamte Informationsmenge beispielsweise eines psychischen Systems in Apparaten einschließen kann.

3.2 Das starke KI Postulat und die Perspektive der Robotik

Moravec bezieht hinsichtlich der Leistungsfähigkeit künftiger Computergenerationen einen ähnlichen Standpunkt, stellt aber in seiner Argumentation sozioökonomische Aspekte in den Vordergrund. Er beschreibt ausführlich das Wachstum der Rechenleistung seit den ersten mechanischen Rechenmaschinen des neunzehnten Jahrhunderts und stellt diese Rechenleistung in Relation zu den Kosten. Die Implementierung einer Rechenmaschine in das ökonomische System findet nur statt, wenn eine spezifische Bits per Dollar per second-Relation besteht. Akzeptiert man die Moravecschen Berechnungen, betragen die Kosten für einen Personalcomputer mit dem Leistungsäquivalent eines menschlichen Gehirnes im Jahr 2030 etwa Eintausend Dollar. Im Anschluß erörtert Moravec Szenarien, die einige soziale Folgen der Entwicklung intelligenter Roboter beschreiben.

Wie gelangt Moravec zu einer angemessenen Indexfindung für die Rechenleistung von Maschinen? Das heute übliche Verfahren besteht darin, ein Standardprogramm auf verschiedenen Computern zu installieren und die Geschwindigkeit zu messen, mit der es die Programmfunktionen ausführt. Das Verfahren hat den Nachteil, daß es nur für die Computer einer Generation anwendbar ist. Je entfernter die Computergenerationen, desto inkompatibler die Systeme. Ein modernes Textverarbeitungsprogramm läuft nicht auf der analytischen Maschine von Charles Babbage. Ein gravierenderes Problem besteht in der Differenzierung von Bits und Nutzinformation. Tipler beispielsweise ignoriert diese Differenz, indem er Bits als kleinste physikalische Einheit der Information faßt und durch bloße Addition der Speicher- und Verarbeitungskapazitäten den Zeitpunkt von Mensch- und Maschinenäquivalenz berechnet. Moravec nimmt in seiner Beschreibung des Performanzzuwachses von Computern explizit bezug auf Shannons Informationstheorie: "Ein Grundgedanke dieser Theorie lautet, daß der Informationsgehalt einer Nachricht in dem Maße zunimmt, in dem ihre Wahrscheinlichkeit, vom Rezipienten aus gesehen, abnimmt... Eine Reihe von Nachrichten haben einen maximalen Informationsgehalt, wenn sie für maximale *Überraschung* sorgen. Mein Index für die Rechenleistung beruht auf demselben Prinzip. Jeder Befehl, der von einer Maschine ausgeführt wird, ist wie eine Nachricht. Je vorhersagbarer die Befehlssequenz, desto kleiner der Nutzen."¹⁰ Dazu ein Beispiel: Ein Programm, das lediglich zu einer beliebigen Zahl jeweils 1 addiert und diese Operation, selbst mit astronomischer Rechenleistung, fortsetzt, leistet praktisch nichts. Ein Programm, daß auf die Frage, wie spät es ist, eine Uhr einblendet, generiert in Moravecs Indizierung der Rechenleistung bereits mehr Informationen, und ein Programm, das auf die gleiche Frage antworten würde: "10.30h, Sie schaffen ihren Zah-

¹⁰ Moravec 1990, S. 92.

nartztermin nur, wenn sie ihre Arbeit jetzt unterbrechen und die Straßenbahnlinie zwei benutzen", stellte abermals einen Informationszuwachs dar. Der Zuwachs an Rechenleistung ist, im Gegensatz zu Tiplers Modellrechnung, nicht mit dem Zuwachs an Information gleichzusetzen. Oder anders: Der Zuwachs an Information ist stets geringer als der Zuwachs an Rechenleistung, denn: "jeder Befehl kann nur eine begrenzte Zahl verschiedener möglicher Operationen angeben und aus einer begrenzten Zahl von Speicherplätzen wählen, die ihrerseits nur eine begrenzte Zahl von Möglichkeiten enthalten. Diese Überraschungsmöglichkeiten lassen sich durch die Formeln der Informationstheorie dergestalt zusammenfassen, daß sie den maximalen Informationsgehalt eines einzigen Computerbefehls ausdrücken." Nach umfangreichen Modellrechnungen kommt Moravec zu dem Schluß, "daß ein typischer Computer, der ein außergewöhnlich effizientes Programm ausführt, pro Operation etwa 50 *Überraschungs-Bits* produziert. Wenn der Computer eine Million Operationen pro Sekunde ausführen kann, liegt seine maximale Rechenleistung bei ungefähr 50 Millionen Bits pro Sekunde. In diesen Einheiten ausgedrückt, brauchte ein dem Menschen gleichwertiger Roboter eine Rechenleistung von ungefähr 10^{14} Bits pro Sekunde."

Unter Zugrundelegung des Shannonschen Informationsbegriff berechnet Moravec eine Vertausendfachung der Rechenleistung alle zwanzig Jahre, seit Beginn dieses Jahrhunderts. Noch drastischer präsentieren sich die Fortschritte der Rechenleistung, wenn man diese in Relation zu den dafür aufzuwendenden Kosten setzt: "Innerhalb von 80 Jahren waren die Rechenkosten um das *Billionenfache* zurückgegangen. Wenn diese Entwicklung anhält, stehen die zehn Teraflop, die für einen menschenähnlichen Computer erforderlich sind, noch vor dem Jahre 2010 in einem 10-Millionen-Dollar-Superrechner und im Jahre 2030 in einem Personalcomputer für 1000 Dollar zur Verfügung."¹¹

Folgt man Moravec, ist eine weitergehende Beschleunigung des Zuwachses an Rechenleistung wahrscheinlicher als eine Verzögerung. Die Gründe dafür liegen in technologischen Durchbrüchen bei der Grundlagenforschung (Nanotechnologie, optische Schaltelemente) und in der weiter steigenden Bereitschaft des ökonomischen Systems, Mittel in die Computerforschung zu investieren. Zu beantworten bleibt die Frage nach der Plausibilität der skizzierten Entwicklung. Warum ist die Konstruktion intelligenter Roboter wahrscheinlich? Moravec antwortet darauf mit einem rigiden Sozialdarwinismus: "Gesellschaften und Wirtschaften sind zweifellos denselben evolutionären Zwängen unterworfen wie biologische Organismen. Früher oder später werden sich diejenigen durchsetzen, die sich am raschesten ausweiten und diversifizieren können. Kulturen stehen untereinander im Wettbewerb um die Ressourcen des ihnen zugänglichen Universums. Wenn Automation effizienter ist als Handarbeit, werden die Organisationen und Gesellschaften, die sich die Automation zu eigen machen, reicher sein und

¹¹ Moravec 1990, S. 99.

besser dazu in der Lage, in schwierigen Zeiten zu überleben und sich unter günstigeren Verhältnissen auszubreiten."¹²

Daß die Entwicklung immer leistungsfähigerer Computer in der Logik einer profitorientierten Produktionsweise liegt, ist ebenso plausibel wie die Tatsache, daß die Ausbreitung von automatisierten Systemen in der Produktion sich beschleunigt. Gründe dafür liegen in der verbilligten Rechenleistung bei gleichzeitiger Verteuerung der menschlichen Arbeitskraft und die Ausweitung der Einsatzmöglichkeiten von Computern. Moravec postuliert das zwangsweise Erreichen eines Punktes, an dem die Leistungsfähigkeit von Maschinen der menschlichen Leistungsfähigkeit äquivalent wird oder sie übertrifft: "Es gibt überhaupt keinen Grund zu der Annahme, daß das Äquivalent menschlicher Leistungsfähigkeit eine Obergrenze für diese Entwicklung darstellt."¹³

Unter intelligenten Maschinen versteht Moravec mobile, lernende, sich selbst reproduzierende Roboter, die für die Menschheit als Gattung eine evolutionäre Bedrohung darstellen: "Doch intelligente Maschinen, mögen sie auch noch so gutartig sein, bedrohen unsere Existenz, weil sie Mitbewohner unserer ökologischen Nische sind. Die Maschinen brauchen nur genauso intelligent wie Menschen zu sein, um in Konkurrenzsituationen außerordentlich erfolgreich zu sein."¹⁴

Noch brisanter ist Moravecs zweite Schlußfolgerung. Maschinen mit einer Rechenleistung in der Größenordnung von 10^{15} Bits sind in der Lage, psychische Systeme komplett zu implementieren. Damit wären sie beliebig reproduzierbar und ihrer biologischen Trägersubstanz entkoppelt. Evolution, soziale wie intrapsychische, würde sich in Computern fortsetzen.

Moravec skizziert umfangreiche Szenarien, die die Übertragung des Bewußtseins eines psychischen Systems in einen Computer betreffen und spekuliert über den Fortgang der Evolution in Computersystemen. Zu seinen Resultaten gelangt er auf der Grundlage eines radikal formulierten Körper-Geist-Dualismus, der die Beziehung zwischen Leib und Seele allenfalls als schwache (und substituierbare) strukturelle Kopplung gelten läßt. Er differenziert zwischen der Position der Körperidentität, die eine kausale Verbindung zwischen sensomotorischen und kognitiven Fähigkeiten postuliert, und seiner Position der Strukturidentität, die eben dies bestreitet: "Die Körper-Identität setzt voraus, daß ein Mensch durch die Substanzen definiert ist, aus denen sein Körper besteht. Nur wenn wir die Kontinuität dieser Körpersubstanzen bewahren, vermögen wir unseren Fortbestand als Individuen zu sichern. Dagegen definiert Strukturidentität das Wesen einer Person...durch Struktur und Prozeß, die in meinem Körper vorkommen, aber nicht durch das Substrat, in dem sich dieser Prozeß manifestiert. Bleibt

¹² Ebd. S. 141.

¹³ Moravec 1990, S. 141.

¹⁴ Ebd.

der Prozeß erhalten, so bleibe auch ich erhalten; der Rest ist Sülze."¹⁵ Die Differenz zwischen Strukturidentität und Körperidentität ist leicht überführbar in das seit Descartes häufig formulierte Leib-Seele-Problem, das bis heute den Kern der philosophischen, fast durchgängig als Kritik vorgetragenen, KI-Rezeption markiert. Im folgenden wird zu zeigen sein, daß das Leib-Seele-Problem in der philosophischen KI-Rezeption als Differenz zwischen einem substantiellen und einem pragmatischen Informationsbegriff reformuliert werden kann. Anhand der Klärung des Verhältnisses unterschiedlicher Konzeptualisierungen von Information kann der Informationsbegriff in eine systemtheoretische Kommunikationstheorie integriert werden, die ihrerseits Information unter den Bedingungen der Technisierung untersucht.

¹⁵ Ebd. S. 163.