

# Aristoteles, Descartes und die Beziehungen zwischen Philosophischer Psychologie und Künstlicher-Intelligenz-Forschung

*Ansgar Beckermann*

Wenn man über die Seele, den Bereich des Psychischen nachdenkt, dann ist es ganz sinnvoll, sich zunächst einmal klar zu machen, daß der eng auf die Phänomene des Denkens und des Bewußtseins bezogene Begriff der Seele, der heute in der philosophischen Diskussion ebenso vorherrscht wie im Weltbild des Common Sense, kaum älter als 300 Jahre ist. Denn dieser Begriff des Seelischen entsteht erst zu Beginn der frühen Neuzeit<sup>1</sup>. In der vorneuzeitlichen und besonders in der antiken Philosophie markierte der Begriff „Seele“, genauer gesagt der Begriff „psyche“, noch einen ganz anderen Unterschied, nämlich den Unterschied zwischen Lebendem und Totem. „apsychos“ heißt in erster Linie „leblo“, „empsychos“ dagegen „belebt“, „lebendig“ und manchmal sogar „lebhaft“. „psyche“, die Seele, ist in der antiken Theorie also in erster Linie das, was das Lebende vom Toten unterscheidet, und Tod dementsprechend der Vorgang, bei dem die Seele den Körper verläßt, bzw. der Zustand, der eintritt, wenn die Seele den Körper verlassen hat. Dieser Zusammenhang ist für die gesamte Antike völlig selbstverständlich. Und so kann auch Platon Sokrates im *Phaidon* ohne weiteres sagen lassen:

„Was, glauben wir, ist nun der Tod? Er ist doch etwas? – Gewiß. – Und er ist doch nichts anderes als die Trennung der Seele vom Körper? Und Totessein heißt: der Körper ist von der Seele getrennt, für sich allein, und die Seele ist, für sich allein, vom Körper getrennt? Tod kann nichts anderes sein als das, oder?“ (Platon, *Phaidon*, 64c2-8)

Besonders prägnant faßt Aristoteles diese antike Auffassung zusammen, wenn er gleich zu Beginn seiner Abhandlung über die Seele formuliert: „... die Seele ist gewissermaßen das erste Prinzip der Lebewesen“ (*De Anima* 402a6-7).

Obwohl uns diese antike Auffassung der Seele heute eher fern liegt, ist sie bei einiger Überlegung doch durchaus nachvollziehbar. Denn der Unterschied zwi-

1) Vgl. zum folgenden auch die parallelen Überlegungen in Beckermann (1986, Kap. 1).

schen dem Lebenden und dem Toten scheint zumindest auf den ersten Blick viel größer zu sein als der zwischen dem mit Geist Begabten und dem „Geistlosen“. Innerhalb des Lebendigen gibt es zwar große Unterschiede von den Pflanzen bis zu den Menschen. Aber diese Unterschiede scheinen graduell zu sein; zumindest gibt es alle möglichen Zwischenstufen. Der Unterschied zwischen Leben und Tod dagegen erscheint viel drastischer und prinzipieller. Was könnte frappierender sein als der Unterschied zwischen einem Lebenden und einem Toten? Ein Mensch, der gerade noch lebte, gehen, sprechen, lachen und weinen und viele andere Dinge tun konnte, liegt auf einmal da, reglos, bewegungslos, all seiner Fähigkeiten beraubt. Äußerlich (und auch im Körper selbst) ist möglicherweise gar nicht viel geschehen. Also muß der Grund für den Unterschied in etwas anderem liegen. Da sich beim Tod der Körper nicht entscheidend verändert, muß es etwas geben, das den Körper beim Sterben verläßt; etwas, was für die verschiedenen Fähigkeiten und Aktivitäten des Lebenden verantwortlich ist und was den Körper dieser Fähigkeiten und Aktivitäten beraubt, wenn es aus ihm weicht. Dieses Etwas ist für die Antike die Seele. Die „psyche“ ist das Prinzip des Lebens und daher auch in erster Linie durch die Eigenschaften gekennzeichnet, die das Lebende vom Toten unterscheiden: Bewegung, Ernährung und Wachstum, Fühlen, Wahrnehmen, Denken.

Die Annahme, daß die Seele das *Etwas* ist, das dem Toten Leben verleihen kann, ist tatsächlich Gemeingut fast aller antiken Philosophen, so sehr sie sich auch in ihren Annahmen über die Natur dieses Etwas unterscheiden mögen. Für die materialistischen Philosophen wie Demokrit, Epikur oder Lukrez ist die Seele ein sehr feiner materieller Stoff, der einerseits über den ganzen Körper verteilt ist, andererseits aber seinen Hauptsitz in der Mitte der Brust hat. Für Platon ist die Seele das eigentliche Selbst eines Menschen, eine nicht materielle Person, die während des Lebens an einen Körper gefesselt ist, nach dem Tode aber frei von allen Behinderungen durch den Körper weiter existiert. Aristoteles schließlich formuliert seine Theorie der Seele mit Hilfe des für seine Philosophie zentralen Begriffspaares Form und Materie. Für ihn ist die Seele die Form des Lebewesens, dessen Körper die zugehörige Materie darstellt. Bei Aristoteles ist damit am wenigsten klar, daß er die Seele für ein Etwas hält. Denn Form könnte vielleicht auch als der Inbegriff all der Eigenschaften verstanden werden, die einen Gegenstand zu einem Gegenstand einer bestimmten Art machen, und Seele entsprechend als der Inbegriff all der Eigenschaften, die ein Lebewesen zu einem bestimmten Lebewesen machen. Aber es gibt sehr viele Stellen, an denen Aristoteles eindeutig klar macht, daß seiner Meinung nach kein wie auch immer gearteter Komplex von Materie *von sich aus* die für ein Lebewesen charakteristischen Fähigkeiten hervorbringen kann. Und deshalb ist für ihn Seele nicht nur der Inbegriff der für ein Lebewesen charakteristischen Fähigkeiten, sondern zugleich auch der *Grund* für diese Fähigkeiten, das, was zur Materie des Körpers hinzu-

kommen muß, damit sich an ihm diese Fähigkeiten zeigen. Für Aristoteles ist die Annahme, daß Materie aus sich heraus diese Fähigkeiten entwickeln könnte, genau so absurd wie die Annahme, daß von selbst ein Haus entsteht, wenn man Steine, Balken und Mörtel an einer Stelle einfach aufeinanderhäuft.

Diese aristotelische Auffassung der Seele war nicht nur für die Philosophie, sondern auch für die antike und mittelalterliche Medizin und Physiologie von großer Bedeutung. Denn von Galen bis Fernel werden die vitalen Prozesse in einem lebenden Körper durch spezielle Fähigkeiten und Kräfte in Gang gehalten, die nicht dem Körper, sondern der Seele zugeschrieben werden. Bei Fernel wird dies besonders deutlich, wenn er schreibt: „Die Leistungen des Körpers gehen nicht von sich selbst noch vom Körper aus“, und „Die Ursache für die Verrichtungen des Körpers ist die Seele“<sup>2</sup>. Wirksam wird die Seele also durch ihre Kräfte und Fähigkeiten, besonders durch ihre drei Hauptvermögen: die *facultas animae naturalis*, die *facultas animae vitalis* und die *facultas animae animalis*. Diesen Hauptvermögen stehen weitere helfende Vermögen zur Seite, der *facultas animae naturalis* z. B. die *facultas attrahens*, die *facultas expellens*, die *facultas continens* und die *facultas concoquens*. Diese Vermögen sind für Ernährung, Wachstum und Ausscheidung verantwortlich. Bei der Ernährung etwa gelangt die Nahrung zuerst in den Magen, wo sie mit Hilfe der *facultas concoquens* in einen Speisesaft verwandelt wird; dieser Saft wird von der dem Magen eigenen *facultas expellens* in die Därme gepumpt und gelangt von dort über die Pfortader vermittels der *facultas attrahens* in die Leber, in der aus dem Speisesaft das Blut entsteht; aus dem Blut entnimmt schließlich eine jeweils eigene *facultas attrahens* das Passende, um es an die einzelnen Glieder anzulagern. Rothschuh faßt den zentralen Gedanken dieser Auffassung so zusammen:

„Die *facultas* begründet jede Aktivität, sie ist zugleich die Quelle der Finalität, der Zweckhaftigkeit ... Die Vermögen sind die Ursachen der *Funktionen* ...“ (S. 19).

Diese Art, physiologische Vorgänge durch die Annahme spezieller *seelischer* Kräfte und Vermögen zu erklären, wird erst zu Beginn der Neuzeit einer radikalen Kritik unterzogen. Im Zusammenhang mit der Grundlegung einer neuen Naturwissenschaft im allgemeinen – insbesondere durch Galilei – werden die Grundannahmen der aristotelischen Weltanschauung fundamental in Zweifel gezogen. Die neue Weltanschauung, zu deren brilliantesten theoretischen Vertretern der Philosoph René Descartes gehört, setzt an ihre Stelle das Paradigma des Mechanismus: Das Verhalten aller Systeme in der Natur soll erklärt werden allein aus der Anordnung der in ihnen vorhandenen Teile und den allgemein in der Natur geltenden Gesetzen, besonders den mechanischen Gesetzen der Kraftübertragung

2) *Universa Medicina*, Liber V, Capitel 1 (zit. nach Rothschuh 1969, S. 17).

durch Stoß. Als Paradebeispiel eines mechanischen Systems gilt dabei die Uhr, deren Verhalten offensichtlich allein von der Anordnung und Gestalt ihrer Räder und Gewichte abhängt.

Gerade Descartes hat versucht, auch die Phänomene des Lebens in dieses mechanistische Weltbild zu integrieren. Er ist nicht nur der Philosoph des *cogito* oder der Vertreter des strengen Dualismus von *res extensa* und *res cogitans*, sondern in besonderem Maße auch ein versierter und zum Teil selbst empirisch forschender Naturwissenschaftler, dessen Grundannahmen den Auffassungen der antiken Atomisten Demokrit und Leukipp viel näher stehen als denen des Aristoteles.

Seine grundsätzliche Kritik an der aristotelischen Seelenlehre hätte Descartes, wenn er die Gelegenheit zu einem Gespräch mit Aristoteles gehabt hätte, dabei in etwa so formulieren können: „Wenn Du, Aristoteles, glaubst, zur Erklärung des Verhaltens und der Fähigkeiten lebender Wesen spezielle Kräfte und Vermögen der Seele postulieren zu müssen, dann liegt das nur daran, daß Du das Funktionieren der Körper dieser Lebewesen nicht richtig verstanden hast. Denn eine genaue Untersuchung der Teile dieser Körper zeigt, daß alle Lebensprozesse und alle die Fähigkeiten, die Du auf das Wirken einer vegetativen oder sensitiven Seele zurückführst, allein aus der Gestalt und Anordnung der in den Körpern der Lebewesen vorkommenden Teile (Organe) erklärt werden können, wenn man nur annimmt, daß diese Teile sich den überall in der Natur geltenden Gesetzen gemäß verhalten.“

Über diese allgemeine Kritik hinaus versucht Descartes jedoch auch im einzelnen für jeden ihm bekannten physiologischen Vorgang tatsächlich eine mechanistische Erklärung zu geben. Für die Vorgänge der Verdauung und Ernährung etwa entwickelt er im *Traité de l'homme* folgende Theorie, in der für spezielle *facultates* kein Platz mehr ist. Nachdem die Nahrung in den Magen gelangt ist, wird sie dort verdaut, und zwar „durch die Einwirkung gewisser Flüssigkeiten, die sich zwischen die Teile der Nahrung schieben, sie trennen, bewegen und erhitzen, so wie es normales Wasser mit ungelöschtem Kalk und wie es Scheidewasser mit Metall macht“. Dabei wirkt mit, daß diese Flüssigkeiten, da sie „durch die Arterien direkt vom Herzen herangebracht werden, der starken Hitze nicht ermangeln“. Außerdem erfolgt die Verdauung auch durch normale Prozesse der Selbstzersetzung, „wie es mit frischem Heu in der Scheune geschieht, wenn man es dort aufeinanderpackt, bevor es trocken ist“. Die Bewegung, in die die Nahrungsteilchen durch die Erhitzung gebracht werden, sowie die Bewegung des Magens und der Eingeweide bewirken aufgrund „der Anordnung kleiner Fäserchen, aus denen diese Eingeweide bestehen“, daß die Nahrungsteilchen im Verlauf der Verdauung langsam den Dünndarm und den Dickdarm passieren, bis am Ende „die größten unter ihnen ausgeschieden werden“. Bei ihrem Weg durch Dünndarm und Dickdarm „treffen die feinsten und bewegtesten Teilchen hier und dort auf

eine Unzahl von kleinen Löchern“, durch die sie auf dem Wege über die Pfortader zur Leber gelangen. Dabei werden diese Teilchen von den größeren wie durch ein Sieb getrennt. Es ist nur „die Kleinheit der Löcher, die sie von den größeren Teilchen scheidet“. Durch die Leber gelangen die kleinen Nahrungsteilchen ins Blut und mit dem Blut in alle Teile des Körpers. Dort können sie leicht an allen „Gliedern hängenbleiben und sich mit ihnen verbinden, sie so ernähren oder sogar wachsen lassen“. Genauer gesagt: Durch ihre Bewegung haben die kleinen Nahrungsteilchen im Blut die Kraft, kleine Teilchen z. B. in den Knochen, im Fleisch, in den Nerven oder in der Haut „ein wenig zur Seite zu stoßen und sich an ihre Stelle zu setzen. Sobald die Arterien wieder zusammenfallen, verbleibt jedes dieser Blutteilchen dort, wo es sich gerade befindet. Und nur aus diesem Grund wird es dort verbunden und vereinigt mit denen, die es berührt“ (Descartes 1969, S. 45–51).

In unseren Augen ist diese Theorie der Verdauung und Ernährung sicher krude. Und sie kann auch kaum besser sein. Denn im Rahmen der mechanistischen Grundannahmen Descartes' läßt sich eine nach heutigen Maßstäben angemessene Idee der Chemie nur schwer entwickeln. Und wie weit kann man in der Physiologie ohne Chemie schon kommen? Aber es ist auch weniger die empirische Adäquatheit als vielmehr die theoretische Striktheit, die Descartes' Physiologie interessant macht. In jeder Einzelheit wird deutlich, daß er die Bezugnahme auf spezielle Vermögen überall peinlich zu vermeiden sucht. Die Verdauung im Magen z. B. führt er nicht auf eine spezifische *facultas concoquens* zurück, sondern auf das Wirken bestimmter Flüssigkeiten und auf die im Magen vorhandene Wärme. Besonders deutlich wird dabei der mechanistische Ansatz, wenn er über das Wirken der Magensäfte schreibt, daß sie „sich zwischen die Teile der Nahrung schieben, sie trennen, bewegen und erhitzen“, wobei anzumerken ist, daß Descartes der Meinung war, schon im zweiten Kapitel des *Traité de la lumière* nachgewiesen zu haben, daß die Entstehung und das Wirken von Wärme ebenfalls als rein mechanische Erscheinungen erklärt werden können.

Ebenso mechanisch wie seine Theorie der Verdauung ist Descartes' Theorie des eigentlichen Ernährungsvorganges. Denn auch hier wirken für ihn keine spezifischen *facultates*, sondern die normalen mechanischen Gesetze. Ernährung, d. h. Eingliederung der kleinen Nahrungsteilchen in das schon vorhandene Gewebe, geschieht, indem diese kleinen, im Blut befindlichen Nahrungsteilchen schon vorhandene Gewebeteilchen einfach beiseite stoßen und sich so an ihre Stelle setzen. Wie immer man diese Annahmen im einzelnen bewerten mag, Descartes selbst war auf jeden Fall der Meinung, eine vollständige mechanische Erklärung aller physiologischen Lebensvorgänge gegeben zu haben. Zumindest schreibt er am Ende des *Traité de l'homme*, er habe gezeigt, daß die Funktionen der Lebewesen „alle von Natur aus allein aus der Disposition ihrer Organe hervorgehen, nicht mehr und nicht weniger, als die Bewegungen einer Uhr oder

eines anderen Automaten von der Anordnung ihrer Gewichte und ihrer Räder abhängen. Daher ist es in keiner Weise erforderlich, hier für diese eine vegetative oder sensitive Seele oder ein anderes Bewegungs- und Lebensprinzip anzunehmen ...“ (S. 136).

Allerdings, auch für Descartes gibt es eine Grenze für die mechanische Erklärbarkeit der Fähigkeiten von Lebewesen. Und diese Grenze liegt für ihn da, wo beim Menschen die Fähigkeiten des Denkens und Sprechens ins Spiel kommen. Leider gibt es zu diesem Punkt nur wenige systematische Texte. Am ausführlichsten sind noch Descartes' Bemerkungen im Teil V des *Discours de la méthode*, in denen er auf offenbar verschollene Teile seines Buches *Le monde* verweist, von dem postum nur die beiden Teile *Traité de la lumière* und *Traité de l'homme* veröffentlicht wurden. Auf jeden Fall betont Descartes in diesen Bemerkungen zunächst noch einmal, daß Tiere seiner Meinung nach vollständig mechanisch erklärt werden können. Menschen jedoch unterscheiden sich seiner Meinung nach in zwei Punkten grundsätzlich von jeder Maschine, d. h. von jedem mechanischen System. Denn erstens können solche Systeme „niemals Worte oder andere Zeichen dadurch gebrauchen, daß sie sie zusammenstellen, wie wir es tun, um anderen unsere Gedanken bekanntzumachen“. Und zweitens werden solche Systeme, auch wenn sie in einigen Punkten sehr gute Leistungen erbringen, „doch zweifellos bei vielem anderen versagen, wodurch offen zutage tritt, daß sie nicht aus Einsicht handeln, sondern nur zufolge der Einrichtung ihrer Organe. Denn die Vernunft ist ein Universalinstrument, das bei allen Gelegenheiten zu Diensten steht, während diese Organe für jede besondere Handlung einer besonderen Einrichtung bedürfen ...“ (Descartes 1960, S. 92f.).

*Sprache* und *Intelligenz* sind also die beiden Merkmale, die den Menschen von jedem mechanischen System unterscheiden, da es sich in beiden Fällen um Fähigkeiten handelt, für die Descartes zufolge mechanische Erklärungen grundsätzlich unmöglich sind. Leider sagt Descartes sehr wenig darüber, warum es seiner Meinung nach für die Fähigkeiten des Denkens und Sprechens keine mechanischen Erklärungen geben kann. Aber vielleicht ist es besonders der im letzten Satz der gerade zitierten Passage angesprochene universale Charakter der Vernunft, der nach Descartes in diesem Zusammenhang ausschlaggebend ist. Auf jeden Fall läßt Descartes keinen Zweifel daran, daß es sich hier seiner Meinung nach um im Rahmen einer mechanistischen Naturwissenschaft nicht erklärbare Phänomene handelt und damit um Phänomene, für die es eine andere Erklärung geben muß. Und diese andere Erklärung kann für ihn nur eine immaterielle Seele sein. Trotz seiner grundsätzlichen Kritik verwendet Descartes somit am Ende eine ähnliche Argumentationsfigur wie Aristoteles. Denn dort, wo es für feststellbare Verhaltensweisen und Fähigkeiten von Lebewesen keine naturwissenschaftliche Erklärung mehr gibt, wird von ihm ebenso wie zuvor von Aristoteles die Seele als *zusätzliches* Erklärungsprinzip in Anspruch genommen. In seiner

Auffassung darüber, was diese Seele ihrem Wesen nach ist, steht Descartes allerdings Platon näher als Aristoteles.

Immerhin scheint die Frage nahe zu liegen, ob man nicht dort, wo Descartes zur Erklärung vermeintlich naturwissenschaftlich nicht erklärbarer Phänomene auf die Annahme einer immateriellen Seele zurückgreift, gegen Descartes genau so argumentieren kann, wie er gegen Aristoteles argumentiert hat. „Wenn Du, Descartes, glaubst, zur Erklärung der Fähigkeiten des Denkens und Sprechens eine immaterielle Seele postulieren zu müssen, dann liegt das nur daran, daß Du das Funktionieren des menschlichen Körpers und insbesondere des menschlichen Gehirns noch nicht richtig verstanden hast. Denn wenn man die Funktionsweise des Gehirns und des peripheren Nervensystems erst einmal richtig verstanden hat, dann wird deutlich, daß auch mentale Phänomene wie das Denken und Sprechen vollständig als neuronale Phänomene erklärt werden können.“

Diese Strategie scheint auf der Hand zu liegen, und in dieser oder ähnlicher Weise hat der nachcartesianische Materialismus auch immer wieder versucht, zumindest Descartes' naturphilosophische Argumente<sup>3</sup> für den Dualismus zu Fall zu bringen. Aber leider ist die Sache doch nicht ganz so einfach. Denn *faktisch* wissen wir einfach immer noch zu wenig über das Funktionieren von Gehirn und ZNS, um die Strategie tatsächlich erfolgreich durchführen zu können. Soweit ich den Forschungsstand überblicke (und ich gestehe, daß mein Überblick nicht sehr groß ist), wissen wir in vielen Fällen nur, welche Teile etwa des Cortex bei bestimmten kognitiven Aktivitäten involviert sind, aber nicht, wie genau die den einzelnen Aktivitäten zugrundeliegenden neuronalen Prozesse aussehen. Und wo – wie z.B. durch die Arbeiten von Hubel und Wiesel über den visuellen Cortex – zumindest zum Teil bekannt ist, was sich im neuronalen Bereich während des Sehvorgangs abspielt, wissen wir nicht, wie die feststellbaren Teilleistungen im einzelnen erbracht werden und was genau sie zum Gesamtvorgang des Sehens beitragen.

Diese Unwissenheit zwingt uns jedoch nicht dazu, nun sofort zu resignieren und den Versuch einer physikalistischen Erklärung mentaler Phänomene von vornherein aufzugeben. Denn wenn wir die Dinge noch einmal genauer betrachten, dann wird deutlich, daß Descartes Aristoteles gegenüber kaum in einer besseren Position war. Auch er wußte keineswegs genau, welche physiologischen Prozesse den verschiedenen Lebensvorgängen im einzelnen zugrundeliegen. Und das war ihm auch durchaus bewußt. Descartes hatte nur einige relativ allgemeine Vorstellungen darüber, wie diese Lebensvorgänge eventuell allein unter Bezugnahme auf die Prinzipien einer mechanistischen Physik erklärt werden könnten.

3) Außer den genannten naturphilosophischen Argumenten führt Descartes für die Existenz einer vom Körper grundsätzlich verschiedenen Seele auch ein metaphysisches Argument an, das ich in (1986) ausführlich analysiert habe.

Und zur Erläuterung dieser Vorstellungen wählt er einen sehr interessanten Weg: Er entwickelt *Modelle*, d. h. er schildert *mögliche Maschinen*, die, obwohl ihr Verhalten einzig und allein von der Gestalt und Anordnung ihrer Teile abhängt, genau dieselben oder wenigstens im Prinzip dieselben Leistungen vollbringen wie pflanzliche oder tierische Lebewesen. Ich habe diesen Aspekt der Argumentation Descartes' bisher unterschlagen. Aber jetzt ist es doch Zeit, ihn explizit zu machen. Sehr instruktiv sind zwei Absätze aus dem Kapitel XVIII, mit dem der *Traité de l'homme* beginnt.

„Ich stelle mir einmal vor, daß der Körper nichts anderes sei als eine Statue oder Maschine aus Erde, die Gott gänzlich in der Absicht formt, sie uns so ähnlich wie möglich zu machen, und zwar derart, daß er ihr nicht nur äußerlich die Farbe und die Gestalt aller unserer Glieder gibt, sondern auch in ihr Inneres alle jene Teile legt, die notwendig sind, um sie laufen, essen, atmen, kurz alle unsere Funktionen *nachahmen* zu lassen, von denen man sich vorstellen könnte, daß sie aus der Materie ihren Ursprung nehmen und lediglich von der Disposition der Organe abhängen“ (Descartes 1969, S. 44; Hervorh. v. V.).

Descartes geht es also nicht (in erster Linie) darum, die wirkliche Physiologie des menschlichen Körpers darzustellen. Er setzt sich vielmehr das Ziel, das Funktionieren einer Maschine zu beschreiben, die alle die Teile enthält, die notwendig sind, um sie alle Funktionen eines lebenden Körpers „nachahmen“ zu lassen – zumindest soweit diese Funktionen nicht die rein menschlichen und mechanisch nicht erklärbaren Fähigkeiten des Denkens und Sprechens betreffen. Und dafür, daß dieses Ziel nicht von vornherein als unerreichbar angesehen werden muß, gibt er folgende Begründung.

„Wir sehen Uhren, kunstvolle Wasserspiele, Mühlen und andere ähnliche Maschinen, die, obwohl sie nur von Menschenhand hergestellt wurden, nicht der Kraft entbehren, sich aus sich selbst auf ganz verschiedene Weise zu bewegen. Und wir mir scheint, könnte ich mir von einer Maschine, die – wie ich einmal annehme – aus der Hand Gottes angefertigt sein soll, nicht so viele Bewegungsarten vorstellen noch ihr so viel kunstvolle Bildung zuschreiben, daß man sich nicht vorstellen könnte, daß sie nicht noch mehr davon besitzen kann.“ (S. 44)

Descartes' Ziel im *Traité de l'homme* ist es also nicht, direkt zu zeigen, daß die vitalen Leistungen aller Lebewesen allein aus der Gestalt und der Anordnung der in ihnen enthaltenen Teile (Organe) erklärt werden können. Er wählt vielmehr einen indirekten Weg, indem er statt dessen versucht, das Funktionieren einer Maschine mit den folgenden Eigenschaften zu beschreiben:



- (1) Die einzelnen Teile der Maschine gleichen den Organen des menschlichen Körpers soweit wie möglich.
- (2) Alle Funktionen der Maschine ergeben sich allein aus den materiellen Eigenschaften dieser Teile.
- (3) Die Maschine erbringt – Denken und Sprechen ausgenommen – dieselben Leistungen wie ein Mensch.

Es liegt deshalb nahe, bei einer Kritik der Descartesschen Thesen über die naturwissenschaftliche Unerklärbarkeit bestimmter mentaler Fähigkeiten einen ähnlichen Weg zu gehen. Denn angesichts unseres immer noch sehr mangelhaften Wissens im Bereich der Neurobiologie ist es vielleicht sinnvoll, wenn wir uns zunächst (noch) nicht das Ziel stecken, direkt nachzuweisen, daß diese Fähigkeiten doch auf bestimmte neuronale Prozesse zurückgeführt werden können, sondern stattdessen das möglicherweise etwas bescheidenere Ziel verfolgen, zu zeigen, daß man – vollständig naturwissenschaftlich erklärbar – Maschinen konstruieren kann, die über dieselben oder doch sehr ähnliche Fähigkeiten verfügen.

Damit sind wir an einem Punkt angekommen, an dem es sinnvoll ist, einen Blick hinüberzuwerfen in einen ganz anderen Wissenschaftsbereich, in die Künstliche-Intelligenz-Forschung. Denn einer inzwischen schon notorischen Bemerkung Marvin Minskys (1968) zufolge ist

„... *Künstliche Intelligenz* die Wissenschaft, in der es darum geht, Maschinen zu konstruieren, die Dinge tun, zu deren Ausführung Intelligenz nötig wäre, wenn sie vom Menschen getan würden.“ (S. V – Übers. v. V.)

In der KI-Forschung geht es also, wenn Minsky recht hat, genau darum, Programme (und zum Teil auch hardware) zu entwickeln, die Computer (und d.h. Maschinen, deren Verhalten sich vollständig physikalisch erklären läßt) befähigen, die mentalen und kognitiven Leistungen zu vollbringen, die Descartes zufolge grundsätzlich nicht mechanisch erklärbar sind. Wenn die KI-Forschung mit diesem Programm Erfolg hat, sind daher zumindest Descartes' naturphilosophische Argumente für den Dualismus widerlegt. Denn wenn im Rahmen dieser Forschung Systeme entwickelt werden können, die die gleichen oder zumindest vergleichbare geistige bzw. kognitive Leistungen wie ein Mensch zu erbringen imstande sind, dann ist damit zumindest im Grundsatz gezeigt, daß die Annahme einer immateriellen Seele zur Erklärung dieser Leistungen nicht notwendig ist. Ich möchte deshalb hier beispielhaft für den Bereich des Sprachverstehens kurz deutlich machen, was die KI-Forschung bisher erreicht hat.

Wenn man fragt, ob es der bisherigen KI-Forschung gelungen ist, Systeme zu entwickeln, die sprechen können wie wir, bzw. in der Terminologie Descartes': die „Worte oder andere Zeichen dadurch gebrauchen, daß sie sie zusammenstellen, wie wir es tun, um anderen unsere Gedanken bekanntzumachen“, dann ist

es zunächst sinnvoll zu klären, was denn als Evidenz für das „semantisch-pragmatisch fundierte Sprachverstehen eines Systems angesehen“ werden kann. In seinem Aufsatz „Natürlichsprachliche Systeme – Eine Einführung in die sprachorientierte KI-Forschung“ formuliert Wahlster (1982, S. 207) folgenden Kriterienkatalog:

- Fragen über ein Sachgebiet oder einen Text beantworten können
- einen aufgabenorientierten Dialog (z.B. Beratungsgespräch) führen oder Konversation betreiben können
- einen Text paraphrasieren, zusammenfassen oder nacherzählen können
- sich gemäß natürlichsprachlich eingegebenen Anweisungen verhalten oder danach handeln können
- einen Text in eine andere Sprache übersetzen können.

Im Hinblick auf diese Kriterien sind der KI-Forschung in den letzten Jahren erstaunliche Erfolge gelungen, wie die in den Kästen 1 und 2 angeführten Beispiele zeigen (beide aus Wahlster 1982, S. 207f.).

Das erste Beispiel gibt einen Dialog wieder, den einer der Autoren des Systems, Dan Bobrow (1977), mit dem System GUS (Genial Understanding System) geführt hat. Dieses System kann mit einem Benutzer einen „aufgabenorientierten“ Dialog führen, in dem eine Flugreise des Benutzers geplant und die damit verbundenen Reservierungen vorgenommen werden. Dabei „können die Flexibilität der Dialogführung und die Korrektheit der vom System ausgeführten Buchungen als Evidenzen für das Sprachverstehen ... gewertet werden“ (Wahlster 1982, S. 209).

<p>SYS: Hallo, meine Name ist GUS. Ich kann Ihnen beim Planen einer einfachen Flugreise helfen. Wohin möchten Sie?</p> <p>BEN: Ich möchte am 28. Mai nach San Diego.</p> <p>SYS: Wann möchten Sie von hier weg?</p> <p>BEN: Um 10 Uhr muß ich in San Diego sein.</p> <p>SYS: Wie wäre es mit dem Flug Nr. 102 der P.S.A., der um 9.15 Uhr ankommt?</p> <p>BEN: Welches ist der nächste Flug?</p> <p>SYS: A. C.-Flug Nr. 310, der um 8.30 Uhr hier losgeht und um 10.00 Uhr ankommt.</p> <p>BEN: Ich nehme den ersten.</p> <p>SYS: Auf welchen Namen soll ich die Reservierung vornehmen, bitte?</p> <p>BEN: Dan Bobrow.</p> <p>SYS: Ich habe folgenden Flug gebucht: P.S.A. Flug Nr. 102 am Mittwoch den 28. Mai von San José nach San Diego, der um 7.30 Uhr startet und 9.15 Uhr ankommt. An welchem Tag möchten Sie zurückkommen?</p> <p>BEN: Am Freitagabend.</p> <p>SYS: Wie wäre es mit dem Flug, der um 19.45 Uhr losgeht?</p> <p>BEN: Prima.</p>
--

*Kasten 1:* Beispieldialog mit dem System GUS

Das zweite Beispiel zeigt die Arbeitsweise des Systems PAM (Plan Applier Mechanism, Wilensky o.J.). Bei diesem System kann der Benutzer nach Eingabe einer Kurzgeschichte z.B. nach Begründungen für die in der Geschichte erwähnten Handlungen fragen. „Das Textverstehen von PAM wird durch Antworten belegt, die in dieser Form nicht explizit im eingegebenen Text enthalten sind. ... Im Paraphrasierungsmodus kann PAM den in der Kurzgeschichte beschriebenen Handlungsablauf aus der Perspektive der beteiligten Akteure nacherzählen, wobei die Nacherzählung wiederum Aussagen enthält, die als Evidenzen für eine semantisch-pragmatische Verarbeitung des eingegebenen Textes gewertet werden können“ (Wilensky o.J.).

Texteingabe

BEN: Peter wollte das Fahrrad von Uwe. Er ging zu Uwe und fragte ihn, ob er es ihm gibt. Uwe lehnte das ab. Peter sagte Uwe, er würde ihm dafür 20 DM geben, aber Uwe war damit nicht einverstanden. Dann sagte Peter, er würde ihm den Arm brechen, wenn er es ihm nicht überlassen würde. Uwe gab ihm das Fahrrad.

Frage-Antwort-Modus

BEN: Warum ging Peter zu Uwe?

SYS: Weil er sein Fahrrad haben wollte.

BEN: Warum gab Uwe Peter sein Fahrrad?

SYS: Weil er nicht verletzt werden wollte.

BEN: Was hatte es für Konsequenzen, daß Peter zu Uwe ging?

SYS: Dies ermöglichte ihm, Uwe um sein Fahrrad zu bitten.

BEN: Was waren die Konsequenzen von Peters Frage an Uwe, ihm Uwes Fahrrad zu geben?

SYS: Uwe sagt ihm, daß Uwe ihm nicht sein Fahrrad geben würde.

Paraphrasierungsmodus

SYS: Wer soll die Geschichte erzählen?

BEN: Peter.

SYS: Ich wollte Uwes Fahrrad haben, also ging ich zu ihm und bat ihn darum. Er sagte mir, daß er es mir nicht geben würde. Also bat ich ihn, es mir für 20 DM zu verkaufen. Da sagte er, daß er es nicht vergeben würde. Ich sagte, wenn er es mir nicht gäbe, würde ich ihm seinen Arm brechen. Er gab mir sein Fahrrad.

SYS: Wer soll die Geschichte erzählen?

BEN: Uwe.

SYS: Peter kam zu mir. Er fragte mich, ob ich ihm mein Fahrrad geben würde. Ich wollte das Fahrrad behalten, also sagte ich, daß ich es ihm nicht geben würde. Da bot er an, es mir für 20 DM abzukaufen. Ich wollte das Fahrrad behalten. Also sagte ich, daß ich es ihm nicht geben würde. Er sagte, wenn ich es ihm nicht gäbe, würde er mir meinen Arm brechen. Ich wollte nicht verletzt werden. Also gab ich ihm mein Fahrrad.

Kasten 2: Beispieldialog mit dem System PAM

Angesichts der zunächst einmal sicher beeindruckenden Leistungen der beiden Systeme GUS und PAM muß man sich jedoch klar machen, daß das Dialogverhalten eines Systems bezüglich eines eingeschränkten Gegenstandsbereichs nicht in jedem Fall ein verlässlicher Indikator dafür ist, daß dieses System tatsächlich Sprache so verstehen und verwenden kann, „wie wir es tun“<sup>4</sup>. Ein lehrreiches Beispiel ist hier das schon 1965 von Joseph Weizenbaum entwickelte System ELIZA. Denn auch der in Kasten 3 wiedergegebene Dialog mit diesem System (Weizenbaum, S. 15f.), in dem der Computer einer jungen Patientin gegenüber den Part eines Psychotherapeuten spielt, der die Rogersche Technik der Gesprächsführung verwendet, sieht auf den ersten Blick recht beeindruckend aus. Tatsächlich sind die Grenzen der „Sprachfähigkeit“ ELIZAs jedoch sehr eng gesteckt (s. dazu Boden 1977, S. 106–111). Dies zeigt sich besonders an zwei Punkten. Erstens nämlich daran, daß ELIZA einfach nicht in der Lage ist, aus der Rolle des Therapeuten hervorzutreten (vgl. den in Kasten 4 angeführten Dialog, Boden 1977, S. 96), der zeigt, was passieren kann, wenn jemand, der einfach nur den Computer benutzen will, dabei zufällig und ohne es zu wissen an ELIZA gerät). Und zweitens daran, daß ELIZA auf den größten sprachlichen Unsinn genauso reagiert wie auf sinnvolle sprachliche Äußerungen. Denn wenn jemand, der sich entschlossen hat, absichtlich Unsinn zu reden, z.B. den Satz eingibt „*Bitten Sie doch den Tisch, einen Ziegel für das Baby aufzuheben*“, dann beschwert sich ELIZA keineswegs, da das System einfach nicht in der Lage ist, diese Eingabe als semantischen Unsinn zu erkennen. Und aus demselben Grund antwortet ELIZA auch auf die Eingabe „*Ich glaube, daß Sie mich bibobabo*“ völlig ungerührt „*Was läßt Sie vermuten, daß ich Sie bibobabo?*“

BEN: Die Männer sind alle gleich.  
 SYS: In welcher Weise?  
 BEN: Sie machen uns immer wegen irgendwas verrückt.  
 SYS: Können Sie ein Beispiel nennen?  
 BEN: Na ja, die Idee zu diesem Gespräch stammt von meinem Freund.  
 SYS: Die Idee zu diesem Gespräch stammt von Ihrem Freund.  
 BEN: Er sagt, daß ich oft deprimiert bin.  
 SYS: Es tut mir leid von Ihnen zu hören, daß Sie deprimiert sind.  
 BEN: Das stimmt, ich bin unglücklich.  
 SYS: Glauben Sie, daß dieses Gespräch Ihnen hilft, nicht mehr unglücklich zu sein?  
 BEN: Ich brauche jedenfalls Hilfe, das steht fest.  
 SYS: Was würde es für Sie bedeuten, Hilfe zu bekommen?  
 BEN: Vielleicht könnte ich lernen, mit meiner Mutter auszukommen.  
 SYS: Erzählen Sie mehr über Ihre Familie.

4) Wahlster (1982) spricht deshalb mit Absicht in diesem Zusammenhang nur von Evidenzen und nicht von Beweisen; vgl. S. 209.

BEN: Meine Mutter erledigt alles für mich.  
 SYS: Wer in Ihrer Familie erledigt noch alles für sie?  
 BEN: Mein Vater.  
 SYS: Ihr Vater.  
 BEN: Sie sind ein wenig wie mein Vater.  
 SYS: Welche Ähnlichkeit sehen Sie?  
 BEN: Sie sind nicht besonders aggressiv, aber ich glaube, Sie möchten nicht, daß ich das beachte.  
 SYS: Was läßt Sie vermuten, daß ich nicht sehr aggressiv bin?  
 BEN: Sie streiten nicht mit mir.  
 SYS: Wieso glauben Sie, daß ich nicht mit Ihnen streite?  
 BEN: Sie haben Angst vor mir.  
 SYS: Macht Ihnen der Gedanke Spaß, daß ich vor Ihnen Angst habe?  
 BEN: Mein Vater hat vor jedem Angst.  
 SYS: Was fällt Ihnen noch ein, wenn Sie an Ihren Vater denken?  
 BEN: Angeber.  
 SYS: Hat das etwas damit zu tun, daß die Idee zu diesem Gespräch von Ihrem Freund stammt?

*Kasten 3: Beispieldialog mit dem System ELIZA*

Diese Unzulänglichkeiten im Dialogverhalten ELIZAs gehen jedoch auf die sehr simple innere Struktur dieses Systems zurück, das letztlich nur mit einer sehr einfachen Methode des *pattern matching* arbeitet. Bei jedem eingegebenen Satz sucht ELIZA zunächst nach Schlüsselwörtern wie *Ich*, *Sie*, *jeder* oder *Vater*. Wenn das System ein solches Schlüsselwort findet, formt es den eingegebenen Satz entsprechend einer mit diesem Schlüsselwort verbundenen Regel um. So reagiert ELIZA z. B. auf den Satz „*Mein Vater hat vor jedem Angst*“ mit der Antwort „*Was fällt Ihnen noch ein, wenn Sie an Ihren Vater denken?*“ oder auch mit der Antwort „*Erzählen Sie mehr über Ihre Familie*“, während das System den Satz „*Ich weiß, daß jeder über mich lacht*“ z. B. mit der Ausgabe „*Gibt es jemand Bestimmten, an den Sie dabei denken?*“ beantwortet. Weiter gibt es spezielle Regeln für die Transformation von *Sie* in *Ich* und von *mich (mir)* in *Sie (Ihnen)*, die z. B. bei der Eingabe „*Sie streiten nicht mit mir*“ zu der Ausgabe „*Wieso glauben Sie, daß ich nicht mit Ihnen streite?*“ führen. Und wenn ELIZA kein Schlüsselwort findet, reagiert das System mit allgemeinen Floskeln wie „*Warum glauben Sie das?*“ oder mit Bemerkungen, in denen auf frühere Themen Bezug genommen wird, wie „*Hat das etwas damit zu tun, daß die Idee zu diesem Gespräch von Ihrem Freund stammt?*“ Es ist klar, daß mit so einfachen Mitteln kein Sprachverhalten erzeugt werden kann, das dem des Menschen auch nur annähernd gleichkommt.

BEN: Glauben Sie, daß ich heute morgen den Computer benutzen kann?  
 SYS: Warum fragen Sie?  
 BEN: Ich könnte vielleicht noch einige weitere Telcomp-Geräte verkaufen.  
 SYS: Warum sind Sie nicht sicher?  
 BEN: Die möglichen Kunden haben noch nie eine Vorführung des Systems gesehen.  
 SYS: Bedeutet das sehr viel für Sie?  
 BEN: Natürlich tut es das.  
 SYS: Sie scheinen in dieser Hinsicht sehr sicher zu sein.  
 BEN: Das ist lächerlich.  
 SYS: Was meinen Sie mit, das ist lächerlich?  
 BEN: Gibt es irgendeinen Grund, warum ich den Computer jetzt nicht benutzen kann?  
 SYS: Erzählen Sie mir mehr über Ihre möglichen Kunden.  
 BEN: Bitte rufen Sie mich an unter 492-1850.

*Kasten 4: Zweiter Beispieldialog mit dem System ELIZA*

Neuere natürlichsprachliche Systeme (NSS) wie GUS und PAM sind allerdings mit solch gravierenden Problemen nicht mehr belastet, da bei diesen Systemen die interne Sprachverarbeitung völlig anders vonstatten geht. Ich kann in diesem Zusammenhang nicht auf die Details eingehen und will deshalb das Innenleben dieser neueren Systeme hier nur ganz kurz andeuten (vgl. Wahlster 1982; Winograd 1984). In Systemen wie GUS und PAM wird jeder eingegebene Satz in einem ersten Schritt einer morphologisch-lexikalischen Analyse unterzogen, als deren Ergebnis jedem Wort dieses Satzes seine lexikalische Kategorie sowie, falls sinnvoll, einige syntaktische Merkmale zugeordnet werden. In einem zweiten Schritt wird der Eingabesatz dann – unter Berücksichtigung der Ergebnisse der morphologisch-lexikalischen Analyse – syntaktisch analysiert. Dabei wird dem Satz, sofern er syntaktisch korrekt ist, seine syntaktische Struktur – z. B. in Form eines Strukturbaums – zugeordnet. Im nächsten Schritt wird dann der eingegebene Satz – diesmal unter Verwendung der Ergebnisse sowohl der morphologisch-lexikalischen als auch der syntaktischen Analyse – in eine interne Repräsentation, in eine Art internen Satz des Systems überführt. Auf diese interne Repräsentation wird schließlich eine Auswertungskomponente angewendet, die die Reaktion des Systems bestimmt. Generell veranlaßt die Auswertungskomponente, daß das System eine bestimmte Handlung ausführt, wobei diese Handlung häufig in der Generierung und Ausgabe eines Satzes besteht.

Noch einmal kurz zusammengefaßt, läßt sich die interne Sprachverarbeitung neuerer NSS also so charakterisieren: Erstens wird jeder eingegebene Satz in eine interne Repräsentation überführt, wobei im Idealfall die morphologisch-lexikalischen, die syntaktischen und die semantischen Komponenten des Systems dafür sorgen, daß allen Sätzen mit gleicher Bedeutung die gleiche Repräsentation zugeordnet wird; auf diese interne Repräsentation wird dann zweitens eine Auswer-

tungskomponente angewendet, die im Idealfall bewirkt, daß das System in einer Weise reagiert, die genau der in der internen Repräsentation gespeicherten Bedeutung des eingegebenen Satzes entspricht.

Es ist klar, daß diese dem System ELIZA gegenüber deutlich verbesserte interne Sprachverarbeitung mithilft, bei neueren NSS ELIZAs Fehler zu vermeiden. Aber trotzdem gibt es eine Reihe von Philosophen, die die Meinung vertreten, daß auch diese neueren NSS immer noch weit davon entfernt sind, Sprache so zu verwenden, „wie wir es tun“. Der bekannteste dieser Philosophen ist wahrscheinlich John Searle, der in mehreren Arbeiten (Searle 1980, 1986) die These vertreten hat, daß sich auch die Sprachfähigkeit neuerer NSS grundsätzlich von den entsprechenden Fähigkeiten des Menschen unterscheidet, weil diese Systeme im Gegensatz zum Menschen einfach *nicht wissen*, was die sprachlichen Ausdrücke *bedeuten*, die man in sie eingibt oder die sie selbst ausgeben. Sein Hauptargument für diese These besteht in der Überlegung, daß alle Leistungen, die diese Systeme erbringen, auch von einem Menschen erbracht werden können, der nichts von der Sprache versteht, in der er entsprechende „Dialoge“ mit der Umwelt führt. Ich kann hier wieder nicht im einzelnen auf Searles Argumente eingehen<sup>5</sup>. Aber so viel will ich doch sagen. Mir scheint, daß Searle recht hat, solange die Systeme, um die es geht, nicht auch über eine perzeptuelle Komponente verfügen, die es ihnen ermöglicht, Wahrnehmungswissen über ihre Umwelt zu erwerben. Wenn jedoch Systeme sowohl eine solche Wahrnehmungskomponente als auch eine sprachverarbeitende Komponente der geschilderten Art besitzen, dann, denke ich, reichen Searles Argumente nicht aus, um zu zeigen, daß diese Systeme nicht wissen, was die von ihnen verarbeiteten sprachlichen Gebilde bedeuten.

Unabhängig von diesen sehr allgemeinen und grundsätzlichen Überlegungen ist jedoch sicher unbestritten, daß auch die Leistungen der eindrucksvollsten NSS immer noch deutlich hinter den entsprechenden menschlichen Leistungen zurückbleiben. Und dies gilt analog auch für alle KI-Systeme in anderen Bereichen. Besonders interessant ist im übrigen, daß es im Bereich *computer vision* bisher überhaupt nicht gelungen ist, ein System zu entwickeln, das auch nur entfernt ähnliche Leistungen erbringen kann wie ein Mensch.

Welche Relevanz können angesichts dieser Sachlage dann aber die Ergebnisse der KI-Forschung für die Probleme der Philosophischen Psychologie haben? Ich will versuchen, meine Antwort auf diese Frage in fünf Punkten zusammenzufassen.

Erstens: Da die bisher entwickelten KI-Systeme in ihrer Leistungsfähigkeit deutlich hinter den entsprechenden Fähigkeiten des Menschen zurückbleiben, kann im Augenblick sicher (noch) nicht die Rede davon sein, daß mit Hilfe die-

5) Ich habe das an anderer Stelle in dem Aufsatz (1988) getan.

ser Systeme die These Descartes', die spezifisch menschlichen Fähigkeiten des Denkens und Sprechens seien mechanisch unerklärbar und bedürften daher zu ihrer Erklärung der Annahme einer immateriellen Seele, gewissermaßen durch einen Gegen-Existenzbeweis endgültig widerlegt werden könnte.

Zweitens: Auf der anderen Seite ist die Leistungsfähigkeit der bisherigen KI-Systeme aber doch so groß, daß man diesen Systemen, wenn man ihnen als Naturgegenständen begegnen würde, sicher einen gewissen Geist nicht absprechen würde. Vielleicht läßt sich daher anhand dieser Systeme auch heute schon das offenbar komplizierte Verhältnis zwischen der *personalen* und der *subpersonalen* Ebene bei der Beschreibung mentaler Systeme, das z. B. Peter Bieri in seinem Beitrag zu diesem Band angesprochen hat, zumindest teilweise ein wenig aufklären. Denn in dem Maße, in dem wir geneigt und vielleicht sogar berechtigt sind, KI-Systeme mit mentalen Begriffen wie „überlegen“, „verstehen“, „meinen“, „erwarten“, „ein Ziel verfolgen“ usw. zu beschreiben, ergibt sich auch bei diesen Systemen das Problem des Verhältnisses der genannten beiden Beschreibungsebenen. Und da wir KI-Systeme im Gegensatz zum ZNS des Menschen und der höheren Tiere zumindest im Prinzip vollständig durchschauen, ergibt sich in diesem Fall eher die Möglichkeit, die Beziehungen zwischen der *personalen* und der *subpersonalen* Ebene im einzelnen zu untersuchen.

Mit diesem Punkt hängt mein dritter Punkt eng zusammen: In gewisser Weise ist es gerade die mangelnde Leistungsfähigkeit der bisherigen KI-Systeme, die diese Systeme für die Philosophische Psychologie besonders interessant macht. Denn wenn diese Systeme auf der einen Seite deutlich hinter den entsprechenden menschlichen Leistungen zurückbleiben, auf der anderen Seite aber doch „Dinge tun, zu deren Ausführung Intelligenz nötig wäre, wenn sie vom Menschen getan würden“, dann stellt sich natürlich die Frage, ob es sich bei diesem Unterschied um einen prinzipiellen oder vielleicht doch nur um einen graduellen Unterschied handelt. Oder anders ausgedrückt: Dann stellt sich die Frage, ob wir diese Systeme mit unserem für Menschen angemessenen mentalen Vokabular beschreiben und ihnen zumindest in gewisser Weise Geist zusprechen dürfen oder ob ihnen irgend etwas fehlt, was für das Besitzen von Geist unerlässlich ist. Der Unterschied zwischen den bisherigen KI-Systemen auf der einen und Menschen auf der anderen Seite ist also für unseren *Begriff des Mentalen* außerordentlich interessant. Denn dieser Unterschied zwingt uns, genauer, als wir es vielleicht bisher getan haben, zu sagen, was es denn ist, das ein Wesen zu einem geistigen Wesen macht, d. h. zu einem Wesen, das wir nicht nur mit subpersonalen, sondern auch mit personalen Begriffen beschreiben und erklären können. Er zwingt uns, genauer über den Begriff des Geistes nachzudenken, dabei die Merkmale des Mentalen präziser zu formulieren und damit unter anderem auch zu entscheiden, ob Geist-Haben ein Alles-oder-nichts-Phänomen ist oder ein Phänomen, das graduelle Unterschiede, also ein Mehr oder Weniger, zuläßt.



Viertens: Ein weiterer Grund, der die bisherigen Versuche der KI-Forschung, Systeme mit zumindest in gewisser Weise intelligenten oder mentalen Fähigkeiten zu entwickeln, für die Philosophische Psychologie interessant macht, liegt darin, daß diese Versuche vielfach zu überraschenden Erkenntnissen darüber geführt haben, worin diese Fähigkeiten eigentlich bestehen. Dies gilt für den Bereich der natürlichsprachlichen Systeme ebenso wie z.B. auch für den Bereich *computer vision*. In diesem letzten Bereich sind die Analysen dessen, was ein System alles können muß, um sehen, d.h. auf dem Weg über visuelle Sensoren Informationen über seine Umwelt gewinnen zu können, sogar besonders eindrucksvoll. Unabhängig von ihren konstruktiven Erfolgen kann die KI-Forschung somit der Philosophischen Psychologie helfen, nicht nur den Begriff des Mentalen im allgemeinen, sondern auch die einzelnen mentalen Fähigkeiten, die unter diesen Begriff fallen, klarer als bisher zu erfassen, indem sie in ihren Analysen dieser Fähigkeiten z.B. deutlicher werden läßt, was es eigentlich heißt, sprechen oder sehen zu können.

Fünftens: Ein weiterer Punkt, der für die Philosophische Psychologie von besonderem Interesse ist, liegt schließlich im generellen Ansatz der KI-Forschung. Denn dieser Ansatz ist durch die Annahme gekennzeichnet, daß intelligente oder mentale Leistungen von *Computern* erbracht werden können, d.h. von *Systemen, die ihrer Natur nach Systeme der algorithmischen Musterverarbeitung sind*. Die Grundidee der KI-Forschung ist also, daß es möglich ist, Systeme zu intelligenten Leistungen zu befähigen, wenn man nur den sensorischen Input dieser Systeme in geeigneter Weise in Symbolmuster transformiert und diese Symbolmuster dann mit Hilfe geeigneter Algorithmen in andere Muster umformt, wobei schließlich die Muster, die sich als Ergebnisse dieser Umformungen ergeben, den motorischen Output der Systeme steuern. Es ist hier nicht möglich, diese Grundidee eingehend zu erläutern. Aber sie spielt in der KI-Forschung überall eine wichtige Rolle, z.B. auch beim Problem der Wissensrepräsentation, das in den letzten Jahren immer mehr zu einem, um nicht zu sagen, zu dem zentralen Problem der KI-Forschung geworden ist. Die Bedeutung dieses Problems ergibt sich aus der Erkenntnis, daß alle kognitiven und linguistischen Fähigkeiten große Mengen von Wissen voraussetzen, das in den entsprechenden Systemen in irgendeiner Form vorhanden, d.h. repräsentiert sein muß. Grundsätzlich kommen dabei jedoch nur zwei Formen der Repräsentation in Frage, die prozedurale und die deklarative Repräsentation von Wissen. Denn auch Wissen kann der Grundidee der KI-Forschung zufolge nur in geeignet strukturierten Mustern oder in entsprechenden Algorithmen zur Veränderung von Mustern gespeichert oder repräsentiert sein.

Die Idee, daß etwa Wissen über die Umwelt in einem intelligenten System in strukturierten Mustern oder in musterverarbeitenden Algorithmen repräsentiert sein muß, hat unter dem Stichwort *Mentale Repräsentationen* inzwischen auch

die neuesten Diskussionen in der Philosophischen Psychologie sehr stark beeinflusst. Denn insbesondere die im Augenblick sehr einflußreiche *Repräsentationale Theorie des Geistes* von J. Fodor (1981), derzufolge intentionale Zustände wie Glauben, Wünschen, Hoffen usw. (also zumindest eine sehr wichtige Klasse von mentalen Zuständen) als Relationen zwischen einem System und bestimmten propositionalen Repräsentationen, d. h. genauer bestimmten Sätzen einer internen Sprache dieses Systems aufgefaßt werden müssen, geht ganz wesentlich auf diese Idee zurück.

Ich kann leider auch auf diesen Punkt nicht ausführlicher eingehen. Aber ich möchte zum Schluß dieser Bemerkungen doch hervorheben, daß die zuletzt angesprochene Überlegung gerade auch für die Neurobiologie sehr interessant ist – zumindest soweit sie sich mit der Frage beschäftigt, was das Gehirn bzw. das ganze ZNS zu den kognitiven Leistungen des Menschen beiträgt. Denn wenn Fodor (1981b) mit seiner Bemerkung Recht hat: „... die Ergebnisse [der gegenwärtigen Arbeit in den kognitiven Wissenschaften, d. V.] lassen vermuten, daß das Konzept der mentalen Repräsentation für empirische Theorien des Geistes grundlegend ist“ (S. 37), dann ergibt sich daraus die interessante neurobiologische Frage, welche neuronalen Strukturen im Gehirn möglicherweise als Repräsentationen von Wissen in Frage kommen und mit Hilfe welcher Prozesse diese Strukturen neuronal verarbeitet werden. Die aus der KI-Forschung stammenden Ideen zur Wissensrepräsentation können somit für die Neurobiologie insofern interessant sein, als sie einen Hinweis darauf geben, wonach man im Gehirn denn überhaupt suchen soll, wenn es darum geht herauszufinden, was dieses Organ zum menschlichen Geist beiträgt.

## Literatur

- Beckermann A (1986) Descartes' metaphysischer Beweis für den Dualismus. Albert, Freiburg München
- Beckermann A (1988) Sprachverstehende Maschinen. Erkenntnis 28:65–85
- Bobrow D et al. (1977) GUS – a frame driven dialog system, in: Artif Intell 8:155–173
- Boden M (1977) Artificial intelligence and natural man. Harvester, Hassocks
- Descartes R (1960) Discourse de la méthode, franz.-deutsch. Meiner, Hamburg
- Descartes R (1969) Über den Menschen. Schneider, Heidelberg
- Fodor J (1981a) RePresentations. MIT Press, Cambridge, Mass
- Fodor J (1981b) Das Leib-Seele-Problem. Spektrum der Wissenschaft, März, 26–37
- Minsky M (1968) Preface, in: Minsky M (ed) Semantic information processing. MIT Press, Cambridge, Mass
- Rothschuh KE (1969) Einleitung. In: Rothschuh KE (Hrsg) Descartes. Über den Menschen. Schneider, Heidelberg
- Searle J (1980) Minds, brains, and programs. Behav Brain Sci 3:417–424, 450–456
- Searle J (1986) Geist, Hirn und Wissenschaft. Suhrkamp, Frankfurt

- Wahlster W (1982) Natürlichsprachliche Systeme – Eine Einführung in die sprachorientierte KI-Forschung. In: Bibel W, Siekmann HJ (Hrsg) Künstliche Intelligenz. Springer, New York Heidelberg Berlin (Informatik-Fachberichte 59)
- Weizenbaum J (1977) Die Macht der Computer und die Ohnmacht der Vernunft. Suhrkamp, Frankfurt
- Wilensky R (o.J.) Understanding goal-based stories. PhD Dissertation, Yale University (Computer Science Department, Rep 140)
- Winograd T (1984) Software für Sprachverarbeitung. Spektrum der Wissenschaft, November, 88–102