

JOHN R. SEARLE

Geist, Gehirn, Programm

Welche psychologische und philosophische Bedeutung sollen wir den in jüngster Zeit unternommenen Versuchen zur Computersimulation menschlicher kognitiver Fähigkeiten beimessen? Bei der Beantwortung dieser Frage finde ich es nützlich, zwischen zwei Formen von artifizieller Intelligenz zu unterscheiden, zwischen einer, wie ich es nennen werde, »starken« und einer »schwachen« oder »behutsamen« AI (artifiziiellen Intelligenz).¹ Der schwachen AI zufolge besteht der Hauptwert, den der Computer für die Untersuchung des Geistes hat, darin, daß er uns ein sehr wirksames Instrument an die Hand gibt. Er ermöglicht uns zum Beispiel eine strengere und präzisere Formulierung und Überprüfung von Hypothesen. Nach der starken AI aber ist der Computer nicht nur ein Instrument bei der Untersuchung des Geistes; vielmehr ist der recht programmierte Computer in Wahrheit selbst ein Geist in dem Sinn, daß man Computern, die mit den richtigen Programmen ausgestattet sind, buchstäblich Verstehen und andere kognitive Zustände zusprechen kann. Weil bei der starken AI der programmierte Computer kognitive Zustände aufweist, sind dort die Programme nicht bloß ein Instrumentarium, das uns die Überprüfung psychologischer Erklärungen erlaubt; vielmehr stellen die Programme selber Erklärungen dar.

Gegen die Ansprüche der schwachen AI habe ich nichts einzuwenden, jedenfalls was den vorliegenden Aufsatz betrifft. Meine Erörterung hier richtet sich gegen jene Ansprüche, die ich als die der starken AI bestimmt habe, speziell gegen den Anspruch, daß der recht programmierte Computer buchstäblich kognitive Zustände aufweise und daß die Programme von daher menschliches Erkennen erklärten. Wenn ich in der Folge von AI spreche, meine ich die starke Version, wie sie in diesen beiden Ansprüchen zum Ausdruck kommt.

Meine Überlegungen werden an die Arbeit von Roger Schank und seinen Kollegen an der Yale Universität (Schank/Abelson, 1977)² anknüpfen, weil ich mit ihr vertrauter bin als mit anderen ähnlichen Vorhaben und weil sie ein sehr prägnantes Beispiel für die Art Arbeiten liefert, die ich untersuchen möchte. Aber keine der Folgerungen, die ich ziehe, ist abhängig von den Besonderheiten der Programme Schanks. Die gleichen Überlegungen ließen sich auf Winograds SHRDLU (Winograd, 1972), Weizenbaums ELIZA (Weizenbaum, 1978) und in der Tat auf jede Maschinensimulation menschlicher geistiger Phänomene à la Turing anwenden.

In aller Kürze und unter Verzicht auf die vielfältigen Einzelheiten läßt sich Schanks Programm folgendermaßen beschreiben: Ziel des Programms ist die Simulation der menschlichen Fähigkeit, Geschichten zu verstehen. Für die Fähigkeit menschlicher Wesen, Geschichten zu verstehen, ist kennzeichnend, daß sie imstande sind, Fragen zu einer Geschichte zu beantworten, auch wenn die Auskunft, die sie geben, gar nicht expliziter Inhalt der betreffenden Geschichte war. Nehmen wir also zum Beispiel an, man erzählt uns die folgende Geschichte: »Ein Mann ging in ein Restaurant und bestellte einen Hamburger. Als der Hamburger kam, war er total verbrannt und nur noch ein zusammengeschmortes Krüstchen, und der Mann stürmte wütend aus dem Lokal, ohne zu bezahlen bzw. ein Trinkgeld dazulassen.« Wenn uns nun jemand fragt: »Hat der Mann den Hamburger gegessen?«, werden wir voraussichtlich antworten: »Nein, hat er nicht.« Desgleichen, wenn man uns die folgende Geschichte erzählt: »Ein Mann ging in ein Restaurant und bestellte einen Hamburger; als der Hamburger kam, gefiel er dem Mann sehr gut; und als der Mann das Lokal verließ, gab er der Serviererin ein dickes Trinkgeld, ehe er die Rechnung bezahlte.« Wenn man uns fragt: »Hat der Mann den Hamburger gegessen?«, werden wir voraussichtlich antworten: »Ja, er hat ihn gegessen.« Nun können Schanks Maschinen gleichfalls Fragen über Restaurants in dieser Manier beantworten. Um dies zu können, ha-

ben sie eine »Repräsentation« der Art von Informationen über Restaurants, über die menschliche Wesen verfügen und durch die sie instande sind, Fragen wie die oben zitierten zu beantworten, wenn Geschichten der behandelten Art erzählt werden. Wenn die Maschine die Geschichte erzählt und dann die Frage gestellt bekommt, druckt sie Antworten von der Art aus, wie wir sie von menschlichen Wesen erwarten würden, denen vergleichbare Geschichten erzählt werden. Verfechter der starken AI behaupten, daß bei dieser Frage-und-Antwort-Folge die Maschine nicht bloß eine menschliche Fähigkeit simuliert, sondern mehr noch: daß (1) von der Maschine sich buchstäblich sagen läßt, daß sie die Geschichte *versteht* und Antwort gibt auf Fragen und daß (2) das, was die Maschine und ihr Programm tun, eine *Erklärung* der menschlichen Fähigkeit darstellt, eine Geschichte zu verstehen und Fragen darüber zu beantworten.

Für beide Behauptungen scheint mir Schanks Arbeit nicht die mindeste Rechtfertigung zu bieten, wie ich im folgenden deutlich zu machen versuchen werde. Ich will damit selbstverständlich nicht sagen, daß Schank selbst diese Behauptungen vertritt.

Eine Möglichkeit der Überprüfung irgendeiner bestimmten Geisttheorie besteht darin, daß man sich fragt, wie es wäre, wenn der eigene Geist tatsächlich auf Grundlage der Prinzipien funktionierte, die die betreffende Theorie für alle Geister voraussetzt. Wenden wir diese Prüfmethode auf das Schank-Programm an, indem wir folgendes Gedankenexperiment anstellen. Nehmen wir an, ich bin in einem Raum eingeschlossen, und man gibt mir einen Packen mit chinesischer Schrift. Nehmen wir weiter an, daß ich (was in der Tat der Fall ist) kein Chinesisch kann, es weder schreiben noch sprechen kann, und daß ich nicht einmal sicher bin, ob ich chinesische Schrift als chinesische Schrift erkennen und von, sagen wir, japanischer Schrift oder sinnlosem Gekritzeln unterscheiden könnte. Chinesische Schrift besteht für mich einfach nur aus sinnlosem Gekritzeln. Nehmen wir nun weiterhin an, daß man

mir nach dem ersten Packen mit chinesischer Schrift einen zweiten Packen mit chinesischen Schriftzeichen gibt, zusammen mit einer Reihe von Anleitungen, wie ich den zweiten Stoß zum ersten in Beziehung setzen soll. Die Anleitungen sind in Englisch abgefaßt, und ich verstehe diese Anleitungen ebenso gut wie jeder andere, dessen Muttersprache Englisch ist. Sie ermöglichen es mir, eine Reihe formaler Symbole in Beziehung zu einer anderen Reihe formaler Symbole zu setzen, und »formal« bedeutet hier nichts weiter, als daß ich diese Symbole ausschließlich an Hand ihrer Form identifiziere. Nehmen wir nun auch noch an, man gibt mir einen dritten Packen chinesischer Symbole, zusammen mit einigen Anweisungen, ebenfalls in Englisch, die es mir ermöglichen, Teile dieses dritten Packens in Beziehung zu setzen zu den zwei ersten Packen; und diese Anleitungen weisen mich an, bestimmte Symbole mit bestimmten Formen in Antwort auf bestimmte Formen, die mir mit dem dritten Packen zugegangen sind, zurückzugeben. Was ich nicht weiß, ist, daß die Leute, die mir all diese Symbole geben, den ersten Packen eine »Schrift«, den zweiten Packen eine »Geschichte« und den dritten Packen »Fragen« nennen. Des weiteren nennen sie die Symbole, die ich ihnen in Antwort auf den dritten Packen zurückgebe, »Antworten auf die Fragen«, und die Reihe von englischsprachigen Anleitungen, die sie mir geben, nennen sie »Programm«. Nun stellen wir uns, einfach um die Sache noch ein bißchen zu komplizieren, vor, daß diese Leute mir auch Geschichten in Englisch geben, die ich verstehe, und daß sie mir dann in Englisch Fragen zu diesen Geschichten stellen und ich ihnen in Englisch antworte. Nehmen wir auch an, daß ich nach einer gewissen Zeit den Anweisungen für das Handtieren mit den chinesischen Symbolen so gut zu folgen lerne und die Programmierer so gut lernen, Programme zu schreiben, daß, von außen betrachtet – d. h. vom Standpunkt eines Menschen aus, der sich außerhalb des Raums befindet, in dem ich eingeschlossen bin –, meine Antworten auf die Fragen absolut ununterscheidbar sind von denen, die einer geben wür-

de, dessen Muttersprache Chinesisch ist. Niemand, der nur meine Antworten sieht, kann erkennen, daß ich kein Wort Chinesisch spreche. Nehmen wir auch an, daß meine Antworten auf die englischen Fragen ununterscheidbar sind von denen, die andere geben würden, deren Muttersprache Englisch ist, was ohne Zweifel der Fall sein wird, einfach deshalb, weil Englisch auch meine Muttersprache ist. Von außen gesehen – vom Standpunkt dessen betrachtet, der meine »Antworten« liest – sind die Antworten auf die chinesischen Fragen und auf die englischen Fragen gleich gut. Aber im Fall des Chinesischen bringe ich, anders als im Englischen, die Antworten dadurch hervor, daß ich mit unverständlichen formalen Symbolen hantiere. Soweit es das Chinesische betrifft, verhalte ich mich einfach wie ein Computer; ich führe kalkulatorische Operationen an formal spezifizierten Elementen aus. In bezug auf das Chinesische bin ich einfach ein verkörpertes Computerprogramm.

Nun behauptet die starke AI, daß der programmierte Computer die Geschichten versteht und daß das Programm in gewissem Sinn menschliches Verstehen erklärt. Wir sind aber jetzt imstande, diese Behauptungen im Licht unseres Gedankenexperiments zu überprüfen.

1. Was die erste Behauptung angeht, so scheint es mir bei unserem Beispiel ganz klar auf der Hand zu liegen, daß ich kein Wort von den chinesischen Geschichten verstehe. Ich habe Eingaben und Ausgaben, die von denen eines Chinesisch als Muttersprache Sprechenden nicht zu unterscheiden sind, und ich mag über jedes beliebige formale Programm verfügen, aber trotzdem verstehe ich nichts. Aus eben diesem Grund versteht auch Schanks Computer nichts von den Geschichten, seien diese nun in Chinesisch, Englisch oder welcher Sprache auch immer abgefaßt; denn im Fall des Chinesischen bin ja ich der Computer, und in den Fällen, wo der Computer ein Computer ist, hat er nichts weiter, als ich habe, wenn ich nichts verstehe.

2. Was die zweite Behauptung angeht, daß das Programm menschliches Verstehen erklärt, so wird deutlich, daß der

Computer und sein Programm fürs Verstehen keine zureichenden Bedingungen schaffen, da ja der Computer und das Programm funktionieren, ohne daß es zu einem Verstehen kommt. Aber schafft das Programm auch nur eine notwendige Voraussetzung fürs Verstehen, oder liefert es wenigstens einen wesentlichen Beitrag dazu? Eine der von den Parteigängern der starken AI aufgestellten Behauptungen ist, daß ich beim Verstehen einer englischsprachigen Geschichte ebendas tue – oder vielleicht ebendas noch mehr tue –, was ich beim Hantieren mit chinesischen Symbolen getan habe. Was den Fall des Englischen, wo ich verstehe, vom Fall des Chinesischen, wo ich nicht verstehe, unterscheidet, ist einfach ein Mehr an formalem Hantieren mit Symbolen. Ich habe diese Behauptung nicht widerlegt, aber daß sie im Blick auf unser Beispiel aberwitzig anmutet, steht außer Zweifel. Was die Behauptung an Plausibilität aufzuweisen hat, schöpft sie aus der Annahme, daß es möglich ist, ein Programm zu entwerfen, das dieselben Eingaben und Ausgaben hat wie ein natürlicher Sprecher, und darüber hinaus wird davon ausgegangen, daß auf einer bestimmten Darstellungsebene auch ein natürlicher Sprecher Verkörperung eines Programms ist. Auf der Basis dieser beiden Voraussetzungen wird dann angenommen, daß Schanks Programm, auch wenn es nicht die ganze Wahrheit übers Verstehen ist, doch aber möglicherweise ein Teil dieser Wahrheit ist. Nun, ich vermute, das ist empirisch denkbar, aber bislang gibt es nicht den mindesten Grund zu glauben, daß es sich tatsächlich so verhält, da ja im Gegenteil unser Beispiel darauf hindeutet – wenn auch zweifellos nicht den Beweis dafür erbringt –, daß das Computerprogramm für mein Verständnis der Geschichte einfach belanglos ist. Im Fall des Chinesischen habe ich alles, was artifizielle Intelligenz mittels eines Programms mir zu geben vermag, und ich verstehe nichts; im Fall des Englischen verstehe ich alles, und es gibt bislang keinen Grund, anzunehmen, daß mein Verständnis das mindeste mit Computerprogrammen zu tun hat, will heißen, mit kalkulatorischen Operationen an rein formal spezifizierten Elementen. Solange das Programm sich

als eine Reihe von kalkulatorischen Operationen an rein formal definierten Elementen bestimmt, legt unser Beispiel den Schluß nahe, daß diese Operationen selbst keine interessensbeziehende Beziehung zum Verstehen haben. Ganz gewiß sind sie keine zureichenden Bedingungen, und es gibt nicht den mindesten Grund, anzunehmen, daß sie eine notwendige Voraussetzung fürs Verstehen sind oder auch nur einen wesentlichen Beitrag dazu liefern. Man beachte, daß unser Argument seine Kraft nicht einfach aus der Überlegung zieht, daß unterschiedliche Maschinen dieselben Ein- und Ausgaben haben können, während sie auf der Grundlage unterschiedlicher formaler Prinzipien funktionieren – das ist hier ganz und gar nicht der Punkt. Der Punkt ist vielmehr, daß, gleichgültig, welche rein formalen Prinzipien man einem Computer eingibt, diese fürs Verstehen nicht ausreichend sind, insofern als ein Mensch imstande ist, eben diese formalen Prinzipien zu befolgen, ohne dabei das mindeste zu verstehen. Keinerlei Grund liegt vor, anzunehmen, daß solche Prinzipien notwendig fürs Verstehen sind oder wenigstens zu ihm beitragen, da ja auch kein Grund zu der Annahme besteht, daß ich, wenn ich Englisch verstehe, überhaupt mit irgendeinem formalen Programm operiere. Was habe ich nun also bei den englischen Sätzen, was ich bei den chinesischen Sätzen nicht habe? Die auf der Hand liegende Antwort lautet, daß ich bei ersteren weiß, was sie bedeuten, während ich bei letzteren nicht die geringste Ahnung habe, was sie bedeuten. Aber was beinhaltet das, und warum sollen wir es, worum auch immer es sich dabei dreht, nicht einer Maschine mitgeben können? Ich werde auf diese Frage später zurückkommen, will aber erst mit dem Beispiel fortfahren. Ich habe Gelegenheit gehabt, dieses Beispiel verschiedenen an artifizieller Intelligenz arbeitenden Wissenschaftlern vorzulegen, und interessanterweise scheinen sie unterschiedlicher Ansicht, was darauf die richtige Antwort ist. Man erhält eine überraschende Vielzahl von Antworten, und im folgenden werde ich mich mit den gängigsten unter ihnen beschäftigen (einzeln aufgeführt am Leitfaden ihrer geographischen Herkunft).

Aber vorher will ich noch einigen verbreiteten Mißverständnissen bezüglich »Verstehen« entgegentreten: In vielen dieser Diskussionen strampeln sich die Betroffenen wegen des Wortes »Verstehen« unglaublich ab. Meine Kritiker weisen darauf hin, daß es viele verschiedene Grade des Verstehens gibt; daß »Verstehen« nicht ein einfaches transitives Verb ist; daß es sogar verschiedene Arten und Ebenen von Verstehen gibt und häufig nicht einmal der Satz vom ausgeschlossenen Dritten eindeutig auf Feststellungen anwendbar ist, die die Form »x versteht y« haben; daß es in vielen Fällen nicht eine einfache Tatsachen-, sondern eine Ermessensfrage ist, ob y von x verstanden wird; und so weiter. Zu all diesen Bemerkungen möchte ich sagen: Na klar, na klar; aber all das hat nichts mit dem zu tun, worum es hier geht. Es gibt klare Fälle, wo »Verstehen« im Sinne des Wortes statthat, und ebenso klare Fälle, wo es nicht statthat; und nichts sonst als diese zwei Sorten Fälle brauche ich im Zusammenhang dieser Argumentation.³ Ich verstehe Geschichten auf englisch; weniger gut kann ich Geschichten auf französisch verstehen; noch weniger gut Geschichten auf deutsch; und auf chinesisch gar nicht. Mein Auto und meine Rechenmaschine andererseits verstehen überhaupt nichts: Das ist nicht ihr Gebiet. Wir legen häufig Autos, Rechenmaschinen und sonstigen Kunsterzeugnissen metaphorisch bzw. analogisierend »Verstehen« oder andere kognitive Prädikate bei, aber solche Zuschreibungen beweisen nichts. Wir sagen: »Die Tür weiß, wann sie aufgehen muß, dank ihrer photoelektrischen Zelle«; »Die Rechenmaschine kann addieren und subtrahieren (versteht sich darauf, ist imstande dazu), aber nicht dividieren«; und »Der Thermostat nimmt Temperaturen wahr«. Der Grund, warum wir diese Zuschreibungen vornehmen, ist durchaus interessant, er hat mit dem Umstand zu tun, daß wir auf Kunsterzeugnisse unsere eigene Intentionalität⁴ übertragen; unsere Werkzeuge sind der verlängerte Arm unserer Zielsetzungen, und deshalb finden wir es natürlich, ihnen metaphorisch Intentionalität zuzuschreiben; aber philosophisch, möchte ich meinen, ist

dergleichen ohne Belang. Der Sinn, in dem eine automatische Tür »Instruktionen versteht«, die von ihrer photoelektrischen Zelle ausgehen, ist mitnichten der Sinn, in dem ich Englisch verstehe. Wenn der Sinn, in dem Schanks programmierte Computer Geschichten verstehen, dem metaphorischen Sinn gleich sein soll, in dem Türen Verständnis beweisen, und nicht gleich dem Sinn, in dem ich Englisch verstehe, könnte man sich die Diskussion sparen. Aber Newell und Simon (1972) schreiben, daß die Art kognitiver Fähigkeit, die nach ihrer Behauptung Computer haben, genau dieselbe wie die menschlicher Wesen ist. Mir gefällt die Offenheit dieses Anspruchs, und mit ihm und seinesgleichen setze ich mich auseinander. Ich vertrete den Standpunkt, daß im wortwörtlichen Sinn der programmierte Computer das versteht, was das Auto und die Rechenmaschine verstehen, nämlich haargenau nichts. Das Verstehen des Computers ist nicht bloß (wie mein Verständnis des Deutschen) bruchstückhaft oder unvollständig; es ist gleich Null.

Nun zu den Repliken:

1. *Die System-Replik (Berkeley)*. »Es ist zwar richtig, daß die einzelne Person, die in dem Raum eingeschlossen ist, die Geschichte nicht versteht, aber Tatsache ist, daß sie nur Teil eines ganzen Systems ist, und das System versteht die Geschichte durchaus. Die betreffende Person hat ein Hauptbuch vor sich, worin die Anleitungen niedergelegt sind, sie hat eine Menge Schmierpapier und Bleistifte, um die Kalkulationen auszuführen, sie hat »Datenbanken« in Form von Garnituren chinesischer Symbole. Verstehen kommt nun nicht dem bloßen einzelnen zu; vielmehr kommt es diesem ganzen System zu, von dem der einzelne ein Teil ist.«

Meine Antwort auf die Systemtheorie ist ganz einfach: Der einzelne möge sämtliche Elemente des Systems sich einverleiben. Er lernt die Anleitungen im Hauptbuch und die Datenbanken aus chinesischen Symbolen auswendig, und er führt alle Kalkulationen im Kopf aus. Er verkörpert nun das ganze

System. Es gibt nichts am System, was er nicht umgreift. Wir können sogar auf den Raum verzichten und den Betreffenden im Freien arbeiten lassen. Nichtsdestoweniger versteht er nichts vom Chinesischen, und das gilt dann um so mehr für das System, da ja das System nichts enthält, was er nicht enthält. Wenn er nicht verstehen kann, dann gibt es auch keinerlei Verstehensmöglichkeit für das System, da das System einfach ein Teil von ihm ist.

Tatsächlich bereitet es mir einige Schwierigkeit, mich auch nur zu dieser Antwort auf die Systemtheorie zu bequemen, weil die Theorie mir, um das gleich zu sagen, so außerordentlich wenig einleuchtet. Der Grundgedanke ist, daß zwar die betreffende Person kein Chinesisch versteht, daß aber irgendwie die *Kombination* dieser Person mit ein bißchen Papier Chinesisch versteht. Ich habe Mühe, mir vorzustellen, wie für jemanden, der nicht ideologisch verblendet ist, dieser Gedanke die mindeste Plausibilität haben kann. Nichtsdestoweniger glaube ich, daß viele, die der Ideologie der starken AI anhängen, im Endeffekt geneigt sein werden, etwas zu vertreten, das dem nahekommt; verfolgen wir die Sache also ein wenig weiter. Einer Version dieser Ansicht zufolge versteht der Mann in dem Beispiel mit dem internalisierten System zwar Chinesisch nicht wie jemand, der es als Muttersprache spricht (weil er, zum Beispiel, nicht weiß, daß die Geschichte von Restaurants und Hamburgern usw. handelt), nichtsdestoweniger aber verfügt »er als ein System des formalen Umgangs mit Symbolen« über ein *wirkliches Verständnis des Chinesischen*. Dasjenige Subsystem des Betreffenden, das das formale Symbolhandhabungssystem fürs Chinesische darstellt, darf dabei nicht mit dem fürs Englische zuständigen verwechselt werden.

Demnach existieren in Wahrheit zwei Subsysteme in dem Betreffenden; das eine versteht Englisch, das andere Chinesisch, und »die beiden Systeme haben schlicht und einfach nicht viel miteinander zu tun«. Aber sie haben, möchte ich antworten, nicht nur nicht viel miteinander zu tun, sie sind sich auch nicht

einmal im entferntesten ähnlich. Das Subsystem, das Englisch versteht (wenn wir so frei sein dürfen, einen Augenblick lang in diesem »Subsystem«-Jargon zu reden), weiß, daß die Geschichte von Restaurants und dem Essen von Hamburgern handelt, daß man ihm Fragen über Restaurants stellt und daß es nach bestem Wissen antwortet, indem es aus dem Inhalt der Geschichte Rückschlüsse zieht, und so weiter. Aber das chinesische System weiß von alledem nichts. Während das englische Subsystem weiß, daß mit »Hamburger« Hamburger gemeint sind, weiß das chinesische System nur, daß auf »kritzel-kritzel« »kratzel-kratzel« folgt. Alles, was es weiß, ist, daß verschiedene formale Symbole an einem Ende eingeführt, entsprechend den in Englisch abgefaßten Anleitungen behandelt und am anderen Ende wieder ausgestoßen werden. Bei meinem ursprünglichen Beispiel ging es mir um nichts anderes als um die Feststellung, daß solcher Umgang mit Symbolen, für sich genommen, nicht ausreicht, um in irgendeinem wörtlichen Sinn Chinesisch zu verstehen, da ja der Betreffende auf »kritzel-kritzel« »kratzel-kratzel« hinschreiben könnte, ohne das mindeste vom Chinesischen zu verstehen. Und dieser Feststellung wird nicht dadurch begegnet, daß man in dem betreffenden Subsysteme annimmt; denn die Subsysteme sind nicht besser dran, als es zuerst der Betreffende war; sie haben noch immer nichts, das dem entfernt gleich wäre, was der Englischsprachige (bzw. das entsprechende Subsystem) hat. Tatsächlich ist in dem von uns beschriebenen Fall das chinesische Subsystem schlicht Teil des englischen Subsystems, ein Teil, dessen Beschäftigung das von Bedeutung freie Hantieren mit Symbolen gemäß englischsprachiger Anleitung ist.

Fragen wir uns, was eigentlich zur Systemreplik motivieren kann; will heißen: Welche *unparteiischen* Beweggründe kann es für die Behauptung geben, das betreffende Subjekt müsse ein Subsystem in sich enthalten, das Geschichten auf chinesisches im Wortsinn versteht? Soweit ich sehe, sind die einzigen Gründe die, daß in dem gegebenen Beispiel ich dieselben Ein- und Ausgaben habe wie jemand, der Chinesisch als Mutter-

sprache spricht, sowie ein Programm, das das eine mit dem anderen verbindet. Aber bei unseren Beispielen ging es ja gerade um den Versuch, zu zeigen, daß das für ein Verstehen in dem Sinn, wie ich Geschichten auf englisch verstehe, nicht ausreichend sein kann, da ja eine einzelne Person, und mithin auch die Reihe von Systemen, aus denen die einzelne Person sich am Ende zusammensetzt, die richtige Kombination aus Eingabe, Ausgabe und Programm haben kann, ohne doch aber in dem einschlägigen buchstäblichen Sinn, in dem ich Englisch verstehe, das mindeste zu verstehen. Der einzige Grund für die Behauptung, es *müsse* in mir ein Subsystem geben, das Chinesisch versteht, ist der Umstand, daß ich ein Programm habe und imstande bin, den Turing-Test zu bestehen; ich kann diejenigen, die Chinesisch als Muttersprache sprechen, hinters Licht führen. Aber die Angemessenheit des Turing-Tests ist genau eine der strittigen Fragen, um die es hier geht. Das Beispiel zeigt, daß es theoretisch zwei »Systeme« gibt, die beide den Turing-Test bestehen können, von denen aber nur das eine zu verstehen vermag; und dagegen läßt sich nun nicht etwa geltend machen, daß, da ja beide den Turing-Test bestehen, sie auch beide notwendig Verstehen beweisen; denn diese Behauptung wird dem Einwand nicht gerecht, daß das System in mir, das Englisch versteht, eine ganze Menge mehr beinhaltet als das System, das Chinesisch einfach nur einem Datenverarbeitungsverfahren unterwirft. Kurz, die Systemreplik bleibt den Beweis einfach schuldig und insistiert ohne Begründung darauf, daß das System ein Verständnis des Chinesischen haben müsse.

Darüber hinaus hat es den Anschein, als müsse die Systemreplik zu Konsequenzen führen, die auch für sich genommen absurd sind. Wenn wir bloß auf Grund des Vorhandenseins einer bestimmten Art von Ein- und Ausgaben, mit einem Programm dazwischen, auf Erkenntnisfähigkeit in uns schließen müssen, dann, so scheint es, erweisen sich alle möglichen nicht-kognitiven Subsysteme als kognitiv. Zum Beispiel gibt es eine Darstellungsebene, auf der mein Magen Datenverar-

beitung betreibt und jede Menge Computerprogramme realisiert, aber ich nehme doch an, wir wollen ihm nicht Verstehen konzedieren (vgl. Pylyshyn, 1985). Aber wenn wir die Systemreplik akzeptieren, weiß ich nicht, wie wir der Konsequenz entgehen sollen zu sagen, daß Magen, Herz, Leber usw. allesamt mit Verstehen begabte Subsysteme sind, da es ja keine fundierte Möglichkeit gibt, den Beweggrund, der einen sagen läßt, das chinesische Subsystem denkt, von dem zu unterscheiden, der einen sagen läßt, der Magen denkt. Als Antwort hierauf genügt es übrigens nicht zu sagen, die Ein- und Ausgaben beim chinesischen System seien Informationen, während die beim Magen Essen und Nahrungsmittel seien, da ja vom Standpunkt der betreffenden Person aus, von meinem Standpunkt aus, weder im Essen noch im Chinesischen Information steckt – das Chinesische besteht einfach aus lauter Gekritzel. Die Information im Fall des Chinesischen existiert nur in den Köpfen der Programmierer und der Übersetzer, und es hindert sie nichts, die Ein- und Ausgaben meiner Verdauungsorgane als Information zu behandeln, wenn sie das wünschen. Dieser letztere Punkt ist auch für einige davon unabhängige Probleme im Zusammenhang mit der starken AI von Bedeutung, und es lohnt sich eine kurze Abschweifung, um das zu explizieren. Wenn die starke AI ein Zweig der Psychologie sein soll, dann muß sie imstande sein, genuin geistige Systeme von denen zu unterscheiden, die das nicht sind. Sie muß imstande sein, die Prinzipien, auf deren Grundlage der Geist funktioniert, von denen zu unterscheiden, auf deren Basis nicht-geistige Systeme funktionieren; sonst wird sie uns über das spezifisch Geistige am Geistigen keinen Aufschluß geben können. Und der Unterschied zwischen »geistig« und »nicht-geistig« kann nicht bloß im Kopf des Betrachters existieren, sondern muß den Systemen selbst inhärent sein; sonst stünde es dem Betrachter frei, Menschen als nicht-geistig und, beispielsweise, Wirbelstürme als geistig anzusehen, wenn ihm das Spaß machte. Aber ganz oft ist in der Literatur über AI der Unterschied dermaßen verwischt, daß dies auf lange Sicht

dem Anspruch der AI, eine Untersuchung des Erkenntnisvermögens zu sein, verderblich werden dürfte. McCarthy z.B. schreibt: »Von Maschinen, die so einfach sind wie Thermostaten, kann man sagen, daß sie Überzeugungen haben, und Überzeugungen zu haben scheint eines der Merkmale der meisten Maschinen zu sein, die in der Lage sind, Probleme zu lösen« [im vorliegenden Band S.185]. Jeder, der meint, die starke AI habe Aussicht, sich als Geisttheorie zu bewähren, sollte sich über die Implikationen dieser Bemerkung Gedanken machen. Wir sind aufgefordert, es als eine Entdeckung der starken AI zu akzeptieren, daß das Stück Metall da an der Wand, das wir für die Regulierung der Zimmertemperatur brauchen, in genau demselben Sinn Überzeugungen hat, wie wir, unsere Ehepartner und unsere Kinder Überzeugungen haben, und daß mehr noch die »meisten« der anderen Maschinen im Zimmer – Telefon, Tonband, Rechenmaschine, Lichtschalter – gleichfalls in diesem buchstäblichen Sinn Überzeugungen haben. Dieser Artikel verfolgt nicht das Ziel, sich mit McCartneys These auseinanderzusetzen, ich will also nur ohne lange Erörterung das Folgende feststellen. Die Erforschung des Geistes fängt mit Tatsachen an wie der, daß Menschen Überzeugungen haben, Thermostaten, Telefone und Rechenmaschinen hingegen nicht. Wenn sich eine Theorie findet, die das widerlegt, ist ein Gegenbeweis gegen die erste Theorie erbracht, und diese ist folglich falsch. Man gewinnt den Eindruck, daß diejenigen von der AI, die solches Zeug schreiben, denken, sie kommen damit durch, weil sie es selber gar nicht wirklich so meinen und sich nicht vorstellen können, daß irgend jemand sonst das tut. Ich schlage vor, daß wir, für einen Augenblick jedenfalls, die Sache ernst nehmen. Machen wir uns eine Minute lang mit aller Anstrengung einmal Gedanken darüber, was alles für den Nachweis nötig wäre, daß das Stück Metall da drüben an der Wand wirkliche Überzeugungen hat. Überzeugungen mit Bezugspunkt, propositionalem Inhalt und Erfüllungskriterien; Überzeugungen, denen die Möglichkeit innewohnt, starke oder schwache Überzeugungen zu

sein; nervöse, unsichere oder zuversichtliche Überzeugungen; dogmatische, rationale oder abergläubische Überzeugungen; blinden Glauben oder zweifelnde Gedanken; jede Art von Überzeugungen. Der Thermostat ist kein Anwärtler darauf. Und ebenso wenig sind es Magen, Leber, Rechenmaschine oder Telefon. Da wir aber ja den Gedanken ernst nehmen, müssen wir bemerken, daß seine Bewahrheitung für den Anspruch der starken AI, eine Wissenschaft vom Geist zu sein, fatale Folgen hätte. Denn jetzt wäre der Geist überall. Wir wollten doch wissen, was den Geist von Thermostaten und Lebern unterscheidet. Und falls McCarthy recht hat, muß die starke AI alle Hoffnung fahrenlassen, das für uns herauszufinden.

2. *Die Roboter-Replik (Yale).* »Angenommen, wir schrieben ein Programm, das anders geartet wäre als Schanks Programm. Angenommen, wir steckten einen Computer in einen Roboter, und dieser Computer würde nicht nur formale Symbole als Eingabe aufnehmen und formale Symbole als Ausgabe abgeben, sondern würde tatsächlich den Roboter in einer Weise funktionieren lassen, die Tätigkeiten wie Wahrnehmen, Gehen, Fortbewegen, Nägeleinschlagen, Essen, Trinken, und was auch immer, sehr nahe käme. Am Roboter würde beispielsweise eine Fernsehkamera angebracht sein, die ihn zum Sehen befähigte, er würde Arme und Beine haben, die ihn zum »Handeln« befähigten, und all dies würde von seinem Computer »Hirn« gesteuert. Anders als Schanks Computer würde ein solcher Roboter echtes Verstehen und andere geistige Zustände aufweisen.«

Als erstes muß einem an der Roboter-Replik auffallen, daß sie stillschweigend einräumt, daß Erkenntnis nicht bloß eine Sache des formalen Hantierens mit Symbolen ist; denn diese Replik fügt ja eine Reihe von ursächlichen Verbindungen mit der Außenwelt hinzu (vgl. Fodor, 1981). Unsere Antwort auf die Roboter-Replik allerdings ist, daß die Hinzufügung solcher »Wahrnehmungs«- und »Bewegungs«-Fähigkeiten in, spe-

ziell gesagt, verstehensmäßiger und, generell gesagt, intentionaler Hinsicht an Schanks ursprünglichem Programm nichts ändert. Das merkt man, wenn man sich klarmacht, daß unser früheres Gedankenexperiment auch auf den Roboter-Fall zutrifft. Nehmen wir an, daß statt des Computers ich im Roboter bin und man mir, so wie im ursprünglichen Fall mit dem Chinesisch, weitere chinesische Symbole mit weiteren englischsprachigen Anweisungen gibt, wie ich chinesisches Symbol mit chinesischem Symbol zusammenfügen und an die Außenwelt chinesische Symbole auswerfen soll. Nehmen wir an, ohne daß ich etwas davon weiß, haben manche der chinesischen Symbole, die durch eine am Roboter angebrachte Fernsehkamera zu mir gelangen, und andere chinesische Symbole, die ich auswerfe, die Funktion, im Roboter befindliche Antriebe zu veranlassen, Beine oder Arme des Roboters in Bewegung zu setzen. Es ist wichtig zu betonen, daß alles, was ich dabei tue, ein Hantieren mit formalen Symbolen ist: Von all dem anderen weiß ich nichts. Ich erhalte »Informationen« durch den »Wahrnehmungs«-Apparat des Roboters, und ich gebe »Anweisungen« an seinen Bewegungsapparat aus, ohne daß ich mir der einen oder der anderen Tatsache bewußt wäre. Ich bin der Homunkulus des Roboters, aber anders als die gewöhnlichen Homunkuli weiß ich nicht, was vorgeht. Ich verstehe nicht das geringste, nur eben die Anleitungen für das Hantieren mit den Symbolen. In diesem Fall möchte ich nun behaupten, daß der Roboter keinerlei intentionale Zustände kennt; er bewegt sich einfach infolge seiner elektrischen Schaltanlage und seines Programms. Und darüber hinaus kenne auch ich, indem ich das Programm verkörpere, keine intentionalen Zustände im hier relevanten Sinn. Ich tue nichts weiter, als formale Anweisungen für das Hantieren mit formalen Symbolen zu befolgen.

3. *Die Gehirnsimulator-Replik (Berkeley und MIT).* »Angenommen, wir entwerfen ein Programm, das nicht Informationen über die Welt nach Art der Informationen in Schanks

Vorlage repräsentiert, sondern das die tatsächliche Abfolge der Neuronenaktivität in den Synapsen des Gehirns von jemandem, dessen Muttersprache Chinesisch ist, simuliert, wenn dieser Geschichten auf chinesisch hört und beantwortet. Die Maschine nimmt als Eingabe chinesische Geschichten und Fragen über diese Geschichten auf, simuliert die formale Struktur wirklicher Chinesengehirne bei der Verarbeitung dieser Geschichten und wirft chinesische Antworten als Ausgabe aus. Wir können uns sogar vorstellen, daß die Maschine nicht bloß mit einer einzigen Programmserie arbeitet, sondern mit einer ganzen Reihe parallellaufender Programme, auf die Art und Weise, wie wahrscheinlich wirkliche menschliche Gehirne funktionieren, wenn sie eine natürliche Sprache verarbeiten. Nun würden wir doch in solch einem Fall sicher sagen müssen, daß die Maschine die Geschichten versteht; und wenn wir uns weigerten, das zu sagen, würden wir dann nicht auch bestreiten müssen, daß diejenigen, die Chinesisch als Muttersprache sprechen, die Geschichten verstehen? Was würde oder könnte auf der Synapsenebene das Programm des Computers vom Programm eines Chinesengehirns unterscheiden?«

Ehe ich auf diese Replik eingehe, möchte ich nur kurz anmerken, daß das für einen Parteigänger der artifiziellen Intelligenz (bzw. für einen Funktionalisten et cetera) eine seltsame Replik ist: Ich dachte, der ganze Witz bei der starken AI sei gerade, daß wir nicht zu wissen brauchen, wie das Gehirn funktioniert, um zu wissen, wie der Geist funktioniert. Die grundlegende Annahme sei, so hatte ich jedenfalls geglaubt, daß es eine Ebene gibt, auf der das Funktionieren des Geistes in nichts als der kalkulatorischen Verarbeitung formaler Elemente besteht und daß diese kalkulatorischen Verarbeitungsförmlichkeiten die Essenz des Geistigen ausmachen und sich in allen möglichen verschiedenen Gehirnprozessen realisieren lassen, ganz ebenso wie irgendein Computerprogramm in verschiedenen Computer-Hardwares realisierbar ist: Der Annahme der starken AI folgend, verhält sich der Geist zum Gehirn wie

das Programm zur Hardware, und auf diese Weise können wir den Geist verstehen, ohne Neurophysiologie betreiben zu müssen. Wenn wir erst wissen müßten, wie das Gehirn funktioniert, ehe wir uns mit der AI beschäftigen könnten, würden wir uns ihretwegen nicht den Kopf zerbrechen. Wie dem auch sei, selbst wenn man der Funktionsweise des Gehirns so außerordentlich nahe käme, würde auch das noch nicht genügen, Verstehen zu erzeugen. Um das deutlich zu machen, stellen wir uns statt eines einsprachigen Menschen, der in einem Raum Symbolkarten durchblättert, diesen Menschen vor, wie er ein ausgetüfteltes System von Wasserleitungen betätigt, die durch Ventile miteinander verbunden sind. Wenn der Betreffende die chinesischen Symbole empfängt, sieht er im englischsprachigen Programm nach, welche Verbindungsventile er auf- oder zudrehen soll. Jede Verbindungsstelle entspricht einer Synapse im Chinesengehirn, und das ganze System ist so aufgebaut, daß, nachdem der Betreffende alle erforderlichen Aktivierungen vorgenommen, will heißen, alle einschlägigen Hähne aufgedreht hat, die chinesischen Antworten am Ausgange des Rohrsystems ausgespuckt werden.

Wo ist nun in diesem System Verstehen? Das System empfängt Chinesisch als Eingabe; simuliert die formale Synapsenstruktur eines Chinesengehirns und produziert Chinesisch als Ausgabe. Aber zweifellos versteht der Mann kein Chinesisch, und ebensowenig tun das die Wasserleitungen, und falls wir versucht sein sollten, den in meinen Augen absurden Standpunkt einzunehmen, daß irgendwie die *Verbindung* von Mann und Wasserleitungen Verstehen beweist, brauchen wir uns nur daran zu erinnern, daß im Prinzip der Betreffende die formale Struktur der Wasserleitungen sich insgesamt einverleiben und all die »Neuronenaktivierungen« in der Vorstellung vornehmen kann. Das Problem beim Gehirnsimulator ist, daß er vom Gehirn die falschen Sachen simuliert. Solange er nur die formale Struktur der Neuronenaktivierungen in den Synapsen simuliert, hat er nicht das am Gehirn simuliert, worauf es ankommt, nämlich dessen kausale Eigenschaften,

seine Fähigkeit, intentionale Zustände hervorzurufen. Und daß die formalen Eigenschaften nicht ausreichen, um die kausalen Eigenschaften darzustellen, zeigt das Wasserleitungsbeispiel: Wir können uns sämtliche formalen Eigenschaften abgehoben vorstellen von den wesentlichen neurophysiologischen kausalen Eigenschaften.

4. *Die Kombinations-Replik (Berkeley und Stanford).* »Für sich genommen mag jede der drei vorangegangenen Repliken als Widerlegung des Gegenbeispiels mit dem chinesischen Zimmer nicht völlig überzeugen, aber wenn man alle drei zusammenfaßt, sind sie als Ganzes erheblich überzeugender und in der Tat schlagend. Man stelle sich einen Roboter vor mit einem hirnförmigen Computer in der Höhlung unter seinem Schädel, man stelle sich den Computer mit allen Synapsen eines menschlichen Gehirns programmiert vor, man stelle sich das ganze Verhalten des Roboters als ununterscheidbar vom menschlichen Verhalten vor, und nun denke man sich das Ganze als ein vereinigtes System und nicht einfach als einen Computer mit Eingaben und Ausgaben. In diesem Fall würden wir doch zweifellos dem System Intentionalität zuschreiben.«

Ich bin ganz und gar derselben Meinung, daß es uns in diesem Fall vernünftig und sogar zwingend vorkommen würde, von der Annahme einer dem Roboter eigenen Intentionalität auszugehen, solange wir nichts weiter über ihn wüßten. Tatsächlich kommt es hierbei nur auf Erscheinung und Verhalten an, und die anderen Elemente der Kombination spielen gar keine Rolle. Wenn wir einen Roboter bauen könnten, dessen Verhalten vom menschlichen Verhalten im großen und ganzen ununterscheidbar wäre, würden wir ihm, solange nicht irgendein Grund dagegenspräche, Intentionalität zuschreiben. Wir müßten dazu nicht erst Kenntnis davon haben, daß sein Computergehirn eine formale Entsprechung des menschlichen Gehirns darstellte.

Aber ich kann einfach nicht sehen, inwiefern damit den An-

sprüchen der AI geholfen ist, und zwar aus folgendem Grund: Der starken AI zufolge ist die Realisierung eines formalen Programms mit den richtigen Ein- und Ausgaben eine zureichende Bedingung, und in der Tat konstitutiv, für Intentionalität. Mit Newell zu sprechen, besteht das Wesen des Geistigen im Funktionieren eines physischen Symbolsystems. Aber wenn wir dem Roboter im vorliegenden Beispiel Intentionalität zuschreiben, hat das mit formalen Programmen nicht das geringste zu tun. Unsere Zuschreibung basiert einfach auf der Annahme, daß, wenn der Roboter hinlänglich wie wir aussieht und sich verhält, wir bis zum Beweis des Gegenteils davon ausgehen würden, er müsse den unseren vergleichbare geistige Zustände haben, die sein Verhalten bewirken und darin zum Ausdruck kommen, und müsse über einen inneren Mechanismus verfügen, der imstande ist, solche Zustände hervorzurufen. Wenn wir wüßten, wie sein Verhalten sich ohne solche Annahmen erklären ließe, würden wir ihm keine Intentionalität zuschreiben, zumal dann nicht, wenn wir wüßten, daß ihm ein formales Programm zugrunde liegt. Und da sind wir genau wieder am Punkt meiner früheren Antwort auf Entgegnung 2.

Angenommen, wir wüßten, daß das Verhalten des Roboters sich zur Gänze daraus erklärt, daß ein Mensch in seinem Innern unverstandene formale Symbole von seinen Sinnesrezeptoren empfängt und unverstandene formale Symbole an seine Antriebsmechanismen ausgibt und daß der Betreffende bei diesem Hantieren mit Symbolen bestimmten Anleitungen folgt. Nehmen wir außerdem an, der Betreffende weiß von der ganzen Sache mit dem Roboter nichts, er weiß nur, welche Handlungen er mit welchen nichtssagenden Symbolen ausführen muß. Unter solchen Umständen ist für uns der Roboter nichts weiter als eine ingeniose mechanische Puppe. Die Annahme, daß die Puppe einen Geist hat, ist nun ebenso unnötig wie ungerechtfertigt, denn es gibt keinen Grund mehr, dem Roboter oder dem System, dessen Teil er ist, Intentionalität zuzuschreiben (außer natürlich die Intentionalität, die

dem mit den Symbolen hantierenden Menschen eignet). Das formale Hantieren mit den Symbolen nimmt seinen Gang, die Eingaben und Ausgaben werden korrekt in Übereinstimmung miteinander gebracht, aber der einzige wirkliche Ort, an dem Intentionalität sich findet, ist der beteiligte Mensch, und der weiß von keinem der hief in Frage kommenden intentionalen Zustände; zum Beispiel *sieht* er nicht, was dem Roboter vor Augen kommt, er *beabsichtigt* nicht, den Roboterarm zu heben, und er *versteht* keine der Bemerkungen, die dem Roboter gegenüber gemacht werden und die er selber macht. Und das gilt aus den obengenannten Gründen ebenso sehr für das System, dem Mensch und Roboter als Bestandteile angehören.

Kontrastieren wir, um diesen Punkt deutlich zu machen, unseren Fall mit solchen Fällen, bei denen uns die Zuschreibung von Intentionalität natürlich vorkommt, z.B. bei Exemplaren bestimmter Primatenarten wie etwa Menschenaffen und kleineren Affen oder bei Haustieren wie etwa Hunden. Dafür, daß uns die Sache hier natürlich vorkommt, gibt es, grob gesprochen, zwei Gründe: Wir können uns das Verhalten des Tiers nicht erklären, wenn wir ihm nicht Intentionalität zuschreiben, und wir können sehen, daß die Tiere aus ähnlichem Stoff gemacht sind wie wir – das da ist ein Auge, das da eine Nase, dies hier ist seine Haut, usw. Wenn das Verhalten des Tiers kohärent ist und wir annehmen, daß ihm dieselben Ursachen wie bei uns zugrunde liegen, kommen wir zu der doppelten Annahme, daß das Tier seinem Verhalten zugrunde liegende geistige Zustände haben muß und daß die geistigen Zustände durch Mechanismen hervorgerufen werden, die aus dem gleichen Stoff wie bei uns gemacht sind. Solange wir keinen Grund hätten, der dagegenspräche, würden wir zweifellos auch beim Roboter ähnliches annehmen, aber sobald wir wüßten, daß das Verhalten des Roboters das Ergebnis eines formalen Programms ist und daß die tatsächlichen kausalen Eigenschaften der physischen Substanz dabei keine Rolle

spielen, würden wir die Annahme von Intentionalität fallenlassen.

Es gibt zwei weitere Antworten auf meinen Beispielsfall, die häufig gegeben werden (und deshalb diskutiert werden müssen), tatsächlich aber an der Sache vorbeigehen.

5. *Die Replik mit dem Geist anderer (Yale).* »Woher wissen Sie, daß andere Menschen Chinesisch oder was auch immer verstehen? Nur aus ihrem Verhalten. Nun kann der Computer die Verhaltenstests (im Prinzip) genauso gut bestehen, wie andere Menschen das können; wenn Sie denen also Erkenntnisvermögen zuschreiben, müssen Sie es im Prinzip auch Computern zuschreiben.«

Dieser Einwand verdient tatsächlich nur eine kurze Abfertigung. Bei dem Problem, das hier erörtert wird, geht es nicht um die Frage, woher ich weiß, daß andere Menschen kognitive Zustände haben, sondern darum, was ich ihnen eigentlich zuschreibe, wenn ich ihnen kognitive Zustände zuschreibe. Die Stoßrichtung unseres Arguments geht dahin, daß es sich dabei nicht einfach nur um Rechenprozesse und deren Output handeln kann, weil Rechenprozesse und ihr Output ohne kognitiven Zustand möglich sind. Sich dumm zu stellen, ist keine Antwort auf unser Argument. In den »Kognitionswissenschaften« sind die Wirklichkeit und Erkennbarkeit des Geistigen ganz ebenso vorausgesetzt wie in den physikalischen Wissenschaften die Wirklichkeit und Erkennbarkeit physischer Objekte.

6. *Die Replik mit den vielen Behausungen (Berkeley).* »Ihr ganzes Argument setzt voraus, daß die AI es bloß mit Analog- und Digitalrechnern zu tun hat. Aber das ist nur zufällig der derzeitige technologische Entwicklungsstand. Wie auch immer diese kausalen Prozesse aussehen, von denen Sie behaupten, daß sie für Intentionalität wesentlich seien (einmal gesetzt, Sie haben recht), irgendwann werden wir in der Lage sein, Apparate zu bauen, die über diese kausalen Prozesse verfügen, und das wird dann die artifizielle Intelligenz sein. Ihr

Argument richtet sich also überhaupt nicht gegen die Fähigkeit der artifiziellen Intelligenz, Erkenntnis zu produzieren und zu erklären.»

Gegen diese Replik habe ich tatsächlich gar nichts einzuwenden, nur muß ich sagen, daß sie das Projekt der starken AI trivialisiert, indem sie als artifizielle Intelligenz alles definiert, was Erkenntnis künstlich produziert und erklärt. Das Interessante an dem ursprünglichen, im Zusammenhang mit der artifiziellen Intelligenz erhobenen Anspruch war demgegenüber gerade, daß es sich dabei um eine präzise, genau bestimmte These handelte: die These, daß geistige Prozesse am Modell des Rechners orientierte Prozesse sind, die an formal definierten Elementen ablaufen. Mein Ziel war es, diese These zu bestreiten. Wenn der Anspruch neu bestimmt wird, so daß er jener These nicht mehr entspricht, greifen meine Einwände nicht mehr, weil es die überprüfbare Hypothese, auf die sie sich beziehen, nicht mehr gibt.

Kehren wir nun zu der Frage zurück, auf die ich eine Antwort zu geben versuchen wollte: Gesetzt also, bei meinem ursprünglichen Beispiel verstehe ich das Englische, während ich das Chinesische nicht verstehe, und gesetzt demnach auch, die Maschine versteht weder Englisch noch Chinesisch, muß es dann nicht aber etwas in mir geben, das macht, daß ich Englisch verstehe, und muß mir entsprechend nicht ein anderes Etwas fehlen, das macht, daß ich Chinesisch nicht zu verstehen mag? Warum soll es nun nicht möglich sein, diese zwei Etwas, egal, worum es sich dabei handelt, einer Maschine einzugeben?

Ich sehe im Prinzip keinen Grund, warum wir nicht einer Maschine die Fähigkeit einpflanzen können sollten, Englisch oder Chinesisch zu verstehen, da ja in einem wesentlichen Sinn unsere Körper mit unseren Gehirnen nichts anderes als solche Maschinen sind. Hingegen gibt es, meine ich, sehr gute Gründe für die Behauptung, daß wir dergleichen keiner Maschine einpflanzen könnten, deren Funktionsweise sich streng im Rahmen von Prozessen hält, die am Modell des Rechners

orientiert und auf formal definierte Elemente beschränkt sind; deren Funktionieren also die Realisierung eines Computerprogramms ist. Nicht, weil ich ein Computerprogramm verkörpere (vermutlich verkörpere ich jede Menge Computerprogramme), verstehe ich Englisch und weise andere Formen von Intentionalität auf, sondern deshalb, soweit wir wissen, weil ich eine bestimmte Art Organismus mit einer bestimmten biologischen (d.h. chemischen und physikalischen) Struktur bin und weil diese Struktur unter bestimmten Bedingungen im ursächlichen Sinne imstande ist, Wahrnehmen, Handeln, Verstehen, Lernen und andere intentionale Zustände hervorzurufen. Vielleicht könnten auch andere physikalische und chemische Vorgänge genau diese Wirkungen zeitigen; vielleicht haben zum Beispiel auch die Marsmenschen Intentionalität, obwohl ihre Gehirne aus einem ganz anderen Stoff bestehen. Das ist ein empirisches Problem, ziemlich ähnlich der Frage, ob die Photosynthese sich mit einer anderen Chemie als der des Chlorophyll bewerkstelligen läßt.

Aber der Hauptpunkt des hier vorgetragenen Arguments ist, daß kein rein formales Modell jemals ausreichen wird für die Darstellung von Intentionalität, weil die formalen Eigenschaften als solche für Intentionalität keine konstitutive Bedeutung haben und weil sie als solche kein kausales Vermögen haben außer demjenigen, bei ihrer Realisierung das jeweils nächste Stadium des formalen Verfahrens nach sich zu ziehen, während und solange die Maschine in Betrieb ist. Und alle sonstigen kausalen Eigenschaften, die die jeweiligen Realisierungen des formalen Modells aufweisen, haben mit dem formalen Modell nichts zu tun; denn wir können ja jederzeit dasselbe formale Modell in eine andere Realisierung stecken, der jene besonderen kausalen Eigenschaften augenscheinlich abgehen. Auch wenn wunderbarerweise diejenigen, die Chinesisch als Muttersprache sprechen, haargenau das Schanksche Programm realisierten, könnten wir nichtsdestoweniger doch dasselbe Programm auch in solche Realisierungen stecken, die Englisch als Muttersprache sprechen, oder in Wasserleitungen

oder in Computer, die allesamt, dem Programm zum Trotz, kein Chinesisch verstehen.

Worauf es bei den Vorgängen im Gehirn ankommt, ist nicht der formale Schatten, den die Synapsensequenz wirft, sondern sind vielmehr die tatsächlichen Eigenschaften dieser Sequenzen. Alle mir bekannten Argumente für die starke Version der artifiziellen Intelligenz legen es darauf an, die Schatten, die das Erkennen wirft, zu umreißen und dann diese Schatten für die Sache selbst zu erklären.

Zum Schluß will ich versuchen, einige der in unserer Beweisführung enthaltenen allgemeinen philosophischen Punkte festzuhalten. Um der Deutlichkeit willen werde ich versuchen, das in der Form einer Frage-Antwort-Folge zu tun, und ich fange mit dieser abgedroschenen Frage an:

»Ist es möglich, daß eine Maschine denkt?«

Die Antwort lautet offensichtlich »ja«. Wir selber sind eben solche Maschinen.

»Gut, aber ist es möglich, daß ein Kunsterzeugnis, eine von Menschen gemachte Maschine, denkt?«

Vorausgesetzt, es ist möglich, künstlich eine Maschine herzustellen, die über ein Nervensystem verfügt, über Neuronen mit Axonen und Dendriten und all dem übrigen, hinlänglich vergleichbar dem, was wir haben, dann scheint die Antwort auf die Frage offensichtlich »ja« lauten zu müssen. Wenn man die Ursachen genau nachbilden kann, kann man auch die Wirkungen nachbilden. Und tatsächlich mag es möglich sein, Bewußtsein, Intentionalität und all das übrige mit Hilfe anderer chemischer Grundstoffe zu erzeugen als denjenigen, aus denen die menschlichen Wesen gemacht sind. Das ist, wie gesagt, eine Frage der Empirie.

»Meinetwegen, aber ist es möglich, daß ein Digitalcomputer denkt?«

Wenn mit »Digitalcomputer« irgend etwas gemeint ist, das eine Darstellungsebene aufweist, auf der es sich zu Recht als die Verkörperung eines Computerprogramms beschreiben

läßt, dann lautet die Antwort selbstverständlich wiederum »ja«; denn wir selber sind Verkörperungen jeder Menge Computerprogramme und haben die Fähigkeit zu denken.

»Aber ist es möglich, daß etwas *allein* kraft des Umstands, daß es ein Computer mit der richtigen Art Programm ist, denken, verstehen usw. kann? Kann die Verkörperung eines Programms, des richtigen Programms natürlich, für sich genommen eine zureichende Bedingung sein für Verstehen?«

So, meine ich, ist die Frage richtig gestellt, ungeachtet dessen, daß sie gewöhnlich mit einer oder mehreren der vorherigen Fragen durcheinandergebracht wird, und auf die so gestellte Frage lautet die Antwort »nein«.

»Warum nicht?«

Weil das formale Hantieren mit Symbolen als solches keinerlei Intentionalität hat; es ist ganz frei von Bedeutung; es ist auch eigentlich gar kein Hantieren mit *Symbolen*, da die Symbole ja gar nichts symbolisieren. Im Jargon der Linguistik gesprochen, haben sie nur eine Syntax, keine Semantik. Was Computer an Intentionalität zu haben scheinen, existiert nur im Geist derer, die sie programmieren und sie benutzen, im Geist derer, die die Eingaben einspeisen und die Ausgaben interpretieren.

Es war das Ziel unseres Beispiels mit dem chinesischen Zimmer, dies zu zeigen, nämlich zu zeigen, daß, wenn man jemanden in das System hineinbringt, der wirklich Intentionalität hat (einen Menschen), und dem Betreffenden das formale Programm eingibt, dieses formale Programm offensichtlich keine zusätzliche Intentionalität mit sich führt. So trägt es etwa zu der Fähigkeit des Betreffenden, Chinesisch zu verstehen, nicht das geringste bei.

Genau das Charakteristikum der AI, das so vielversprechend schien – die Unterscheidung zwischen Programm und Programmrealisierung –, erweist sich für den Anspruch, daß eine Simulation ein Duplikat sein könne, als verhängnisvoll. Die Unterscheidung zwischen dem Programm und seiner Realisierung in der Hardware scheint der Unterscheidung zwi-

schen der Ebene der geistigen Funktionen und der der Gehirnfunktionen parallel. Und wenn wir die Ebene der geistigen Funktionen als ein formales Programm darstellen könnten, dann, so scheint es, könnten wir das Wesentliche am Geist darstellen, ohne sei's introspektive Psychologie, sei's Neurophysiologie des Gehirns betreiben zu müssen. Aber die Gleichung »Geist verhält sich zu Gehirn wie Programm zu Hardware« geht in mehreren Punkten nicht auf, unter anderem in den folgenden:

Erstens hätte die Unterscheidung zwischen Programm und Realisierung zur Folge, daß ein und dasselbe Programm alle möglichen verrückten Realisierungen finden könnte, die keine Form von Intentionalität hätten. Weizenbaum (1978, Kap. 2) zeigt zum Beispiel ausführlich, wie sich mit Hilfe einer Rolle Klopapier und einem Stapel kleiner Steinchen ein Computer bauen läßt. Ähnlich läßt sich das Programm zum Verstehen von chinesischen Geschichten in einer Folge von Wasserleitungen, einer Reihe von Ventilatoren oder jemandem, der als einzige Sprache Englisch spricht, realisieren, ohne daß eine von diesen Realisierungen dadurch die Fähigkeit erlangte, Chinesisch zu verstehen. Erst einmal sind Steine, Klopapier, Lüftung und Wasserleitungen nicht die richtige Art Material für Intentionalität – nur etwas, das dieselben kausalen Kräfte hat wie ein Gehirn, kann Intentionalität haben –, und auch wenn der Englischsprechende die richtige Art Stoff für Intentionalität aufweist, läßt sich doch leicht sehen, daß er durch Auswendiglernen des Programms keine zusätzliche Intentionalität erhält, da ja das Auswendiglernen ihm kein Chinesisch beibringt.

Zweitens ist das Programm rein formal, und die intentionalen Zustände sind nicht in dieser Weise formal. Sie sind durch ihren Inhalt, nicht durch ihre Form bestimmt. Die Überzeugung zum Beispiel, daß es regnet, ist nicht als ein bestimmtes formales Schema, sondern als ein geistiger Inhalt definiert, mit Erfüllungskriterien, einem Bezugspunkt und dergleichen. Tatsächlich weist die Überzeugung als solche gar kein formales

Schema in diesem syntaktischen Sinn auf, da ja ein und dieselbe Überzeugung in unterschiedlichen linguistischen Systemen auf unzählig verschiedene Weise ausgedrückt werden kann. Drittens werden, wie bereits erwähnt, geistige Zustände und Ereignisse buchstäblich durch das Funktionieren des Gehirns hervorgebracht, während das Programm nicht in dieser Weise die Hervorbringung des Computers ist.

»Also, wenn Programme keine konstitutive Rolle für die geistigen Prozesse spielen, warum haben dann so viele Menschen geglaubt, daß es sich genau andersherum verhält? Mindestens das verlangt nach einer Erklärung.«

Darauf weiß ich wirklich keine Antwort. Der Gedanke, Simulationen könnten die Sache selbst sein, hätte schon deshalb fragwürdig erscheinen müssen, weil der Computer ja durchaus nicht nur geistige Prozesse simuliert. Niemand erwartet, daß Computersimulationen eines Brands der Alarmstufe fünf die Umgebung des Computers in Schutt und Asche legen oder daß die Computersimulation eines heftigen Regengusses uns alle naß werden läßt. Warum um alles in der Welt sollte jemand meinen, daß eine Computersimulation des Verstehens tatsächlich etwas versteht? Man hört manchmal sagen, einen Computer dazu zu bringen, daß er Schmerz empfindet oder sich verliebt, wäre schrecklich schwer, aber Schmerz empfinden bzw. lieben ist nicht schwerer oder leichter als erkennen oder sonst etwas. Zum Simulieren braucht man nur die richtigen Eingaben und Ausgaben sowie ein Programm dazwischen, das erstere in letztere transformiert. Das ist alles, was dem Computer für all seine Aktivitäten zur Verfügung steht. Die Simulation mit einem Duplikat zu verwechseln ist immer derselbe Fehler, egal, ob es sich um Schmerz, Liebe, Erkenntnis, Brände oder Regengüsse handelt.

Nichtsdestoweniger gibt es Gründe, warum die AI den Eindruck erwecken mußte – und bei manchen Menschen vielleicht immer noch erweckt –, daß sie in gewissem Sinn geistige Phänomene reproduziert und damit erklärt. Und ich glaube, wir werden diesen falschen Eindruck nicht eher aus der Welt

schaffen können, als bis wir diese Gründe, auf denen er beruht, voll und ganz freigelegt haben.

Der erste und wichtigste Grund ist ein verworrener Gebrauch des Begriffs »Datenverarbeitung«: Viele Leute in der Kognitionswissenschaft glauben, daß das menschliche Gehirn mit seinem Geist etwas macht, das sich »Datenverarbeitung« nennt, und analog dazu macht auch der Computer mit seinem Programm Datenverarbeitung: Brände und Regengüsse hingegen machen keinerlei Datenverarbeitung. So steht also der Computer, obwohl er die formalen Eigenschaften jedes beliebigen Prozesses simulieren kann, in einer besonderen Beziehung zu Geist und Gehirn, weil, wenn der Computer richtig programmiert ist, im Idealfall sogar mit demselben Programm wie das Gehirn, die Datenverarbeitung in beiden Fällen identisch ist, und diese Datenverarbeitung macht tatsächlich das Wesen des Geistigen aus. Aber das Problem mit diesem Argument ist, daß es auf einer Zweideutigkeit des Begriffs »Daten« beruht. In dem Sinn, in dem die Menschen »Daten verarbeiten«, wenn sie, sagen wir, über Fragen der Arithmetik nachdenken oder wenn sie Geschichten lesen und Fragen dazu beantworten, macht der programmierte Computer keine »Datenverarbeitung«. Was er vielmehr tut, ist, mit formalen Symbolen zu hantieren. Der Umstand, daß für den Programmierer und für denjenigen, der die Computerausgabe interpretiert, die Symbole Dinge in der Welt repräsentieren, liegt ganz und gar außerhalb der Reichweite des Computers. Der Computer, noch einmal gesagt, hat eine Syntax, aber keine Semantik. Wenn man also in den Computer »2 plus 2 gleich ?« eintippt, wird er »4« ausdrucken, aber er wird nicht die leiseste Ahnung davon haben, daß »4« 4 bzw. überhaupt etwas bedeutet. Und das Problem ist dabei nicht etwa dies, daß ihm gewisse Informationen zweiter Ordnung für die Interpretation seiner Symbole erster Ordnung fehlen, sondern vielmehr dies, daß seine Symbole erster Ordnung, soweit es den Computer selbst betrifft, bar jeder Interpretation sind. Alles, was der Computer hat, sind immer wieder nur Symbole. Die Ein-

führung des Begriffs der »Datenverarbeitung« kreiert mithin ein Dilemma: Entweder wir fassen den Begriff »Datenverarbeitung« so, daß zum Verarbeitungsprozeß Intentionalität als Bestandteil dazugehört, oder wir tun das nicht. Wenn ersteres, dann leistet der programmierte Computer keine Datenverarbeitung, sondern hantiert nur mit formalen Symbolen. Wenn letzteres, dann leistet er zwar Datenverarbeitung, aber nur in dem Sinn, in dem Rechenmaschinen, Schreibmaschinen, Mägen, Thermostaten, Regengüsse und Orkane Datenverarbeitung leisten: d. h., sie lassen sich auf einer bestimmten Ebene als etwas beschreiben, das am einen Ende Informationen aufnimmt, diese umwandelt und als Ausgabe Informationen auswirft. Aber in diesem Fall bleibt es den äußeren Beobachtern überlassen, die Ein- und Ausgaben als Informationen im üblichen Sinn zu interpretieren. Und eine Ähnlichkeit zwischen Computer und Gehirn auf Grund einer vergleichbaren Datenverarbeitung wird dadurch nicht hergestellt.

Zweitens überdauert in einem Großteil der AI ein Rest Behaviorismus oder Operationalismus. Weil entsprechend programmierte Computer Eingabe-Ausgabe-Muster aufweisen können, die denen der Menschen ähnlich sind, sind wir versucht, im Computer Geisteszustände vorzusetzen, die den menschlichen Geisteszuständen vergleichbar sind. Aber wenn wir erst einmal eingesehen haben, daß es für ein System sowohl begrifflich als auch empirisch möglich ist, in einem bestimmten Bereich über bestimmte Fertigkeiten zu verfügen, ohne doch die mindeste Intentionalität aufzuweisen, sollten wir instande sein, mit dieser Versuchung fertigzuwerden. Die Rechenmaschine auf meinem Schreibtisch weist rechnerische Fertigkeiten auf, aber keine Intentionalität, und in diesem Aufsatz habe ich zu zeigen versucht, daß ein System über ein Eingabe- und Ausgabevermögen verfügen könnte, das dem entspräche, über das ein Chinesisch als Muttersprache Sprechender verfügt, ohne daß es deshalb Chinesisch verstünde, ganz egal, wie es programmiert wäre. Der Turing-Test ist typisch für die Tradition einer unverfrorenen be-

havioristischen und operationalistischen Haltung, und wenn diejenigen, die auf dem Gebiet der AI arbeiten, sich vollständig vom Behaviorismus und Operationalismus lösen, würde das, wie ich glaube, die Gefahr, daß Simulation und Duplikation durcheinandergebracht werden, zum großen Teil beseitigen.

Drittens verbindet sich mit diesem Rest von Operationalismus eine Residualform des Dualismus; in der Tat ergibt die AI einen Sinn nur unter der dualistischen Voraussetzung, daß in Sachen Geist das Gehirn keine konstitutive Rolle spielt. Worauf es bei der starken AI (wie auch beim Funktionalismus) ankommt, sind Programme, und Programme sind unabhängig davon, in welchen Maschinen sie realisiert werden; in der Tat könnte, soweit es die AI betrifft, ein und dasselbe Programm seine Realisierung in einer elektronischen Maschine, einer Cartesianischen Geistsubstanz oder einem Hegelschen Weltgeist finden. Die einzigartige, höchst überraschende Entdeckung, die ich bei der Diskussion dieser Probleme gemacht habe, ist, daß die meisten derer, die auf dem Gebiet der AI arbeiten, geradezu entsetzt von meiner Vorstellung sind, gegebene geistige Phänomene beim Menschen könnten abhängig sein von gegebenen physikalisch-chemischen Eigenschaften der gegebenen menschlichen Gehirne. Aber wenn man einen Augenblick darüber nachdenkt, sieht man, daß ich gar keinen Grund hatte, überrascht zu sein; denn wenn man nicht irgendeine Art Dualismus akzeptiert, hat das mit der starken AI verfolgte Projekt keine Chance. Das Projekt zielt auf eine Reproduktion und Erklärung des Geistigen durch den Entwurf von Programmen ab; aber wenn der Geist nicht gleichermaßen begrifflich und empirisch unabhängig vom Gehirn ist, ist das Projekt undurchführbar; denn die Programme sind von jeder Realisierung vollständig unabhängig. Ohne die Überzeugung, daß der Geist sowohl begrifflich als auch empirisch vom Gehirn abtrennbar ist – d.h. ohne ausgeprägten Dualismus –, kann man nicht hoffen, durch die Ausarbeitung und Abwicklung von Programmen Geistiges zu reproduzieren, da

ja die Programme unabhängig von Gehirnen oder jeder sonstigen besonderen Verkörperung sein müssen. Wenn geistige Operationen in einem am Modell des Rechners orientierten Operieren mit formalen Symbolen bestehen, dann folgt daraus, daß sie keine besondere Beziehung zum Gehirn unterhalten; die einzige Beziehung wäre dann die, daß zufällig das Gehirn eine der unendlich vielen Arten von Maschinen ist, die das Programm verkörpern können. Diese Art Dualismus ist nicht die traditionelle Cartesianische, der zufolge es zweierlei *Substanzen* gibt, sondern sie ist Cartesianisch nur in dem Sinn, daß sie behauptet, zwischen dem spezifisch Geistigen am Geist und den besonderen Eigenschaften des Gehirns gebe es keine innere Beziehung. Dieser zugrundeliegende Dualismus wird dadurch verdeckt, daß die Literatur zur AI voll von Bannflüchen wider den »Dualismus« ist; was die Verfasser anscheinend nicht bemerken, ist, daß ihre eigene Position eine starke Form des Dualismus voraussetzt.

»Ist es möglich, daß eine Maschine denkt?« Ich selbst bin der Ansicht, daß *nur und ausschließlich* einer Maschine zu denken möglich ist und in der Tat nur sehr speziellen Arten von Maschinen, nämlich Gehirnen und solchen Maschinen, die über dieselben kausalen Kräfte wie Gehirne verfügen. Und das ist der Hauptgrund, warum die starke AI uns so wenig übers Denken sagen kann, kann sie uns doch nichts über Maschinen sagen. Ihrem eigenen Selbstverständnis nach handelt sie von Programmen, und Programme sind nicht Maschinen. Was immer Intentionalität sonst noch sein mag, sie ist ein biologisches Phänomen, und sie weist wahrscheinlich die gleiche Abhängigkeit von der spezifischen Biochemie ihrer Entstehungsbedingungen auf wie die Milchbildung, die Fotosynthese oder irgendein anderes Phänomen. Niemand würde meinen, wir könnten Milch und Zucker dadurch herstellen, daß wir eine Computersimulation der wesentlichen Vorgänge bei der Milchbildung oder der Fotosynthese ablaufen lassen, aber wo es um den Geist geht, sind viele bereit, an solch ein Wunder zu glauben, auf Grund eines tiefsitzenden und hartnäckigen Dua-

lismus: Der Geist, so glauben sie, ist eine Sache formaler Prozesse und ist auf eine Weise unabhängig von ganz spezifischen materialen Ursachen, wie Milch und Zucker das nicht sind.

In Verteidigung dieses Dualismus wird häufig die Erwartung geäußert, das Gehirn werde sich als ein Digitalrechner herausstellen (die frühen Computer wurden ja oft als »Elektronengehirne« bezeichnet). Aber auch das hilft nicht weiter. Natürlich ist das Gehirn ein Digitalrechner. Da alles ein Digitalrechner ist, sind es auch die Gehirne. Der entscheidende Punkt ist, daß die kausale Fähigkeit des Gehirns, Intentionalität hervorzubringen, nicht darin aufgeht, daß das Gehirn ein Computerprogramm verkörpert, da es ja in bezug auf jedes beliebige Programm möglich ist, daß etwas dieses Programm verkörpert, ohne deshalb doch irgendwelche geistigen Zustände aufzuweisen. Was immer das Gehirn tut, um Intentionalität hervorzubringen, sein Tun kann nicht in der Verkörperung eines Programms aufgehen, da kein Programm, für sich genommen, für Intentionalität ausreicht.⁵

Anmerkungen

- 1 Der englische Begriff »artificial intelligence« (»AI«), der in den übrigen Beiträgen des vorliegenden Bandes als »Künstliche Intelligenz« (»KI«) übersetzt wird, ist in der hier abgedruckten Übersetzung als »artifizielle Intelligenz« (»AI«) wiedergegeben. (Anm. der Hrsg.)
- 2 Die von Searle abgekürzt zitierte Literatur findet sich in den Literaturhinweisen des vorliegenden Bandes (S. 271–278) vollständig verzeichnet. (Anm. der Hrsg.)
- 3 Im übrigen bedeutet »Verstehen« sowohl das Vorhandensein bestimmter geistiger (intentionaler) Zustände als auch das Wahrsein (die Gültigkeit, den Erfolg) dieser Zustände. Für die Zwecke dieser Erörterung sind wir nur mit dem Vorhandensein der Zustände faßt.

- 4 Intentionalität ist definitionsgemäß jene Eigenschaft bestimmter Geisteszustände, durch die diese sich auf Gegenstände oder Angelegenheiten in der Welt richten oder beziehen. Überzeugungen, Begierden und Absichten sind demnach intentionale Zustände; unbestimmte Formen der Angst und Depression sind keine.
- 5 Ich schulde einer ziemlich großen Zahl von Leuten Dank dafür, daß sie diese Dinge mit mir diskutiert und sich geduldig bemüht haben, meine Unwissenheit auf dem Gebiet der artifiziellen Intelligenz zu überwinden. Insbesondere möchte ich Ned Block, Hubert Dreyfus, John Haugeland, Roger Schank, Robert Wilensky und Terry Winograd Dank sagen.