

Philosophie / Informatik
Sekundarstufe II

Können Computer denken?

„Künstliche Intelligenz“ als Thema
für einen fächerübergreifenden Unterricht

Teil 2: Von der Philosophie des Denkens zur Computer-Philosophie

Amt für Schule
1994



Freie und Hansestadt Hamburg
Behörde für Schule, Jugend und Berufsbildung

Können Computer denken?

„Künstliche Intelligenz“ als Thema
für einen fächerübergreifenden Unterricht

Teil 2: Von der Philosophie des Denkens zur Computer-Philosophie

Fachreferenten:	Dr. Uwe Heinrichs Ulrich Polzin	Amt für Schule S 13 Amt für Schule S 13/33
Verfasser:	Reinhard Golecki	Gymnasium Klosterschule
Hamburg, 1994		(PDF-Version März 2000)

Inhaltsverzeichnis

Teil 2: Von der Philosophie des Denkens zur Computer-Philosophie	1
1 Das Computermodell des Geistes	2
2 Mentale Repräsentation	5
2.1 HOBBS: Denken ist Rechnen	5
2.2 DESCARTES: Geist und Materie	7
2.3 LOCKE: Operieren mit Ideen	11
2.4 HUME: Die Mechanik des Geistes	14
2.5 KANT: Verstand, Urteilskraft, Vernunft	17
2.6 RYLE: Gegen das „Gespenst in der Maschine“	26
3 Idealsprachliche Rekonstruktion	31
3.1 LEIBNIZ: Kalkül des vernünftigen Denkens	31
3.2 FREGE: Die Wende zur Sprache	33
3.3 RUSSELL: Logischer Atomismus	38
3.4 WITTGENSTEIN: Sätze als logisches Bild der Tatsachen	40
3.5 CARNAP: Rationale Nachkonstruktion der Wirklichkeit	45
4 Funktionalismus	53
4.1 TURING: Die Universalmaschine	53
4.2 PUTNAM: Die funktionale Ebene	56
4.3 FODOR: Die Sprache des Geistes	58
5 Intermezzo: Großtantes Alien-Fax	64
6 Verstehen, Intentionalität, Bewußtsein	68
6.1 SEARLE: Ursprüngliche Intentionalität	68
6.2 DENNETT: Intentionale Systeme	74
7 Zwischenbilanz: Können Computer denken?	79
Literaturhinweise	90
Literaturverzeichnis	95

Teil 2

Von der Philosophie des Denkens zur Computer-Philosophie

*In der Schöpfung des Elektronengehirns aber
gibt der Mensch seine eigene Reflexion an den Gegenstand ab
und lernt in diesem Spiegel seiner selbst
seine Funktion in der Welt begreifen.*

Gotthard Günther

Die jeweils aktuellen technischen Geräte und Erfindungen haben auch in der Vergangenheit als Bild und Metapher für den Geist und die Denkprozesse gedient: das Hirn arbeite wie eine Mühle, wie ein Webstuhl, wie eine Hydraulikanlage, wie eine Telefonzentrale. Heute bestimmt die Computermetapher bis hinein in die Alltagssprache unser Denken über das Denken. Allerdings ist der Computer mehr als nur ein weiteres Glied in der genannten Kette. So wie vorher die mechanische Uhr ist der Computer Symbol für eine ganze Epoche; war jene – auch in der Gestalt mechanischer Puppen und Tiere – in der Nachfolge der Arbeiten DESCARTES' und NEWTONS im Rahmen des mechanistischen Weltbildes technisch realisiertes Modell für die Deutung und Beherrschung sowohl der unbelebten wie belebten Natur, für alles Materielle,¹ so ist der Computer heute im „Informationszeitalter“ Modell für den damals ausgeklammerten Bereich des Geistes, für die Objektivierung (oder vielleicht besser Intersubjektivierung) des Subjektiven. Ging es damals, im zweiwertigen Weltbild von Subjekt und Objekt, von Geist und Materie, um die Abgrenzung des Menschen von der Natur, um die Besonderheit des menschlichen *Lebens* (mit dem Ergebnis, in Bezug auf den *Körper* die Sonderstellung des Menschen aufzugeben und nur ihm eine *Seele* zuzusprechen), so geht es heute, nach der Entdeckung, daß sich Information weder restlos als Geist noch als Materie erklären läßt, um die Frage, ob dem *Geist* des Menschen eine exklusive Sonderstellung zukommt (wobei sich aus biologischer Sicht das *Hirn* des Menschen als nur graduell verschieden von denen höherer Säuger erwiesen hat).

Aus der Sicht einiger „Häretiker“ sind Menschen und Computer nur unterschiedliche Arten der umfassenderen Gemeinschaft „informationsverarbeitender Systeme“, ähnlich wie wir vorher lernen mußten, daß Tintenfische, Schweine und Menschen nur unterschiedliche Arten lebender Wesen sind, zwischen denen auf der physiologischen Ebene kein wesentlicher Unterschied besteht. Nach dieser Auffassung ist der Mensch nicht „das vernünftige

¹Ein wenig mehr dazu habe ich in GOLECKI (1991) ausgeführt.

unter den Tieren,“ sondern „das Tier unter den Vernunftwesen, d. h. dasjenige Wesen, dessen mentale Zustände kontingenterweise in biologischer ‚wetware‘ realisiert sind.“² Im Spannungsfeld dieser neuen Fragestellung sind mit dem Computer ähnliche Hoffnungen und Ängste verbunden wie damals mit den feinmechanischen Puppen; so gesehen ist der Computer aber auch Metapher und Modell für den Glauben an die Berechenbarkeit und Beherrschung der Welt durch *Informationsverarbeitung*, analog dem früheren und nun schwindenden Glauben an die *Mechanik*.

1 Das Computermodell des Geistes

Für die am Paradigma der Informationsverarbeitung orientierten Kognitionsforscher ist der Computer noch mehr als das, nämlich Modell des menschlichen Geistes. Der Computer und seine Programmierung haben erstens eine heuristische Funktion für Theorien über Kognition, die erfolgreiche Technik kann Anhaltspunkte dafür liefern, wie Menschen vergleichbare Probleme lösen. Umgekehrt lassen sich die durch diese Analogie gewonnenen Theorien in Computerprogramme umsetzen, und in einer ganz eigenen Mischung aus gedanklichem und empirischem Experiment ergibt sich dadurch die Möglichkeit, die Kohärenz und Adäquatheit dieser Theorien zu testen, indem man prüft, ob die Programme die gleichen Leistungen (und die gleichen Fehler) zeigen wie wir Menschen. Zweitens kann man dann weiter die These vertreten, daß Computer nicht nur nützliche Werkzeuge zur Erforschung des Denkens, sondern auch Modelle *für* den menschlichen Geist seien, weil das Denken tatsächlich so ablaufe wie Computerprogramme, nur auf der physikalischen Ebene anders realisiert.

Zentral für diese These ist der Begriff der *Symbolverarbeitung*. „Symbole“ sind „physikalische Muster (wie zum Beispiel Kreidemarkierungen auf einer Tafel), die als Komponenten von Symbolstrukturen (bisweilen ‚Ausdrücke‘ genannt) auftreten können.“ Ein „Symbolsystem“ verfügt „über eine Anzahl einfacher Prozesse, die über Symbolstrukturen operieren – Prozesse, die Symbole erzeugen, verändern, kopieren und zerstören.“ Ein „physikalisches Symbolsystem“ ist dann „eine Maschine, die in ihrer Bewegung durch die Zeit eine evolviende Kollektion von Symbolstrukturen erzeugt.“ Die Systeme sind „physikalisch“ in dem Sinne, „daß es sich um Dinge aus der realen Welt handelt, hergestellt aus Glas und Metall (Computer) oder aus Fleisch und Blut (Gehirne). In der Vergangenheit waren wir mehr daran gewöhnt, uns die Symbolstrukturen der Mathematik und der Logik als abstrakt und körperlos vorzustellen – wenn man von Papier, Bleistift und Verstand absieht, die notwendig waren, um sie ins Leben zu rufen. Die Computer haben die Symbolsysteme aus dem platonischen Reich der Ideen in die empirische Welt aktueller Prozesse verlagert, die in Maschinen oder Gehirnen oder in einer Verbindung beider ablaufen.“³ Zusammengefaßt ergibt sich daraus – als vorherrschendes Paradigma der *kognitiven KI* – die *Hypothese der physikalischen Symbolsysteme*:

Ein physikalisches Symbolsystem hat die notwendigen und hinreichenden Mittel für allgemeine intelligente Handlungen.⁴

²Geert KEIL (1993a, S. 331 f., 1993b, S. 195 f.). „Mit der Definition des Menschen als *animal rationale*“ ist das „übrigens durchaus verträglich,“ nur kehrt man damit „das Verhältnis zwischen *genus proximum* und *differentia specifica* um“ (a. a. O.).

³Herbert A. SIMON (1990), S. 19 f.

⁴Allen NEWELL & Herbert A. SIMON (1992), S. 61. Mit dem Ansatz der Symbolverarbeitung konkur-

Der kognitiv orientierte Teil der KI-Forscher (und das waren am Anfang alle) will Maschinen mit echter, nicht bloß simulierter Intelligenz bauen, sie beantworten die Frage „Können Computer denken?“ mit einem überzeugten „Ja!“, weil nach ihrer Überzeugung auch *wir* eine Art Computer sind. Ich habe bisher kaum eine Frage erlebt, bei der sich wie bei dieser zwei Fraktionen unversöhnlich gegenüberstehen, die die jeweilige Antwort für völlig evident und kaum des Nachdenkens wert halten, mit großem Nachdruck und viel Eifer den Standpunkt vertreten, daß nur absolut verbohrt Menschen das nicht einsehen könnten, und die – wie bei einer „Gestalt-Switch“ Figur – buchstäblich nicht sehen können, wie die Sichtweise der anderen zustande kommt. Auffällig ist weiter, daß die Auseinandersetzung häufig gemäß den Grenzen der schon in der Einleitung (Teil 1) genannten „Zwei Kulturen“ verläuft. Es drängt sich der Eindruck auf, „daß es in dieser Debatte weniger um eine Abwägung der Argumente geht als um eine Art ‚Kulturkampf‘, wobei die Alternative von den Befürwortern der Analogisierung gerne als die Auseinandersetzung zwischen Hinterwäldlern und Aufgeklärten geführt wird. Die Gegner der Analogisierung sehen demgegenüber die Analogisierungsdebatte als eine Auseinandersetzung zwischen Humanismus und pathologischer Fortschrittsgläubigkeit.“⁵

Diejenigen, die (ggf. aus guten Gründen) an die Sonderstellung menschlichen Denkens glauben, können gegenüber Vertretern des „Computermodells des Geistes“ drei Haltungen einnehmen:

- Sie können sich zurücklehnen, ein wenig gelangweilt an die Decke schauen und bei sich denken „... ja, ja, kennen wir schon, erst waren es die Mühlen, dann die automatischen Webstühle und nun die Computer. Was wird die nächste Mode sein?“
- Sie können sich vorbeugen und sofort mit einer beeindruckenden Reihe *Apriori*-Argumente beweisen, daß Computer niemals werden denken können bzw. daß sie niemals zu diesem oder jenem fähig sein werden (unter der Gefahr, sich damit ähnlich lächerlich zu machen wie die eifrigen Proponenten der KI mit ihren positiven Prognosen; ersatzweise kann dann der Hinweis erfolgen, daß es nur so aussehe, als ob der Computer es könne, nicht aber *wirklich* so sei).

rieren seit einigen Jahren *subsymbolische* Ansätze („Konnektionismus“, „künstliche neuronale Netze“). Einen ersten Überblick dazu geben Christel KEMKE (1988) und Joachim DIEDERICH (1991). Eine umfassende Darstellung des gegenwärtigen Standes der Forschung bieten Patricia S. CHURCHLAND & Terrence J. SEJNOWSKI (1997). Eine andere Variante sind die „Agenten“ der „Gesellschaft des Geistes“ von Marvin MINSKY (1990).

⁵Heiner HASTEDT (1988), S. 146. Darüber hinaus glaube ich, daß die extremen Befürworter und Gegner gleichermaßen dem für die Durchsetzung des Projekts der Neuzeit so wichtigen Dualismus verhaftet sind, daß also die paradoxe Situation besteht, daß sich beide Parteien auf verschiedene Seiten der gleichen Medaille berufen. Wenn die gesamte „Natur“ für berechenbar, beherrschbar, verfügbar erklärt wird, so werden zwangsläufig die dafür entwickelten Methoden und Verfahren früher oder später auch als geeignet für die Tätigkeit unseres Gehirns erscheinen; wenn der autonome Mensch als „Geist“ aus dieser „Natur“ herausgenommen und ihr gegenübergestellt wird, muß dieser Versuch mit allen Kräften abgewehrt werden. So ist dann die Verwirrung über die KI unvermeidlich, wenn von Gegnern und Befürwortern der KI Begriffe und Konzepte aus der Welt der Mechanik in einer mit diesem Inventar nicht mehr zu deutenden Welt verwendet werden. In der KI geht es nur in einem sehr übertragenen Sinne um die „Mechanisierung“ des Denkens, und ihre Pointe wird verfehlt, wenn diese Redeweise zu wörtlich genommen wird. Eine darüber hinausgehende Debatte der KI (die auch den Seitenblick auf andere nachmechanistische Konzepte nicht unterläßt: Evolution, Selbstorganisation, Autopoiese, deterministisches Chaos usw.) kann dagegen einen wichtigen Beitrag zur Selbstaufklärung leisten.

- Sie können aufstehen, ein wenig in Bibliotheken stöbern, sich einige Programme und Computer ansehen, auf einem gemeinsamen Spaziergang sich die Gründe für den anderen Standpunkt erklären lassen, Kritik anmelden und dann in einen Dialog eintreten, der offen ist für eine Veränderung der Sichtweisen auf beiden Seiten.

Ich plädiere für die letztere Haltung, allerdings ist eine – neben der Bewertung erzielter Resultate – dafür nötige Diskussion der Grundannahmen und leitenden Vorstellungen der kognitiven KI schwieriger, als man auf den ersten Blick vermuten würde. Das liegt erstens daran, daß sie nur selten explizit angegeben werden⁶ und daß man zweitens nicht ohne weiteres von einem gemeinsamen „Bildungsgut“ der meist anglo-amerikanischen Proponenten und kontinental-europäischen Opponenten ausgehen kann. Drittens erweist sich beim näheren Hinsehen, daß das scheinbar monolithische Lager der kognitiven KI durchaus zerstritten ist und sich auch auf recht unterschiedliche Traditionen stützt. Viertens sind dabei die Bezüge so vielfältig und auf so viele Disziplinen verteilt, daß eine Zusammenschau nur schwer möglich ist. Das „Computermodell des Geistes“ hat in diesen Disziplinen nicht nur seine Wurzeln, sein Erfolg beruht u. a. auch darauf, hier Verbindungen geschaffen und seinerseits wichtige Impulse für diese Fachwissenschaften gegeben zu haben.⁷

Bei der Diskussion der Frage, ob Computer denken können, ist es deshalb sicher nützlich, sowohl dem *Begriff des Denkens* als auch – in Umkehrung einer Frage von Marvin MINSKY zu „Why do many people think computers *can*?“ – den *Wurzeln der KI* in der Philosophie nachzugehen, also einerseits die Beiträge aus der philosophischen Tradition, auf die sich die Proponenten der KI berufen können, ebenso sorgfältig zu prüfen wie andererseits den Beitrag der KI zur Weiterentwicklung der Philosophie des Geistes.

Dabei wird auf den folgenden Seiten zunächst (Abschnitt 2) die Entwicklung des „*mentalistischen Paradigmas*“ vom Beginn der Neuzeit bis in die Gegenwart dargestellt (bei der auch anders als manchmal vermutet die empiristische Tradition eine große Rolle spielt), zum einen, weil es unser Alltagsverständnis vom Denken entscheidend geprägt hat, zum anderen berufen sich sowohl Verfechter wie auch Gegner der KI gleichermaßen auf Teile dieses Paradigmas (die nicht reduzierbare, eigenständige „mentale“ Ebene des regelgeleiteten Operierens mit symbolischen Repräsentationen der Außenwelt einerseits, die untrennbare Verbindung von (Selbst-) Bewußtsein, Denken und Person sein andererseits). Weiter

⁶Ein Grund dafür ist sicher der von Johann GALTUNG (1983, S. 325) treffend und humorvoll beschriebene Unterschied in den intellektuellen Stilen hier und in den USA, pointiert in den typischen Fragen als Reaktion auf eine These: „teutonische“ Version: „Wie können Sie das zurückführen/ableiten?“, „sachsonische“ (US-) Version: „How do you operationalize it?“ („gallische“ Version: „Kann man das auch auf gut französisch sagen?“, „nipponische“ Version: „Wer ist ihr Meister?“). Mir ist durchaus bewußt, daß die vorliegende Schrift ein sehr teutonisches Unternehmen ist.

⁷Allen NEWELL und Herbert SIMON (1992, S. 62 ff.) nennen (das geschieht nur selten) als Ahnen ihrer „physikalischen Symbolsystem-Hypothese“ die Philosophen und Logiker Gottlob FREGE, Alfred North WHITEHEAD und Bertrand RUSSELL, Alan TURING, Emil POST und Alonzo CHURCH, den Mathematiker und Designer der heute vorherrschenden Computerarchitektur John von NEUMANN, den Pionier einer elektronischen Modellierung des Gehirns Warren MCCULLOCH, den Begründer der „Informationstheorie“ Claude SHANNON sowie den Entwickler der Programmiersprache LISP John MCCARTHY.

Marvin MINSKY (1990, S. 19) bezieht sich dagegen oder darüber hinaus bei seiner Theorie der „Gesellschaft des Geistes“ auf die Entdecker der chemischen Struktur der DNS James WATSON und Francis CRICK, auf den Mathematiker Kurt GÖDEL und auf die „Psychologen“ Sigmund FREUD und Jean PIAGET.

Für eine umfassende Darstellung dieser neben der Philosophie weiteren „Ahnen und Geschwistern“ der KI – Psychologie, Linguistik, Neurobiologie – vgl. Howard GARDNER (1989).

werden dann das Konzept einer *idealen Sprache* als rationale Rekonstruktion und als Instrument des Denkens von LEIBNIZ bis hin zum Logischen Empirismus (Abschnitt 3) und danach – als Zusammenführung von Teilen dieser beiden Stränge – der *Funktionalismus* (Abschnitt 4) vorgestellt.

Der Funktionalismus ist eine gegenwärtig sehr einflußreiche Richtung in der Philosophie des Geistes, die das Anliegen der KI dadurch stärkt, daß sie Denken als im Wortsinne *Berechnen* auffaßt, daß sich also der Geist zum Hirn so verhält wie das Programm eines Computers zu seiner Hardware; diese *Computational theory of mind* stützt sich auf

- die Annahme, daß Denken auf dem (meist unbewußten) Operieren mit mentalen Repräsentationen beruht (Abschnitt 2);
- die Annahme, daß dieses Operieren wie bei den in der Logik entwickelten „Symbolischen Maschinen“ bloß „formal“, „syntaktisch“ erfolgt (Abschnitt 3);
- die These, daß jede Art solcher „Berechnung“ durch programmierte Digitalcomputer erfolgen kann und daß ihre sehr unterschiedliche technische (aber funktional gleiche) Realisierbarkeit nahelegt, auch beim Denken zwischen der (eigentlich wichtigen) logisch-funktionalen und der (kontingenten) physisch-strukturellen Ebene zu unterscheiden (Abschnitt 4).

Nach einer kleinen Geschichte, die sich an SEARLES Gedankenexperiment vom *Chinese room* anlehnt (Abschnitt 5), wird die daran anschließende Diskussion um die *Intentionalität* als zentralem Charakteristikum des Mentalen dargestellt (Abschnitt 6).

2 Mentale Repräsentation

2.1 Hobbes: Denken ist Rechnen

Mit einigem Recht kann man Thomas HOBBS (1588–1679) als philosophischen „Großvater der KI“⁸ bezeichnen. Angeregt durch den von ihm sehr bewunderten GALILEI versuchte er, dessen Prinzipien der Naturerklärung (atomistische Mechanik und Mathematik, d. h. insbesondere die „resolutiv-kompositive“, also auseinandernehmend-zusammensetzende Geometrie EUKLIDS als Methode für Darstellung und Beweis) auch auf den Bereich des Geistes und der Gesellschaft zu übertragen. Berühmt wurde er für seine Staatsphilosophie, in unserem Zusammenhang ist aber interessanter, was er über das Denken sagte.

Zunächst erklärt er, wie unsere *Empfindungen* entstehen: „Ursache der Empfindung ist der äußere Körper oder Objekt, der auf das jeder Empfindung entsprechende Organ drückt, entweder unmittelbar wie beim Schmecken und Fühlen, oder mittelbar wie beim Sehen, Hören und Riechen. Dieser Druck setzt sich durch die Vermittlung der Nerven und anderer Stränge und Membranen des Körpers nach innen bis zu dem Gehirn und Herzen fort und verursacht dort einen Widerstand oder Gegendruck oder ein Bestreben des Herzens, sich davon freizumachen.“⁹ Unsere Empfindungen sind also durch Objekte der Außenwelt verursachte innere Bewegungsmuster der Materie, die allerdings nur mittelbar den Objekten

⁸So John HAUGELAND (1987), S. 19.

⁹*Leviathan*, Frankfurt a. M. 1984 (Erstausgabe 1651), 1. Kap., S. 11.

entsprechen. Wie schon GALILEI (und vor ihm die griechischen Atomisten) unterscheidet er die primären Qualitäten der Objekte (Ausdehnung, Gestalt, Anzahl, Bewegung) von den sekundären (Farbe, Ton, Geruch, Hitze, Kälte, Härte usw.) unserer Sinnlichkeit (ein Pullover *ist* also nicht rot, seine materielle Beschaffenheit verursacht durch Medien vermittelte Empfindungen, die wir „rot“ *nennen*¹⁰), das Objekt und unsere Vorstellung davon sind also zwei verschiedene Dinge.

Vorstellungen, Erinnerungen und Erfahrungen sind für HOBBS die verebbenden oder überlagerten Nachwirkungen einzelner oder immer wieder aufgetretener Empfindungen (so wie die Wellen am Elbufer, nachdem ein Schiff vorbei fuhr). Soweit unterscheidet sich der Mensch nicht vom Tier, der wesentliche Unterschied liegt erst in der Fähigkeit zur Sprache, im Gebrauch von Namen oder Benennungen und ihrer Verknüpfung. Solche sprachlichen Zeichen, die Worte, haben dabei zwei Funktionen: erstens dienen sie als Merkzeichen der Erinnerung, zweitens als Mitteilungszeichen für den Austausch von Vorstellungen und Gedanken zwischen den Menschen, Sprache ist also eine Übertragung der Folge von Gedanken in eine Folge von Wörtern. Wahrheit besteht dann „in der richtigen Anordnung der Namen bei unseren Behauptungen“, „*wahr* und *falsch* sind Attribute der Sprache, nicht von Dingen.“¹¹ In einem etymologischen Analogieschluß bereitet er dann seine Definition von Denken vor:

Die Römer nannten Geldrechnungen *rationes* [...], und was wir in Rechnungen oder in der Buchführung als *Posten* bezeichnen, nannten sie *nomina*, das heißt *Namen*. Und von da her scheinen sie die Bedeutung des Wortes *ratio* auf die Fähigkeit des Rechnens in allen anderen Gebieten ausgedehnt zu haben. Die Griechen besitzen für *Sprache* und *Vernunft* nur das Wort *λόγος*. Nicht, daß sie gedacht hätten, es gebe keine Sprache ohne Vernunft, sondern umgekehrt kein Denken ohne Sprache. Und den Denkkakt nannten sie *Syllogismus*, was Zusammenzählen der Folgerungen aus einer Aussage mit der einer anderen bedeutet. [...] *Denken* heißt nichts anderes als sich eine Gesamtsumme durch *Addition* von Teilen oder einen Rest durch *Subtraktion* einer Summe von einer anderen vorstellen. Geschieht dies durch Wörter, so ist es ein Vorstellen dessen, was sich aus den Namen aller Teile für den Namen des Ganzen, oder aus den Namen des Ganzen und eines Teiles für den Namen des anderen Teiles ergibt. [...] Kurz: Wo *Addition* und *Subtraktion* am Platze sind, da ist auch *Vernunft* am Platze, und wo sie nicht am Platze sind, hat *Vernunft* überhaupt nichts zu suchen. [...] Denn *Vernunft* in diesem Sinne ist nichts anderes als *Rechnen*, das heißt Addieren und Subtrahieren, mit den Folgen aus den allgemeinen Namen, auf die man sich zum *Kennzeichnen* und *Anzeigen* unserer Gedanken geeinigt hat.¹²

In diesen fast dreieinhalb Jahrhunderte alten Zeilen, in diesem Verständnis von *Rationalität* als im Wortsinne *Berechnung*, lassen sich schon viele zentrale Annahmen der heutigen Kognitionswissenschaft und speziell der KI-Forschung finden: Denken als regelgeleitete Manipulation von Symbolen, die einerseits „Objekte“ der Außenwelt repräsentieren, andererseits physikalisch im Gehirn realisiert sind und sich funktional äquivalent so in ein anderes

¹⁰Wie problematisch es ist, Farbe („objektiv“) als „Attribut von Dingen“ oder („subjektiv“) als „bloße Empfindung“ zu deuten, belegen eindrucksvoll FRANCISCO J. VARELA & EVAN THOMPSON (1992, S. 218–252), da Farbkategorien sowohl „empirisch geprägt“ sind als auch „unserer gemeinsamen biologischen und kulturellen Welt“ angehören (sie plädieren für den „Mittleren Weg“ der „wechselseitigen Spezifikation“). Überhaupt bildet – wegen der Vielfalt der beteiligten Disziplinen (Hirnforschung, Psychologie, Künstliche Intelligenz, Linguistik, Philosophie und Anthropologie) – „die Erforschung der Farbe einen Mikrokosmos der Kognitionswissenschaft“. Vgl. zu diesem Problem auch DANIEL C. DENNETT (1994), S. 481 ff.

¹¹HOBBS, a. a. O., 4. Kap., S. 27 f.

¹²A. a. O., 4. Kap., S. 29, 5. Kap., S. 32.

Symbolsystem – die Sprache – übertragen lassen, daß die schlußfolgernden und anderen Operationen innerhalb der Sprache das gleiche Ergebnis wie im Gehirn (und eine Entsprechung in der „Wirklichkeit“) haben. Allerdings konnte sich HOBBS' Sicht des Denkens zu seiner Zeit nicht durchsetzen, die strikt materialistische und deterministische Deutung der Wahrnehmungs- und Denkvorgänge, das damit verbundene Leugnen der Willensfreiheit, die Vorstellung, auch nur eine (wenn auch kompliziertere) mechanische Maschine zu sein, paßte nicht zur Aufbruchstimmung des autonomen Subjekts der Neuzeit; so – als passive Unterworfenheit – war Francis BACONS berühmtes, programmatisches Motto der Wissenschaft der Neuzeit „Man beherrscht die Natur nur, indem man sich ihren Gesetzen unterwirft“ nicht gemeint, so war schwerlich „die Macht und Herrschaft des menschlichen Geschlechts über die Gesamtnatur zu begründen“,¹³ das ging nur, wenn man sich selbst davon ausnahm (zudem mußte er eine detaillierte Erklärung des Entstehens bedeutungsvoller, komplexer Symbole und ihrer „Berechnung“ aus der Mechanik des Gehirns schuldig bleiben – was angesichts vergleichbarer Probleme der heutigen Hirnforschung kein Vorwurf sein soll).

2.2 Descartes: Geist und Materie

Anders als HOBBS war sein Zeitgenosse René DESCARTES (1596–1650) nicht nur ein großer Bewunderer der Mathematik, sondern selbst ein bedeutender Mathematiker. Er entwickelte die Analytische Geometrie, d. h. die Darstellung von Raumpunkten durch Zahlen (Koordinaten) und von Lagebeziehungen und Bewegungen durch (arithmetische) Gleichungen. Ihm gelang damit ein sehr wichtiger Schritt auf dem Weg zu einer rein mathematisch-mechanischen Naturbeschreibung und -erklärung. Die Mathematik – wiederum im wesentlichen der axiomatische Aufbau der Geometrie durch EUKLID – war ihm auch Vorbild für die von ihm propagierte *Methode* (sicherer Weg) „des richtigen Vernunftgebrauchs und der Wahrheitsforschung in den Wissenschaften.“ Der sichere Weg zur Erkenntnis ist für DESCARTES gegeben durch die strenge Beachtung der folgenden Regeln:

Die erste besagte, niemals eine Sache als wahr anzuerkennen, von der ich nicht evidentermaßen erkenne, daß sie wahr ist: d. h. Übereilung und Vorurteile sorgfältig zu vermeiden und über nichts zu urteilen, was sich meinem Denken nicht so klar und deutlich darstellte, daß ich keinen Anlaß hätte, daran zu zweifeln.

Die zweite, jedes Problem, das ich untersuchen würde, in so viele Teile zu teilen, wie es angeht und wie es nötig ist, um es leichter zu lösen.

Die dritte, in der gehörigen Ordnung zu denken, d. h. mit den einfachsten und am leichtesten zu durchschauenden Dingen zu beginnen, um so nach und nach, gleichsam über Stufen, bis zur Erkenntnis der zusammengesetztesten aufzusteigen, ja selbst in Dinge Ordnung zu bringen, die natürlicherweise nicht aufeinander folgen.

Die letzte, überall so vollständige Aufzählungen und allgemeine Übersichten aufzustellen, daß ich versichert wäre, nichts zu vergessen.¹⁴

Diese analytisch-deduktive Methode galt für DESCARTES anders als bei GALILEI nicht nur für den Bereich der Naturerkenntnis, er benutzte sie in einem radikalen Prozeß der Selbstanwendung auch zur Untersuchung der Erkenntnismittel selbst. Ohne eine solche Gewißheit der Erkenntnis, ohne einen absolut sicheren Ausgangspunkt für das Denken,

¹³ *Neues Organ der Wissenschaften*, Darmstadt 1981 (Erstausg. 1620), 1. Buch, Aphorismus 129 (S. 96).

¹⁴ *Von der Methode*, Zweiter Teil, 7.-10., Hamburg 1960 (Erstausgabe 1637).

wäre für DESCARTES jede wissenschaftliche Untersuchung ohne Fundament. So brachte er mit seinem Verständnis von Meta-Physik, von der *prima philosophia*, ein neues Thema in die Philosophie ein – die moderne Erkenntnistheorie als Frage nach der Reichweite, Sicherheit und Gewißheit von Wahrnehmung und Denken.¹⁵

Auf der Suche nach einem absolut gewissen Ausgangspunkt begann er gemäß der ersten Regel an allem zu zweifeln, was er bisher für wahr hielt, und alles, was nur leisen Anlaß zum Bezweifeln gab, so anzusehen, als sei es falsch. Insbesondere verwarf er dabei die Sinne als Quelle sicherer Erkenntnis (Sinnestäuschungen, Traum, der scheinbare Umlauf der Sonne ...). Aber auch wenn man an allem zweifelt, so muß es doch jemanden geben, die oder der da zweifelt, und jeder weitere Zweifel verstärkt noch diese Gewißheit über ein denkendes Wesen als Träger des Zweifels. Mit seinem berühmten *cogito, ergo sum* hat er den gesuchten, klar und deutlich zu erkennenden Ausgangspunkt: „Das Denken ist's, es allein kann von mir nicht getrennt werden. [...] Ich bin also genau nur ein denkendes Wesen [*res cogitans*], d. h. Geist [*mens*], Seele [*animus*], Verstand [*intellectus*], Vernunft [*ratio*] – [...] Ein denkendes Wesen! Was heißt das? Nun, – ein Wesen, das zweifelt, einsieht, bejaht, verneint, will, nicht will und das sich auch etwas bildlich vorstellt und empfindet.“¹⁶ Kurz: „Unter Denken verstehe ich alles, was derart in uns geschieht, daß wir uns seiner unmittelbar aus uns selbst bewußt sind. [...] Alle Bewußtseinsarten lassen sich [...] auf zwei zurückführen; [...] die Wirksamkeit des Verstandes [...] oder die Wirksamkeit des Willens. Das Wahrnehmen, das Einbilden und das reine Denken sind nur verschiedene Arten des Vorstellens, und das Begehren, Ablehnen, Behaupten, Verneinen und Zweifeln sind verschiedene Arten des Wollens.“¹⁷

Nach DESCARTES' Analyse gibt es zwei grundverschiedene, nach ihren Hauptattributen benannte Substanzen, die besagte *res cogitans* und weiter die *res extensa*. Die wesensbestimmende, nur ihr zukommende Eigenschaft der letzteren Substanz (deren reale Existenz außerhalb unseres Bewußtseins er über einen traditionellen Gottesbeweis postuliert) ist das Ausgedehntsein, denn nur das, nicht etwa Farbe, Härte, Geruch usw., bleibt unverändert, wenn ich beispielsweise ein Stück Wachs ans Feuer halte (und das kann ich durch reines *Nachdenken* erkennen), nur ihr – und nicht der *res cogitans* – kann man sinnvoll Eigenschaften wie Gestalt, Lage, Größe, Anzahl, Bewegung usw. zusprechen (gemäß der von ihm gewählten Methode gerade die primären, *geometrischen* Qualitäten, im Gegensatz zu den sekundären, sinnlichen). Von diesen Eigenschaften und den Grundprinzipien ihrer Verknüpfung finden wir in unserem Geist „angeborene Ideen“ (Vorstellungen) vor, die für uns klar und deutlich erkennbar sind. Zwar können wir uns irren, das liegt dann aber nicht an den Kräften unseres Verstandes, sondern daran, daß „unser Wille frei ist und wir nach Willkür vielem zustimmen oder nicht zustimmen können“, wenn wir uns also ent-

¹⁵Richard RORTY (1981, S. 149 ff.) weist sehr überzeugend darauf hin, daß dies eine nachträgliche systematische Deutung und nicht das zeitgenössische Selbstverständnis ist. Die Erkenntnistheorie als – auch für die anderen Wissenschaften – *fundamentale* philosophische Disziplin separierte und etablierte sich erst nach KANT. Vorher war sie Teil einer Argumentation im übergeordneten praktischen Interesse, bei DESCARTES, um die neuzeitlichen Wissenschaften und ihr anderes Naturverständnis durchzusetzen und gegen die Kirche zu verteidigen, bei LOCKE zur Durchsetzung eines den Grundsätzen des Liberalismus verpflichteten Staates, bei KANT, um die ethische Autonomie des Subjekts und den Vorrang der praktischen Vernunft zu begründen und so die Eigenständigkeit der Philosophie gegenüber der NEWTONSchen Wissenschaft zu sichern. In der vorliegenden Schrift sind erkenntnistheoretische Fragestellungen nur insofern interessant, als sie Aufschluß über die philosophischen Bemühungen geben, dem *Denken* auf die Spur zu kommen.

¹⁶*Meditationes de prima philosophia / Meditationen über die Grundlagen der Philosophie*, Zweite Meditation 6. u. 8., Hamburg 1959 (Erstausgabe 1641).

¹⁷*Die Prinzipien der Philosophie*, Erster Teil, 9. u. 32., Hamburg 1965 (Erstausgabe 1644).

schließen, uns an die Regeln zu halten, kann kein Irrtum auftreten. Auf dieser Grundlage unternimmt er es dann, mittels reiner Vernunftkenntnis in einem groß angelegten Werk die gesamte Körperwelt in den Begriffen der Mechanik (Korpuskeln, Stoß und Druck, d. h. anders als später NEWTON ohne „leeren Raum“ und Fernwirkungen) zu erklären. Am Ende resümiert er stolz, nun sei „die ganze Einrichtung der Welt aus wenigen Prinzipien hergeleitet [. . .], aus den obersten und einfachsten Prinzipien der menschlichen Erkenntnis in ununterbrochener Folge abgeleitet worden“.¹⁸

Mit seinem radikalen, revolutionären Neubeginn, mit seiner Einsetzung des einzelnen Subjekts als Legitimationsinstanz jeder Erkenntnis hat DESCARTES den wesentlichen Keim geschaffen für unseren modernen Bewußtseinsbegriff (auch wenn seine Gleichsetzung von Bewußtsein mit dem gesamten Bereich des Mentalen heute so nicht mehr allgemein akzeptiert wird). Er setzt sich damit ab vom aristotelischen Begriff der Seele (die sich nach ARISTOTELES beim Menschen nur graduell, durch den Besitz höherer und erweiterter Fähigkeiten – z. B. zur Erkenntnis von Universalien – von der anderer Lebewesen – die auch wahrnehmen und empfinden können – unterscheidet), der sich am Phänomen des *Lebens*, am Gegensatz von belebter und unbelebter oder natürlicher und künstlicher Welt orientierte. Dagegen setzt er die Dualität von Geist und Körper als *Innen* und *Außen* des Bewußtseins. Trotz des offensichtlichen und ungelösten Problems, wie denn diese beiden grundverschiedenen Substanzen aufeinander wirken können (Leib-Seele-Problem), wurde diese strikte Trennung von Subjekt und Objekt, verbunden mit der Sicht der körperlichen Außenwelt als mechanistisches und deterministisches „Uhrwerk“ sowie der festen Überzeugung, alle wesentlichen Komponenten und Prozesse dieser Objektwelt in der Symbolik der Mathematik adäquat beschreiben und durch Anwendung von Gesetzen der Mathematik auch vorhersagen und damit beherrschen zu können, in der Folge zur Standardtheorie des mechanistischen Weltbildes. Für DESCARTES gab es keinen Unterschied zwischen der „künstlichen“ Mechanik und der „natürlichen“ Physik¹⁹ wie bei ARISTOTELES, keinen zwischen den damals modernen mechanischen Spielzeugen einerseits und Pflanzen, Tieren, dem Körper des Menschen andererseits.²⁰ Die entscheidende Trennlinie war für ihn die *res cogitans*, Bewußtsein und die Fähigkeit zum Denken, die er Tieren und mechanischen Automaten in gleicher Weise absprach:

¹⁸A. a. O., Erster Teil, 39., Vierter Teil, 205. u. 206.

¹⁹„Denn es gibt in der Mechanik keine Gesetze, die nicht auch in der Physik gälten, von der sie nur ein Teil oder eine Unterart ist, und es ist daher der aus diesen und jenen Rädern zusammengesetzten Uhr ebenso natürlich, die Stunden anzuzeigen, als es dem aus diesem oder jenem Samen aufgewachsenen Baum natürlich ist, diese Früchte zu tragen. So wie nun die, welche in der Betrachtung der Automaten geübt sind, aus dem Gebrauche einer Maschine und einzelner ihrer Teile, die sie kennen, leicht abnehmen, wie die anderen Teile, die sie nicht sehen, gemacht sind, so habe auch ich versucht, aus den sichtbaren Wirkungen und Teilen der Naturkörper zu ermitteln, wie ihre Ursachen und unsichtbaren Teilchen beschaffen sind.“ (a. a. O., Vierter Teil, 203.)

²⁰„Dies wird dem keineswegs sonderbar vorkommen, der weiß, wie viele verschiedene *Automaten* oder bewegungsfähige Maschinen menschliche Geschicklichkeit zustandebringen kann, und dies unter Verwendung nur sehr weniger Einzelteile verglichen mit der großen Anzahl von Knochen, Muskeln, Nerven, Arterien, Venen und all den anderen Bestandteilen, die sich im Leibe jedes Tieres finden. Er wird diesen Leib für eine Maschine ansehen, die aus den Händen Gottes kommt und daher unvergleichlich besser konstruiert ist und weit wunderbarere Getriebe in sich birgt als jede Maschine, die der Mensch erfinden kann“ (*Von der Methode*, Fünfter Teil, 9.). Hier wird aus der Metapher ein Modell: „Der Automat ist sozusagen der Platzhalter des verschwundenen Lebens. Was einmal in der spektakulären Absicht geschaffen worden war, Lebendiges mit mechanischen Mitteln zu imitieren, wird nun zum theoretischen Repräsentanten eines Lebens, das keines mehr sein soll“ (A. SUTTER (1988), S. 52). Dasselbe kann man heute vom *Computer* und vom *Geist* sagen.

Wenn es Maschinen mit den Organen und der Gestalt eines Affen oder eines anderen vernunftlosen Tieres gäbe, so hätten wir gar kein Mittel, das uns nur den geringsten Unterschied erkennen ließe zwischen dem Mechanismus dieser Maschinen und dem Lebensprinzip dieser Tiere; gäbe es dagegen Maschinen, die unseren Leibern ähnelten und unsere Handlungen insoweit nachahmten, wie dies für Maschinen wahrscheinlich möglich ist, so hätten wir immer zwei ganz sichere Mittel zu der Erkenntnis, daß sie deswegen keineswegs wahre Menschen sind. Erstens könnten sie nämlich niemals Worte oder andere Zeichen dadurch gebrauchen, daß sie sie zusammenstellen, wie wir es tun, um anderen unsere Gedanken bekanntzumachen. Denn man kann sich zwar vorstellen, daß eine Maschine so konstruiert ist, daß sie Worte und manche Worte sogar bei Gelegenheit körperlicher Einwirkungen hervorbringt, die gewisse Veränderungen in ihren Organen hervorrufen, wie zum Beispiel, daß sie, berührt man sie an irgendeiner Stelle, gerade nach dem fragt, was man ihr antworten will, daß sie, berührt man sie an einer anderen Stelle, schreit, man täte ihr weh und ähnliches; aber man kann sich nicht vorstellen, daß sie die Worte auf verschiedene Weisen zusammenordnet, um auf die Bedeutung alles dessen, was in ihrer Gegenwart laut werden mag, zu antworten, wie es der stumpfsinnigste Mensch kann. Das zweite Mittel ist dies: Sollten diese Maschinen auch manches ebensogut oder vielleicht besser verrichten als irgendeiner von uns, so würden sie 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 Handlung einer besonderen Einrichtung bedürfen; was es unwahrscheinlich macht, daß es in einer einzigen Maschine genügend verschiedene Organe gibt, die sie in allen Lebensfällen so handeln ließen, wie uns unsere Vernunft handeln läßt.²¹

Der strikte Dualismus DESCARTES', die völlige Trennung von Geist und Materie, sein mechanistisches Weltbild, beeinflussen bis heute unser Denken über das Denken (und die Ablehnung der Vorstellung einer künstlichen Intelligenz, besonders im kontinentalen Europa). Es mag daher erstaunen, daß er zu einer Klärung der Prozesse des Denkens nicht viel beigetragen hat. Trotz (oder wegen: weil sich das Subjekt der *res cogitans* unmittelbar bewußt ist, bedarf es keiner genaueren wissenschaftlichen Klärung²²) seines *cogito* als Antwort auf die radikale Frage nach der Gewißheit jeder Erkenntnis widmet sich der Haupt-

²¹ *Von der Methode*, Fünfter Teil, 10. Entscheidende, exklusive und nicht mechanisierbare Merkmale menschlichen Denkens sind für ihn also das Erfassen der *Bedeutung* in der Sprache und die *Universalität* der Vernunft, die Flexibilität unserer Intelligenz. DESCARTES' Erklärungen organischer Vorgänge waren nicht wirklich überzeugend, die damaligen Kenntnisse der Physiologie waren völlig unzureichend. Ein wichtiges Argument gegen die „Aristoteliker“ zur Unterstützung seiner These von der mechanischen Erklärbarkeit der gesamten Objektwelt war für ihn die Beschreibung tatsächlicher oder fiktiver leistungsanaloger Automaten, die er für den Bereich des Mentalen für unmöglich hielt. In unseren Zeiten, wo man glaubt, Verstehen von Bedeutung auf Verarbeitung von Information reduzieren zu können, und in der mit dem Computer eine Universalmaschine zur Verfügung steht, sehen das manche anders, tatsächliche oder fiktive leistungsanaloge Computerprogramme lieferten demnach heute modellhaft Argumente gegen das *Apriori*-Verdikt der Cartesianer. „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. [...] Denn wenn im Rahmen dieser Forschung [KI] 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.“ (A. BECKERMANN (1989), S. 111 und 113)

²² „Im Gegensatz zu der Weisheitstradition [...] wird das mentale Innenleben nicht differenziert als Teil des sozialen Verkehrs analysiert und bedacht, sondern allein als innerlicher Repräsentationsraum für die Geschehnisse der Welt im allgemeinen. Dies markiert das Paradox, daß sich bei dem Philosophen, der dem subjektiven Bewußtsein eine Zentralstellung gibt, sehr wenig über das subjektive Bewußtsein lernen läßt.“

teil seines Werkes der methodisch gesicherten, mathematisch-mechanischen Erklärung der Naturerscheinungen, der Objektwelt, nicht so sehr der Untersuchung des menschlichen Verstandes.

2.3 Locke: Operieren mit Ideen

Das hingegen war die selbstgestellte Aufgabe von John LOCKE (1632–1704). Angesichts vieler fruchtloser Debatten kam ihm „der Gedanke, daß wir einen falschen Weg eingeschlagen hätten und vor Beginn solcher Untersuchungen notwendig unsere eigenen geistigen Anlagen prüfen und zusehen müßten, mit welchen *Objekten* sich zu befassen unser Verstand tauglich sei.“²³ Diese Objekte des Verstandes bezeichnet er als *Ideen*, sie stehen für „*Phantasma, Begriff, Vorstellung, oder was immer es sei, das den denkenden Geist beschäftigen kann.*“ Anders als DESCARTES bestreitet er die Existenz angeborener Ideen, nicht einmal für die obersten Prinzipien der Logik, die Sätze von der Identität und vom verbotenen Widerspruch, läßt er das gelten, u. a. weil „alle Kinder und Idioten nicht im geringsten eine Vorstellung oder einen Gedanken von diesen Sätzen haben.“²⁴ Der Annahme angeborener Ideen stellt er das Credo des Empirismus entgegen:

Nehmen wir also an, der Geist sei, wie man sagt, ein unbeschriebenes Blatt, ohne alle Schriftzeichen, frei von allen Ideen; wie werden ihm diese dann zugeführt? Wie gelangt er zu dem gewaltigen Vorrat an Ideen, womit ihn die geschäftige schrankenlose Phantasie des Menschen in nahezu unendlicher Mannigfaltigkeit beschrieben hat? Woher hat er all das *Material* für seine Vernunft und für seine Erkenntnis? Ich antworte darauf mit einem einzigen Worte: aus der *Erfahrung*. Auf sie gründet sich unsere gesamte Erkenntnis, von ihr leitet sie sich schließlich her. Unsere Beobachtung, die entweder auf äußere sinnlich wahrnehmbare Objekte gerichtet ist oder auf innere Operationen des Geistes, die wir wahrnehmen und über die wir nachdenken, liefert unserm Verstand das gesamte *Material* des Denkens. Dies sind die beiden Quellen der Erkenntnis, aus denen alle Ideen entspringen, die wir haben oder naturgemäß haben können.²⁵

Die erste Quelle, die Sinneswahrnehmung der Außenwelt, nennt LOCKE *Sensation*. Ihr verdanken wir die meisten unserer Ideen wie z. B. *gelb, kalt, weich, süß*. Die andere Quelle – durch sie als „inneren Sinn“ gewinnen wir Ideen wie *wahrnehmen, denken, zweifeln, glauben, schließen, erkennen, wollen* – nennt er *Reflexion*, „die Wahrnehmung der Operationen des eigenen Geistes in uns [...]. Den Ausdruck *Operationen* gebrauche ich hier in einem weiten Sinne, da er nicht nur die aktiven Einwirkungen des Geistes auf seine Ideen bezeichnet, sondern auch bestimmte, bisweilen durch sie herbeigeführte passive Zustände, wie zum Beispiel die aus irgendeinem Gedanken entspringende Zufriedenheit oder Unruhe.“²⁶ Aus dem Material der einfachen Ideen, die der Geist bloß passiv aufnimmt (sei es

Descartes scheint zwar an der *res cogitans* als Gegenbegriff zur *res extensa* und aus erkenntnistheoretischen Gründen interessiert, aber nicht am individuellen Bewußtsein als solchem. Damit hängt auch zusammen, daß es sich beim Cartesianischen Ich keineswegs um ein biographisches oder überhaupt individuelles Ich handelt, sondern um ein systematisches Ich der allgemeinen Erkenntnistheorie.“ (H. HASTEDT (1991b), S. 647f.)

²³ *Versuch über den menschlichen Verstand*, Sendschreiben an den Leser, Hamburg 1981, S. 7 (Erstausgabe 1689)

²⁴ A. a. O., Einleitung, 8., Erstes Buch, I, 5.

²⁵ A. a. O., Zweites Buch, I, 2.

²⁶ A. a. O., I, 4.

aus einem oder mehreren Sinnen, sei es aus der Reflexion oder aus beiden), kann der Geist durch aktive, selbständige Handlungen weitere, komplexe Ideen bilden:

Diese Tätigkeiten, bei denen der Geist seine Macht über seine einfachen Ideen entfaltet, sind vornehmlich die folgenden drei: 1. Das Kombinieren mehrerer einfacher Ideen zu einer zusammengesetzten. Auf diese Weise entstehen sämtliche *komplexe Ideen*. 2. Die zweite Tätigkeit besteht darin, zwei Ideen, seien es einfache oder komplexe, so zusammenzustellen, daß man sie zu gleicher Zeit überblickt, ohne sie doch zu einer einzigen zu verschmelzen. Auf diese Weise erlangt der Geist alle seine *Ideen von Relationen*. 3. Die dritte Tätigkeit besteht in der Trennung einer Idee von allen anderen Ideen, die sie in ihrer realen Existenz begleiten. Dies Verfahren nennen wir Abstraktion; dadurch werden alle *allgemeinen Ideen* des Geistes gebildet.²⁷

Schärfer noch als bei DESCARTES ist für LOCKE das *Bewußtsein* der Ideen und Denkopoperationen zentraler Begriff, nur seine fortdauernde Identität liefert die Berechtigung, von einer *Person*²⁸, von einem *Ich* zu sprechen (wobei er anders als DESCARTES nicht notwendig eine *res cogitans* postuliert): „Das *Ich* ist das bewußt denkende Wesen, gleichviel aus welcher Substanz es besteht (ob aus geistiger oder materieller, einfacher oder zusammengesetzter), das für Freude und Schmerz empfindlich und sich seiner bewußt ist, das für Glück und Unglück empfänglich ist und sich deshalb soweit um sich selber kümmert, wie jenes Bewußtsein sich erstreckt.“²⁹

Zwischen seiner Untersuchung über den Ursprung und die Zusammensetzung der Ideen (Zweites Buch) und den Umfang und die Zuverlässigkeit der Erkenntnis (Viertes Buch) widmet sich LOCKE mit einer damals ungewöhnlichen Ausführlichkeit der Sprache, weil er feststellen mußte, daß „unsere Erkenntnis zu den Wörtern in einer so engen Beziehung steht, daß nur wenige klare und zutreffende Aussagen über die Erkenntnis möglich sind, ohne vorher genau zu erforschen, was die Wörter leisten und in welcher Art sie die Dinge bezeichnen.“³⁰ Wie schon HOBBS war er strikter Nominalist, für ihn sind Wörter „sinn-

²⁷A. a. O., XII, 1. Diese Unterscheidung in völlig passiv empfangene einfache Ideen und aktiv gebildete komplexe Ideen läßt sich kaum halten. Schon früh wurde eingewandt, daß auch eine „pure“ Farbvorstellung sich aus der Wahrnehmung eines Objektes und aus dem Vergleich mit einem anderen Objekt nur durch („aktives“) Weglassen anderer Eigenschaften der Objekte ergibt, d. h. durch Abstraktion (und daß somit diese Fähigkeit des Verstandes schon vor jeder Idee vorhanden sein muß).

²⁸„Meiner Meinung nach bezeichnet dieses Wort ein denkendes, verständiges Wesen, das Vernunft und Überlegung besitzt und sich selbst als sich selbst betrachten kann. Das heißt, es erfaßt sich als dasselbe Ding, das zu verschiedenen Zeiten und an verschiedenen Orten denkt. Das geschieht lediglich durch das Bewußtsein, das vom Denken untrennbar ist und, wie mir scheint, zu dessen Wesen gehört. Denn unmöglich kann jemand wahrnehmen, ohne *wahrzunehmen*, daß er es tut. Wenn wir etwas sehen, hören, riechen, schmecken, fühlen, überlegen oder wollen, so wissen wir, daß wir das tun. Das gilt jederzeit hinsichtlich unserer gegenwärtigen Sensationen und Wahrnehmungen; jeder wird dadurch für sich selbst zu dem, was er *sein eigenes Ich* nennt“ (a. a. O., XXVII, 9.). Ob allerdings irgendwer oder irgendwas – sei es Mensch, Tier oder Maschine – ein solches „Wesen“ ist, kann von außen weder *a priori* noch empirisch entschieden werden, letztlich handelt es sich um einen Akt der *Anerkennung*: Person „ist ein juristischer Ausdruck, der sich auf Handlungen und ihren Lohn bezieht; er findet also nur bei vernunftbegabten Wesen Anwendung, für die es Gesetze geben kann und die glücklich und unglücklich sein können.“ (XXVII, 26.)

²⁹A. a. O., XXVII, 17.

³⁰A. a. O., Drittes Buch, IX, 21. Er fährt dort fort: „Denn die Erkenntnis, deren Gegenstand die Wahrheit ist, hat es stets mit Sätzen zu tun. Obwohl ihr letztes Ziel zwar in den Dingen selbst liegt, so sind doch die Wörter dabei so sehr als Vermittler notwendig, daß sie von unserer allgemeinen Erkenntnis kaum trennbar zu sein scheinen. Wenigstens drängen sie sich so zwischen unseren Verstand und die Wahrheit, die dieser betrachten und begreifen will, daß Dunkelheit und Verworrenheit, einem Medium gleich, durch das sichtbare Gegenstände hindurchdringen, nicht selten unseren Blick trüben und den Verstand irreführen.“

lich wahrnehmbare Kennzeichen der Ideen [...]; die Ideen, für die sie stehen, machen ihre eigentliche und unmittelbare Bedeutung aus.“ Und dabei gilt: „ihre Bedeutung ist auf die Ideen dessen beschränkt, der sie gebraucht, und für nichts anderes können sie als Zeichen dienen.“³¹ Neben mehrdeutiger Verwendungsweise (d. h. die Sprecher verwenden gleiche Namen für unterschiedliche bzw. unterschiedliche Namen für gleiche eigene Ideen) ist für ihn die hauptsächliche Quelle für den „Mißbrauch der Wörter“, „der ihre Bedeutung unweigerlich verdunkelt und verwirrt“, daß die Menschen insgeheim voraussetzen, „daß ihre Wörter auch Kennzeichen der Ideen im Geiste anderer sind, mit denen sie sich unterhalten“ und „daß die Wörter auch die Realität der Dinge vertreten.“³²

Zusammenfassend gilt: „Die Erkenntnis scheint mir nichts anderes zu sein als die Wahrnehmung des Zusammenhangs und der Übereinstimmung oder der Nichtübereinstimmung und des Widerstreits zwischen irgendwelchen von unseren Ideen. Allein darin besteht die Erkenntnis.“ Und: „Wahrheit im eigentlichen Sinne des Wortes scheint mir nichts anderes zu bedeuten als die Verbindung und Trennung von Zeichen, je nachdem die durch sie bezeichneten Dinge miteinander übereinstimmen oder nicht. Unter Verbindung oder Trennung von Zeichen verstehe ich hier das, was man mit einem andern Namen Satz nennt.“³³

Nach LOCKE können wir uns unser Denken als eine Art Theater vorstellen, zu dem wir als bewußtes Ich über einen inneren Sinn (*the mind's eye*) exklusiven Zugang haben und in dem irgendwelche „Ideen“ genannte Akteure ein Stück aufführen, von dem wir in einer eigenartigen Personalunion gleichzeitig Zuschauer, Mitspieler, Autor, Regisseur und Kritiker sind.³⁴ Über diese Aufführung können wir möglichst akribische Protokolle verfassen, um sie dann mit anderen auszutauschen, von denen wir aufgrund unseres Stückes annehmen, daß sie über ein ähnliches Theater verfügen. Aber wovon handelt eigentlich das Stück, wovon sprechen die Akteure?

³¹A. a. O., II, 1. u. 8. Damit wendet er sich gegen die Lehre von der realen Existenz der sogenannten Universalien („Realisten“). Wir bestimmen „die Arten der Dinge nach *unseren* abstrakten Ideen [...]; wenn sich aber jemand nach den angeblich *realen* Wesenheiten richten will, so wird er, fürchte ich, in Verlegenheit geraten; er wird niemals imstande sein, genau zu wissen, wann etwas aufhört, der Art Pferd oder Blei anzugehören.“ (III, 13.) Die Außenwelt gibt uns keine „natürliche“ Ordnung der Dinge vor, die Wahl der Namen ist prinzipiell völlig willkürlich, allerdings gibt es Konventionen, wenn in einer menschlichen Gemeinschaft „der Sprachgebrauch bekannte Namen mit bestimmten Ideen verknüpft hat.“ (VI, 51.) Dabei ist es möglich, „daß die Bewohner eines Landes durch Sitten und Lebensweise veranlaßt worden sind, gewisse komplexe Ideen zu bilden und ihnen Namen zu geben, die andere niemals zu spezifischen Ideen zusammenfaßten.“ (V, 8.)

³²A. a. O., II, 4. u. 5. Er plädiert daher für einen sorgfältigeren und bewußteren Sprachgebrauch, ist aber Realist genug, nicht „zu glauben, daß sich ein einzelner Mensch anmaßen könnte, eine vollkommene Reform aller Sprachen der Welt, ja auch nur seiner eigenen Muttersprache durchzuführen, ohne sich lächerlich zu machen.“ (XI, 2.) Allerdings wäre es „wünschenswert, daß Männer, die in physikalischen Untersuchungen bewandert und mit den verschiedenen Arten der natürlichen Körper vertraut sind, diejenigen einfachen Ideen aufzeichneten, in denen ihrer Beobachtung zufolge die Einzelwesen einer Art regelmäßig übereinstimmen.“ (XI, 25.) Dieser Vorschlag wurde gut zweihundert Jahre später wieder aufgegriffen (in der Tat fast nur von Männern, vgl. unten S. 46).

³³A. a. O., Viertes Buch, I, 2. u. V, 2.

³⁴Eigentlich kann LOCKE dieses Bild nicht dem Leser vermitteln wollen: die sinnlichen *Eindrücke* wirken bei ihm auf das weiße Papier des Verstandes ein wie die Düsen meines Druckers, und auf *dieser Ebene* will er das Denken erklären. Durch die Hintertür führt er aber das Ich als *Betrachter* dieser Zeichen ein, d. h. als ein Etwas, das Eindrücke nicht nur *hat*, sondern auch noch *beurteilt* (R. RORTY (1981, S. 158 ff.) nennt das seine „Kontamination von Erklärung und Rechtfertigung“). Um dieser Verwirrung zu entgehen, beschränkte HUME sich auf die Erklärung, KANT hingegen wollte auch Rechtfertigung vor dem „Tribunal der reinen Vernunft“ und verteilte daher die Rollen und Aufgaben präziser.

Zwar gilt für LOCKE das Prinzip: „Der Geist hat bei allem Denken und Folgern kein anderes unmittelbares Objekt als seine eigenen Ideen; er betrachtet nur sie und kann nur sie betrachten. Daher ist es offenbar, daß es unsere Erkenntnis lediglich mit unseren Ideen zu tun hat.“³⁵ Aber er sieht selbstverständlich selbst, damit dem Einwand ausgesetzt zu sein, daß er damit „nur an einem Luftschloß baue“, daß so „die Visionen eines Enthusiasten und die Schlußfolgerungen eines nüchternen Mannes gleichermaßen gewiß“ seien. Er antwortet darauf, daß er den Nachweis gebracht habe, daß der Geist die einfachen Ideen „unter keinen Umständen selbst in sich erzeugen kann. Sie müssen daher notwendigerweise das Produkt von Dingen sein, die auf natürlichem Wege auf den Geist einwirken“.³⁶ Dieses Ergebnis ging schon als Voraussetzung in seine Beschreibung der Sensation ein: „Wenn unsere Sinne mit bestimmten sinnlich wahrnehmbaren Objekten in Berührung treten, so führen sie dem Geist eine Reihe verschiedener Wahrnehmungen von Dingen zu, *die der mannigfach verschiedenen Art entsprechen*, wie jene Objekte auf die Sinne einwirken.“³⁷ Das widerspricht klar dem genannten Prinzip, aber seine Bewunderung für das Werk von BOYLE, HUYGENS und insbesondere NEWTON war zu groß, als daß er es zu bloßen Hirngespinnsten machen wollte.³⁸ Mit seinem programmatischen Verzicht auf Spekulationen und apriorische Prinzipien, mit seiner Betonung der Erfahrung und des „gesunden Menschenverstandes“, mit seiner Forderung nach gemeinsamer Klärung der sprachlichen Mittel (damit man nicht um bloße Laute streitet), hatte LOCKE eine große Wirkung auf die gesamte anglo-amerikanische Tradition der Philosophie des Geistes bis hin zur Gegenwart.

2.4 Hume: Die Mechanik des Geistes

David HUME (1711–1776) ging LOCKES Weg konsequent zu Ende. Er wollte „die geheimen Triebfedern und Prinzipien entdecken [...], durch welche die Operationen des menschlichen Geistes ausgelöst werden“, so wie NEWTON „die Gesetze und Kräfte bestimmt hat, die den Umlauf der Planeten regieren und lenken“ und wie es „in anderen Bereichen der Natur vollbracht“ wurde.³⁹ Begrifflich schärfer als LOCKE unterschied er bei unseren Vorstellungen (*perceptions*) zwischen den unmittelbaren „Eindrücken“ (*impressions*), „wenn wir hören, sehen, fühlen, lieben, hassen, begehren oder wollen“, und den schwächeren „Gedanken“ (*thoughts, ideas*), „welche die weniger lebhaften Perzeptionen sind, deren wir uns bewußt sind, wenn wir auf eine der obenerwähnten Wahrnehmungen oder Gemütsbewegungen reflektieren“, als „Abbilder unserer Eindrücke“. Das ist „der ganze Stoff des Denkens“, die „ganze schöpferische Kraft des Geistes [besteht] nur in dem Vermögen, das uns durch die Sinne und Erfahrung gegebene Material zu verbinden, zu transponieren, zu vermehren oder zu verringern“, und – so lautete seine Herausforderung an die philosophische Welt – „diejenigen, die behaupten wollten, daß dieser Satz nicht allgemein und ausnahmslos wahr sei,“ sollten „zu seiner Widerlegung [...] jene Vorstellung vorweisen,

³⁵A. a. O., I, 1.

³⁶A. a. O., IV, 1. u. 4.

³⁷A. a. O., Zweites Buch, I, 3. (Hervorhebung von mir)

³⁸„Bisher ist es noch niemandem gelungen, eine Philosophie zu erfinden, die sowohl glaubwürdig ist als auch keinen Widerspruch in sich enthält. LOCKE strebte Glaubwürdigkeit an und erreichte sie auf Kosten der Konsequenz. Die meisten großen Philosophen haben das Gegenteil getan. Eine Philosophie ohne innere Konsequenz kann nicht ganz wahr sein, wohl aber kann eine in sich konsequente Philosophie vollkommen falsch sein.“ (RUSSELL (1975), S. 623)

³⁹*Eine Untersuchung über den menschlichen Verstand*, I. Abschn. Stuttgart 1982, S. 28 (Erstausg. 1748).

die [...] nicht aus dieser Quelle stammt.“⁴⁰ Was läßt sich aus diesem Stoff schneiden, wie kann das Material verbunden werden, was kann gesagt werden über „die Assoziation der Vorstellungen“? „Für mich ergeben sich nur drei Prinzipien der Vorstellungsverknüpfung, nämlich *Ähnlichkeit*, raum-zeitliche *Berührung* und *Ursache* oder *Wirkung*.“⁴¹

Weiter unterscheidet er: „Alle Gegenstände menschlichen Denkens und Forschens lassen sich naturgemäß in zwei Arten gliedern, nämlich in *Vorstellungsbeziehungen* (*Relations of Ideas*) und in *Tatsachen* (*Matters of Fact*).“ Zur ersten Art gehören – neben direkten Vergleichen von Eindrücken wie *rot* ist nicht *grün* – vor allem die Sätze der Mathematik wie „drei mal fünf ist der Hälfte von dreißig gleich“ (die Mathematik handelt nicht von der Wirklichkeit, sie spricht nur von dem, was sie sich selbst geschaffen hat). „Sätze dieser Art lassen sich durch bloße Denktätigkeit entdecken, unabhängig davon, ob irgendwo im Weltall etwas existiert.“ Die Wahrheit von Tatsachen ist nie so evident, ihr Gegenteil ist – rein logisch gesehen – immer möglich: „*Daß die Sonne morgen nicht aufgehen wird*, ist ein nicht minder einsichtiger Satz und enthält keinen größeren Widerspruch als die Behauptung, *daß sie aufgehen wird*.“ Wenn es also nicht möglich ist, seine Falschheit im strengen Sinne zu beweisen, so gilt es, „das Wesen jener Evidenz zu untersuchen, die uns jedes wirklich Existierenden und jeder Tatsache versichert, die über das gegenwärtige Zeugnis der Sinne oder die Angaben unseres Gedächtnisses hinausgehen. [...] Alle Tatsachen betreffenden Vernunftbegründungen scheinen auf der Beziehung von *Ursache und Wirkung* zu beruhen. [...] Wollen wir somit eine zufriedenstellende Erklärung für das Wesen jener Evidenz der Gewißheit von Tatsachen erlangen, haben wir zu untersuchen, wie wir zur Erkenntnis von Ursache und Wirkung kommen.“⁴² Nun folgt seine berühmte Analyse der Kausalität, der „skeptische Zweifel an den Verstandestätigkeiten“ und die „skeptische Lösung dieser Zweifel“:

Ich wage es, den Satz als allgemeingültig und keine Ausnahme duldend aufzustellen, daß die Kenntnis dieser Beziehung in keinem Falle durch Denkkakte *a priori* gewonnen wird, sondern ausschließlich aus der Erfahrung stammt, indem wir feststellen, daß gewisse Gegenstände immerdar miteinander verbunden sind. Man lege einem noch so klugen und fähigen Menschen einen Gegenstand vor; ist ihm dieser gänzlich fremd, wird er – trotz sorgfältigster Untersuchung seiner sinnfälligen Qualitäten – nicht fähig sein, irgendeine seiner Ursachen oder Wirkungen zu entdecken. [...] Die Schlüsse, die [die Vernunft] aus der Betrachtung eines Kreises zieht, sind die gleichen, die sie auf Grund einer Prüfung aller Kreise der Welt ziehen würde. Aber niemand, der nur einen Körper gesehen hat, wie er sich nach dem Stoß durch einen anderen bewegt, könnte schließen, daß sich jeder andere Körper, nach ähnlichem Stoß, bewegen wird. Alle Erfahrungsschlüsse sind somit Folgen der Gewohnheit, nicht der Vernunft. [...] Aller Glaube an Tatsachen oder wirkliche Existenz stammt lediglich von einem dem Gedächtnis oder den Sinnen gegenwärtigen Gegenstand und einer gewohnheitsmäßigen Verbindung zwischen diesem und irgendeinem anderen Gegenstand; oder, mit anderen Worten: Nachdem man gefunden hat, daß in vielen Fällen zwei Arten von Gegenständen – Feuer und Hitze, Schnee und Kälte – immer in Zusammenhang standen, wird der Geist, wenn Feuer oder Schnee sich erneut den Sinnen darbieten, aus

⁴⁰A. a. O., II. Abschnitt, S. 32 ff.

⁴¹A. a. O., III. Abschnitt, S. 39. Hier lohnt ein Vergleich der „Arten der Übereinstimmung oder Nichtübereinstimmung der Ideen“ bei LOCKE: „1. *Gleichheit oder Verschiedenheit*. 2. *Beziehung*. 3. *Koexistenz oder notwendiger Zusammenhang*. 4. *Reale Existenz*.“ (LOCKE, a. a. O. Viertes Buch, I, 3.) Man sieht, die „Reale Existenz“ der Substanzen ist – nach BERKELEY – bei HUME gestrichen. Auf den Unterschied bei 3. wird im folgenden Absatz eingegangen.

⁴²A. a. O., IV. Abschnitt, S. 41 ff.

Gewohnheit dazu gebracht, Hitze oder Kälte zu erwarten und zu *glauben*, daß es eine derartige Qualität gibt und sie sich nun bei eingehenderer Beschäftigung entdecken wird. Dieser Glaube ist das notwendige Resultat, wenn man den Geist in eine solche Lage bringt. Es ist ein seelischer Vorgang, der in dieser Lage ebenso unvermeidlich ist wie das Gefühl der Liebe, wenn wir Wohltaten empfangen, oder des Hasses, wenn uns Unrecht widerfährt. Alle diese Vorgänge sind eine Art natürlicher Instinkte, die keine Vernunfttätigkeit, d. h. kein Denk- oder Verstandesprozeß jemals hervorzubringen oder zu verhindern vermag.“⁴³

Wenn ich zum achten Mal in meinem Leben erlebe, daß auf einen Blitz ein Donner folgt, kann ich dieser gesammelten Erfahrung nicht mehr darüber entnehmen, wie oder warum Blitze Donner bewirken, als beim ersten Mal, hinzugekommen ist aber meine Erwartung, daß auch künftig auf einen Blitz der Donner folgt, daß sich auch hier die Natur als gleichförmig erweist. Auf die Frage, wie ich diese Erwartung rechtfertige, kann ich schlecht darauf verweisen, daß – wie die Erfahrung zeige – in der Natur gleiche (oder ähnliche) Ursachen gleiche (oder ähnliche) Wirkungen haben, denn das stand ja gerade in Frage. Erkenntnistheoretisch ist das natürlich eine Katastrophe,⁴⁴ über die Realität der Außenwelt, über die Existenz anderer denkender Wesen, ja, selbst über die durchgehende und einheitliche Identität des Ichs läßt sich nichts Sicheres sagen, alles, was wir haben, ist ein „Bündel von Perzeptionen“ und die Macht der Gewohnheit. Interessanter in unserem Zusammenhang sind aber die damit verbundenen Aussagen über die „Mechanik“ des Denkens. HUMES skeptische Analyse nimmt Abschied vom lieb gewordenen und mit Leidenschaft verteidigten Selbstverständnis eines aktiven Geistes, der durch bewußte und gerichtete Tätigkeit „Macht über seine einfachen Ideen entfaltet“. Die Perzeptionen, Eindrücke und Ideen, bewegen sich im Geiste nach bestimmten Gesetzen und auf Grund bestimmter Kräfte, ganz ähnlich *wie* sich die Planeten bewegen, *warum* das geschieht, ist eine sinnlose Frage.⁴⁵

⁴³A. a. O., IV. Abschnitt, S. 43f., V. Abschnitt, S. 63 u. 66f. Auch LOCKE behandelte schon die „Assoziation der Ideen“ als Folge der Gewohnheit, allerdings als *schlechte* Gewohnheit und Zufall, auf die er Vorurteile, Denkfehler und bestimmte Arten des Wahnsinns zurückführte (LOCKE, a. a. O., Zweites Buch, XXXIII). Diese gilt es aufzuklären, indem sie durch echtes *Wissen* ersetzt werden, das sich auf die Wahrnehmung des *notwendigen* Zusammenhangs zurückführen läßt. (Viertes Buch, I, 6.)

⁴⁴Und es nützt selbstverständlich nichts, wenn nun ein Physiker die *Kausalität* von Blitz und Donner als *notwendige* (und nicht bloß *zeitliche*) Folge mit Luftverdrängungen durch elektrische Entladungen und deren Wirkungen auf mein Trommelfell erklären will, denn wie bei jeder empirischen Wissenschaft sind dafür allgemeine, aus der Erfahrung durch *Induktion* gewonnene Gesetze nötig, und die Rechtfertigung von Induktionsschlüssen steht ja gerade in Frage. Bertrand RUSSELL bezeichnet HUMES Analyse als „Bankrott der Vernünftigkeit des achtzehnten Jahrhunderts“ und fährt fort: „Es ist deshalb wichtig, festzustellen, ob im Rahmen einer völlig oder hauptsächlich empirischen Philosophie eine Antwort auf Hume überhaupt möglich ist. Wenn nein, dann besteht zwischen geistiger Gesundheit und Geisteskrankheit kein Unterschied. Der Irrsinnige, der sich für ein „verlorenes Ei“ hält, ist nur deshalb zu verurteilen, weil er in der Minderheit ist, oder sagen wir lieber – da wir nicht unbedingt von der Demokratie ausgehen müssen –, weil die Regierung anderer Meinung ist als er. Das ist eine trostlose Auffassung, und wir müssen hoffen, ihr auf irgendeine Weise entrinnen zu können.“ (RUSSELL (1975), S. 682f.) Eine solche Antwort – nämlich eine Rechtfertigung der wahrheitserweiternden Induktion als Drittes neben bloßer, unmittelbarer Erfahrung und wahrheitserhaltender Deduktion – steht bis heute aus. Auch die mathematische Wahrscheinlichkeitstheorie hilft hier nicht weiter (denn die Aussage, daß bei einem gleichmäßigen Würfel die Chancen für eine Vier eins zu sechs sind, ist eine *relation of ideas*, keine *matter of fact*), erst recht nicht die (empirische) Statistik.

⁴⁵HUME ist damit Vorläufer einer Stimulus-Response-Psychologie, im weiteren Sinne jeder „naturalistischen“ Erklärung des Denkens, bei der wir „nicht mehr Herr im eigenen Haus“ sind – bis hin zur Soziobiologie. Hier tritt ein Problem auf: nach HUME sind das Feststellen von Kausalbeziehungen und die Berechtigung induktiver Schlüsse ein der Kraft der Gewohnheit geschuldeter Glaube, als das „notwendige

DESCARTES und LOCKE unternahmen den (in der Differenz häufig überschätzten) Versuch, die Autonomie des Menschen gegenüber Gott und der (als mechanisch gedachten) Natur zu sichern. Gemeinsam postulierten sie einen eigenen, nur dem Subjekt zugänglichen Bereich des Bewußtseins, des „Geistes“; unterschiedlich sind jeweils die bevorzugten Methoden und die einfachsten und gewissesten „Bausteine“ dieses Bereichs: die klaren und deutlichen „eingeborenen Ideen“ des Verstandes einerseits und die „einfachen Ideen“ der unmittelbaren Sinneswahrnehmung andererseits. Ihre jeweiligen Nachfolger versuchten, diesen Weg zu Ende zu gehen, und landeten doch wieder bei Gott oder der „Mechanik des Geistes“ als letzter Instanz. Den Rationalisten in der Tradition DESCARTES' gelang es zwar, als Leistung des Verstandes und weitgehend ohne Rückgriff auf die von ihnen als ungenau angesehene Erfahrung, aus wenigen Axiomen mit deduktiver Strenge ein beeindruckendes Gebäude in der Sprache der Mathematik und Mechanik zu errichten, sie mußten sich aber die reale Existenz des so Beschriebenen mit (z. T. abenteuerlichen) Glaubenssätzen über eingeborene Ideen sowie die Existenz und das Wirken Gottes erkaufen. Auf der anderen Seite bauten die Empiristen in der Tradition LOCKES auf die Erfahrung und den *Common sense* sowie auf die Analyse der Sprache, blieben dabei aber gefangen im skeptischen Zweifel, in der abgeschlossenen Welt des individuellen Bewußtseins (wobei auch dessen Kontinuität und Einheit mit kaum mehr als der Gewohnheit zu rechtfertigen war). Für den philosophischen Ehrgeiz war diese Situation äußerst unbefriedigend, und so gab es einige Versuche, sozusagen *the best of both worlds* zu vereinigen und zugleich die jeweiligen negativen Konsequenzen und Aporien zu vermeiden.

2.5 Kant: Verstand, Urteilskraft, Vernunft

Berühmt wurden die Bemühungen von Immanuel KANT (1724–1804). Den Empiristen gestand er zu:

Daß alle unsere Erkenntnis mit der Erfahrung anfangt, daran ist gar kein Zweifel; denn wodurch sollte das Erkenntnisvermögen sonst zur Ausübung erweckt werden, geschähe es nicht durch Gegenstände, die unsere Sinne rühren und teils von selbst Vorstellungen bewirken, teils unsere Verstandesfähigkeit in Bewegung bringen, diese zu vergleichen, sie zu verknüpfen oder zu trennen und so den rohen Stoff sinnlicher Eindrücke zu einer Erkenntnis der Gegenstände zu verarbeiten, die Erfahrung heißt? *Der Zeit nach* geht also keine Erkenntnis in uns vor der Erfahrung vorher und mit dieser fängt alle an.⁴⁶

Erkenntnis (*cognitio*) ist für KANT eine Vorstellung (*repraesentatio*), und zwar eine bewußte Vorstellung (*perceptio*), die sich auf ein Objekt bezieht (im Gegensatz zur subjektiven Empfindung (*sensatio*)). Meine Sinneseindrücke allein stellen noch keine Erkenntnis

Resultat“, wenn die Wahrnehmung eines Ereignisses und die in uns nach den Gesetzen der Ideenassoziation entstandene Erwartung auf die eines weiteren zusammentreffen; d. h. auf der *psychischen* Ebene nimmt er durchaus Kausalität an und erwartet, daß dieser Mechanismus auch künftig wirkt. (Das ist ein wenig so, als wenn jemand sagen würde, er oder sie habe nun das Geheimnis um Kausalität und Induktion endgültig entschleiert, es sei nicht anderes als die zunehmende Bildung und Verstärkung von Neuronenverbindungen, und auf Nachfrage, woher er oder sie das denn so genau wisse, antwortete: „Na ja, bei mir sind die Verbindungen von Neuronen entsprechend gebildet und verstärkt worden.“) Dieses Selbstanwendungsproblem besteht für viele Theorien (was haben Richard DAWKINS' Gene davon, wenn ihre „Kampfmaschine“ solche Bücher schreibt), die Reflexion der Reflexion führt zu diesen „seltsamen Schleifen“, die wir seit Douglas R. HOFSTADTER so lieben.

⁴⁶Kritik der reinen Vernunft, B 1, Darmstadt 1975 (Erstausgabe 1781, ²1787)

dar, erst muß ich in das gleichzeitig auf mich einstürmende Rohmaterial ordnend, abgrenzend, formend eingreifen, und zwar *denkend*, durch *Begriffe*, um mittels eines Merkmals, das mehreren Dingen gemeinsam sein kann, verschiedene Vorstellungen unter einer gemeinschaftlichen zu ordnen und so eine *Einheit* herzustellen, um in dieser Flut aus Farben, Tönen und Gerüchen z. B. eine *Billardkugel* zu identifizieren. Es gibt nach KANT also „zwei Stämme der menschlichen Erkenntnis [...], nämlich Sinnlichkeit und Verstand, durch deren ersteren uns Gegenstände *gegeben*, durch den zweiten aber *gedacht* werden.“ Im Gegensatz zur passiven, Vorstellungen bloß rezeptiv empfangenden Sinnlichkeit ist der *Verstand* „dagegen das Vermögen, Vorstellungen selbst hervorzubringen oder die *Spontaneität* des Erkenntnisses [...]. Keine dieser Eigenschaften ist der anderen vorzuziehen. [...] Gedanken ohne Inhalt sind leer, Anschauungen ohne Begriffe sind blind. [...] Nur daraus, daß sie sich vereinigen, kann Erkenntnis entspringen.“⁴⁷

Werden Begriffe in einem zweiten Schritt auf auf Anschauungen bezogene Begriffe bezogen, so entstehen *Urteile* genannte Vorstellungen von Vorstellungen, deren sprachlicher Ausdruck im einfachsten Fall Sätze sind, in denen von einem Subjekt ein Prädikat ausgesagt wird: *Die Billardkugel ist rot*. Das Verhältnis vom Subjekt zum Prädikat ist nun nach KANT „auf zweierlei Art möglich. Entweder das Prädikat *B* gehört zum Subjekt *A* als etwas, was in diesem Begriffe *A* (versteckter Weise) enthalten ist; oder *B* liegt ganz außer dem Begriff *A*, ob es zwar mit demselben in Verknüpfung steht. Im ersten Fall nenne ich das Urteil *analytisch*, in dem anderen *synthetisch*. [...] Die ersteren könnte man auch *Erläuterungs-*, die anderen *Erweiterungsurteile* heißen“.⁴⁸ Der genannte Satz drückt mit dieser Unterscheidung ein synthetisches Urteil aus, der Satz *Die Billardkugel ist rund* hingegen ein analytisches. Nur die Synthesis, nicht die Analyse, ist eine echte Erweiterung der Erkenntnis. Traditionell werden weiter empirische, aus der Erfahrung gewonnene Urteile *a posteriori* unterschieden von solchen *a priori*, die mit unbedingter Notwendigkeit und strenger Allgemeinheit gelten, die somit von jeder Erfahrung unabhängig sind, also durch Erfahrung auch nicht widerlegt werden können.

Unstrittig ist nun, daß es *synthetische Urteile a posteriori* („Diese Billardkugel ist rot“) und *analytische Urteile a priori* („Die Kugel ist rund“) gibt. Aber – und das ist seine Pointe und das unterscheidet ihn von seinen Vorgängern – nach KANT gibt es auch *synthetische Urteile a priori*! Als Beispiele nennt er u. a. Sätze aus der Mathematik („ $7 + 5 = 12$ “, „Eine gerade Linie zwischen zwei Punkten ist die kürzeste“) und der „reinen“ Naturwissenschaft („In allen Veränderungen der körperlichen Welt bleibt die Quantität der Materie unverändert“), denn diese Sätze gelten nicht nur notwendig und allgemein, sie erweitern auch unsere Erkenntnis, denn „der Begriff der Summe von 7 und 5“ enthält nichts weiter „als die Vereinigung beider Zahlen in eine einzige, wodurch ganz und gar nicht gedacht wird, welches diese einzige Zahl sei,“ denn der „Begriff vom *Geraden* enthält nichts von Größe, sondern nur eine Qualität“, denn „in dem Begriffe der Materie denke ich mir nicht die Beharrlichkeit, sondern bloß ihre Gegenwart im Raume durch die Erfüllung desselben.“ Diese Sätze sind KANTS Antwort auf HUMES Herausforderung, sie sind nicht bloße *Relations of Ideas* und keine Produkte der Gewohnheit. Wie ist nun solche echte Erweiterung der Erkenntnis, die notwendig, allgemeingültig, unabhängig von der Erfahrung ist, kurz: „*Wie sind synthetische Urteile a priori möglich?*“ Diese Klärung ist die „eigentliche Aufgabe der reinen Vernunft“,⁴⁹ der Schlüssel dazu findet sich in der Fortsetzung des ersten

⁴⁷A. a. O., B 29, B 75.

⁴⁸A. a. O., B 10 f.

⁴⁹A. a. O., B 15 ff., B 19.

o. g. Zitates:

Wenn aber gleich alle unsere Erkenntnis *mit* der Erfahrung anhebt, so entspringt sie darum doch nicht eben alle *aus* der Erfahrung. Denn es könnte wohl sein, daß selbst unsere Erfahrungserkenntnis ein Zusammengesetztes aus dem sei, was wir durch Eindrücke empfangen, und dem, was unser eigenes Erkenntnisvermögen (durch sinnliche Eindrücke bloß veranlaßt) aus sich selbst hergibt, welchen Zusatz wir von jenem Grundstoffe nicht eher unterscheiden, als bis lange Übung uns darauf aufmerksam und zur Absonderung desselben geschickt gemacht hat.

KANT war geübt und geschickt genug, den zusätzlichen Eigenbeitrag unseres Erkenntnisvermögens abzusondern. „Anschauung und Begriffe [...], die Elemente aller unserer Erkenntnis [...], sind entweder rein oder empirisch. *Empirisch*, wenn Empfindung (die die wirkliche Gegenwart des Gegenstandes voraussetzt) darin enthalten ist; *rein* aber, wenn der Vorstellung keine Empfindung beigemischt ist. Man kann die letztere die Materie der sinnlichen Erkenntnis nennen. Daher enthält reine Anschauung lediglich die Form, unter welcher etwas angeschaut wird, und reiner Begriff allein die Form des Denkens eines Gegenstandes überhaupt. Nur allein reine Anschauungen oder Begriffe sind *a priori* möglich, empirische nur *a posteriori*.“⁵⁰ Reine Anschauung, die bloße Form der Erscheinungen, gewinnt KANT zunächst durch Isolierung der Sinnlichkeit, durch Absonderung alles dessen, was der Verstand durch seine Begriffe denkt, und dann durch Abtrennung dessen, was zur Empfindung gehört. Übrig bleiben die Vorstellung des *Raumes* als Form des äußeren Sinnes, mit dem wir uns Gegenstände als außer uns vorstellen, und die Vorstellung der *Zeit* als Form des inneren Sinnes, des Anschauens unserer selbst und unseres inneren Zustandes, mit dem wir uns das Zugleichsein oder Aufeinanderfolgen vorstellen. Raum und Zeit sind nach KANT weder Begriffe (also durch Abstraktion, d. h. durch Weglassen anderer Eigenschaften entstanden) noch empirischen Ursprungs, sie sind Vorstellungen *a priori*, die als „Bedingung der Möglichkeit der Erscheinungen“ aller äußeren (Raum) bzw. „aller Erscheinung überhaupt“ (Zeit) notwendigerweise zu Grunde liegen. „Zeit und Raum sind demnach zwei Erkenntnisquellen, aus denen *a priori* verschiedene synthetische Erkenntnisse geschöpft werden können, wie vornehmlich die reine Mathematik in Ansehung der Erkenntnisse vom Raume und dessen Verhältnissen ein glänzendes Beispiel gibt.“⁵¹

Reine Begriffe als Formen des Denkens *a priori* erhält KANT aus der Tafel möglicher Urteilsformen (in Anlehnung an ARISTOTELES, durch dessen Arbeit er die Logik für im wesentlichen abgeschlossen hielt), denn Begriffe sind ja die Prädikate möglicher Urteile. So korrespondieren mit den Urteilsformen die *Kategorien* als reine Verstandesbegriffe, aufgeteilt nach *Quantität* (Einheit, Vielheit, Allheit), *Qualität* (Realität, Negation, Limitation), *Relation* (Inhärenz und Subsistenz, Kausalität und Dependenz, Gemeinschaft) und *Modalität* (Möglichkeit – Unmöglichkeit, Dasein – Nichtsein, Notwendigkeit – Zufälligkeit).⁵²

⁵⁰A. a. O., B 74f.

⁵¹A. a. O., B 39, B 50, B 55. Gelegentlich wird gesagt, die Physik des 20. Jahrhunderts habe KANT in diesem Punkt empirisch widerlegt. Mir scheint dagegen die Physik gewisse Probleme zu haben, die in Bezug auf die Zeit (d. h. Vergangenheit und Zukunft) völlig symmetrischen Gleichungen der klassischen Mechanik, der Relativitätstheorie und der Quantentheorie einerseits, den „Pfeil der Zeit“ der Thermodynamik andererseits und die zeitliche Gerichtetheit und Abhängigkeit unserer Wahrnehmung und unseres Denkens plausibel zusammenzubringen (vgl. Roger PENROSE (1991), S. 295 ff., 337, 345 ff., 362, 432 ff.). Ein anderer Hinweis darauf, daß für unser Denken die Zeit mehr als nur ein empirisches Problem darstellt, sind die Schwierigkeiten einer temporären bzw. nicht-monotonen Logik und die Umsetzung in Computerprogramme. Zum Problem der Zeitlichkeit vgl. auch Winfried D'AVIS (1994).

⁵²A. a. O., B 106.

Begriffe lassen sich als eine Art Regeln (nämlich zur Unterscheidung und Zusammenfassung) auffassen, um aber im Einzelfall zu wissen, welche Regel, d. h. welcher Begriff angewandt werden soll, bedarf es der *Urteilkraft* als „Vermögen, unter Regeln zu *subsumieren*, d. i. zu unterscheiden, ob etwas unter einer gegebenen Regel stehe oder nicht.“ Wollte man dafür wieder Regeln angeben, würde man das Problem nur verschieben, „und so zeigt sich, daß zwar der Verstand einer Belehrung und Ausrüstung durch Regeln fähig, Urteilkraft aber ein besonderes Talent sei, welches gar nicht belehrt, sondern nur geübt sein will. [...] Dieses ist auch der einzige und große Nutzen der Beispiele, daß sie die Urteilkraft schärfen.“⁵³ Um wirklich Erkenntnis zu liefern, um nicht „leer“ zu sein, müssen Begriffe und damit auch die Kategorien auf sinnliche Anschauungen bezogen werden. Als Vermittelndes zwischen den Kategorien und der Zeit als Anschauungsform des inneren Sinnes steht der Urteilkraft nach KANT zu jeder Kategorie ein *Schema* als Vorstellung „von einem allgemeinen Verfahren der Einbildungskraft, einem Begriff sein Bild zu verschaffen“, als „der sinnliche Begriff eines Gegenstandes in Übereinstimmung mit der Kategorie“ zur Verfügung,⁵⁴ eine Art inneres, schematisches, eine Zeitvorstellung enthaltendes Bild (z. B. ist das Schema der Quantität die *Zahl* als Vorstellung der sukzessiven Addition von Einheiten, das der Kausalität die *regelmäßige Abfolge*, das der Notwendigkeit das *Dasein zu jeder Zeit*).

Die reinen Anschauungsformen und die Kategorien, vermittelt durch die Schemata im Medium der Zeit, sind KANTS Antwort auf die Frage, wie synthetische Urteile der reinen Naturwissenschaft *a priori* möglich sind, sie liefern das „System der Grundsätze des reinen Verstandes“. Sie sind der gesuchte Eigenbeitrag unseres Erkenntnisvermögens zu jeder Erfahrungserkenntnis, aber sie limitieren auch den Bereich des objektiven, also gegenständlichen Wissens. Erkenntnis ist die Anwendung von Begriffen auf Anschauungen, in der sinnlichen Anschauung verfügen wir nur über *Erscheinungen*, über durch unsere Vorstellungen von Raum und Zeit notwendig geformte Empfindungen, das sind die „gegebenen Gegenstände“, die der Verstand mittels der Kategorien denken und damit erst *konstituieren* kann. Die „Bedingungen der *Möglichkeit der Erfahrung* überhaupt sind zugleich Bedingungen der *Möglichkeit der Gegenstände der Erfahrung* und haben darum objektive

⁵³ A. a. O., B 171 ff. Daß man Urteilkraft nicht lehren kann, daß man sie nicht in Regeln fassen kann, wurde den Entwicklern von Expertensystemen schmerzlich bewußt (und auch das „Rahmenproblem“ der KI, also die Entscheidung, welche Veränderungen einer Situation relevant sind und welche nicht, oder welches *frame* oder *script* in der konkreten Situation das angemessene ist, hängt damit zusammen). Dem Wert von Beispielen wird neuerdings durch *fallbasierte* (statt regelbasierte) Expertensysteme Rechnung getragen (vgl. Abschnitt 3.3 im Teil 1).

⁵⁴ A. a. O., B 179 f., B 186. Gelegentlich wird angemerkt, daß die Schemata und ihre Funktion ein wenig dunkel bleiben; das sah KANT ähnlich: „Dieser Schematismus unseres Verstandes in Ansehung der Erscheinungen und ihrer bloßen Form ist eine verborgene Kunst in den Tiefen der menschlichen Seele, deren wahre Handgriffe wir der Natur schwerlich jemals abraten und sie unverdeckt vor Augen legen werden.“ (B 180 f.) Auch im Zusammenhang mit empirischen Begriffen untersucht hier KANT ein bis heute aktuelles Problem. Um in der Vielfalt meiner Sinneseindrücke einen Vogel *als* Vogel zu erkennen, benötige ich eine Art geistiges Bild, um zwischen dem Begriff „Vogel“ (der viel mehr beinhaltet und viel abstrakter ist, als die sichtbare Gestalt) und der konkreten Anschauung vermitteln zu können; dabei muß andererseits dieses „Bild“ hinreichend *verschwommen* sein (und wohl auch nicht wirklich ein *Bild*), um eine „Resonanz“ sowohl mit Kolibris als auch mit Adlern, Hühnern, Pinguinen . . . zu haben. Die Ähnlichkeit des KANTSCHEN Schemas mit der „Familienähnlichkeit“ (statt exakter Definitionen) bei WITTGENSTEIN ist offensichtlich; aber zur Familie gehört auch das in der KI bisher nicht befriedigend gelöste Problem der mit einem Begriff verknüpften „stereotypen“ Beschreibung (vgl. HAUGELAND (1987), S. 171 ff.), d. h. eine Festlegung dessen, was man im „Normalfall“ erwartet (z. B. durch *frames* und *scripts*), ohne daß beim Fehlen einzelner Merkmale gleich alles zusammenbricht (denn „Charly, der Pinguin“ kann nun mal nicht fliegen . . .).

Gültigkeit in einem synthetischen Urteile *a priori*.“⁵⁵ D. h. – und das ist seine „kopernikanische Revolution“ – „die Gegenstände müssen sich nach unserer Erkenntnis richten,“ nicht die Erkenntnis nach den Gegenständen. Die Natur (der Naturwissenschaften) als gesetzmäßige Ordnung der Gegenstände unterliegt den Gesetzen unserer Erkenntnis, in den Wissenschaften *geben* wir der Natur ihre Gesetze (und machen dadurch „allererst eine Natur möglich“). Da Erfahrung gerade dadurch entsteht, daß der „rohe Stoff“ der Sinne durch die Kategorien geformt wird, ist es dann keine Überraschung, daß beispielsweise das Kausalitätsprinzip notwendig und unabhängig von der Erfahrung gilt.⁵⁶

Auch wenn wir uns die Gegenstände und Gesetze der Natur nach KANT – salopp gesagt – selbst machen, so bedeutet das keineswegs, daß sie bloße Hirngespinnste sind. Ausgehend vom Faktum, daß wir Erfahrungen machen, fragte KANT nach den Bedingungen, die unser Erkenntnisvermögen *a priori*, notwendigerweise und allgemein, bei allen Menschen in gleicher Weise, erfüllen muß, um überhaupt Erfahrungserkenntnis haben zu können (er nennt diese Untersuchung und ihre Ergebnisse *transzendental*, d. h. die Grenze sicheren Wissens, des Gebietes möglicher Erfahrung bestimmend, im Gegensatz zu „transzendent“, d. h. diese Grenze überschreitend). Diese Bedingungen sind allgemeingültig und nicht bloß subjektiv, sie gelten für alle Menschen und in Bezug auf alle Gegenstände der Erfahrung, sie haben also eine intersubjektiv gesicherte Objektivität. Empirische Realität können allerdings immer nur (in der Einheit der Begriffe gedachte) Erscheinungen (*Phaenomena*), Dinge „für uns“, nicht Dinge „an sich“ (als unabhängig von der sinnlichen Anschauung bloß vom Verstand gedachte „Gegenstände“: *Noumena*) haben; darüber, wie Dinge unabhängig von unserem Erkenntnisvermögen beschaffen sind, können wir keine Aussagen machen (außer der, daß sie keine bloßen Einbildungen sind, daß sie also unabhängig von uns existieren als ein „Etwas“, das unsere Sinne rührt).⁵⁷

⁵⁵A. a. O., B 197. Man beachte, daß vor „Bedingungen der *Möglichkeit der Gegenstände* . . .“ kein „die“ steht. Ein zeitgenössisches Wiederanknüpfen an KANT, das in voller Kenntnis des *Linguistic turn* und der naturalistischen Gegenangriffe erfolgt, stammt von Hilary PUTNAM (1982, S. 78, 81, 82): „[E]in Zeichen, das von einer bestimmten Gemeinschaft von Zeichenbenutzern auf bestimmte Weise verwendet wird, kann *innerhalb des Begriffsschemas dieser Zeichenbenutzer* bestimmten Gegenständen entsprechen. Unabhängig von Begriffsschemata existieren keine ‚Gegenstände‘. Wir spalten die Welt in Gegenstände auf, indem wir dieses oder jenes Beschreibungsschema einführen.“ Er bestreitet nicht, „daß es für unser Wissen *Input* durch Erfahrung gibt;“ er bestreitet jedoch, „daß es *Inputs* gibt, *die ihrerseits nicht durch unsere Begriffe geformt sind*,“ und er bestreitet, „daß es *Inputs* gibt, *die nur eine einzige Beschreibung zulassen, die unabhängig ist von allen begrifflichen Entscheidungen*.“ „Unsere Auffassungen . . . sind . . . eng mit unseren psychischen Gegebenheiten verknüpft. Sie sind abhängig von unserer biologischen und kulturellen Situation und keineswegs ‚wertfrei‘. . . Sie beschreiben eine Art von Objektivität – *etwas, das für uns Objektivität ist* –, auch wenn es nicht die Objektivität des Gottesgesichtspunkts ist. Was wir haben, sind Objektivität und Rationalität nach Menschenmaß; sie sind besser als nichts.“

⁵⁶A. a. O., B XVI, B 263. KANT gibt also HUME gegen LOCKE darin recht, daß das Kausalitätsprinzip nicht aus der Erfahrung gewonnen werden kann (und dennoch wie bei LOCKE notwendige Verbindungen der Vorstellungen schaffen kann), und will mit seiner Arbeit die menschliche Vernunft durch die Klippen der „Schwärmerei“ von LOCKE und des „Skeptizismus“ von HUME bringen (B 127 f.).

⁵⁷Gegen den skeptischen Idealismus DESCARTES’, der aus dem für ihn allein gewissen „Ich denke“ mit nur geringerer Zuverlässigkeit auf die Existenz äußerer Dinge geschlossen hat, wendet er ein, daß Denken und Bewußtsein notwendig mit Zeitbestimmung verbunden ist und daß diese (neben Veränderung) auch etwas Beharrliches erfordert, daß also „mein Dasein in der Zeit durch dieses Beharrliche allererst bestimmt werden kann“; somit gilt: „das Bewußtsein meines eigenen Daseins ist zugleich ein unmittelbares Bewußtsein des Daseins anderer Dinge außer mir.“ (a. a. O., B 275 ff.) Der „dogmatische“ Idealismus von BERKELEY, den er für unvermeidlich hält, „wenn man den Raum als Eigenschaft, die den Dingen an sich selbst zukommen soll, ansieht“, ist für ihn durch seine transzendente Ästhetik (also der Bestimmung des Raums und der Zeit als Formen unserer Anschauung) erledigt (B 69 ff., B 274), allerdings gibt er damit BERKELEY soweit recht, daß es keinen Sinn macht, in der oben bei HOBBS und DESCARTES angedeuteten Weise zwischen

Die durch die synthetischen Urteile *a priori* begrenzte Sphäre der Erscheinungen als Raum möglicher Erfahrungen ist das ganze „Land der Wahrheit“ (allerdings „umgeben von einem weiten und stürmischen Ozeane, dem eigentlichen Sitz des Scheins“). HUME nennt er einen „Geographen der menschlichen Vernunft“, für den das Gebiet des möglichen Wissens als unbestimmbar weit ausgebreitete Ebene durch einen empirischen Horizont begrenzt war, dessen Grenze er nicht genau bestimmen konnte. Für KANT (der sich eher als Architekten sieht) war dagegen die Vernunft eine Sphäre, deren Begrenzung und Inhalt durch die synthetischen Sätze *a priori* genau bestimmt werden kann. Von dem auf dem „stürmischen Ozeane . . . auf Entdeckungen herumschwärmenden Seefahrer“ der dogmatischen Schulphilosophie unterscheidet sich HUME dadurch, daß er „sein Schiff, um es in Sicherheit zu bringen, auf den Strand (den Skeptizismus) setzte, da es denn liegen und verfaulen mag, statt dessen es bei mir darauf ankommt, ihm einen Piloten zu geben, der, nach sicheren Prinzipien der Steuermannskunst, die aus der Kenntnis des Globus gezogen sind, mit einer vollständigen Seekarte und einem Kompaß versehen, das Schiff sicher führen könne, wohin es ihm gut dünkt“⁵⁸

Erkenntnis bezieht sich auf Erfahrung, ist also auf dieses Gebiet beschränkt, aber „denken kann ich, was ich will, wenn ich mir nur nicht selbst widerspreche“ (auch wenn ich dann „bloß mit Vorstellungen gespielt“ habe),⁵⁹ das Denken läuft in seiner Entdeckungslust also immer Gefahr, sich auf dem Ozeane zu verirren oder sich Nebelbänke als neue Länder zu lügen. Hierzu – aber auch zu geordneten Expeditionen mit dem Gewinn echter neuer Erkenntnisse – werden wir befähigt durch ein weiteres, dem Verstand übergeordnetes Erkenntnisvermögen, die *Vernunft*. Die Irrfahrten kommen zustande, weil wir „durch Fragen belästigt“ werden, die wir nicht abweisen, aber auch nicht beantworten können, da sie unser Vermögen übersteigen.⁶⁰ In der uns räumlich und zeitlich unbegrenzt erscheinenden Welt suchen wir ein Ganzes, in der nicht abbrechenden Kette von Ursachen und Wirkungen eine letzte Ursache, wir versuchen, „zu dem bedingten Erkenntnis des Verstandes das Unbedingte zu finden, womit die Einheit desselben vollendet wird.“⁶¹ Die Vernunft ist die Fähigkeit zum (nicht bloß deduktiven) Schließen, d. h. zu mittelbaren Urteilen als Relationen von Urteilen, ihr Gegenstand sind nicht Anschauungen (wie beim Verstand mit seinen Kategorien), sondern Begriffe und Urteile. „Die Einheit aller möglichen empirischen Verstandeshandlungen systematisch zu machen, ist ein Geschäft der Vernunft, so wie der Verstand das Mannigfaltige der Erscheinungen durch Begriffe verknüpft und unter empirische Gesetze bringt.“⁶² Dieses Geschäft betreibt die Vernunft mittels (transzendentaler) *Ideen*, reiner Vernunftbegriffe, die (im Gegensatz zu den an Anschauungen gebundenen Verstandesbegriffen) die Möglichkeit der Erfahrung übersteigen (z. B. *Seele, Welt, Gott*). Nun besteht – im spekulativen Gebrauch der Vernunft – immer die Gefahr, daß man eine mit den Ideen korrespondierende Realität annimmt und deren Beschaffenheit zu ergründen sucht. Eines der wesentlichen Anliegen KANTS war der Nachweis, daß man dabei einem bloßen Schein erliegt und das „Land der Wahrheit“ verlassen hat. Durch die Ideen werden keine Gegenstände der Erfahrung geschaffen (konstituiert), ihre positive und nützliche Funktion – im hypothetischen Gebrauch der Vernunft – ist es, so zu denken,

„primären“ und „sekundären“ Qualitäten zu unterscheiden.

⁵⁸A. a. O., B 295, B 788 ff., *Prolegomena*, A 17. Diese Metapher der Steuermannskunst wurde später als *Kybernetik* von Norbert WIENER und anderen wieder verwendet (vgl. dazu Bernhard DOTZLER (1987)).

⁵⁹*Kritik der reinen Vernunft*, B XXVI Anm., B 195.

⁶⁰A. a. O., A I.

⁶¹A. a. O., B 364.

⁶²A. a. O., B 692.

als ob sie Realität hätten, um so Begriffe zu ordnen und „den Verstand zu einem gewissen Ziele zu richten“; sie sind „heuristische Grundsätze“, *regulative* „Prinzipien des systematischen Verstandesgebrauchs im Felde der Erfahrung“.⁶³ Allerdings „ist die letzte Absicht der weislich uns versorgenden Natur bei der Einrichtung unserer Vernunft eigentlich nur aufs Moralische gestellt“; als *praktische* Vernunft gibt sie uns Gesetze, die „Gesetze der Freiheit sind, und welche sagen, was *geschehen soll*, ob es gleich vielleicht nie geschieht, und sich darin von *Naturgesetzen*, die nur davon handeln, *was geschieht*, unterscheiden; weshalb sie auch praktische Gesetze genannt werden.“⁶⁴

Was bedeutet das alles für den Begriff der Seele, für die *res cogitans*, für das Denken, das Ich und das Bewußtsein? KANT untersuchte die Bedingungen, die wir notwendigerweise als erfüllt annehmen müssen, damit wir so etwas wie Erfahrung und Erkenntnis überhaupt haben können. Einen Gegenstand erkennen hieß für ihn, als aktive Leistung des Verstandes mittels eines Begriffes ordnend, verbindend und abgrenzend die Mannigfaltigkeit der Anschauungen (wie mit einer Form das Material des Teiges beim Plätzchenbacken) in *einer* Vorstellung zu vereinen, somit die „Vorstellung der *synthetischen* Einheit des Mannigfaltigen“ zu erzeugen; das ist ein *Bewußtseinsakt*, das: „*Ich denke*, muß alle meine Vorstellungen begleiten *können*; denn sonst würde etwas in mir vorgestellt werden, was gar nicht gedacht werden könnte, welches eben soviel heißt als: die Vorstellung würde entweder unmöglich oder wenigstens für mich nichts sein.“⁶⁵ Das ist aber auch ein Akt des

⁶³A. a. O., B 672, 691, 799. Ein solches auf Einheit gerichtetes Prinzip ist beispielsweise wirksam, wenn beim DESY und anderswo nach einer Theorie der Vereinigung von Gravitation, elektromagnetischer Kraft, starker und schwacher Wechselwirkung gesucht wird. Weiter ist nach KANT (insbesondere im Bereich der organischen Natur, wo die Mechanik offensichtlich an ihre Grenzen stößt) auch die Annahme eines teleologischen Naturzweckes (dessen Existenz weder bewiesen noch widerlegt werden kann) ein nützliches Prinzip zur Erweiterung der empirischen Kenntnisse (B 716). In einem späteren Werk erörtert KANT dieses Problem ausführlicher und erklärt die mechanische Physik für außer Stande, das Problem der *Selbstorganisation* zu lösen. Anders als bei einer Uhr (deren Zweck der Mensch setzt) kann es bei einer Zelle Zellkern und Zellwand nur geben, weil es die Zelle gibt (und die kann es nur geben, weil es Kern und Wand gibt), anders als bei einer Uhr kann ein Polyp ihm abgeschnittene Teile ersetzen; ein Organismus unterscheidet sich von einer Maschine dadurch, daß „er ohne die Kausalität der Begriffe von vernünftigen Wesen außer ihm möglich“ ist, daß bei ihm „ein jeder Teil, so wie er nur *durch* alle übrigen da ist, auch als *um der anderen* und des Ganzen *willen* existierend“ gedacht werden muß, ein solches Wesen hat nicht „lediglich *bewegende* Kraft, sondern es besitzt in sich *bildende* Kraft“, und „nur dann und darum wird ein solches Produkt [der Natur] als *organisiertes* und *sich selbst organisierendes Wesen* ein *Naturzweck* genannt werden können. [...] Genau zu reden, hat also die Organisation der Natur nichts Analogisches mit irgendeiner Kausalität, die wir kennen“; allerdings kann man „umgekehrt einer gewissen Verbindung,“ nämlich der „Umbildung eines großen Volks zu einem Staat“, „durch eine Analogie mit den genannten unmittelbaren Naturzwecken Licht geben“ (*Kritik der Urteilskraft*, § 65).

⁶⁴*Kritik der reinen Vernunft*, B 829 f.

⁶⁵A. a. O., B 130 ff. RORTY (1981, S. 167 ff.) sieht hier eine „Kontamination von Prädikation und Synthesis“ und die vergebene Chance zu einer bedeutend früheren pragmatischen Wende: „Wäre Kant von der Einsicht, daß man ‚das einzelne Urteil‘ nicht mit ‚der Einzelheit eines sinnlich Gegebenen‘ gleichsetzen darf (und natürlich auch nicht mit der eines verstandesmäßig Gegebenen), dazu fortgeschritten, Erkenntnis als eine Relation zwischen Personen und Propositionen aufzufassen, so hätte er den Synthesisbegriff nicht benötigt.“ „Wäre das ‚Erkenntnisproblem‘ in der Begrifflichkeit von Relationen zwischen Propositionen und ihrem Gewißheitsgrad formuliert worden, statt in der Begrifflichkeit vermeintlicher *Komponenten* von Propositionen – wir wären vermutlich nicht mit unserer heutigen Vorstellung von der ‚Geschichte der Philosophie‘ geboren worden. [...] Wie dem auch sei, Kant vollzog diese pragmatische Wende jedenfalls nicht.“ Ohne KANT mit mehr befrachten zu wollen als er tragen kann, sei dennoch der Hinweis erlaubt, daß ihm dieser Gedanke nicht völlig fern war: „Wahrheit aber beruht auf der Übereinstimmung mit dem Objekte, in Ansehung dessen folglich die Urteile eines jeden Verstandes einstimmig sein müssen [...]. Der Probestein des Fürwahrhalten, ob es Überzeugung oder bloße Überredung sei, ist also äußerlich die Möglichkeit, dasselbe mitzuteilen und das Fürwahrhalten für jedes Menschen Vernunft gültig zu befinden“

Selbstbewußtseins (von KANT *Apperzeption* genannt), der Vorstellung der durchgängigen Einheit des Bewußtseins, der *einen* Vorstellung, daß sie alle „insgesamt *meine* Vorstellungen“ sind, „denn sonst würde ich ein so vielfarbiges verschiedenes Selbst haben, als ich Vorstellungen habe, deren ich mir bewußt bin“; diese *reine, ursprüngliche, transzendente* „synthetische Einheit der Apperzeption“ ist für KANT die letzte Bedingung für die Möglichkeit des Denkens, sie ist „der höchste Punkt, an dem man allen Verstandesgebrauch, selbst die ganze Logik und nach ihr die Transzendental-Philosophie heften muß, ja dieses Vermögen ist der Verstand selbst.“⁶⁶ *Denken* (als Vermögen des Verstandes) – so kann man zusammenfassen – ist demnach für KANT die „Handlung, die Synthesis des Mannigfaltigen, welches ihm anderweitig in der Anschauung gegeben worden, zur Einheit der Apperzeption zu bringen“.⁶⁷ KANTS transzendente Bestimmung des *Ich* zieht zugleich auch die Grenze zwischen dem empirischen Ich (zu dessen Bewußtsein und Erforschung das transzendente schon vorausgesetzt werden muß) und der metaphysischen *Seele*, deren Erforschung die Möglichkeit jeder Erfahrung übersteigt und die nur als regulative Idee der Vernunft ihre Berechtigung hat (diese Unterscheidung ermöglicht auch erst die praktische Vernunft und die Moral).⁶⁸

KANTS Transzendentalphilosophie der theoretischen und praktischen Vernunft mit ihren Elementen Sinnlichkeit (und reine Anschauungsformen), Verstand (und Kategorien), Urteilskraft (und Schemata), Vernunft (und regulative Ideen)⁶⁹ für den kognitiven Bereich

(B 848; „Proberstein“ für das subjektive Glauben an ein Urteil ist für ihn „das *Wetten*“ (B 852)).

⁶⁶A. a. O., B 134. Aus ihr ergibt sich dann auch die „transzendente Deduktion der reinen Verstandesbegriffe“. Aus dieser *notwendigen Einheit* folgt dann auch die *Objektivität* von Urteilen, also z. B. daß ein Körper schwer *ist* (d. h. nicht, daß diese Vorstellungen notwendig zueinander gehören – es ist ja ein empirisches Urteil –, sondern daß sie „vermöge der *notwendigen Einheit* der Apperzeption in der Synthesis der Anschauungen zu einander“ gehören); ginge es nur um subjektive Vorstellungen „nach Gesetzen der Assoziation“, könnte man nur sagen: „wenn ich einen Körper trage, so fühle ich einen Druck der Schwere“ (B 142).

⁶⁷A. a. O., B 145.

⁶⁸KANT wendet sich damit gegen die Möglichkeit einer „rationalen Psychologie“ (wie sie in der „Schulphilosophie“ vor ihm betrieben wurde), in der über die Seele als „Substanz“ spekuliert wird. Ausgehend vom Faktum der Erfahrung analysiert er ihre Bedingungen *a priori*, weitergehende Aussagen lassen sich über eine „Quelle“ von Sinnlichkeit, Verstand und Vernunft nicht machen (auch nicht, daß deren genannten Fähigkeiten *angeboren* seien, denn das wäre eine *empirische* Aussage). Der empirischen Psychologie (die sich auf Anschauungen stützen muß), die zu seiner Zeit „noch nicht so reich“ war, daß „sie allein ein Studium“ ausmacht, aber „doch zu wichtig, als dass man sie [...] anderwärts anheften sollte“, wollte er dagegen Gastrecht und Fürsorge in der Philosophie geben, bis dieser aufgenommene „Fremdling [...] in einer ausführlichen Anthropologie, (dem Pendant zu der empirischen Naturlehre,) seine eigene Behausung wird beziehen können.“ (B 877)

Die scharfe Unterscheidung zwischen dem empirische Ich als *Phaenomenon* und der transzendenten Seele als *Noumenon* ist für die praktische Vernunft und die Begründung der Moral grundlegend. „Praktisch ist alles, was durch Freiheit möglich ist“ (diese „*freie Willkür*“, die nicht bloß „durch sinnliche Antriebe“ bestimmt wird und Moral erst möglich macht, unterscheidet für ihn den Menschen vom Tier, B 828 ff.), sonst ist es *Natur*, und damit der Kausalität unterworfen. Nur durch diese Unterscheidung in Gegenstand der Erfahrung einerseits und Freiheit des Willens als regulative Idee andererseits lassen sich nach KANT die bekannten – durch falsche Gleichsetzung entstandenen – Paradoxien der Willensfreiheit (vgl. B 472 ff.) vermeiden, insofern gezeigt wird, daß über die *Seele* kein Wissen möglich ist: er mußte „also das *Wissen* aufheben, um zum *Glauben* Platz zu bekommen“ (B XXX).

⁶⁹Zur nochmaligen Illustration dieser Unterscheidung in KANTS eigenen Worten (nicht nur im Bereich der KI kann man ohne Militär anscheinend nicht auskommen): „Der unter gemessenen Befehlen stehende Haus- oder Staatsdiener braucht nur Verstand zu haben; der Offizier, dem für das ihm aufgetragene Geschäfte nur die allgemeine Regel vorgeschrieben und nun überlassen wird, was in vorkommendem Falle zu tun sei, selbst zu bestimmen, bedarf Urteilskraft; der General, der die möglichen Fälle beurteilen und für sie sich die Regel selbst ausdenken soll, muß Vernunft besitzen“ (*Anthropologie in pragmatischer Hinsicht*,

des Denkens und Erkennens sowie Gefühle (Lust und Unlust) und Wille für die Bestimmung von Handlungen stellt einen Höhepunkt der von DESCARTES begonnenen und von LOCKE fortgesetzten Bemühungen dar, das sich selbst bewußte Subjekt zum Mittel- und Ausgangspunkt der Philosophie und der Wissenschaften zu machen (er wußte im Gegensatz zu seinen Vorgängern, daß man nicht beides haben kann: eine von unserem Erkennen unabhängige Realität der Außenwelt *und zugleich* das in Fragen des Denkens und der Moral autonome, willensfreie und sich selbst gewisse Subjekt der Neuzeit; er entschied sich – aus praktischem Interesse – für die Autonomie). Diese Bewußtseinsphilosophie stellt nicht mehr (wie die antike Philosophie) die Frage nach dem Sein des Seienden in den Mittelpunkt, ihr Gegenstand sind die geistigen *Vorstellungen*, ihre Ausgangsfrage ist „Was kann ich wissen?“; das *mentalistische* Paradigma hat das ontologische abgelöst.⁷⁰

Im Alltagsbewußtsein hat dieses über Jahrhunderte wirksame mentalistische Paradigma (trotz inzwischen sinkenden Sterns) zwei recht feste Überzeugungen hinterlassen:

- Wir verfügen im Geist über Vorstellungen, die bestimmte Aspekte der Außenwelt repräsentieren (nicht unbedingt schlicht abbilden), und Denken besteht darin, daß wir *mit diesen Vorstellungen nach gewissen Regeln* (teils zwangsläufig, teils willkürlich, bewußt oder unbewußt) *operieren*. Das versetzt uns in die Lage, auf bestimmte Einflüsse und Anforderungen der Umwelt „intelligent“ zu reagieren, d. h. auch wieder auf sie einzuwirken (wobei besonders wirksame und interessante, dem Denken am stärksten verwandte Einflüsse und Handlungen sprachlicher Art sind).⁷¹
- (*Selbst-*) *Bewußtsein, Denken können und Person sein* (zum einen als moralisches, verantwortliches Subjekt, zum anderen damit als „eine oder einer von uns“) sind – nebst der Fähigkeit zur Sprache – nur verschiedene, eng miteinander verbundene Aspekte dessen, was den Menschen zum Menschen macht und ihn grundsätzlich vom Tier unterscheidet.⁷²

BA 118). Nun ja, in humanistischer Hinsicht spricht dann allerdings wenig gegen eine Rationalisierung der Verwaltungs- und Büroarbeit durch Datenverarbeitung und viel für eine Ausbreitung der Vernunft in allen Bereichen.

⁷⁰Vgl. Herbert SCHNÄDELBACH (1991). Ob sich das dort als zur Zeit gültig angegebene *linguistische* Paradigma wirklich als epochaler Nachfolger etablieren wird, scheint mir nicht ausgemacht. Für mindestens ebenso plausibel halte ich die Deutung, daß sich der *Linguistic turn* als ein Übergang herausstellen wird von dem Versuch, die Probleme des Mentalismus (letztlich immanent) zu lösen durch eine Akzentverschiebung hin zu der intersubjektiv kontrollierbaren sprachlichen Darstellung der Vorstellungen („WITTGENSTEIN I“), und der Entdeckung, daß sich ein solches Unternehmen notwendig verfangen muß in der Lebenswelt, in der nicht exakten Bestimmbarkeit und historisch-kulturellen Bedingtheit der gemeinsamen Praxis als Hintergrund („WITTGENSTEIN II“), also als ein Durchgangsstadium zu einem *pragmatischen* Paradigma (angeregt durch eine Untersuchung der Sprach*praxis* analog der Entstehung des mentalistischen Paradigmas im Zusammenhang mit der damaligen Etablierung der Mechanik als *Wissenschaft des Seienden*). Beide Ansätze stehen heute in Konkurrenz zu einem *Naturalistic turn* (vgl. Geert KEIL (1993a), Bernhard IRRGANG (1993)).

⁷¹Man sieht, das „Grundschema der Datenverarbeitung“: *Eingabe – Verarbeitung – Ausgabe* gibt es um einiges länger als Computer. DESCARTES, LOCKE und KANT als Vorläufer und Mitbegründer einer Sicht des Menschen als „informationsverarbeitendes System“?

⁷²Empirisch und moralisch wird das *so* begründete Festhalten an einer Sonderstellung des Menschen gegenüber Tieren immer schwieriger. Abgesehen davon ist nicht ohne weiteres klar, welche notwendige Verbindung es beispielsweise geben soll zwischen der bewußten Empfindung eines Schmerzes als Schmerz, der Fähigkeit, im Schach unter verschiedenen Zügen den auszuwählen, der am schnellsten zum Matt führt, und der legitimen und legalen Möglichkeit, einem Chauffeur (im Gegensatz zum Autopiloten) einen im Rahmen der Halterhaftung erhaltenen Bußgeldbescheid weiterzureichen. Die genannte Gleichsetzung findet

2.6 Ryle: Gegen das „Gespenst in der Maschine“

Die Bewußtseinsphilosophie, mit ihrer radikalen Trennung von Geist und Materie, hat neben der (theoretischen) Lösung ihrer selbst gestellten Probleme (wie läßt sich ein sicheres Fundament dafür schaffen, die Natur so zu erforschen und zu berechnen, daß sie für die Menschen beherrschbar wird? und: wie läßt sich Moral so begründen, daß Menschen Verantwortung nur gegenüber sich selbst und Ihresgleichen haben?) auch neue Fragen und Rätsel hinterlassen. Dazu gehören:

- Wenn man (wie z. B. DESCARTES) annimmt, daß es zwei radikal getrennte Substanzen, den „Geist“ und die körperliche, materielle, physikalisch kausal abgeschlossene (Erhaltungssätze) Welt des „Ausgedehnten“ gibt, dann ist schwer verständlich, wie diese beiden Welten miteinander interagieren können; und *daß* sie interagieren, scheint ja offensichtlich, wenn bohrende Zahnschmerzen die Laune verderben oder wenn man den Entschluß faßt aufzustehen – ohne eine solche Interaktion sind Wahrnehmungen und Handlungen kaum erklärbar („Leib-Seele-Problem“).
- Wenn man (wie LOCKE und KANT) annimmt, daß (beispielsweise *mein*) Denken auch ein aktives Operieren mit Vorstellungen gemäß ihrer Bedeutung ist, wer oder was ist dann dieser kluge Operateur und Regisseur im inneren Theater des Bewußtseins (*ich* bin es sicher nicht, denn *mein* Denken soll ja dadurch erst entstehen und erklärt werden); ja, und wer oder was ist dann der innere Operateur meines Operateurs . . . ? („Homunkulus-Problem“)
- Wenn man dagegen (wie HOBBS und vielleicht HUME) annimmt, daß dieser „Denken“ genannte Prozeß rein „mechanisch“ (d. h. ohne Beachtung der Bedeutung) abläuft, dann ist schwer zu erklären, wie dieser Prozeß (im Gegensatz z. B. zu den Planetenbewegungen) von etwas anderem handeln, sich auf etwas anderes beziehen, kurz *Intentionalität* haben kann („Paradoxon des mechanischen Verstandes“ und „Problem der ursprünglichen Bedeutung“⁷³).

Neben vielen Versuchen, die genannten Probleme *innerhalb* des mentalistischen Paradigmas zu lösen, tauchte nach einiger Zeit der Gedanke auf, daß der Grund für die Schwierigkeiten in den Grundannahmen des Paradigmas selbst liegt. Folgenreich und viel diskutiert war der Frontalangriff von Gilbert RYLE (1900–1976). Gemäß seiner These entstand das Leib-Seele-Problem dadurch, „daß in den drei Jahrhunderten seit dem Beginn des naturwissenschaftlichen Zeitalters die logischen Kategorien zur Einordnung der Begriffe der geistigen Vermögen und Tätigkeiten falsch ausgewählt wurden.“⁷⁴ DESCARTES’ radikalen Dualismus, die Annahme, „daß es zwei verschiedene Arten von Existenz oder Sein gibt“, verspottet er als „Dogma vom Gespenst in der Maschine“; diese Lehre, so RYLE, „besteht

sich sowohl bei KI-Gegnern – wenn z. B. argumentiert wird, daß Computer nie werden *wirklich* denken können, weil ihnen Bewußtsein fehlt (was, da ja letztlich solche Urteile nur auf einer Übertragung der durch Tradition und Sprache gefärbten Selbsterfahrung beruhen können, nur bedeuten kann: sie sind nicht „einer von uns“) – als auch bei einigen extremen KI-Propagandisten, wenn etwa Bürgerrechte für Computer gefordert werden, weil sie Denkleistungen vollbringen.

⁷³Vgl. HAUGELAND (1987), S. 33 ff.

⁷⁴*Der Begriff des Geistes*, Stuttgart 1969 (Erstausgabe: *The Concept of Mind*, 1949), S. 4.

aus einem einzigen großen Irrtum, einem Irrtum von ganz besonderer Art, nämlich einer Kategorienverwechslung,⁷⁵ und er illustriert das an einem berühmt gewordenen Beispiel:

Ein Ausländer kommt zum erstenmal nach Oxford oder Cambridge, und man zeigt ihm eine Reihe von Colleges, Bibliotheken, Sportplätzen, Museen, Laboratorien und Verwaltungsgebäuden. Nach einiger Zeit fragt er: ‚Aber wo ist denn die Universität?‘ [...] Der Irrtum des Ausländers lag in seiner unschuldigen Annahme, es sei richtig, [...] von ‚der Universität‘ so zu sprechen, als bezeichneten die Worte ‚die Universität‘ ein weiteres Mitglied der Klasse, zu der jene anderen obenerwähnten Einheiten auch gehören. Er reihte die Universität irrtümlich in dieselbe Kategorie ein, zu der diese anderen Institutionen gehören. [...] Die Unterschiede zwischen dem Körperlichen und Geistigen wurden so als Unterschiede innerhalb des gemeinsamen Rahmens der Kategorien ‚Ding‘, ‚Material‘, ‚Eigenschaft‘, ‚Zustand‘, ‚Vorgang‘, ‚Veränderung‘, ‚Ursache‘ und ‚Wirkung‘ dargestellt. Geister sind Dinge, aber Dinge von anderer Art als Körper; geistige Vorgänge sind Ursachen und Wirkungen, aber Ursachen und Wirkungen anderer Art als Körperbewegungen. Und so weiter. So wie der Ausländer erwartete, die Universität werde ein zusätzliches Gebäude sein, ähnlich wie ein College, aber auch recht anders, so haben die Antimechanisten Geister als zusätzliche Zentren von Kausalvorgängen dargestellt, ähnlich wie Maschinen aber auch recht anders. Ihre Theorie war eine paramechanische Hypothese.⁷⁶

Einher mit diesem Dogma geht nach RYLE eine systematische Fehleinschätzung der Rollen von Können (*knowing how*) und Wissen (*knowing that*). Wir „neigen dazu, alle anderen Begriffe für geistiges und seelisches Verhalten mit Hilfe von Begriffen der Erkenntnis zu definieren. [...] Die griechische Idee, Unsterblichkeit sei dem theoretisierenden Teil der Seele vorbehalten,“ führte zu der „Idee, das Vermögen der Wahrheitserkenntnis sei die wesentliche Eigenschaft des Geistes. Andere menschliche Fähigkeiten konnten nur dann als geistig angesehen werden, wenn sich zeigen ließ, daß sie irgendwie durch das verstandesmäßige Begreifen von wahren Sätzen gelenkt seien.“ Intelligentes und geistig aufmerksames Tun „ist nach dieser Legende immer das Tun von zwei Dingen, nämlich erstens gewisse passende Sätze oder Vorschriften erwägen und zweitens das in die Praxis umsetzen, was diese Sätze

⁷⁵A. a. O., S. 9, 13. Kategorienfehler bestehen in der Verwendung von Ausdrücken eines bestimmten logischen Typs an Stellen, an denen nur andere logische Typen sinnvoll sind (so ist der Satz „Quadratzahlen sind traurig“ nicht falsch, sondern sinnlos).

⁷⁶A. a. O., S 14f. u. 18f. (RYLES Vergleich wirkt bei einer Wanderung durch Oxford oder Cambridge sehr viel überzeugender als bei einer durch – z. B. – Bochum). Damit will er keineswegs den Mechanisten oder anderen Monisten Recht geben, seine These ist, „daß sowohl Idealismus als Materialismus die Antwort auf eine falsch gestellte Frage sind.“ Wenn man – zulässigerweise – sagt „Geister existieren“ und „Körper existieren“, dann zeigen diese Ausdrücke „nicht verschiedene Arten von Existenz [...] vielmehr zwei verschiedene Bedeutungen des Wortes ‚existieren‘ an“ (S. 23). An einer anderen Stelle macht er die Metapher des „inneren Auges“ für diesen Fehler verantwortlich: „Als Galileis und Descartes’ Darstellungen der mechanischen Welt die Rettung des Geistes vor der Mechanisierung zu fordern schienen, und zwar dadurch, daß er als eine zweite Welt bildend dargestellt wurde, da wurde das Bedürfnis verspürt zu erklären, wie der Inhalt dieser Geisterwelt ausfindig gemacht werden könne, und zwar wieder ohne die Hilfe der Sinneswahrnehmung. Die Metapher vom ‚Licht‘ schien besonders geeignet, da ja die Galileische Wissenschaft vorwiegend mit der optisch entdeckten Welt befaßt war. ‚Bewußtsein‘ wurde also eingeführt, um in der psychischen Welt dieselbe Rolle zu spielen, die das Licht in der mechanischen spielte. [...] Dasselbe Modell wurde auch wieder von Locke in seiner Beschreibung der beobachtungsartigen Prüfung verwendet, die der Geist zeitweilig seine laufenden Zustände und Vorgänge unterziehen kann. Er nannte diese vermeintliche innere Wahrnehmung ‚Reflexion‘ (unsere ‚Introspektion‘), wobei er das Wort ‚Reflexion‘ dem wohlbekanntesten optischen Phänomen der Widerspiegelung von Gesichtern in Spiegeln entnahm. Der Geist kann seine eigenen Operationen in dem von ihm selbst abgegebenen ‚Licht‘ ‚sehen‘ oder ‚betrachten‘. Der Mythos vom Bewußtsein ist ein Stück Paraoptik.“ (S. 214)

oder Vorschriften anbefehlen.“ RYLE dagegen ist angetreten, „die intellektualistische Doktrin richtigzustellen, die Intelligenz mit Hilfe von Wahrheitserkenntnis zu definieren sucht, statt Wahrheitserkenntnis mit Hilfe von Intelligenz.“ Er will zeigen, „daß die intellektualistische Legende falsch ist, und daß die Beschreibung einer Handlung als intelligent nicht den doppelten Vorgang von Erwägen und Ausführen bedingt“, eine solche „Handlung hat eine besondere Art oder Ausführung, nicht besondere Vorgänger.“ Er will somit zeigen, „daß wir, wenn wir von Leuten sagen, sie betätigen ihre geistigen Fähigkeiten, uns nicht auf okkulte Vorfälle beziehen, deren Auswirkungen sich im Benehmen und in Äußerungen zeigen; wir beziehen uns vielmehr auf das Benehmen und die Äußerungen selbst.“⁷⁷

Aus dieser Sicht ist Wissen (wie auch andere geistige Fähigkeiten) nur eine *Disposition*, eine (erworbene) Fertigkeit, unter bestimmten Umständen, beim Auftreten bestimmter Ereignisse, in bestimmter Weise sich zu verhalten und zu handeln. So gesehen hat die „traditionelle Theorie des Geistes [...] die Typenunterscheidung zwischen Disposition und Betätigung in einen mythischen Zwiespalt zwischen unbeobachtbaren geistigen Ursachen und beobachtbaren physischen Wirkungen mißdeutet.“ Dagegen setzt RYLE sein Vorhaben, mittels einer Analyse der Logik der (Umgangs-) Sprache im Bereich des Geistigen – u. a. durch die sorgfältige Unterscheidung der Typen der „Dispositionswörter“ (in der Art von „zerbrechlich“ oder „in Wasser löslich“) und der „Episodenwörter“ – Sätze über psychische Zustände und Ereignisse als „dispositionale Aussagen“, d. h. als Sätze über öffentlich zugängliche, intersubjektiv überprüfbare Verhaltens- und Handlungsmöglichkeiten zu rekonstruieren. Dadurch erhalten wir nach RYLE die Möglichkeit, von einer Person nur noch *einen* Lebenslauf zu schreiben, statt des Rekurses auf „eine zweite Gruppe von Schattenoperationen“ „in irgend jemandes Bewußtseinsstrom“, Erklärungen im Bereich des Geistigen sind dann zwar anders, aber nicht geheimnisvoller als (und vom ähnlichen Muster wie) „Das Glas zerbrach, als es von einem Stein getroffen wurde, weil es zerbrechlich war“. Danach folgen mehrere Hundert Seiten der Durchführung des Vorhabens, durch Einordnung in die richtigen Kategorien die paramechanische Theorie des Geistes und insbesondere das traditionelle Dogma, „der Geist oder die Seele arbeitet in drei unreduzierbar verschiedenen Modi, im Modus des Erkennens, im Modus des Affekts und im Modus des Strebens“, kurz die „Dreifaltigkeitstheorie des Geistes“ als ein „theoretisches Kuriosum“ zu entlarven; angefangen vom „Willensakt“ (ein „künstlicher Begriff“, der zudem noch wie „Phlogiston“ und „Lichtaether“ jede Nützlichkeit verloren hat) bis zum „Verstand“.⁷⁸

⁷⁷ A. a. O., S. 27 f., 32, 29, 33, 36 u. 26. „Der entscheidende Einwand gegen die intellektualistische Legende ist also dieser. Das Erwägen von Sätzen ist selbst eine Tätigkeit, die mehr oder weniger intelligent, mehr oder weniger dumm ausgeführt werden kann. Aber wenn zur intelligenten Ausführung einer Tätigkeit eine vorhergehende theoretische Tätigkeit nötig ist, und zwar eine, die intelligent ausgeführt werden muß, dann wäre es logisch unmöglich, daß irgend jemand in diesen Zirkel eindringen könnte.“ (S. 34)

⁷⁸ A. a. O., S. 38, 154 ff., 60 f. u. 78 f. Nach Meinung der meisten (ihm im Prinzip wohlgesonnenen) Kritiker aus dem Lager der analytischen Philosophie gelingt ihm das recht gut für Meinungen, Wünsche, Stimmungen, Absichten, Einstellungen usw. Problematisch ist es hingegen bei momentanen Bewußtseinszuständen und Empfindungen (Schmerzen, Sinnesqualitäten, geistige Bilder). Insbesondere seine These „Das, was ich über mich selbst herausfinden kann, ist von derselben Art wie das, was ich über andere herausfinden kann, und die Methoden, es herauszufinden, sind ungefähr dieselben“ (S. 209) ist sehr umstritten (zur Diskussion dieser Einwände vgl. RORTY (1981), S. 116 ff.). Das führte innerhalb der analytischen Philosophie des Geistes u. a. zur Gegenposition der Identitätstheorie (die grob gesagt mentale Zustände mit neuronalen identifiziert) und zu ausgedehnten Debatten, in denen „es laufend um Marsmenschen, Gehirne in Gläsern und den ewigen Feuern der C-Fibern geht“ (HASTEDT (1991b), S. 673) und die hier nicht nachgezeichnet werden können (siehe dazu z.B. auch BIERI (1993), LANZ (1993)). Weiter ist umstritten, wieweit eine „dispositionale“ Bestimmung von „Wissen“ ohne solche „nichtdispositionalen“ Begriffe wie „Wahrheit“ auskommt (vgl. dazu BRÜGGEN (1974), S. 1735 ff.).

Nach RYLE brauchen wir zur Erklärung des Erkennens von Gegenständen keine „Geschichten über äußerst schnelle und unbemerkte Schlüsse, oder daß Begriffe heraufgeholt und visuellen Daten über den Kopf gestülpt werden“, es werden einfach – durch erlernte „Wahrnehmungsrezepte“ – erworbene „Erwartungsdispositionen“ befriedigt (ähnlich, wie „wissen, wie eine Melodie geht“ heißt „eine Gruppe von auditiven Erwartungsdispositionen erworben haben“, und „eine Melodie wiedererkennen“ heißt „einen erwarteten Ton nach dem anderen hören“; nicht viel anders ist das beim Erkennen von Fingerhüten und Kühen, nur wird das dadurch verschleiert, daß hier alles viel schneller geht als beim Hören des „Donauwalzer“). Für RYLE sind wir damit viele Probleme der klassische Erkenntnistheorie losgeworden: „Die Beschreibung, wie Kinder Wahrnehmungsrezepte lernen, bringt kein größeres erkenntnistheoretisches Rätsel mit sich als die Beschreibung, wie Jungen radfahren lernen. Sie lernen durch die Praxis, und wir können die Arten von Übungen angeben, die den Lernvorgang fördern.“⁷⁹

Entsprechend heißt für RYLE der Besitz einer abstrakten Idee (z. B. „Höhenlinie“) nichts anderes, als daß jemand bestimmte „Aufgaben und noch eine unbestimmte Reihe anderer verwandter Aufgaben bewältigen kann, regelmäßig bewältigt oder in diesem Augenblick bewältigt.“ Eine Theorie haben heißt nicht, innerlich etwas tun, sondern „bereit sein, sie mitzuteilen oder anzuwenden, wenn sich die Gelegenheit dazu ergibt“, ein Urteil macht jemand, wenn er nach erlernten Regeln und Standards „ein Stück Theorie, das er schon besitzt, in Aussageform bringt,“ ein Schluß wird gezogen, „wenn jemand zum öffentlichen oder privaten Gebrauch sagt oder schreibt, ‚Das, ergo jenes‘ [...], vorausgesetzt, er sagt oder schreibt es im Wissen, er sei dazu berechtigt.“ Ein wichtiger Grund dafür, warum Erkenntnistheoretiker bei der Analyse von „Idee“, „Urteil“, „Schließen“ usw. immer nach geheimnisvollen Tätigkeiten im inneren Raum des Geistes suchten statt nach Dispositionen und öffentlichen Handlungen, liegt für RYLE in der ungenügenden Unterscheidung zwischen „Aufgabenzeitwörtern“ (rennen, reisen ...) und „Erfolgszeitwörtern“ (siegen, ankommen ...). Er setzt dagegen, „daß die Wörter ‚Urteil‘, ‚Deduktion‘, ‚Abstraktion‘ und so weiter rechtmäßig zur Klassifikation der Denkergebnisse gehören und daß sie mißdeutet werden, wenn sie als die Bezeichnung jener Handlungen verwendet werden, aus denen das Nachdenken besteht. [...] Sie sind Ausdrücke für Schiedsrichter, nicht für Biographen.“ Er hofft gezeigt zu haben, „daß diese von Erkenntnistheoretikern postulierten, für sich bestehenden intellektuellen Prozesse paramechanische Dramatisierungen der klassifizierten Elemente schon aufgestellter und auseinandergesetzter Theorien sind.“⁸⁰

RYLES Ersetzung mentaler Begriffe durch solche über Dispositionen und beobachtbare Ereignisse hat Ähnlichkeiten mit dem (psychologischen) Behaviorismus (vgl. Unterrichtsmodul 1), allerdings setzt er sich von einem dort häufig vertretenen Reduktionismus genauso ab wie vom mentalistischen Paradigma, vgl. S. 449 ff.), deshalb wird seine Position auch „Logischer Behaviorismus“ genannt (eine andere Bezeichnung – die mehr auf RYLES Position rekurriert, daß „Geist“ zwar nicht als eigene Substanz existiert, wohl aber das Reden über Mentales zu einer ganz anderen Sorte „Sprachspiel“ gehört als das Reden über Physisches – ist „Sprachdualismus“). Zur Diskussion darüber, wieweit RYLES Leugnen (Entlarven als Scheinproblem) eines Leib-Seele-Problems wirklich zu dessen Lösung beiträgt, vgl. CARRIER & MITTELSTRASS (1989, S. 283 ff.) und HASTEDT (1988, S. 83 f. u. 98 ff.).

⁷⁹ *Der Begriff des Geistes*, S. 312–316.

⁸⁰ A. a. O., S. 422 f., 392, 407, 412, 414, 390 f. u. 398. Den historischen Grund für diese Verwechslung sieht er zum einen in der Orientierung an der Metapher des Sehens (anders als „schauen“ genaugenommen ein „visuelles Erfolgswort“), weiter in der Anlehnung des Logikunterrichts an die axiomatische Euklidische Geometrie (statt ans Rechnen) sowie in einer einengenden Sicht der Sprache als im wesentlichen *didaktische* Sprache. (415 f., 419 f., 424 ff.)

Zusammengefaßt ist für RYLE die gesamte Bewußtseinsphilosophie und Erkenntnistheorie ein historisch bedingter Irrweg: der „Geist“ ist für ihn nur der Sammelname für ein Bündel von Dispositionen und Fertigkeiten, über die man in bestimmter Weise reden sollte, nicht eine oder gar die grundlegende Fähigkeit des Menschen selbst; als Nachfolger für die Erkenntnistheorie schlägt er erstens eine Theorie der Wissenschaften als systematisches Studium des Gefüges schon gebauter Theorien, zweitens eine Theorie des Lernens, der Entdeckung und der Erfindung vor.

Bei allen berechtigten Einwänden besteht die Stärke des Ansatzes von RYLE darin, den ganzen Menschen (und nicht bloß seinen Geist) in den Mittelpunkt zu stellen, mit seiner biologischen, kulturellen und individuellen Geschichtlichkeit, in seiner Wechselwirkung mit der Außenwelt im Rahmen einer Gemeinschaft mit anderen Menschen, ohne andererseits einem szientistischen Reduktionismus das Wort zu reden.⁸¹ Für die kognitive KI – insbesondere für den Teil, der sich als empirische Variante des eben unter „zweitens“ genannten Nachfolgers der Erkenntnistheorie versteht und deshalb auch mehr auf das Lernen als auf das Programmieren einer fertigen Intelligenz setzt – hat RYLES Konzept eine Bedeutung dadurch, daß der Blick weg vom nur privat zugänglichen und geheimnisvollen Bewußtsein hin zu den *Denkleistungen* und den sie ermöglichenden Dispositionen, Fähigkeiten und Fertigkeiten, gelenkt wird.⁸²

Die letzten Absätze – Philosophie als logische Analyse der Umgangssprache – waren chronologisch ein Vorgriff (und auch eine kleine Abweichung) auf dem Weg zur kognitiven KI. Nach KANT verzweigten sich die Motive und Wege des vorher einheitlichen Projekts einer die Wissenschaften fundierenden Erkenntnistheorie (andere – wie DEWEY, HEIDEGGER und der späte WITTGENSTEIN – verwarfen das Projekt ganz). Der von DESCARTES und LEIBNIZ aufgeworfene methodologisch-normative Aspekt einer an der Strenge und Gewißheit der Mathematik orientierten Naturwissenschaft führte im ausgehenden 19. Jahrhundert zu einer neuen Blüte der *Logik*. Etwa gleichzeitig etablierte sich im Anschluß an die Assoziationstheorien von LOCKE und HUME eine empirische *Psychologie*. In der Philosophie mochte man sich mit der Unerkennbarkeit des „Dings an sich“ nicht abfinden. Ein Ausweg wurde – ausgehend vom sich selbst, Gott und die Welt denkenden Bewußtsein – in einer „Phänomenologie des Geistes“, in den Höhenflügen und Abgründen des deutschen Idealismus gesucht (die zumindest die Dimension der Geschichtlichkeit des Geistes ins Be-

⁸¹ „Der Mensch braucht durch die Behauptung, er sei kein Gespenst in einer Maschine, nicht zu einer Maschine degradiert zu werden. Er könnte schließlich doch eine Art Lebewesen sein, nämlich ein höheres Säugetier. Es muß noch der verwegene Sprung zu der Hypothese gewagt werden, daß er vielleicht ein Mensch sei.“ (a. a. O., S. 451)

⁸² Aus dieser Sicht einer „Erkenntnistheorie“ gleicht die Frage „Wie ist Erkenntnis möglich?“ der Frage „Wie sind Telefone möglich?“ und bedeutet soviel wie „Wie kann man etwas konstruieren, was *das* kann?“ (RORTY (1981), S. 171). Es ist schon auffallend, welche Ähnlichkeiten zu RYLE im berühmten, 1950 veröffentlichten Aufsatz seines Oxbridge-Kollegen Alan TURING (1967, 1987b) zu finden sind – insbesondere bei seiner Konstruktion des Tests zur Denkfähigkeit einer Maschine, bei seiner Diskussion des „Bewußtseinsarguments“ und bei seinen Bemerkungen über lernende Maschinen.

Konstrukteure einer künstlichen Intelligenz müssen allerdings die *black box* nicht nur öffnen, sondern auch selbst bauen; mit nicht zuviel Gewalt lassen sich so auch Produktionsregeln und trainierte Neuronale Netze als technische Umsetzung von RYLES Dispositionen verstehen. Besonders deutlich wird die Akzentverschiebung von KANT zu RYLE (von der transzendentalen Einheit und den Repräsentationen zum „Bündel der Dispositionen“) bei Marvin MINSKY, wenn er beispielsweise schreibt: „*Geist ist einfach, was Gehirne tun*. [...] Es ist gleichgültig, was Agenten *sind*, es ist nur wichtig, was sie *tun* – und womit sie verbunden sind.“ und „*Praktische Vernunft ist keine einfache Angelegenheit. Sie ist vielmehr eine unermesslich große Gesellschaft harterworbener praktikabler Ideen – einer Vielzahl im Leben erlernter Regeln und Erwartungen, Dispositionen und Tendenzen, Abwägungen und Kontrollen.*“ (1990, S. 287 und 22)

wußtsein hob). Als Reflex auf den vermeintlich oder tatsächlich zunehmenden Verlust des Realitäts- und Praxisbezuges dieser Philosophie suchte ein anderer Ansatz – vorzugsweise auf der anderen Seite des Kanals und des Atlantiks wirksam, also dort, wo man an den Schulen mehr und mit größerer Sympathie als hier LOCKE und HUME liest, und dort, wo das Projekt der KI entstand – die Lösung darin, die Wahrheit von Urteilen nicht mehr in der mehr oder weniger gesicherten Widerspiegelung oder Konstitution der Außenwelt in Bewußtseinsakten, sondern in einem intersubjektiv überprüfbareren, direkten Vergleich der Außenwelt mit den von „subjektiven“ Ungenauigkeiten gereinigten *idealsprachlichen* Rekonstruktionen dieser Urteile zu gründen.

3 Idealsprachliche Rekonstruktion

3.1 Leibniz: Kalkül des vernünftigen Denkens

Wenn man Charaktere oder Zeichen finden könnte, die geeignet wären, alle unsere Gedanken ebenso rein und streng auszudrücken, wie die Arithmetik die Zahlen oder die analytische Geometrie die Linien ausdrückt, könnte man offenbar bei allen Gegenständen, soweit sie dem vernünftigen Denken unterworfen sind, das tun, was man in der Arithmetik und der Geometrie tut.

Denn alle Forschungen, die vom vernünftigen Denken abhängen, würden durch die Umwandlung dieser Charaktere und eine Art Kalkül zustande kommen, was die Erfindung schöner Dinge ganz leicht machen würde. Denn es würde nicht nötig sein, sich den Kopf ebenso zu zerbrechen, wie man heute gezwungen ist zu tun, und man würde trotzdem sicher sein, alles was hier zu tun sein würde, tun zu können *ex datis*.

Zudem würde man jeden von dem überzeugen, was man gefunden oder erschlossen hätte, da es leicht sein würde, den Kalkül zu prüfen, sei es, indem man ihn nachvollzieht, sei es indem man einige Proben versucht, ähnlich solchen, wie es die Neunerprobe in der Arithmetik ist. Und wenn jemand an dem, was ich vorgebracht haben würde, zweifelte, würde ich zu ihm sagen: „Rechnen wir, mein Herr!“, und Feder und Tinte nehmend, würden wir uns bald aus der Verlegenheit ziehen.⁸³

Mit diesen Sätzen formulierte Gottfried Wilhelm LEIBNIZ (1646–1716) programmatisch die Vision einer künstlichen, idealen Sprache, in der es „keine Mehrdeutigkeiten und Zweideutigkeiten“ gibt, und in der man „alles, was man verständlich sagen wird,“ angemessen sagen kann; sie würde ein Hilfsmittel sein, „das nicht weniger dazu dienen wird, die Vernunft zu steigern wie das Fernrohr dazu dient, das Sehen zu vervollkommen.“⁸⁴ Nötig wäre dazu ein System von Grundsymbolen, die den nicht weiter zerlegbaren Bestandteilen der Gedanken eineindeutig zugeordnet sind, und Verknüpfungsregeln für diese Symbole derart, daß die komplexen Zeichenreihen komplexen Gedanken entsprechen (*characteristica universalis*). Hätte man weiter gewisse Regeln zum Operieren mit diesen Symbolen, einen „Kalkül“ (*calculus ratiocinator*), so daß die Ergebnisse der Operationen den Ergebnissen „vernünftigen Denkens“ (z. B. logisches Schließen) entsprechen, so wäre der Traum in Erfüllung gegangen: alle vernünftigen Gedanken ließen sich durch schematisches Operieren mit Symbolen finden (*ars inveniendi*); und alle Streitfragen wären so zu lösen, wie man im Restaurant die Rechnung prüft (*ars iudicandi*).

⁸³G. W. LEIBNIZ, *Allgemeine Wissenschaft* (Vorrede), in: LEIBNIZ, *Fragmente zur Logik*, ausgew., übers. u. erläutert v. F. SCHMIDT, Berlin 1960, S. 90 f.

⁸⁴A. a. O., S. 92.

LEIBNIZ selbst und einige andere lieferten Ansätze zu einer solchen idealen Sprache, zu einem groß angelegten Forschungsprogramm entwickelte sich dieses Projekt aber erst etwa zweihundert Jahre später, weniger als philosophisches Anliegen, sondern im Dienste der Mathematik. Das Selbstverständnis der Mathematik als exakte, empirieunabhängige Wissenschaft erlitt durch das Auftreten von Antinomien einen empfindlichen Schlag, Beweisverfahren und Begriffe, die aus einer finiten bzw. sehr anschaulichen Mathematik stammten, versagten bei der Anwendung auf unendlich viele Objekte. In dem Bemühen um die Sicherung der Grundlagen der Mathematik wurden als Hilfsmittel logische Kalküle des deduktiven Schließens entwickelt, um die häufig nur intuitiv oder inhaltlich begründeten Beweisverfahren narrensicher und streng überprüfbar zu machen. Als *symbolische* und *formale* Logik erlebte dadurch die Logik, von der noch KANT dachte, sie sei durch ARISTOTELES im wesentlichen abgeschlossen, eine neue Blüte.

Kalküle sind streng genommen nichts anderes als Herstellungsverfahren von Zeichenreihen aus bestimmten Grundzeichen nach bestimmten Regeln. Zwar kann man den Zeichen eine gewisse Bedeutung geben, somit den Kalkül als formale Sprache *interpretieren* (und meist werden Kalküle auch in Hinblick auf die intendierte Bedeutung entworfen, und erst dadurch werden sie zu *Symbolen*), aber die Pointe von Kalkülen ist gerade, daß das Operieren im Kalkül lediglich auf die äußere Gestalt, die *Form*, und nicht auch auf diese Bedeutung Bezug nimmt, von ihr völlig unabhängig ist. Festgelegt wird nur, welche Zeichen zum Kalkül gehören und auf welche Weise Zeichen zu Zeichenreihen formiert und Zeichenreihen in andere Zeichenreihen transformiert werden dürfen (z. B. durch Löschen, Verschieben, Ersetzen einzelner Zeichen). Daß solche Transformationsregeln – etwa: von $A \vee B$ und $\neg A$ darf zu B übergegangen werden (wobei A und B Mitteilungszeichen für wiederum nach bestimmten Regeln herstellbare Zeichenreihen sind) – in bestimmter Weise interpretiert werden können – etwa als formale Darstellung der Gedankenkette: „Peter ist mit Mary oder mit Jane unterwegs.“ „Mit Mary ist er nicht zusammen.“ „Dann ist er mit Jane unterwegs.“ –, ist zwar erwünscht, ein *Wissen* um diese Bedeutung der Zeichen ist aber, anders als bei der inhaltlichen Gedankenkette, für die korrekte Anwendung der Regel völlig unnötig. Die Kunst besteht nun darin, Kalküle als „symbolische Maschinen“⁸⁵ so zu entwerfen, daß im schlichten Vollzug der rein *syntaktischen* Regeln die intendierte *Semantik* automatisch mitvollzogen wird (und dann z. B. die gesamte Ableitung ein *Beweis* für die zuletzt erzeugte Symbolkette ist) nach dem „FORMALISTENMOTTO: Wenn man auf die Syntax achtet, wird die Semantik selbst auf sich achten.“⁸⁶

Vorbild und Anlaß für die Entwicklung der formalen Logik war zwar anfangs die Mathematik, aber die sich dabei zunächst einstellenden großen Erfolge ermutigten eine Reihe von an dieser Entwicklung maßgeblich beteiligten Philosophen dazu, die ursprüngliche Intention des von LEIBNIZ skizzierten Programmes wieder aufzunehmen und die neu entwickelten Methoden auf alle Bereiche des „vernünftigen Denkens“ auszudehnen. Von der Rekonstruktion der Gedanken und der Repräsentationen der Welt in einer idealen Sprache (statt in der privaten Sphäre des Bewußtseins), für jeden nach strengen Regeln aufs

⁸⁵Vgl. Sybille KRÄMER (1988).

⁸⁶HAUGELAND (1987), S. 92. Durch die rein syntaktische Darstellung der Regeln kann das Rechnen in Kalkülen auch durch wirkliche symbolmanipulierende Maschinen (Computer) ausgeführt werden, wenn die Anwendung der Regeln nicht freigestellt, sondern in *Algorithmen* eindeutig an bestimmte Bedingungen geknüpft und damit automatisierbar ist und der Algorithmus in ein *Programm* für einen Computer umgesetzt wird. Mehr zur Unterscheidung von Syntax und Semantik, zu Algorithmen und zur formalen Logik im Unterrichtsmodul 2.

einfachste überprüfbar, versprach man sich einen entscheidenden Fortschritt in der Philosophie bis hin zur endgültigen Klärung einer Reihe ihrer traditionellen Probleme und darüber hinaus, in einer neuen Synthese aus Rationalismus und Empirismus, wesentliche Impulse auch für die Erfahrungswissenschaften.

3.2 Frege: Die Wende zur Sprache

Gottlob FREGE (1848–1925) war – durch seinen Einfluß auf RUSSELL, WITGENSTEIN und CARNAP – ein wichtiger Wegbereiter des „linguistischen Paradigmas“ der Philosophie, der (sprach-) analytischen Philosophie, mit ihrer Grundüberzeugung, daß eine Klärung philosophischer Probleme nur durch eine philosophische Analyse der Sprache erreicht werden kann. Allerdings war sein Einfluß in dieser Richtung eher unfreiwillig, sein ursprüngliches Hauptinteresse galt nicht der Sprache, sondern der philosophischen Klärung der Grundlagen der Logik und Mathematik (mit seinem in der *Begriffsschrift* entwickelten Logikkalkül wird er vielfach als Begründer der modernen formalen Logik angesehen). Er war entschiedener Gegner eines damals verbreiteten Psychologismus, der die Gesetze der Logik als Gesetze des Denkens auf die Psychologie zurückführen wollte. Dagegen vertrat er eine vollständig objektive Auffassung der Logik und wollte weiter die (nichtgeometrische) Mathematik nur auf der Logik gründen, d. h. gegen KANT (dem er vorwarf, nicht genügend zwischen subjektiven und objektiven Vorstellungen unterschieden zu haben) ihre „analytische und apriorische Natur“ nachweisen.⁸⁷

Zur Durchführung dieses Programms entwickelte er nicht nur die formale Logik (um Beweise von bloß subjektiven Einsichten und Intuitionen freizuhalten), er unternahm auch eine tiefgreifende logische Analyse der mathematischen Begrifflichkeit (Variable, Konstante, Funktion usw.), deren Ergebnisse er dann auch auf andere Bereiche übertrug. Bei seiner Untersuchung ließ er sich von drei methodologischen Grundsätzen leiten: „Es ist das Psychologische von dem Logischen, das Subjektive von dem Objektiven scharf zu trennen; Nach der Bedeutung der Wörter muß im Satzzusammenhange, nicht in ihrer Vereinzelung gefragt werden; Der Unterschied zwischen Begriff und Gegenstand ist im Auge zu behalten.“ Der zweite Grundsatz hängt eng mit dem ersten zusammen, denn wenn man ihn „unbeachtet läßt, ist man fast genötigt, als Bedeutung der Wörter innere Bilder oder Taten der einzelnen Seele zu nehmen und damit auch gegen den ersten zu verstoßen. Was den dritten Punkt betrifft, so ist es nur Schein, wenn man meint, einen Begriff zum Gegenstande machen zu können, ohne ihn zu verändern.“⁸⁸ Den ersten Grundsatz verdeutlicht er an anderer Stelle:

Wahrheiten zu entdecken, ist Aufgabe aller Wissenschaften: der Logik kommt es zu, die Gesetze des Wahrseins zu erkennen. [...] Aus den Gesetzen des Wahrseins ergeben sich nun Vorschriften für das Fürwahrhalten, das Denken, Urteilen, Schließen.

⁸⁷*Die Grundlagen der Arithmetik: Eine logisch mathematische Untersuchung über den Begriff der Zahl*, Stuttgart 1987 (Erstausgabe 1884), § 109, S. 138. „Die Vorstellung im subjektiven Sinne ist das, worauf sich die psychologischen Assoziationsgesetze beziehen; sie ist von sinnlicher, bildhafter Beschaffenheit. Die Vorstellung im objektiven Sinne gehört der Logik an und ist wesentlich unsinnlich [...] Die subjektive Vorstellung ist oft nachweisbar verschieden in verschiedenen Menschen, die objektive für alle dieselbe. Die objektiven Vorstellungen kann man einteilen in Gegenstände und Begriffe. Ich werde, um Verwirrung zu vermeiden, ‚Vorstellung‘ nur im subjektiven Sinne gebrauchen. Dadurch, daß Kant mit diesem Worte beide Bedeutungen verband, hat er seiner Lehre eine sehr subjektive, idealistische Färbung gegeben und das Treffen seiner wahren Meinung erschwert.“ (§ 27 Anm., S. 59)

⁸⁸A. a. O., Einleitung, S. 23 f.

Und so spricht man wohl auch von Denkgesetzen. Aber hier liegt die Gefahr nahe, Verschiedenes zu vermischen. Man versteht vielleicht das Wort „Denkgesetz“ ähnlich wie „Naturgesetz“ und meint dabei das Allgemeine im seelischen Geschehen des Denkens. Ein Denkgesetz in diesem Sinne wäre ein psychologisches Gesetz. Und so kann man zu der Meinung kommen, es handle sich in der Logik um den seelischen Vorgang des Denkens und um die psychologischen Gesetze, nach denen es geschieht. Aber damit wäre die Aufgabe der Logik verkannt; denn hierbei erhält die Wahrheit nicht die ihr gebührende Stellung. Der Irrtum, der Aberglaube hat ebenso seine Ursachen wie die richtige Erkenntnis. Das Fürwahrhalten des Falschen und das Fürwahrhalten des Wahren kommen beide nach psychologischen Gesetzen zustande. Eine Ableitung aus diesen und eine Erklärung eines seelischen Vorganges, der in ein Fürwahrhalten ausläuft, kann nie einen Beweis dessen ersetzen, auf das sich dieses Fürwahrhalten bezieht. [...] Um jedes Mißverständnis auszuschließen und die Grenze zwischen Psychologie und Logik nicht verwischen zu lassen, weise ich der Logik die Aufgabe zu, die Gesetze des Wahrseins zu finden, nicht die des Fürwahrhaltens oder Denkens. In den Gesetzen des Wahrseins wird die Bedeutung des Wortes „wahr“ entwickelt. [...] Das Wort „wahr“ erscheint sprachlich als Eigenschaftswort. Dabei entsteht der Wunsch, das Gebiet enger abzugrenzen, auf dem die Wahrheit ausgesagt werden, wo überhaupt Wahrheit in Frage kommen könne.⁸⁹

Vorstellungen sind nur einem Menschen zugänglich, sie sind privat, letztlich nicht vergleichbar, „bloß subjektiv“; Vorstellungen (auch die von KANT „Urteile“ genannten) kommen daher für FREGE als Kandidaten nicht in Betracht. Vielmehr ist es der *Gedanke*, „bei dem überhaupt Wahrheit in Frage kommen kann.“ Gedanken sind – Vorbild ist da in einem ganz platonischen Verständnis der Mathematik der pythagoreische Lehrsatz – „zeitlos wahr, unabhängig davon wahr, ob irgend jemand ihn für wahr hält.“ Sie „sind weder Dinge der Außenwelt noch Vorstellungen. Ein drittes Reich muß anerkannt werden. Was zu diesem gehört, stimmt mit den Vorstellungen darin überein, daß es nicht mit den Sinnen wahrgenommen werden kann, mit den Dingen aber darin, daß es keines Trägers bedarf, zu dessen Bewußtseinsinhalte es gehört. [...] Wenn man einen Gedanken faßt oder denkt, so schafft man ihn nicht, sondern tritt nur zu ihm, der schon vorher bestand, in eine gewisse Beziehung, die verschieden ist von der des Sehens eines Dinges und von der des Habens einer Vorstellung. [...] Beim Denken erzeugen wir nicht die Gedanken, sondern wir fassen sie. Denn das, was ich Gedanken genannt habe, steht ja im engsten Zusammenhange mit der Wahrheit. Was ich als wahr anerkenne, von dem urteile ich, daß es wahr sei ganz unabhängig von meiner Anerkennung seiner Wahrheit, auch unabhängig davon, ob ich daran denke. Zum Wahrsein eines Gedankens gehört nicht, daß er gedacht werde. ‚Tatsachen! Tatsachen! Tatsachen!‘ ruft der Naturforscher aus, wenn er die Notwendigkeit einer sicheren Grundlegung der Wissenschaft einschärfen will. Was ist eine Tatsache? Eine Tatsache ist ein Gedanke, der wahr ist. Als sichere Grundlage der Wissenschaft aber wird der Naturforscher sicher nicht etwas anerkennen, was von den wechselnden Bewußtseinszuständen von Menschen abhängt. Die Arbeit der Wissenschaft besteht nicht in einem Schaffen, sondern in einem Entdecken von wahren Gedanken.“⁹⁰

⁸⁹Der *Gedanke. Eine logische Untersuchung*. in: G. FREGE, *Logische Untersuchungen* (herausgegeben von G. PATZIG), Göttingen 1966 (Erstveröffentlichung 1918), S. 30 f.

⁹⁰A. a. O., S. 33, 43 f., 50. Genauer unterscheidet er „1. das Fassen des Gedankens – das Denken, 2. die Anerkennung der Wahrheit eines Gedankens – das Urteilen, 3. die Kundgebung dieses Urteils – das Behaupten“ (S. 35). Das von ihm postulierte „dritte Reich“ (ein unverschuldet schrecklicher Name) entspricht der „Welt 3“ bei Karl POPPER. Heutzutage zieht man es in der Regel vor, nicht einen eigenen, mit starken metaphysischen Annahmen belasteten Bereich der „objektiven Gedanken“ zu postulieren. Wenn man

Damit ist sie geschehen, die „Verstoßung der Gedanken aus dem Bewußtsein.“⁹¹ Untersuchungsgegenstand sind nun nicht mehr Vorstellungen, sondern Gedanken. Wie können wir Gedanken untersuchen, wie erhalten wir Zugang zu dieser besonderen Welt? „Ich bin hier nicht in der glücklichen Lage eines Mineralogen, der seinen Zuhörern einen Bergkristall zeigt. Ich kann meinen Lesern nicht einen Gedanken in die Hände geben mit der Bitte, ihn von allen Seiten recht genau zu betrachten. Ich muß mich begnügen, den an sich unsinnlichen Gedanken in die sinnliche sprachliche Form gehüllt dem Leser darzubieten. Dabei macht die Bildlichkeit der Sprache Schwierigkeiten. Das Sinnliche drängt sich immer wieder ein und macht den Ausdruck bildlich und damit uneigentlich. So entsteht ein Kampf mit der Sprache, und ich werde genötigt, mich noch mit der Sprache zu befassen, obwohl das hier nicht meine eigentliche Aufgabe ist.“ Fast widerwillig – wegen der im Gegensatz zur Mathematik notorischen Mehrdeutigkeit und Kontextabhängigkeit der Sprache – vollzieht hier FREGE die *Wende zur Sprache*. „Demnach kann ich sagen: der Gedanke ist der Sinn eines Satzes, ohne damit behaupten zu wollen, daß der Sinn jedes Satzes ein Gedanke sei. Der an sich unsinnliche Gedanke kleidet sich in das sinnliche Gewand des Satzes und wird uns damit faßbarer. Wir sagen, der Satz drückt einen Gedanken aus.“⁹² Weiter gilt (und hier kommt der zweite o. g. Grundsatz ins Spiel), daß wir mit begrenzten sprachlichen Mitteln „unübersehbar viele Gedanken“ ausdrücken können, und das wäre „nicht möglich, wenn wir in den Gedanken nicht Teile unterscheiden könnten, denen Satzteile entsprechen, so daß der Aufbau des Satzes als Bild gelten könnte des Aufbaues des Gedankens. Freilich sprechen wir eigentlich in einem Gleichnisse, wenn wir das Verhältnis von Ganzem und Teil auf den Gedanken übertragen. Doch liegt das Gleichnis so nahe und trifft im Ganzen so zu, daß wir das hier und da vorkommende Hinken kaum als störend empfinden.“⁹³

Unterschiedliche sinnlich gegebene Sätze (Äußerungen, Geschriebenes) können denselben Gedanken ausdrücken: „Peter liebt Mary“, „Mary wird von Peter geliebt“, „Peter loves Mary“ (woher wissen wir, daß derselbe Gedanke ausgedrückt wird? Nun, würde FREGE sagen, man kann nicht die Wahrheit oder Falschheit eines Satzes anerkennen ohne zugleich auch die Wahrheit oder Falschheit der anderen Sätze anzuerkennen), korrekte Umsetzungen vom Aktiv ins Passiv, fremdsprachliche Übersetzungen „berühren den Gedanken nicht“. Gleiches gilt nach FREGE für gewisse „Färbungen und Beleuchtungen, welche Dichtkunst und Beredsamkeit dem Sinne zu geben suchen“, die durch Worte wie „leider“ oder „gottlob“ erzeugt werden (oder auch durch die Wahl von „Roß“, „Gaul“, „Mähre“ statt „Pferd“), und „Winke in der Rede“ durch Worte wie „schon“, „noch“ oder „aber“ (statt „und“). In einer bestimmten Situation gesprochene Sätze enthalten Worte wie „heute“, „hier“, „ich“. „In allen solchen Fällen ist der bloße Wortlaut, wie er schriftlich festgehalten werden kann, nicht der vollständige Ausdruck des Gedankens, sondern man bedarf zu dessen richtiger Auffassung noch der Kenntnis gewisser das Sprechen begleitender Umstände, die dabei als Mittel des Gedankenausdrucks benutzt werden. Dazu können auch Fingerzeige,

als „subjektiv“ den Bereich charakterisiert, der von den Vorstellungen einzelner Menschen abhängt, und als „objektiv“ den, der (als „Dingwelt“) unabhängig von allen Menschen und ihren Vorstellungen ist, so macht es durchaus Sinn, einen weiteren Bereich zu unterscheiden, der zwar von der Existenz von Menschen, aber nicht nur je einzelner abhängt. Zu diesem meist „intersubjektiv“ genannten Bereich gehört sicher die Sprache – und wohl auch die durch sie ausgedrückten, öffentlich überprüfbaren Gedanken.

⁹¹M. DUMMETT (1988), S. 32 ff. Dort ist auch eine genauere Untersuchung dazu zu finden, wie aus dieser Wende zum *Gedanken* eine zur *Sprache* wurde. Vgl. auch HOLENSTEIN (1992).

⁹²*Der Gedanke*, S. 40 Anm., S. 33.

⁹³*Logische Untersuchungen. Dritter Teil: Gedankengefüge*, in: G. FREGE, *Logische Untersuchungen* (herausgegeben von G. PATZIG), Göttingen 1966 (Erstveröffentlichung 1923), S. 72.

Handbewegungen, Blicke gehören. [...] Erst der durch die Zeitbestimmung ergänzte und in jeder Hinsicht vollständige Satz drückt einen Gedanken aus. Dieser ist aber, wenn er wahr ist, nicht nur heute oder morgen, sondern zeitlos wahr.“ Hinzu kommt, daß in der Umgangssprache verschiedene Worte dieselbe für die Wahrheit eines Gedanken(gefüge)s wesentliche logische Beziehung („aber“, „und“, „sowie“) bzw. dieselben Worte verschiedene logische Beziehungen („ist“, „oder“) ausdrücken können (das war für FREGE ein Grund neben anderen, sich eine eigene logische Kunstsprache zu schaffen). „Eben darin sehe ich die größte Schwierigkeit der Philosophie, daß sie für ihre Arbeiten ein wenig geeignetes Werkzeug vorfindet, nämlich die Sprache des Lebens, für deren Ausbildung ganz andere Bedürfnisse mitbestimmend gewesen sind, als die der Philosophie. So ist auch die Logik genötigt, aus dem, was sie vorfindet, sich erst ein brauchbares Werkzeug zurechtzufileilen. Auch für diese Arbeit findet sie zuerst nur wenig brauchbare Werkzeuge vor.“ Man sieht, es ist nicht ganz einfach, den gedanklichen Kern aus einem Satz herauszupräparieren.⁹⁴

Um den durch einen Satz ausgedrückten Gedanken, den „Sinn“ eines Satzes, zu bestimmen, entwickelt FREGE eine Bedeutungstheorie, eine philosophische *Semantik*. Ausgangspunkt seiner Analyse sind Identitätsaussagen der Form $a = b$, bei denen „a“ und „b“ für Bezeichnungen für einzelne (im weitesten Sinne) Gegenstände, kurz *Eigennamen* stehen; solche Sätze können – anders als Sätze der Form $a = a$, die analytisch und *a priori* gültig sind – „oft sehr wertvolle Erweiterungen unserer Erkenntnis enthalten.“ Seine Frage ist nun, wovon genau die Identität ausgesagt wird. Die sprachlichen Ausdrücke selbst – z. B. „der Ehemann von Hannelore Renner“ und „der gegenwärtige Bundeskanzler“ – können es nicht sein, die Buchstabenfolgen sind offensichtlich verschieden. Die durch die Ausdrücke bezeichneten „Dinge“ kommen auch nicht in Frage, denn sonst „schiene $a = b$ von $a = a$ nicht verschieden sein zu können, falls nämlich $a = b$ wahr ist“; wir hätten dann die (analytische) Aussage, daß Helmut Kohl Helmut Kohl ist. „Eine Verschiedenheit kann nur dadurch zustande kommen, daß der Unterschied des Zeichens einem Unterschiede in der Art des Gegebenseins des Bezeichneten entspricht. [...] Es liegt nun nahe, mit einem Zeichen (Namen, Wortverbindung, Schriftzeichen) außer dem Bezeichneten, was die Bedeutung des Zeichens heißen möge, noch das verbunden zu denken, was ich den Sinn des Zeichens nennen möchte, worin die Art des Gegebenseins enthalten ist.“ Außer Zeichen und Gegenstand gibt es also ein Drittes, den Sinn des Ausdrucks, möglicherweise „einseitig“ und „abhängig vom Standorte“, aber nicht „subjektiv“: von „der Bedeutung und dem Sinne des Zeichens ist die mit ihm verknüpfte Vorstellung zu unterscheiden.“ Kurz: „Wir drücken mit einem Zeichen dessen Sinn aus und bezeichnen mit ihm dessen Bedeutung.“⁹⁵

Unterschiedliche Zeichen können denselben Sinn ausdrücken, sinnverschiedene Ausdrücke

⁹⁴ *Über Sinn und Bedeutung*, in: G. FREGE, *Funktion, Begriff, Bedeutung* (herausgegeben von G. PATZIG), Göttingen ³1969 (Erstveröffentlichung 1892), S. 45; *Der Gedanke*, S. 36 ff., 52; *Gedankengefüge*, S. 83. Nach FREGE kann auch ein Frage- oder Befehlssatz einen Gedanken „fassen“, weiter beschäftigt er sich auch mit der hypothetischen und indirekten Rede (und somit mit dem Unterschied von *Gebrauchen* und *Erwähnen*). Da es ihm aber wesentlich um das Urteilen und Behaupten geht, ist der Standardausdruck eines Gedankens der (indikativische) Aussagesatz. Heute ist es eher üblich, *Sätze* (grammatisch korrekte Wortfolgen), *Aussagen* und *Propositionen* (gemeinsamer Sinn sprachlich unterschiedlicher Aussagen, entspricht etwa FREGES „Gedanken“) zu unterscheiden.

⁹⁵ *Über Sinn und Bedeutung*, S. 40 f., 43, 45 f. Sein Beispiel ist „Morgenstern“ und „Abendstern“ als Bezeichnungen für die Venus (bei meinem Beispiel handelt es sich genau genommen um *Kennzeichnungen*). FREGES Unterscheidung wurde zwar vielfach (z. T. modifiziert) übernommen, zur Verwirrung können aber sehr unterschiedliche Begrifflichkeiten beitragen. So wird für „Sinn“ (engl. *sense, meaning*) auch „Bedeutung“ und „Intension“, für „Bedeutung“ (engl. *reference, denotation*) auch „Bezug“, „Referenz“ und „Extension“ verwendet.

dieselbe Bedeutung haben. Weiter gibt es Ausdrücke, die einen Sinn, aber keine Bedeutung haben („der gegenwärtige König von Frankreich“). In einer idealen Sprache hätte jeder Ausdruck einen bestimmten Sinn, „aber die Volkssprachen erfüllen diese Forderung vielfach nicht“. Dennoch wird der „Sinn eines Eigennamens . . . von jedem erfaßt, der die Sprache oder das Ganze von Bezeichnungen hinreichend kennt, der er angehört“ (wie, sagt FREGE leider nicht, und das Problem beschäftigt uns bis heute). „Wir können nun drei Stufen der Verschiedenheit von Wörtern, Ausdrücken und ganzen Sätzen erkennen. Entweder betrifft der Unterschied höchstens die Vorstellungen, oder den Sinn aber nicht die Bedeutung, oder endlich auch die Bedeutung. [...] Von den Vorstellungen und Anschauungen soll im folgenden nicht mehr die Rede sein; sie sind hier nur erwähnt worden, damit die Vorstellung, die ein Wort bei einem Hörer erweckt, nicht mit dessen Sinne oder dessen Bedeutung verwechselt werde.“⁹⁶

Nach diesen Vorarbeiten mag es überraschen, daß FREGE als Ausdruck eines Gedankens den Sinn eines Satzes und nicht dessen Bedeutung ansieht. Er argumentiert so: „Nehmen wir einmal an, der Satz habe eine Bedeutung! Ersetzen wir nun in ihm ein Wort durch ein anderes von derselben Bedeutung, aber anderem Sinne, so kann dies auf die Bedeutung des Satzes keinen Einfluß haben. Nun sehen wir aber, daß der Gedanke sich in solchem Falle ändert“. Denn der Gedanke des Satzes „Der Ehemann von Hannelore Renner ißt am liebsten Saumagen“ ist verschieden von dem des Satzes „Der gegenwärtige Bundeskanzler ißt am liebsten Saumagen“, jeweils nur unvollständig informierte Zeitgenossen könnten den einen darin ausgedrückten Gedanken für wahr, den anderen für falsch halten. „Der Gedanke kann also nicht die Bedeutung des Satzes sein, vielmehr werden wir ihn als den Sinn aufzufassen haben.“⁹⁷

In Sätzen können Ausdrücke vorkommen, die Sinn, aber nicht Bedeutung haben. Dann hat auch der Satz nur einen Sinn, nicht aber eine Bedeutung („Der gegenwärtige König von Frankreich ist kahl“). Was ist nun aber, sofern vorhanden, die Bedeutung eines ganzen Satzes? „Wer eine Bedeutung nicht anerkennt, der kann ihr ein Prädikat weder zu- noch absprechen. Nun wäre aber das Vordringen bis zur Bedeutung des Namens überflüssig; man könnte sich mit dem Sinne begnügen, wenn man beim Gedanken stehenbleiben wollte. [...] Warum wollen wir denn aber, daß jeder Eigenname nicht nur einen Sinn, sondern auch eine Bedeutung habe? Warum genügt uns der Gedanke nicht? Weil und soweit es uns auf seinen Wahrheitswert ankommt. [...] So werden wir dahin gedrängt, den *Wahrheitswert* eines Satzes als seine Bedeutung anzuerkennen. Ich verstehe unter dem Wahrheitswerte eines Satzes den Umstand, daß er wahr oder daß er falsch ist.“⁹⁸

Wenn „die Bedeutung eines Satzes sein Wahrheitswert ist, so muß dieser unverändert

⁹⁶A. a. O., S. 42, 45 f. Mit (im weitesten Sinne) Eigennamen allein lassen sich noch keine ganzen Sätze bilden. Benötigt werden noch Ausdrücke für Begriffe bzw. Beziehungen. FREGE faßt die Bedeutung solcher Ausdrücke als spezielle Funktionen mit „leeren Stellen“, als „ungesättigt“ auf („... ist rot“, „... liebt ...“); erst wenn diese Leerstellen durch zulässige Eigennamen als Argumente ausgefüllt werden, kann ein sinnvoller Satz entstehen (symbolisch z. B. $\Phi(a), \Psi(a, b)$). Der Vorteil dieser (in der Logik völlig eingebürgerten) Sicht gegenüber der traditionellen Unterscheidung in Subjekt und Prädikat ist u. a. eine sehr viel durchsichtigere Behandlung von Existenz- und Allaussagen.

⁹⁷A. a. O., S. 46 f.

⁹⁸A. a. O., S. 47 f. Entsprechend definiert er auch: „ein Begriff ist eine Funktion, deren Wert immer der Wahrheitswert ist. [...] Wir können demnach als Begriffsumfang den Wertverlauf einer Funktion bezeichnen, deren Wert für jedes Argument ein Wahrheitswert ist.“ (*Funktion und Begriff* (1891), in: *Funktion, Begriff, Bedeutung*, S. 28) Diese Auffassung von Begriffen ist weit entfernt von der LOCKES als Namen von Ideen.

bleiben, wenn ein Satzteil durch einen Ausdruck von derselben Bedeutung, aber anderem Sinne ersetzt wird. Und das ist in der Tat der Fall“, das LEIBNIZ zugeschriebene Prinzip des *salva veritate* ist gültig.⁹⁹ Alle wahren Sätze haben damit dieselbe Bedeutung, ebenso alle falschen. „Wir sehen daraus, daß in der Bedeutung des Satzes alles einzelne verwischt ist. Es kann uns also niemals auf die Bedeutung eines Satzes allein ankommen; aber auch der bloße Gedanke gibt keine Erkenntnis, sondern erst der Gedanke zusammen mit seiner Bedeutung, d. h. seinem Wahrheitswerte. Urteilen kann als Fortschreiten von einem Gedanken zu seinem Wahrheitswerte gefaßt werden. [...] Man könnte auch sagen, Urteilen sei Unterscheiden von Teilen innerhalb eines Wahrheitswertes. Diese Unterscheidung geschieht durch Rückgang zum Gedanken. Jeder Sinn, der zu einem Wahrheitswerte gehört, würde einer eigenen Weise der Zerlegung entsprechen.“¹⁰⁰

Denken als Fassen eines Gedankens, Tatsachen als wahre Gedanken, der Aufbau eines Satzes als Bild des Aufbaus eines Gedankens, Analyse von „gereinigten“, in einer idealen Sprache formulierten Sätzen im Hinblick auf Sinn und Wahrheit als Funktion ihrer Bestandteile – das ist nun der Weg, dem Denken – als Fassen eines Gedankens – auf die Spur zu kommen, das ist das Programm, das FREGE seinen Nachfolgern aufgegeben hat.

3.3 Russell: Logischer Atomismus

FREGES Interesse als Philosoph der Logik und Mathematik galt wie gesagt dem Analytischen und Apriorischen (und einer Ablehnung der synthetischen Urteile *a priori* KANTS für den Bereich der Mathematik). Dieses Interesse teilte auch Bertrand RUSSELL (1872–1970). Sein zusammen mit Alfred North WHITEHEAD verfaßtes monumentales Werk *Principia Mathematica* unternahm wie FREGES *Grundgesetze der Arithmetik* den Versuch, die Mathematik vollständig auf die Logik zurückzuführen und dabei die von ihm bei FREGE entdeckte Möglichkeit von Antinomien durch eine Hierarchie logischer Typen zu vermeiden (auch wenn dieser Versuch nicht gelang bzw. nicht akzeptiert wurde, hatte dieses Werk anders als das von FREGE großen Einfluß auf die Entwicklung der formalen Logik und ihrer Notation), darüber hinaus wollte er aber die dabei entwickelten Mittel und Methoden auch für den Bereich der Erkenntnistheorie fruchtbar machen.

RUSSELL, der als Student unter dem Dogmatismus des britischen Neo-Hegelianismus litt, hatte eine herzliche Abneigung gegen die Lehre, daß nur das Ganze wahr sei und daß das Denken die Welt erzeugt, gegen KANTS Lehre von Raum und Zeit als notwendige Anschauungsformen und der synthetischen Erkenntnis *a priori*, gegen alles in seinen Augen unnötig Metaphysische (und auch gegen die „merkwürdige“ und „schattenhafte“ Existenz von FREGES „Drittem Reich“). Er wollte zurück zu einer Philosophie, die schlicht die Vorstellung erlaubte, daß das Gras grün ist und daß die Sterne auch da sind, wenn niemand hinschaut; er wollte wieder anknüpfen an den britischen Empirismus mit seinen vertrau-

⁹⁹Als Ausnahmen von diesem Prinzip diskutiert er die indirekte Rede und – das wird uns weiter unten (S. 59) im Zusammenhang mit „Propositionalen Einstellungen“ wieder begegnen – daß-Sätze nach Wörtern wie „meinen“, „hoffen“, „bedauern“ usw. (*Über Sinn und Bedeutung*, S. 51 ff.)

¹⁰⁰A. a. O., S. 49f. FREGE sieht hier eine wechselseitige Verschränkung von Sinn und Wahrheit. Die genauen Verhältnisse und Abhängigkeiten von Sinn, Bedeutung und Wahrheit sind bis heute umstritten (FREGE hielt es für „wahrscheinlich, daß der Inhalt des Wortes ‚wahr‘ ganz einzigartig und undefinierbar ist.“ *Der Gedanke*, S. 32). Häufig wird, wenn nach den Wahrheitsbedingungen von *Propositionen* gefragt wird, die Klärung des Sinns eines Satzes davon unabhängig als gegeben schon vorausgesetzt (vgl. dazu M. DUMMETT (1988), S. 24 ff.).

enerweckenden, der Sinneswahrnehmung unmittelbar zugänglichen einfachen *Tatsachen*. Allerdings unterschied er sich in einem wesentlich von LOCKE und HUME, nämlich in der Frage, „ob Urteile aus sogenannten Ideen bestehen“, er vertrat den Standpunkt,

daß diese Theorie grundsätzlich verkehrt ist. Sie scheint zu besagen, daß es eine geistige Entität gibt, die als Idee von etwas bezeichnet werden kann, das außerhalb des Geistes der Person existiert, die diese Idee hat, und daß, da ein Urteil ein geistiger Vorgang ist, dessen Bestandteile Bestandteile des Geistes der urteilenden Person sein müssen. Dadurch werden die Ideen zu einem Vorhang zwischen uns und den Dingen außer uns – wir erreichen in Wirklichkeit in der Erkenntnis nie die Dinge, die wir angeblich erkennen, sondern nur die Ideen dieser Dinge. Die Relation zwischen Geist, Idee und Objekt nach dieser Theorie ist ganz und gar unverständlich, und kein Untersuchungsergebnis bestätigt, soweit ich sehen kann, das Eindringen der Idee zwischen Geist und Objekt. Ich habe den Verdacht, daß diese Auffassung durch eine Abneigung gegen Relationen und die Meinung genährt wird, der Geist könne keine Objekte erkennen, wenn es nicht etwas *in* ihm gibt, das als Zustand des Erkennens des Objekts bezeichnet werden könnte. Eine solche Ansicht führt jedoch ins Uferlose, da die Relation zwischen Idee und Objekt mit Hilfe der Annahme erklärt werden muß, daß die Idee selbst eine Idee des Objekts besitzt etc. ad infinitum. [...] Ein Urteil besteht nach unseren Feststellungen nicht aus geistigen Bestandteilen, die Ideen genannt werden, sondern ist ein Vorkommnis, dessen Bestandteile ein Geist und bestimmte Objekte sind, nämlich Individuen und Universalien.¹⁰¹

Relationen sind Gegenstände der Logik, und so machte sich RUSSELL an die Aufgabe, mittels *logischer* (und nicht grammatischer) *Analyse* die Idee der Idee als überflüssig zu erweisen und andere Rätsel des Denkens und der Erkenntnis zu lösen.¹⁰² Dabei läßt er sich von der „Überzeugung des gesunden Menschenverstandes“ leiten, „daß es viele einzelne Dinge gibt und daß die augenscheinliche Mannigfaltigkeit in der Welt nicht nur aus scheinbaren Bestandteilen einer einzigen unteilbaren Realität besteht.“ Für ihn gibt es „unendlich viele Arten von Elementen. Es gibt Individuen, Qualitäten und Relationen verschiedener Ordnung. [...] Die einzige andere Art von Objekten, die uns in der Welt begegnen, sind die Dinge, die wir Tatsachen nennen. Die Tatsachen aber sind die Dinge, die mit Hilfe von Aussagen behauptet oder verneint werden.“ Tatsachen sind also „diejenigen Dinge, die eine Aussage entweder wahr oder falsch machen.“ Den so durch logische Analyse ergänzten Empirismus nennt er „*logischen* Atomismus, weil die Atome, zu denen ich als den letzten unzerlegbaren Bestandteilen bei der Analyse kommen möchte, nicht

¹⁰¹ *Erkenntnis durch Bekanntschaft und Erkenntnis durch Beschreibung* (Original: *Knowledge by Acquaintance and Knowledge by Description*, 1910–11), in: B. RUSSELL, *Die Philosophie des Logischen Atomismus* (Hrsg. J. SINNREICH), München 1979, S. 74f., 81.

¹⁰² Als Musterbeispiel einer logischen Analyse galt RUSSELL und seinen Anhängern seine Theorie der *Beschreibungen* (Kennzeichnungen). Für ihn ist „der gegenwärtige König von Frankreich“ nicht der *Name* einer (schattenhaft existierenden) Entität, sondern eine *Aussagefunktion* (nämlich „x ist gegenwärtig König von Frankreich, und für alle y gilt: wenn y gegenwärtig König von Frankreich ist, dann ist y mit x identisch“), die erst zu einer Aussage wird, wenn x bestimmt wird; auf diese Weise wird der Ausdruck als Bestandteil des Urteils (und somit auch die Notwendigkeit einer mit ihm korrespondierenden Idee) „weg-analysiert“, die Annahme einer unbekanntem Entität wurde durch die logische Konstruktion aus bekannten Entitäten ersetzt (und die Analyse hat die genauere logische Form unter der einfacheren grammatischen Oberfläche aufgezeigt). Eine echte Kennzeichnung (*definite description*), die logisch einen Eigennamen vertreten kann, liegt für RUSSELL demnach nur vor, wenn Existenz und Eindeutigkeit gesichert ist (anders als für FREGE ist für ihn „Der gegenwärtige König von Frankreich ist kahl“ nicht bedeutungslos – d. h. ohne Wahrheitswert –, sondern *falsch*).

physikalische, sondern logische Atome sind. Von diesen Atomen sind die einen die Dinge, die ich Individuen nenne – kleine Farbflecke, Töne, kurzlebige Dinge – die anderen die Prädikate, Relationen etc.“¹⁰³ Dieser „moderne analytische, von mir skizzierte Empirismus unterscheidet sich von dem Empirismus Lockes, Berkeleys und Humes dadurch, daß er die Mathematik einbezieht und eine brauchbare logische Technik entwickelt. Das ermöglicht es ihm, bei gewissen Problemen zu endgültigen Lösungen zu kommen, die mehr wissenschaftlichen als philosophischen Charakter tragen. [...] Für mich steht es außer Frage, daß nur durch Methoden dieser Art philosophische Erkenntnisse – soweit überhaupt möglich – erlangt werden können; und ebensowenig zweifle ich daran, daß durch diese Methoden viele alte Probleme vollständig lösbar werden.“¹⁰⁴

Eine solche „wissenschaftliche“ Philosophie (über die „Welt“ als Menge aller Tatsachen) ist nach RUSSELLS Überzeugung möglich durch die Konstruktion einer logisch perfekten Sprache. „In einer solchen Sprache würden die Wörter einer Aussage mit Ausnahme von solchen wie ‚oder‘, ‚nicht‘, ‚wenn‘ und ‚dann‘, die eine andere Funktion erfüllen, je einem Bestandteil der betreffenden Tatsache korrespondieren. In einer logisch perfekten Sprache gäbe es für jedes einfache Objekt nur ein einziges Wort. Was nicht einfach ist, würde durch Kombinationen aus Wörtern für einfache Dinge ausgedrückt, so daß jedem Wort der Kombination ein Bestandteil des betreffenden Komplexes korrespondiert. Eine solche Sprache wäre vollkommen analytisch und würde auf den ersten Blick zeigen, welche logische Struktur die behauptete oder negierte Tatsache besitzt. Die in den *Principia Mathematica* entwickelte Sprache will eine solche Sprache sein.“ Eine solche Analyse und Darstellung hätte zum Ergebnis, „daß in einer logisch korrekten Symbolik eine fundamentale Übereinstimmung zwischen der Struktur einer Tatsache und ihres Symbols herrscht und daß die Komplexität eines Symbols mit der der symbolisierten Tatsache sehr genau übereinstimmt.“¹⁰⁵ Wenn dann die atomaren Aussagen nur aus Bestandteilen bestehen, mit deren Bedeutung wir „bekannt“ sind, und die Wahrheit komplexer Aussagen nur von der Wahrheit der atomaren Teilaussagen abhängt, ist das Ziel erreicht: unsere gesamte Erkenntnis ist dann entweder empirisch oder folgt logisch aus empirischen Erkenntnissen. Und: statt der obskuren mentalen Repräsentation haben wir eine öffentlich zugängliche Repräsentation der Welt in einer idealen Sprache.

3.4 Wittgenstein: Sätze als logisches Bild der Tatsachen

In das Zentrum philosophischer Untersuchungen rückt die Sprache bei Ludwig WITTGENSTEIN (1889–1951). In seiner frühen Schrift faßt er seine These über das Verhältnis von

¹⁰³ *Philosophie des logischen Atomismus* (Original: *The Philosophy of Logical Atomism*, 1918–19), in: B. RUSSELL, *Die Philosophie des Logischen Atomismus* (Hrsg. J. SINNREICH), München 1979, S. 178, 226, 179. Sinnesdaten als „Individuen“ gab RUSSELL später auf (in der Einleitung zur zweiten Auflage der *Principia Mathematica* (1924) werden nur noch „das“ und „jenes“ als Beispiele für Individuen genannt).

¹⁰⁴ RUSSELL (1975), S. 843. Später karikierte RUSSELL selbst Methode und Ziel der „logischen Analyse“, indem er von einem Dialog zwischen einem Ladenbesitzer und einem Angestellten im hinteren Teil des Ladens berichtete, der sich entspannt, als RUSSELL auf einer Wanderung sich in diesem Laden nach dem Weg erkundigte: „Der Herr möchte den kürzesten Weg nach Winchester wissen.“ „Winchester?“ „Ja.“ „Weg nach Winchester?“ „Ja.“ „Der kürzeste Weg?“ „Ja.“ „Weiß ich nicht.“ (zitiert nach H. HERRING in AYER (1970), S. 212).

¹⁰⁵ *Philosophie des logischen Atomismus*, S. 196 f. Eine solche Sprache wäre für unsere alltäglichen Bedürfnisse „unglaublich unbequem. Es ist daher geradezu eine Gnade, daß das nicht der Fall ist“ (S. 195). Im Vorwort zur Publikation dieser Vorlesungen kündigt er an, der Artikel befasse sich „mit gewissen Ideen, die von meinem Freunde und früheren Schüler Ludwig Wittgenstein stammen.“

Sprache, Denken und Wirklichkeit so zusammen:

Der Satz ist ein Modell der Wirklichkeit, so wie wir sie uns denken.¹⁰⁶

Diese These erscheint auf den ersten Blick recht unscheinbar, ihre Voraussetzungen, ihr präziser Inhalt und ihre weitreichenden Implikationen bedürfen der Erläuterung. Zu den Voraussetzungen gehören einige *ontologische* Annahmen. WITTGENSTEIN vertritt hier wie RUSSELL den logischen Atomismus. Gewisse „Gegenstände bilden die Substanz der Welt“ (was genau diese „Gegenstände“ sind, *sagt* WITTGENSTEIN nicht (und könnte es auch nicht), aber in seinen Sätzen *zeigt* sich, daß er an so etwas wie logische Individuen, Begriffe und Relationen denkt), es ist diesen Gegenständen „wesentlich, der Bestandteil eines Sachverhaltes sein zu können“, die „Konfiguration der Gegenstände bildet den Sachverhalt“ (solche „Sachverhalte“ sind z. B., daß diese Billardkugel rot ist, daß Peter Mary liebt, daß Hamburg zwischen Lübeck und Bremen liegt, „Gegenstände“ sind dann nicht nur diese Billardkugel, Peter und Bremen, sondern auch die Eigenschaft Rot und die Relation ... liegt zwischen ... und ...). Sachverhalte sind *mögliche* Konfigurationen von Gegenständen, wenn sie auch *wirklich* diese Konfiguration haben, so spricht WITTGENSTEIN vom „Bestehen von Sachverhalten“, von einer (positiven) „Tatsache“. Weiter ist wesentlich, daß die durch die Konfiguration der „Gegenstände“ gebildeten „Sachverhalte“ in gewissem Sinne elementar, atomar sind, d. h. nicht aufeinander zurückgeführt werden können: „Die Sachverhalte sind von einander unabhängig. Aus dem Bestehen oder Nichtbestehen eines Sachverhaltes kann nicht auf das Bestehen oder Nichtbestehen eines anderen geschlossen werden“ (ohne weitere Analyse ist somit keineswegs klar, ob die genannten Beispiele in diesem Sinne atomar sind). Unter dieser Voraussetzung bildet die Gesamtheit der Sachverhalte einen „logischen Raum“ von so vielen Dimensionen, wie es (atomare) Sachverhalte gibt, und in jeder Dimension können zwei Werte (positiv, negativ) angenommen werden. Die Gesamtheit aller (positiven) Tatsachen (und damit auch alle Sachverhalte, die nicht bestehen) bilden die „Welt“, die „gesamte Wirklichkeit“: „Das Bestehen und Nichtbestehen von Sachverhalten ist die Wirklichkeit.“¹⁰⁷

Modelle sind Bilder der Wirklichkeit, aber von ganz anderer Art als z. B. Photographien. Modelle können vereinfachte, auf das für wesentlich Gehaltene beschränkte Bilder sein; aber für dieses Wesentliche müssen sehr präzise Bedingungen erfüllt sein. In einem Modell (z. B. Drahtmodelle oder Strukturformeln für Moleküle in der Chemie) muß jedem Gegenstand einer bestimmten Art (Atome, Bindungen) des Originals genau ein Element einer bestimmten Art (Holzkugeln bzw. Buchstaben(folgen), Drahtstücke bzw. Striche) des Modells entsprechen. Weiter muß jeweils die „Möglichkeit seines Vorkommens in Sachverhalten“, die „Form“ eines Gegenstandes im Modell wiedergegeben werden (Drähte bzw. Striche müssen immer genau zwei Kugeln bzw. Buchstaben verbinden, keine Kugel bzw. Buchstabe darf ohne Verbindung sein, zu Kugeln einer bestimmten Farbe und Größe bzw. zu bestimmten Buchstaben muß eine bestimmte Anzahl von Verbindungen führen usw.): „Was jedes Bild, welcher Form immer, mit der Wirklichkeit gemein haben muß, um sie überhaupt – richtig oder falsch – abbilden zu können, ist die logische Form, das ist, die Form der Wirklichkeit.“ Durch „die Form der Abbildung“ stellt das Bild „dar, was es darstellt“ (und was „das Bild darstellt, ist sein Sinn“), „unabhängig von seiner Wahr- oder

¹⁰⁶ *Tractatus logico-philosophicus*, Frankfurt a. M. ⁵1968 (Erstveröffentlichung 1921), 4.01. Arbeitstitel dieser Schrift war weniger ironisch-bombastisch (und eher FREGE als SPINOZA grüßend) schlicht *Der Satz*.

¹⁰⁷ A. a. O., 2.021, 2.011, 2.0272, 2, 2.061, 2.062, 1.13, 2.063, 2.06.

Falschheit“ („Um zu erkennen, ob das Bild wahr oder falsch ist, müssen wir es mit der Wirklichkeit vergleichen. Aus dem Bild allein ist nicht zu erkennen, ob es wahr oder falsch ist. Ein a priori wahres Bild gibt es nicht.“). Eine Strukturformel kann aber formal völlig korrekt sein und dennoch *falsch*. Der „Zusammenhang der Elemente des Bildes“ muß auch die „Art und Weise, wie die Gegenstände im Sachverhalt zusammenhängen“, die „Struktur“ des Sachverhaltes (für die die Form nur die Möglichkeit ist) wiedergeben, erst dann handelt es sich bei dem Bild um ein zutreffendes Modell (man spricht dann auch von einer *isomorphen Abbildung*).¹⁰⁸

Auch die möglichen Konfigurationen sinnlich wahrnehmbarer (Laut- oder Schrift-) Zeichen sind Sachverhalte, konkrete Äußerungen von Sätzen also Tatsachen. Als solche können sie, wenn die entsprechenden Bedingungen erfüllt sind, auch Modelle von bestehenden Sachverhalten sein. Der Satz „Peter liebt Mary“ besteht aus den Elementen „Peter“, „Mary“ und „... liebt ...“, wobei letzteres (im Gegensatz zu „Peter“ und „Mary“) zwei ungesättigte Leerstellen hat, die durch Elemente der Art von „Peter“ und „Mary“ gefüllt werden („Nur insoweit ist der Satz ein Bild der Sachlage, als er logisch gliedert ist.“). Auf Grund dieser *Form* ist der Satz (und z. B. auch „aRb“) ein *mögliches* Modell für den Sachverhalt, daß Peter Mary liebt (der ja aus der Konfiguration zweier Personen und einer zweistelligen (unsymmetrischen) Relation gebildet wird), auf Grund der *Struktur* („Peter“ steht *vor* „liebt“, „Mary“ *dahinter*; „Nicht: ‚Das komplexe Zeichen ‚aRb‘ sagt, daß a in der Beziehung R zu b steht‘, sondern: *Daß* ‚a‘ in einer gewissen Beziehung zu ‚b‘ steht, sagt, *daß* aRb.“) ist er ein *zutreffendes* Modell („Der Satz *zeigt* seinen Sinn“, er „*zeigt*, wie es sich verhält, *wenn* er wahr ist. Und er *sagt*, *daß* es sich so verhält.“), somit ein Modell der Wirklichkeit, wenn der Sachverhalt eine Tatsache ist.¹⁰⁹

Selbstverständlich ist eine solche Analyse nicht immer so einfach wie bei „Peter liebt Mary“ (man versuche es beispielsweise mal mit dem letzten Satz). „Die Umgangssprache ist ein Teil des menschlichen Organismus und nicht weniger kompliziert als dieser. Es ist menschenunmöglich, die Sprachlogik aus ihr unmittelbar zu entnehmen. Die Sprache verkleidet den Gedanken. [...] Die stillschweigenden Abmachungen zum Verständnis der Umgangssprache sind enorm kompliziert. [...] So entstehen leicht die fundamentalsten Verwechslungen (deren die ganze Philosophie voll ist).“ WITTGENSTEIN intendierte daher eine künstliche Idealsprache: „Um diesen Irrtümern zu entgehen, müssen wir eine Zeichensprache verwenden, welche sie ausschließt, indem sie nicht das gleiche Zeichen in verschiedenen Symbolen, und Zeichen, welche auf verschiedene Art bezeichnen, nicht äußerlich auf die gleiche Art verwendet. Eine Zeichensprache also, die der *logischen* Grammatik – der logischen Syntax – gehorcht. (Die Begriffsschrift Freges und Russells ist eine solche Sprache, die allerdings noch nicht alle Fehler ausschließt.)“¹¹⁰

Den einfachsten, atomaren Tatsachen entsprechen als Modell einfachste Sätze: „Der einfachste Satz, der Elementarsatz, behauptet das Bestehen eines Sachverhaltes.“ Nun gibt

¹⁰⁸A. a. O., 2.0141, 2.18, 2.22, 2.221, 2.223–2.225, 2.15, 2.032, 2.033 (nebenbei: „Vorstellung“ wird im Englischen häufig genauso übersetzt wie „Darstellung“: *representation*).

¹⁰⁹A. a. O., 4.032, 3.1432, 4.022.

¹¹⁰A. a. O., 4.002, 3.324, 3.325. Im Gegensatz zu anderen hat WITTGENSTEIN gegen die Umgangssprache keineswegs den Einwand, daß sie vage und unpräzise sei („Alle Sätze unserer Umgangssprache sind tatsächlich, so wie sie sind, logisch vollkommen geordnet.“ 5.5563), sie ist nur zu kompliziert. Später gab er das Projekt einer Idealsprache nicht nur auf, er entwickelte sich auch zu einem der schärfsten Kritiker an der Möglichkeit und des Sinns eines solchen Projekts, d. h. an der Annahme, daß sich die „Welt“ auf „eine und nur eine vollständige“ (3.25) Art in einzelne, von einander unabhängige, kontextinvariante „Tatsachen“ zerlegen läßt, die ihrerseits wieder auf diese Weise aus einfachsten „Gegenständen“ zusammengesetzt sind.

es Sätze, für die das nicht gilt, z. B. „Der schöne Peter liebt Mary, nicht Jane“. Hier wird von mehreren Sachverhalten gesprochen, nämlich davon, daß die Sachverhalte, daß Peter schön ist und daß Peter Mary liebt, bestehen und daß der Sachverhalt, daß Peter Jane liebt, nicht besteht. Der komplexe Satz ist genau dann wahr, wenn die Teilsätze „Peter ist schön“ und „Peter liebt Mary“ wahr sind und der Teilsatz „Peter liebt Jane“ falsch ist, der Wahrheitswert des komplexen Satzes ergibt sich eindeutig aus den Wahrheitswerten der Teilsätze: „Der Satz ist eine Wahrheitsfunktion der Elementarsätze.“¹¹¹

Wüßten wir den Wahrheitswert aller Elementarsätze, so wüßten wir auch den Wahrheitswert aller Sätze; und: wir hätten ein perfektes Modell „der Tatsachen im logischen Raum“, der tatsächlichen „Welt“ im Rahmen möglicher Welten. WITTGENSTEIN erläutert das in einem Gleichnis: „Ein Bild zur Erklärung des Wahrheitsbegriffes: Schwarzer Fleck auf weißem Papier; die Form des Fleckes kann man beschreiben, indem man für jeden Punkt der Fläche angibt, ob er weiß oder schwarz ist. Der Tatsache, daß ein Punkt schwarz ist, entspricht eine positive – der, daß ein Punkt weiß (nicht schwarz) ist, eine negative Tatsache.“ Mit einer idealen Sprache könnten wir solche Flecke (und damit den Zustand der Welt in jedem Augenblick bzw. Ausschnitte davon) gleichsam *punktgenau* beschreiben. Mit den formalen Sprachen, die wir tatsächlich benutzen – seien es die Formeln der Mechanik oder irgendwelche Programmiersprachen –, ist das nur vergrößernd möglich: ich kann der Beschreibung des Fleckes „beliebig nahe kommen, indem ich die Fläche mit einem entsprechend feinen quadratischen Netzwerk bedecke und nun von jedem Quadrat sage, daß es weiß oder schwarz ist. Ich werde auf diese Weise die Beschreibung der Fläche auf eine einheitliche Form gebracht haben. Diese Form ist beliebig, denn ich hätte mit dem gleichen Erfolge ein Netz aus dreieckigen oder sechseckigen Maschen verwenden können. [...] Den verschiedenen Netzen entsprechen verschiedene Systeme der Weltbeschreibung. Alle Sätze der Weltbeschreibung müssen aus einer Anzahl gegebener Sätze ... auf eine gegebene Art und Weise erhalten werden. [...] Welches Gebäude immer du aufführen willst, jedes muß du irgendwie mit diesen und nur diesen Bausteinen zusammenbringen.“¹¹² (Allerdings: Quadrate sind keine Punkte, in einer solchen *digitalen* Beschreibung wird, wie bei den Pixeln auf meinem Monitor, eine diagonale Linie immer eine Treppe sein.)

Die Elemente eines Satzes bilden nicht einfach eine Konfiguration wie die Sterne des Sternbildes Orion, *wir* bilden sie; sie sind nicht einfach ein Modell der Tatsachen, *wir gebrauchen* sie *als* Modell: „Wir benützen das sinnlich wahrnehmbare Zeichen (Laut- oder Schriftzeichen etc.) des Satzes als Projektion der möglichen Sachlage.“ Insofern ist der Satz ein Modell der Wirklichkeit, *so wie wir sie uns denken*: die „Projektionsmethode ist das Denken des Satz-Sinnes“, im „Satz drückt sich der Gedanke sinnlich wahrnehmbar aus.“ Das „angewandte, gedachte Satzzeichen“, „der sinnvolle Satz“, „der Gedanke“, ist dann das „logische Bild der Tatsachen“. Denken ist also, so kann man zusammenfassen, nach WITTGENSTEIN ein Abbilden mit Zeichen gemäß der logischen (Elementar-) Form der Welt. Und

¹¹¹A. a. O., 4.21, 5. Mein Beispiel ist nur illustrativ gemeint, es ist sehr fraglich, daß „Peter ist schön“, „Peter liebt Mary“ usw. im Sinne WITTGENSTEINS Elementarsätze sind (ihr Kennzeichen ist, „daß kein Elementarsatz mit ihm in Widerspruch stehen kann.“ 4.211): Sätze wie „Peter ist häßlich“, „Peter verachtet Mary“ etc. stehen zu den Sätzen im Widerspruch, ohne schlichte Verneinungen zu sein.

Eine in der Logik übliche Darstellung des komplexen Satzes der Art $aS \wedge aRb \wedge \neg aRc$ wäre nach WITTGENSTEIN kein Modell der Wirklichkeit: „Mein Grundgedanke ist, daß die ‚logischen Konstanten‘ nicht vertreten. Daß sich die *Logik* der Tatsachen nicht vertreten läßt“ (4.0312). Die „logischen Konstanten“ sind keine Namen für irgendwelche die Wirklichkeit repräsentierenden Ideen, allerdings – das war HOBBS' Fehler – ermöglichen erst sie ein „Denken ist Rechnen“.

¹¹²A. a. O., 4.063, 6.341.

den Sinn („seine Übereinstimmung, und Nichtübereinstimmung mit den Möglichkeiten des Bestehens und Nichtbestehens der Sachverhalte“) eines Satzes „verstehen, heißt, wissen was der Fall ist, wenn er wahr ist.“¹¹³

Damit ist der Raum der „Gesamtheit der wahren Sätze“, „die gesamte Naturwissenschaft“ abgesteckt: „Ein Sachverhalt ist denkbar“ heißt: Wir können uns ein Bild von ihm machen.“ (allerdings: „Wir fühlen, daß selbst, wenn alle *möglichen* wissenschaftlichen Fragen beantwortet sind, unsere Lebensprobleme noch gar nicht berührt sind. Freilich bleibt dann eben keine Frage mehr; und eben dies ist die Antwort.“). Der Philosophie bleibt als Zweck „die logische Klärung der Gedanken“, denn alles „was überhaupt gedacht werden kann, kann klar gedacht werden. Alles was sich aussprechen läßt, läßt sich klar aussprechen.“ Damit ist aber auch dem Denken eine Grenze gezogen, die Philosophie „soll das Denkbare abgrenzen und damit das Undenkbare. Sie soll das Undenkbare von innen durch das Denkbare begrenzen. [...] Wir können nichts Unlogisches denken, weil wir sonst unlogisch denken müßten. [...] Die Logik ist transcendental. [...] *Die Grenzen meiner Sprache* bedeuten die Grenzen meiner Welt. Die Logik erfüllt die Welt; die Grenzen der Welt sind auch ihre Grenzen“ („... und was jenseits der Grenze liegt, wird einfach Unsinn sein“).¹¹⁴

WITTGENSTEIN hat damit das Denken auf die Logik zurückgeführt, nicht umgekehrt. „Die Logik ist keine Lehre, sondern ein Spiegelbild der Welt“, sie stellt ihr „Gerüst“ dar, wir können ihr folgen, sie aber nicht begründen („Die Logik muß für sich selber sorgen“), ja, noch nicht einmal darstellen. Der sinnvolle Satz, der Gedanke, ist das logische Bild der Tatsachen, die „Form der Abbildung aber kann das Bild nicht abbilden; es weist sie auf. [...] Der Satz kann die gesamte Wirklichkeit darstellen, aber er kann nicht das darstellen, was er mit der Wirklichkeit gemein haben muß, um sie darstellen zu können – die logische Form. Um die logische Form darstellen zu können, müßten wir uns mit dem Satz außerhalb der Logik aufstellen können, das heißt außerhalb der Welt. Der Satz kann die logische Form nicht darstellen, sie spiegelt sich in ihm. [...] Was *sich* in der Sprache ausdrückt, können *wir* nicht durch sie ausdrücken. Der Satz *zeigt* die logische Form der Wirklichkeit. [...] Es gibt allerdings Unaussprechliches. Dies *zeigt* sich, es ist das Mystische.“¹¹⁵

¹¹³A. a. O., 3.11, 3.1, 3.5, 4, 3, 4.2, 4.024. Hier wird deutlich, daß für den WITTGENSTEIN des *Tractatus* – anders als in seiner Spätphilosophie, wo er neben dem Beschreiben noch eine Unzahl weiterer „Sprachspiele“ annimmt – nur deskriptive, beschreibende Sätze einen *Sinn* haben, also *sinnvoll* sind (die *Bedeutung* ganzer Sätze kommt im *Tractatus* nicht vor, nur die Bedeutung einzelner „Namen“ (der Satz ist kein Name) als der durch sie vertretene „Gegenstand“, 3.144, 3.203, 3.22).

¹¹⁴A. a. O., 4.11, 3.001, 6.52, 4.112, 4.116, 4.114, 3.03, 6.13, 5.6, 5.61, Vorwort. „Wittgenstein verlegt die Grenzen der theoretischen Vernunft an die Grenzen der Sprache. Während Kant ‚der theoretischen Vernunft möglich‘ für einen *engeren* Begriff als ‚logisch möglich‘ hielt, sind nach Wittgenstein diese beiden Begriffe *identisch*. [...] Es ist für Wittgensteins Anschauung wesentlich, daß die logische Analyse der Sprache, wie er sie auffaßt, eine Art transzendentaler Deduktion im Sinne Kants ist mit dem Ziel, die apriorische Form der Erfahrung anzugeben, die durch jede sinnvolle Sprache ‚gezeigt‘ wird und daher nicht ‚gesagt‘ werden kann. Von diesem Standpunkt aus könnte der Traktat eine ‚Kritik der reinen Sprache‘ genannt werden.“ (STENIUS (1969), S. 285, 287)

¹¹⁵*Tractatus*, 6.13, 6.124, 5.473, 2.172, 4.12, 4.121, 6.522. „Meine Sätze erläutern dadurch, daß sie der, welcher mich versteht, am Ende als unsinnig erkennt, wenn er durch sie – auf ihnen – über sie hinausgestiegen ist. (Er muß sozusagen die Leiter wegwerfen, nachdem er auf ihr hinaufgestiegen ist.)“ (6.54) „Der Versuch, den Begriff des logischen Bildes verständlich zu machen, führt Wittgenstein in eine Ontologie, durch die schließlich die Tatsache, daß wir sprechen und denken können, rätselhafter wird als sie es vor dem Versuch war, sie zu erklären. [...] Er hat ein Problem durch ein neues ersetzt: Wie sollen jene Gegenstände durch ihre logische Form das Denken zustande bringen?“ (STEINVORTH (1985), S. 114, 118).

3.5 Carnap: Rationale Nachkonstruktion der Wirklichkeit

Rudolf CARNAP (1891–1970), der bei FREGE in Jena u. a. die *Principia Mathematica* von WHITEHEAD/RUSSELL studierte¹¹⁶ und fünf Jahre in Wien lehrte (wo er auch Kontakt mit WITTGENSTEIN hatte), war in den dreißiger Jahren die zentrale Figur der empirischen und analytischen Philosophie und hatte – durch seine Lehrtätigkeit in Chicago, Harvard, Princeton und Los Angeles nach seiner Übersiedlung im Jahre 1936 – von allen bisher genannten Philosophen sicher den größten Einfluß auf die Entwicklung der Philosophie in den USA.¹¹⁷ CARNAP gehörte dem Wiener Kreis an, einer Vereinigung von Philosophen und Wissenschaftlern verschiedener Fachrichtungen, deren Interesse an Philosophie außer durch sprachlogische und erkenntnistheoretische Fragen auch noch *wissenschaftstheoretisch* bestimmt war:

Die alte Naturphilosophie wurde durch die Philosophie der Naturwissenschaften ersetzt. Die neuere Philosophie befaßt sich nicht mit der Entdeckung von Tatsachen und Gesetzen (das ist die Aufgabe des empirischen Wissenschaftlers) und auch nicht mit der Formulierung einer metaphysischen Theorie der Welt. Stattdessen richtet sie ihre Aufmerksamkeit auf die Wissenschaft selbst, indem sie die verwendeten Begriffe und Methoden, die möglichen Resultate, die Aussageformen und die Arten von Logik, die man verwenden kann, untersucht.¹¹⁸

Der Wiener Kreis verstand unter Philosophie „Wissenschaftslogik“ im Dienste der Wissenschaften (sie hat also nicht mehr wie bei KANT einen Vorrang durch Rechtfertigung der Möglichkeit von Wissenschaft, seine synthetischen Urteile *a priori* wurden nun – nach den *Principia Mathematica* und der Relativitätstheorie – als entweder nicht synthetisch oder nicht apriorisch abgelehnt), ihr Ziel war es aber auch, das Wissenschaftliche präzise abzugrenzen vom Unwissenschaftlichen, von Wertsetzungen, Kunst, Metaphysik, Religion und Aberglauben.¹¹⁹ Eine Trennung von Natur- und Geisteswissenschaften wurde abgelehnt, angestrebt wurde eine *Einheitswissenschaft* nach dem Leitbild der experimentellen Physik mit einer einheitlichen Sprache. Auch Disziplinen wie Psychologie und Soziologie sollten sich (wie etwa im Behaviorismus) ausschließlich auf intersubjektiv überprüfbare Beobachtungen der „Körperwelt“ und mathematisch-logische Methoden stützen, die (Sprache der) Psychologie sollte auf die (Sprache der) Biologie, diese auf die (Sprache der) Physik

¹¹⁶Lustigerweise zusammen mit Wilhelm FLITNER, ohne den (CARNAP war einziger Interessent) nach der Devise *tres faciunt collegium* die Veranstaltung nicht stattgefunden hätte (vgl. SEIFFERT (1992), S.13).

¹¹⁷In den oben (S. 4) zitierten Selbstauskünften über die geistigen „Ahnen und Geschwister“ der KI fehlt zwar sein Name (vielleicht weil einer der schärfsten Kritiker der frühen Vorstellungen von CARNAP Rudolf CARNAP war), aber der KI-Pionier Herbert SIMON studierte bei ihm in Chicago (vgl. GARDNER (1989), S. 35).

¹¹⁸*Einführung in die Philosophie der Naturwissenschaften*, Frankfurt a. M.–Berlin 1986 (Orig. *Philosophical Foundations of Physics*, New York 1966), S. 187.

¹¹⁹„Nehmen wir irgendein Buch zur Hand, z. B. über Theologie oder Schulmetaphysik, so laßt uns fragen: Enthält es eine abstrakte Erörterung über Größe und Zahl? Nein. Enthält es eine auf Erfahrung beruhende Erörterung über Tatsachen und Existenz? Nein. So übergebe man es den Flammen, denn es kann nichts als Sophisterei und Blendwerk enthalten.“ CARNAP bezieht sich in seinem 1934 veröffentlichten Aufsatz *Die Aufgabe der Wissenschaftslogik* (in: SCHULTE & MCGUINNESS (1992), S. 106) ausdrücklich auf diesen Schlußsatz aus HUMES *Untersuchung über den menschlichen Verstand*, und auch A. J. AYER (1972, S. 69) zitiert in seinem 1936 erschienenen, vom Wiener Kreis inspirierten Buch diese Passage mit den Worten: „was ist dies anderes als eine rhetorische Version unserer eigenen These, daß ein Satz, der nicht entweder eine formal wahre Proposition oder eine empirische Hypothese ausdrückt, der wissenschaftlichen Bedeutung entbehrt?“

zurückgeführt werden, gelingt das nicht, so handelt es sich nicht um Wissenschaft (ein wahrlich einheitstiftendes, aber auch extrem reduktionistisches Programm). Sie knüpften an am strikten Empirismus und glaubten – ähnlich wie RUSSELL – seine Schwächen („Vorhang der Ideen“, Induktionsproblem) durch konsequente Anwendung der neu entwickelten formalen Logik überwinden zu können (*Logischer Empirismus*); sie hofften offenbar, daß, nachdem mit der Logik eine *Deduktionsmaschine* bereitgestellt wurde, sich eine ähnlich leistungsfähige und sichere *Induktionsmaschine* entwickeln ließe (in der Tat hat CARNAP in seiner zweiten Lebenshälfte auch dazu grundlegende Arbeiten verfaßt, ohne allerdings das Problem befriedigend lösen zu können¹²⁰).

Wie WITTGENSTEIN verstanden sie unter dem Sinn eines Satzes die Methode seiner Verifikation, d. h. sinnvoll waren für sie nur solche Sätze, die einen „empirischen Gehalt“ haben, „sachhaltig“ sind, für die es also (zumindest prinzipiell) ein unmittelbares, intersubjektiv gültiges Verfahren gibt, ihren Wahrheitsgehalt empirisch zu überprüfen (je nach konzeptioneller Akzentuierung „Beobachtungssatz“, „Protokollsatz“ oder „Basissatz“ genannt), oder solche, die sich auf solche Sätze zurückführen lassen. Daneben gibt es noch sinnlose Sätze, die analytischen Sätze, also Sätze, die auf Grund von Definitionen gelten oder tautologisch sind, alle anderen Sätze sind unsinnig. Damit lautet das *empiristische Sinnkriterium*: Ein Satz ist dann und nur dann sinnvoll, wenn er aus einer widerspruchsfreien Menge von Beobachtungssätzen logisch abgeleitet werden kann und kein analytischer Satz ist.¹²¹ Zur Illustration: „Auf der Vega gibt es grüne Aschenbecher.“ ist in diesem Sinne sinnvoll, „Junggesellen sind unverheiratet.“ ist sinnlos, „Die Würde des Menschen ist unantastbar. Sie zu achten und zu schützen ist Verpflichtung aller staatlichen Gewalt.“ ist unsinnig – und selbstverständlich ebenso „Ein Satz ist dann und nur dann sinnvoll, wenn er aus einer widerspruchsfreien Menge von Beobachtungssätzen logisch abgeleitet werden kann und kein analytischer Satz ist.“

CARNAP war von Anfang an fasziniert von den Möglichkeiten künstlicher und formaler Sprachen zur Klärung von Gedanken und zur Verständigung von Menschen aus unterschiedlichen Bereichen (der studierte Physiker ging zur Philosophie über, weil sein Dissertationsvorhaben, „eine Formalisierung der speziellen Relativitätstheorie in der Notation der *Principia*“, am physikalischen Institut in Jena nicht zugelassen wurde)¹²²; er lernte (und benutzte) Esperanto und interessierte sich auch für *Interlingua* (eine auch etwas spöttisch *Standard Average European* genannte Kunstsprache,¹²³ deren Idee heute auch eine gewisse Rolle als Hilfsmittel bei Übersetzungen mittels Computer (vgl. Unterrichtsmodul 2) spielt, und er verfaßte grundlegende Arbeiten zur logischen Syntax, zur extensionalen und

¹²⁰„Heute deutet eine Unzahl negativer Resultate – darunter einige imponierende Überlegungen von Goodman – darauf hin, daß es eine vollständige *formale* induktive Logik *nicht* geben kann.“ (PUTNAM (1982), S. 170)

¹²¹Diese erste Fassung des empiristischen Sinnkriteriums mußte – u. a. wegen der Nichtverifizierbarkeit von Naturgesetzen als logische Allsätze und wegen der Nichtzurückführbarkeit von Dispositionsbegriffen – mehrfach modifiziert und aufgeweicht werden (vgl. STEGMÜLLER (1978), S. 380 ff., 397 ff., 461 ff.).

In seiner Schrift *Überwindung der Metaphysik durch logische Analyse der Sprache* (Erstveröffentlichung 1931) „entlarvt“ CARNAP Sätze aus HEIDEGGERS Antrittsvorlesung (z. B. „*Das Nichts selbst nichtet.*“) mit Hilfe der Prädikatenlogik als sinnlose „metaphysische Scheinsätze“, indem er zeigte, daß sie sich nicht in eine „logisch korrekte Sprache“, in die „Schreibweise der Logistik“ übersetzen lassen. Noch heute könnte man die Antwort auf die Frage, ob das nun gegen HEIDEGGER oder gegen die Prädikatenlogik spricht, als eine Art Lackmus-Test zur Unterscheidung von KI-Anhängern und -Gegnern verwenden (für ein aktuelles Beispiel vgl. MOHNHAUPT & REHKÄMPER (1990), S. 40).

¹²²PUTNAM (1982), S. 170.

¹²³Vgl. CARNAP (1993), S. 109.

intensionalen Semantik sowie zur Struktur und Darstellung von Wissenschaftssprachen. Kurz: anders als der WITTGENSTEIN des *Tractatus* glaubte CARNAP durchaus, daß man sinnvoll über die Sprache sprechen kann.

In seinem ersten großen Werk ging er das gewaltige Unternehmen an, mit „der rationalen Nachkonstruktion von Begriffen aller Erkenntnisgebiete auf der Grundlage von Begriffen, die sich auf das unmittelbar Gegebene beziehen“, die im britischen Empirismus von LOCKE bis RUSSELL immer wieder aufgestellte „These, daß es grundsätzlich möglich sei, alle Begriffe auf das unmittelbar Gegebene zurückzuführen“, nicht etwa nur mit weiteren Argumenten zu stützen, sondern „zum ersten Mal den Versuch zu unternehmen, ein Begriffssystem der behaupteten Art wirklich aufzubauen“. Die Durchführung dieser „Aufgabe war nur möglich dank der neuen Logik,“ die entwickelt wurde „vor allem durch Frege, Whitehead und Russell“, einer „Logik der Beziehungen.“ Er sah darin die Erfüllung der Aufgabe, „eine Synthese des alten Empirismus mit dem alten Rationalismus herzustellen“ mit ihrer übereinstimmenden Ansicht, „daß alle Begriffe und Urteile aus der Zusammenwirkung von Erfahrung und Vernunft hervorgehen“, mit ihrer gemeinsamen These: „die Sinne liefern das Material der Erkenntnis, die Vernunft verarbeitet das Material in ein geordnetes System der Erkenntnis.“¹²⁴

CARNAPS Ziel war die Ableitung aller empirischen, wissenschaftlichen Begriffe aus einigen wenigen Grundbegriffen, ein „*Konstitutionssystem*“, d. h. ein „*Einheitssystem aller Begriffe*“, „eine stufenweise Ordnung der Gegenstände derart, daß die Gegenstände einer jeden Stufe aus denen der niederen Stufen konstituiert werden“, gemäß der von RUSSELL übernommenen (und als Motto dem Werk vorangestellten) Devise: „The supreme maxim in scientific philosophising is this: Wherever possible, logical constructions are to be substituted for inferred entities.“ Als „Grundelemente“ wählte er „*Elementarerlebnisse*“, „gewisse *Stellen des Erlebnisstromes*“, von denen „Aussagen gemacht werden können von der Art, daß eine solche Stelle zu einer bestimmten anderen in einer bestimmten Beziehung stehe“ (ohne allerdings zu behaupten, „der Erlebnisstrom könne eindeutig in solche Stellen *zerlegt* werden“, vielmehr war für ihn – in Anlehnung an KÖHLERS „*Gestalttheorie*“ – das „Gegebene“ die „*Erlebnisse selbst in ihrer Totalität und geschlossenen Einheit*“). Als Grundbegriff nahm er eine zweistellige, asymmetrische, „*Ähnlichkeitserinnerung*“ genannte „Grundrelation“ („x und y sind Elementarerlebnisse, die durch Vergleich einer Erinnerungsvorstellung von x mit y als teilähnlich erkannt sind, d. h. als in einem Erlebnisbestandteil annähernd übereinstimmend“). Diese Grundrelation bildet den einzigen (undefinierten) Grundbegriff des Systems („nicht die Grundelemente; diese werden erst aus den Grundrelationen (als deren Feld) konstituiert“).¹²⁵

Die eben erfolgte Charakterisierung der Grundrelation und der Grundelemente war nur eine „Übersetzung“ in „eine Sachverhaltsangabe in realistischer Sprache“, die nur „das

¹²⁴ *Der logische Aufbau der Welt*, Hamburg ²1961 (Erstveröffentlichung 1928), S. IX f. (Vorwort zur zweiten Auflage). In der Tat lohnt ein Vergleich mit den von DESCARTES aufgestellten methodischen Regeln (vgl. oben S. 7).

¹²⁵ A. a. O., §§ 1, 3, 67, 108 u. 75. CARNAP wählte hier eine „eigenpsychische“ Basis, konstituierte dann die „physischen“ und darauf die „fremdpsychischen“ und „geistigen Gegenstände“ (der Kulturwissenschaften). Mit diesem „methodischen Solipsismus“ (§ 64) war keine Parteinahme im Streit zwischen Realismus, Idealismus und Phänomenalismus verbunden, die „*Konstitutionstheorie stellt das ihnen gemeinsame, neutrale Fundament dar. Sie divergieren erst im metaphysischen Gebiet*“ (§ 178). Später plädierte er dann für eine „Basis im Physischen“, weil dann „eine größere intersubjektive Übereinstimmung besteht“ (S. XI). „Synthetische Urteile *a priori*“ kommen bei CARNAP „nicht vor“ (§ 106), bemerkenswert ist aber, daß sein Grundbegriff einen Vergleich *in der Zeit* erfordert (vgl. oben S. 19).

Verständnis der einzelnen Konstitutionen“ erleichtern soll (die dabei benutzten Begriffe aus der „Sprache der Psychologie“ müssen ja erst noch auf einer späteren Stufe konstituiert werden und dürfen – wenn kein *circulus vitiosus* vorliegen soll – dafür nicht schon vorausgesetzt werden). Eigentliche „Systemsprache“ ist die „*symbolische Sprache der Logistik*.“ Mit ihrer Hilfe werden durch explizite Definitionen aus der Grundrelation (Er) die weiteren Begriffe als Klassen und Relationen konstituiert (Beispiel: $Ae =_{Df} Er \cup \check{E}r \cup Er^0$; „Teilähnlichkeit“ als Vereinigung von Er, ihrer „Konversen“ und der „Identität im Felde von Er“), alle so definierten Gegenstände sind dann „*zurückführbar*“ auf die Gegenstände unterer Stufen. Auf diese Weise, durch „Strukturbeschreibungen“ (statt „Eigenschaftsbeschreibungen“) und rein „extensional“, wollte CARNAP die These belegen, „*daß die Gegenstände aller Wissenschaften aus denselben Grundgegenständen durch bloße Anwendung der Stufenformen Klasse und Relation konstituiert seien*“ und daß es somit möglich sei, „zu einer intersubjektiven, *objektiven Welt* zu gelangen, die begrifflich erfaßbar ist und zwar als eine identische für alle Subjekte.“¹²⁶

Ein solcher Nachweis für die „stolze These“, „daß die (rationale) Wissenschaft nicht nur jeden Gegenstand zu ihrem Objekt machen könne, sondern auch nirgends auf eine Grenze, auf eine grundsätzlich unbeantwortbare Frage stoße“ (die sich nach CARNAP unter Verweis auf WITTGENSTEIN durchaus verträgt „mit der demütigen Einsicht, daß wir auch mit der Beantwortung sämtlicher Fragen nicht etwa die vom Leben uns gestellte Aufgabe schon gelöst haben würden“),¹²⁷ wäre nicht nur die Verwirklichung der angestrebten Synthese aus Empirismus und Rationalismus, sondern auch des LEIBNIZ-Programmes einer idealen sprachlichen Repräsentation allen „vernünftigen“ Denkens in einem Kalkül, in den bloß syntaktischen Regeln der formalen Logik (und damit in ähnlich leistungsfähigen Programmiersprachen wie LISP und PROLOG).

Besonders deutlich zeigt sich CARNAPS Konstitutionssystem als Vorläufer des Projekts der KI durch die Verwendung einer vierten Sprache, der „*Sprache der fiktiven Konstruktion*“ mit ihren „*Operationsvorschriften für ein konstruktives Verfahren*.“ Bei uns Menschen liegt das „Gegebene“ im „Bewußtsein niemals als bloßes, unverarbeitetes Material vor, sondern immer schon in mehr oder weniger verwickelten Bindungen und Gestaltungen. Die Erkenntnissynthese, die Verarbeitung des Gegebenen zu Gebilden, zu Vorstellungen der Dinge, der ‚Wirklichkeit‘, geschieht meist unabsichtlich, nicht nach bewußtem Verfahren.“ Dagegen „ergeben sich aus dem Zweck der Konstitutionen als rationaler Nachkonstruktion der Gegenstandserkenntnis“, unter der Abstraktion einer „Trennung zwischen dem reinen Gegebenen und den synthetischen Komponenten“, eine Reihe von „*Fiktionen*“ unter der Rahmenannahme, „daß wir die Aufgabe haben, einem bestimmten Subjekt, das wir mit A bezeichnen wollen, die Operationen Schritt für Schritt vorzuschreiben, durch die A zur Konstruktion gewisser Schemata (der ‚Bestandslisten‘) gelangt, die den einzelnen, zu konstituierenden Gegenständen entsprechen“.¹²⁸

¹²⁶A. a. O., §§ 7, 75, 95, 110, 35, 10 ff., 43, 2 u. 42. Eine explizite Durchführung dieses Programmes geschieht nur für die unteren Stufen, die höheren werden (in Wortsprache) nur skizziert. Später gab er die „These der Extensionalität“ auf und er kam zu der „Erkenntnis, daß die Zurückführung höherer Begriffe auf niedere nicht immer in der Form von expliziten Definitionen möglich ist“ (S. XII f.).

¹²⁷A. a. O., § 183.

¹²⁸A. a. O., §§ 99–101. Ob CARNAP selbst schon diese Verbindung zur *maschinellen* Nachkonstruktion des Denkens und Erkennens sah, sei dahingestellt; für ihn lag der Wert dieser Sprache in ihrer Anschaulichkeit und in ihrer heuristischen Funktion: läßt sich nämlich „eine konstitutionale Definition in eine solche Operationsvorschrift übersetzen, so sind wir sicher, daß die Konstitution rein extensional ist, wie es die Konstitutionstheorie von jeder Konstitution verlangt.“ (a. a. O., § 99)

Für A gilt die „*Fiktion der zeitlichen Trennung des Gegebenen von der Verarbeitung*“ und dafür müssen „dem A bestimmte Fähigkeiten zugeschrieben werden,“ über die *wir* so nicht verfügen, und es müssen „ihm auch bestimmte Kenntnisse abgesprochen“ werden, die für uns selbstverständlich sind. So wird angenommen, „daß A von dem Gegebenen nichts vergißt“ und daß „jedes Elementarerlebnis *als ein identisches festgehalten* wird, so daß es bei der Verarbeitung mehr als einmal aufgegriffen und dabei jedesmal als *dasselbe* festgestellt werden kann“ (z. B. mit Hilfe einer Numerierung). Den – in der Sprache der formalen Logik formulierten – Konstitutionsstufen der Gegenstände unserer Erfahrung entsprechen dann genau definierte Operationen von A mit den Elementen seiner „Bestandslisten“ und der Zusammenfassung zu neuen Listen. (z. B. stellt A die Bestandsliste der oben genannten „Relation Ae auf, indem er alle Paare der Er-Liste in sie einträgt, ferner auch die umgekehrten Paare (d. h. außer a,b stets auch b,a) und schließlich alle Identitätspaare von Gliedern der Liste (a,a; b,b; usw.). Die *Rückübertragung* besteht hier darin, daß A in jeder der früher [...] aufgestellten Gegenstandsbeschreibungen eines Er-Gliedes (also eines Elementarerlebnisses) an Hand der Ae-Liste vermerkt, zu welchen anderen es in der Relation Ae steht.“) Die Bestandsliste der Grundrelation Er, als Basis aller Konstitutionen, muß zwar hinreichend reich sein, wie sie aber genau aussieht, wissen wir nicht, diese „Liste ist nur dem A, nicht uns bekannt“, „diese Fiktion zwingt uns, die Konstitutionen als Operationsvorschriften unabhängig vom individuellen Subjekt zu formulieren.“ „Andererseits ist wohl uns, nicht aber dem A der Sinn der Grundbeziehung bekannt“, semantisches Verstehen wird nicht vorausgesetzt, nur syntaktisches Operieren.¹²⁹

Eine Liste geordneter Paare von als „teilähnlich“ erkannten „Elementarerlebnissen“ (durch Sensoren gewonnen oder als Daten gegeben), ein hinreichend großer und zuverlässiger Speicher, eine interne Repräsentation z. B. durch Zahlen und die Fähigkeit zur Symbolmanipulation gemäß eindeutig vorgegebenen Operationsregeln, das ist nach CARNAP alles, was für „*eine rationale Nachkonstruktion des gesamten*, in der Erkenntnis vorwiegend intuitiv vollzogenen *Aufbaues der Wirklichkeit*“ nötig ist. Dabei gilt zwar, „daß die Konstitution nicht die Form des wirklichen Erkenntnisprozesses wiederzugeben hat, sondern nur als rationale Nachkonstruktion zu demselben Ergebnis zu führen braucht“, aber dabei „müssen selbstverständlich alle Inhalte, die wirklich in den Erlebnissen vorkommen auch konstitutiv zum Vorschein kommen“, „außer Sinnesempfindungen“ auch „*Gefühle [...] Gedanken, Wollungen* oder was sonst immer“.¹³⁰

Das LEIBNIZ-Programm einer Nachkonstruktion allen „vernünftigen“ Denkens „ex datis“, aus dem unmittelbar „Gegebenen“, in einer idealen Sprache, als Kalkül mit bloß syntaktischen Regeln der Symbolmanipulation, mit der Möglichkeit, im Inneren des Geistes verborgene, subjektiv gefärbte und häufig nur intuitiv vollzogene Prozesse des Denkens öffentlich mit narrensicheren Mitteln überprüfbar und automatisch generierbar zu machen, erlebte – in enger Verbindung mit der Entwicklung der formalen Logik, mit dem Logizismus von FREGE, mit dem logischen Atomismus und der Bildtheorie der Sprache bei RUSSELL und WITTGENSTEIN, mit dem wissenschaftstheoretischen Programm des Logischen

¹²⁹ A. a. O., §§ 101, 110, 108 u. 102.

¹³⁰ A. a. O., §§ 100, 81 u. 85. Die mehr ingenieurmäßig ausgerichteten Konstrukteure von Expertensystemen setzen auf dieselben Mittel, sind aber in den Zielen heute bescheidener. Nicht ganz so bescheiden ist die sprachorientierte KI-Forschung: „Angestrebt wird eine vollständig operationalisierte, extreme Form der Referenzsemantik, die bis auf die sensorische Ebene ‚durchgeschaltet‘ wird. Stets wird mit einer expliziten, maschinell handhabbaren und vollständigen Repräsentation der Diskurswelt gearbeitet.“ (Wolfgang WAHLSTER (1982), S. 213)

Empirismus bei CARNAP und im Wiener Kreis – in der ersten Hälfte dieses Jahrhunderts einen verheißungsvollen Aufschwung, erlitt aber spätestens seit der Jahrhundertmitte auch entscheidende Dämpfer.¹³¹

- Das Vorhaben, die Mathematik auf die Logik zurückzuführen, hat sich als nicht durchführbar erwiesen. Um dabei den klassischen Satzbestand der Mathematik zu erhalten, müssen logische Grundsätze angenommen werden, die keineswegs evident und somit im hohen Maße *ad hoc* sind.
- Der formale Aufbau der Logik machte es gerade möglich, prinzipielle Grenzen formaler Verfahren zu beweisen. Kurt GÖDEL zeigte, daß jedes zumindest für die elementare Zahlentheorie hinreichend ausdrucksstarke formale System wesentlich unvollständig ist (d. h. es gibt wahre Sätze, die formal nicht beweisbar sind) und daß seine Widerspruchsfreiheit nicht mit den formalen Mitteln dieses Systems (geschweige denn mit schwächeren) bewiesen werden kann. Alonzo CHURCH wies nach, daß schon der minimale, interessante Teil der formalen Logik, die Prädikatenlogik erster Stufe, unentscheidbar und somit keine *ars iudicandi* ist. Die von ihm und Alan TURING vorgeschlagenen formalen Präzisierungen des Begriffes der Berechenbarkeit und des Algorithmus führten zu der Entdeckung weiterer algorithmisch unlösbarer Probleme. Auf Leopold LÖWENHEIM und Thoralf SKOLEM geht der Nachweis zurück, daß formale Systeme wesentlich verschiedene Modelle haben können, daß es also nicht möglich ist, (unendliche) Gegenstandsbereiche in eindeutiger Weise durch die Struktureigenschaften formaler Systeme zu charakterisieren (eine weitere Folge ist z. B., daß es kein Axiomensystem mit nur endlich vielen Axiomen zur eindeutigen Definition der natürlichen Zahlen geben kann).
- Gerade der Aufbau der Logik als formales System, als Kalkül (üblicherweise durch „Axiome“ und „Ableitungsregeln“) hat den Blick für die Grundlagen der Logik geschärft. Unterschiedliche Axiome und Ableitungsregeln können zu denselben, aber auch zu verschiedenen Logiksystemen führen. Neben durchaus erwünschten Erweiterungen (z. B. modale, deontische, temporäre und andere nichtextensionale und nichtmonotone Logiken) betrifft das auch den Kern der wahrheitsdefiniten, zweiwertigen, extensionalen, deduktiven, kurz: „klassischen“ Logik. Die Auszeichnung eines Systems als *die* Logik bedarf dann einer Rechtfertigung außerhalb der Logik, insbesondere wenn (z. B. von den „Konstruktivisten“) andere, aber ähnlich leistungsfähige Systeme vorgeschlagen werden. Eine Entscheidung kann dann nur noch im Hinblick auf die intendierte Anwendung getroffen werden, man wird dann nicht nur Schlüsse verwerfen, weil sie nicht den Regeln entsprechen, sondern auch Regeln, weil sie nicht die erwünschten Schlüsse erlauben. Die Logik erhält damit eine *pragmatische* Basis, und eine Rückführung des Denkens auf eine für sich selbst sorgende oder durch die Beschaffenheit der „Welt“ gerechtfertigte Logik wird unmöglich.
- Der Streit um die Gültigkeit des ehrwürdigen *tertium non datur* ist nur ein Beispiel für die Fragwürdigkeit der grundlegenden Annahme einer klaren Trennung zwischen

¹³¹Neben den schon angedeuteten Revisionen im späteren Werk von RUSSELL, WITTGENSTEIN und CARNAP beziehe ich mich im folgenden außer auf die ausdrücklich genannten Personen auch auf die Arbeiten von u. a. John L. AUSTIN, Paul FEYERABEND, Thomas S. KUHN, Imre LAKATOS, Paul LORENZEN, Karl R. POPPER, Willard Van Orman QUINE und Hilary PUTNAM. Für eine etwas ausführlichere Darstellung der logischen und sprachphilosophischen Probleme vgl. das Unterrichtsmodul 2.

analytischen und empirischen Sätzen. Die Existenz eindeutiger Abgrenzungskriterien wird zunehmend in Zweifel gezogen zu Gunsten der These, daß es sich jeweils nur um Überzeugungen mit einer unterschiedlichen Bereitschaft der Revidierbarkeit handelt. (Für jemanden, der zu einer für alle Menschen identischen „objektiven Welt“ gelangen will durch Zurückführung aller Gegenstände auf eine unmittelbare Erfahrungsbasis kraft Logik und Definitionen, ist ein solches Ergebnis natürlich eine Katastrophe; unter anderem läßt sich dann nur noch schwer trennen, was innerhalb einer Theorie der Empirie und was der Sprache geschuldet ist.)

- Für das Herauspräparieren des unterschiedlichen sprachlichen Formulierungen oder in verschiedenen Sprachen formulierten gemeinsamen „Gedankens“ oder der jeweils vorliegenden „Proposition“ sind präzise Synonymitätsbedingungen nötig. Das ist bisher (für natürliche Sprachen) nicht überzeugend gelungen, und es gibt gute Gründe dafür, daß es auch nie gelingen wird. Damit ist nicht nur die Möglichkeit eindeutiger Übersetzungen (und ihre automatische Herstellung) in Frage gestellt, sondern auch die eindeutige und operationalisierbare Angabe der „logischen Form“ (mit einfachsten oder „atomaren“ Bestandteilen) eines Satzes (und auch die eindeutige „Übersetzung“ in jede andere formale oder „ideale“ Sprache).
- Außer in terminologisch streng reglementierten Zusammenhängen gelingt es kaum, einen Begriff durch eine Liste von (notwendigen und hinreichenden) Eigenschaften eindeutig zu definieren. Ob etwas ein Stuhl, ein Vogel, ein Witz, ein Yuppie ist oder nicht, entscheiden wir nicht nach solchen Listen, vielmehr sehen wir eine kaum zu präzisierende „Familienähnlichkeit“, wir vergleichen mit „irgendwie“ vorhandenen „Prototypen“.
- Die Konzentration auf Aussagesätze und ihre logische Verknüpfung, auf die deskriptive, beschreibende, behauptende Funktion der Sprache, verstellte den Blick dafür, daß wir mit Sätzen auch Handlungen ganz anderer Art vollziehen (bitten, versprechen, drohen, in die Irre führen, verhöhnen, schmeicheln, flirten, grüßen, ernennen, taufen usw. usf.), daß sich der Gebrauch von Sprache nicht auf die Kodierung und den Transport von „Informationen“ reduzieren läßt. Alle diese „Sprechakte“ und „Sprachspiele“ haben ihre eigene „Logik“, ihr „Sinn“ und ihre „Bedeutung“ erschließen sich nicht aus dem Satz allein (selbst wenn er „indikatorfrei“ ist).
- Die „klassische“ semantische Sicht, daß die Bedeutung sprachlicher Zeichen letztlich kontextinvariant durch den Bezug auf von der Sprache unabhängige (konkrete oder abstrakte) „Gegenstände“ der „Welt“ bestimmt wird (somit die Sprache ein Abbild der „Welt“ ist), wird zunehmend von der Ansicht verdrängt, daß die Bedeutung eines Zeichens die Regeln seines Gebrauchs sind (statt daß der Gebrauch sich nach der Bedeutung richtet). Sprache wird damit zu einer „sozialen Kunstfertigkeit“, die Semantik läßt sich nicht mehr trennscharf von der Pragmatik abgrenzen, Sinn- und Bedeutungsinstanz ist nicht mehr der einzelne Satz, sondern die gesamte Sprache der einzelnen Menschen und ihrer Sprachgemeinschaft (einschränkend muß allerdings gesagt werden, daß zur Zeit Sinn und Bedeutung der semantischen Prädikate „Sinn“ (auch: „Bedeutung“), „Bedeutung“ (auch: „Bezug“) und „Wahrheit“ sowie ihr Zusammenhang alles andere als klar sind).
- Beobachtungen werden mitbestimmt durch unsere Erwartungen und Interessen (und unsere biologische Ausstattung), Beobachtungen werden erlebt, formuliert, mitge-

teilt und überprüft gemäß sprachlicher Konventionen. Alle unsere Beobachtungen, auch unsere „einfachen“ Sinneswahrnehmungen sind demnach sprach- und theoriegetränkt. Auch wenn man KANTS Theorie der apriorischen Anschauungsformen und Verstandeskategorien nicht folgt, läßt sich mit guten Gründen bezweifeln, daß es ein „objektiv“ unmittelbar „Gegebenes“ gibt.

- Bei der Verifikation oder Falsifikation eines empirisch gehaltvollen Satzes steht genau genommen nie dieser einzelne Satz allein auf dem Prüfstand, vielmehr geht es um die Gesamtheit der diesen Satz prognostizierenden Theorie (einschließlich ihres logischen und mathematischen Apparates), der zugrundeliegenden sprachlichen Konventionen und der Randbedingungen; dafür, was im Falle einer Falsifikation davon modifiziert werden muß, lassen sich keine allgemeinen Kriterien angeben (und im Falle der Verifikation kann bis auf den Satz selbst alles andere falsch sein).
- Nach der Einsicht, daß das Verifikationsprinzip für wissenschaftliche Theorien und generelle Sätze schon aus logischen Gründen unhaltbar ist (und somit die Wahrheit einer Theorie niemals endgültig bewiesen werden kann), wurden andere Kriterien dafür gesucht, wissenschaftliche Vernunft, Regeln und Methoden wissenschaftlichen Erkenntnisfortschritts, vom Unwissenschaftlichen zu trennen. Bisher ist es nicht gelungen, überzeugende Kriterien dafür anzugeben, wann es „vernünftig“ ist, in einer bestimmten Situation angesichts vorliegender „Daten“ eine Theorie zugunsten einer konkurrierenden aufzugeben (u. a. weil konkurrierende Theorien im gewissen Sinne nicht über dasselbe sprechen, also auch nicht gemeinsame „Daten“ haben). Genauere wissenschaftshistorische Untersuchungen zeigen zudem, daß neue Theorien kaum aus bloß internen, „rein“ wissenschaftlichen Gründen akzeptiert werden, ohne daß deshalb Wissenschaft weniger erfolgreich ist („wissenschaftlich“ ist demnach das, was eine in ein soziales und historisches Umfeld eingebettete Gemeinschaft von Wissenschaftlern als Begründungen akzeptiert, eine von der Philosophie bereitgestellte allgemeine „Wissenschaftslogik“ wird dabei kaum vermisst).
- Neuere Untersuchungen zur Selbstorganisation und Emergenz (mit einer immer stärkeren Stützung der Thesen, daß das Ganze mehr ist als die Summe seiner Teile und daß das Ganze das Verhalten seiner Teile bestimmt) lassen es schließlich als äußerst fraglich erscheinen, daß es je gelingen wird, in den Bereichen Leben und Geist eine auch nur „naturwissenschaftliche“ Einheitswissenschaft zu etablieren (von den Kultur- und Geisteswissenschaften ganz zu schweigen). Phänomene aus diesen Bereichen erfordern wohl eine jeweils eigene Beschreibungs- und Erklärungsebene, die nicht vollständig auf den physikalisch-chemischen Mikrobereich reduziert („zurückgeführt“) werden kann.

Das alles hat dazu geführt, daß – außerhalb der KI¹³² – kaum noch jemand an den Logischen Empirismus und sein Programm der rationalen Nachkonstruktion *allen* „vernünft-

¹³² „Ich hoffe gezeigt zu haben, daß die Künstliche-Intelligenz-Forschung in dem Sinne in der Tradition Rudolf Carnaps steht, in welchem nicht nur Carnaps ‚Aufbau‘ als Vorwegnahme der Idee einer formalen Wissensrepräsentation gelten kann, sondern auch beide Forschungsprogramme zu ein und demselben Paradigma zu rechnen sind. Dieses Paradigma verbindet KI und Carnap methodisch durch den formalistischen Ansatz und die darin implizierte Idee, der Semantik auf rein syntaktischem Wege Genüge zu tun. [...] In diesem Sinne kann die KI als Fortsetzung des Programmes von Carnap und des Wiener Kreises mit neuen Mitteln aufgefaßt werden. Die Verheißung des Erfolges liegt m. E. in diesen Mitteln.“ (KESE (1988), S. 64f.)

tigen“ Denkens und Wissens durch direkte „Konfrontation“ mit der „Wirklichkeit“ und Ableitung in einer idealen Sprache glaubt. Die *technische* KI (vgl. Abschnitt 3 im Teil 1) zeigt sich durch dieses negative Resultat kaum berührt. Ähnlich wie die beispielsweise in Mathematik, Physik, Statistik, Spieltheorie und Wirtschaftswissenschaften schon vor der Existenz von Computern entwickelten „symbolischen Maschinen“ sollen hier Computerprogramme lediglich *genau bestimmte*, fest umrissene menschliche Denkprozesse leistungsanalog ersetzen (bzw. an Präzision, Ausdauer und Schnelligkeit übertreffen). Die Symbolmanipulation nach ausschließlich syntaktischen Regeln ist dabei nur das Mittelstück eines umfassenderen Denk- und Reflexionsprozesses, der beim Menschen beginnt und beim Menschen endet (*wir* sorgen für den richtigen *input* und *wir* interpretieren den *output* und *wir* haben den Computer und das Programm daraufhin entworfen). Die Manipulation der Symbole geschieht in pragmatischer, methodischer Absicht gemäß einer von uns intendierten Semantik, die Bedeutung der Symbole und somit die „Intelligenz“ der Maschine ist eine bloß *verliehene* (sowenig wie wir erwarten, daß Schreibpapier und Kugelschreiber irgend etwas davon „verstehen“, was wir mit ihrer Hilfe berechnen oder beweisen, und sowenig wir erwarten, daß ein Taschenrechner die Bedeutung der im Display angezeigten Symbole „verstehet“, sowenig müssen wir das bei einem Expertensystem in Bezug auf seine „Wissensbasis“ und „Produktionsregeln“ erwarten).

Anders ist die Situation für die *kognitive* KI. Hier geht es nicht um eine umfassende *Nach*konstruktion von Denken und Wissen, hier werden Computer mit *eigener*, nicht bloß simulierter, verliehener Intelligenz angestrebt (insofern ist das Programm der kognitiven KI noch radikaler als das des Logischen Empirismus). Der Optimismus der kognitiven KI beruht dabei auf der – über die *methodischen Vorschläge* von LEIBNIZ und CARNAP weit hinausgehenden – *empirischen These*, daß auch wir letztlich informationsverarbeitende, symbolmanipulierende „Maschinen“, somit auch eine Art Computer sind. So rückt – angereichert durch die inzwischen erreichten logischen, (sprach)philosophischen und technischen Mittel und Ergebnisse – wieder die Frage in den Mittelpunkt des Interesses, wie *wir* es denn im Unterschied zu den oben genannten symbolischen Maschinen leisten, im Gebrauch mentaler und sprachlicher Repräsentationen Intentionalität und Verstehen zu realisieren.

4 Funktionalismus

4.1 Turing: Die Universalmaschine

1937 veröffentlichte Alan TURING (1912–1954) einen zunächst wenig beachteten Aufsatz zu einem Problem der mathematischen Grundlagenforschung und – über den ursprünglichen

„Die Hauptformel des Tractatus $[\bar{p}, \bar{\xi}, N(\bar{\xi})]$ – sie drückt die fortgesetzte Anwendung auf alle Werte von ξ aus und meint die fortgesetzte logische Verknüpfung aller Elementarsätze – ließe sich in LISP umschreiben und wäre dann, angewendet auf die Beobachtungsfakten der Naturwissenschaft, das Programm für ein formales, vollständiges Weltbild – mehr: für eine ewig unfertige, aber im Prinzip fertigstellbare wissenschaftliche Tatsachenbeschreibung der Welt, wie sie dem Tractatus vorschwebte Ohne sich dessen bewußt zu sein, hält eine verbreitete positivistische Grundeinstellung diese Form der Weltbeschreibung für ausreichend und vollständig, geradezu für allein richtig die Wirklichkeit ausmachend. Der Computer erscheint in dieser Sicht als Werkzeug für die Ausführung dieses Programms. [...] Und es wird noch eine Weile dauern, bis man sagen können wird, die Zukunft der *Philosophischen Untersuchungen* zur Überwindung der Grenzen des *Computer-Tractatus* hat begonnen“ (Heinz ZEMANEK (1993), S. 88, 105).

Kontext hinaus – mit einer nachhaltigen Langzeitwirkung für die Informatik, Philosophie und Psychologie. Das Problem hatte David HILBERT der Fachwelt aufgegeben (und es war eine nur auf die Mathematik bezogene und somit etwas bescheidenere Fassung der *ars iudicandi* von LEIBNIZ): Läßt sich ein allgemeines, eindeutiges und endliches Entscheidungsverfahren – in heutiger Terminologie kurz: ein Algorithmus – dafür finden, ob eine beliebige (in der Sprache eines Kalküls vorgelegte) mathematische Aussage beweisbar (d. h. im Kalkül ableitbar) ist oder nicht?¹³³ TURING bewies – wie kurz vorher schon Alonzo CHURCH – in seinem Aufsatz, daß es kein solches Verfahren geben kann. Noch folgenreicher als dieses Resultat waren aber für ein Nachdenken über das Denken die dabei von ihm benutzten Mittel.

Der Nachweis, daß es kein solches Verfahren geben kann, daß also alle denkmöglichen Verfahren versagen werden, erfordert als generelle, allgemeingültige Aussage eine präzise Definition dieser Verfahren. Intuitiv war klar, worum es ging: Nach endlich vielen Schritten sollte durch die eindeutige und zwangsläufige Anwendung genau definierter Regeln, ohne weiteres Nachdenken, Intuition, Phantasie, also ganz *mechanisch*, ein eindeutiges Ergebnis vorliegen, *berechnet* worden sein. Um diese intuitive Vorstellung zu präzisieren, überlegte TURING, auf welche elementaren Operationen das allgemeine Vorgehen eines menschlichen Rechners (engl.: *computer*) reduziert werden kann, der (durch regelgeleitete Symbolmanipulation wie z. B. schriftliches Dividieren) Zahlen oder Formeln berechnet:

Rechnungen werden für gewöhnlich in der Weise ausgeführt, daß bestimmte Symbole auf ein Stück Papier geschrieben werden. Wir wollen annehmen, daß dieses Stück Papier kariert ist, wie das Rechenheft eines Kindes. Beim elementaren Rechnen wird zuweilen die Zweidimensionalität des Papiers ausgenutzt. Aber diese Verwendungsart ist keineswegs unvermeidlich, und ich denke, man wird sich darüber einig sein, daß die Zweidimensionalität des Papiers für die Rechnung nicht wesentlich ist. Ich gehe daher davon aus, daß die Rechnung auf eindimensionalem Papier ausgeführt wird, d. h. auf einem durch Felder unterteiltem Band. [...] Das Verhalten des Rechnenden wird zu jedem Zeitpunkt durch die wahrgenommenen Symbole und durch seinen momentanen „Geisteszustand“ bestimmt. [...] Wir stellen uns die vom Rechnenden durchgeführten Operationen in „einfache Operationen“ aufgeteilt vor, die so elementar sind, daß es schwer fällt, sie sich noch weiter aufgespalten vorzustellen. Jede dieser Operationen besteht in irgendeiner Veränderung im physikalischen System, das vom Rechnenden und seinem Band gebildet wird. Wir kennen den Zustand des Systems, wenn wir die Symbolfolge auf dem Band, die Symbole, die vom Rechnenden (gegebenenfalls mittels eines besonderen Befehls) wahrgenommen werden, und den Geisteszustand kennen. Wir nehmen an, daß in einer einfachen Operation nicht mehr als ein Symbol verändert wird. Alle anderen Veränderungen können auf einfache Veränderungen dieser Art zurückgeführt werden. [...] Die einfachen Operationen müssen neben diesen Symbolveränderungen auch Distributionswechsel der wahrgenommenen Felder umfassen. [...] Es ist möglich, daß einige dieser Veränderungen notwendig eine Veränderung

¹³³Als Lösung für die durch die Entdeckung von Antinomien beim Aufbau der Mengenlehre zu Beginn dieses Jahrhunderts ausgelöste Grundlagenkrise der Mathematik schlug HILBERT vor, die Mathematik durch streng formalisierte Kalküle aufzubauen und diesen Aufbau durch streng finite und unangreifbare *metamathematische* Mittel abzusichern. Das sogenannte „HILBERT-Programm“ umfaßte dabei erstens den Nachweis der *Vollständigkeit* der Kalküle (d. h. *alle* wahren Sätze der Mathematik sind im Kalkül ableitbar), zweitens den der *Widerspruchsfreiheit* (*nur* die wahren sind ableitbar), drittens – und darum geht es hier bei TURING – die *Entscheidbarkeit*. Sechs Jahre vor TURINGs Arbeit bewies Kurt GÖDEL schon die Unerfüllbarkeit der beiden ersten Forderungen für Kalküle, die mindestens die elementare Arithmetik umfassen. Zur Grundlagenkrise der Mathematik und die unterschiedlichen Ansätze zur Lösung vgl. Christian THIEL (1972), Kapitel 3. Den Zusammenhang zur Informatik stellt Wolfgang COY (1993) her.

des Geisteszustands implizieren. Als allgemeinste einfache Operationen müssen daher die beiden folgenden gelten:

- (A) Eine mögliche Symbolveränderung . . . zusammen mit einer möglichen Veränderung des Geisteszustands.
- (B) Eine mögliche Veränderung . . . wahrgenommener Felder zusammen mit einer möglichen Veränderung des Geisteszustandes.

Die tatsächlich durchgeführte Operation wird . . . durch den Geisteszustand des Rechnenden und durch die wahrgenommenen Symbole bestimmt. Insbesondere bestimmen sie den Geisteszustand des Rechnenden, nachdem die Operation ausgeführt worden ist.¹³⁴

TURING präzierte nun den Begriff der „mechanischen“ Berechenbarkeit, indem er eine automatische *Maschine* ersann, die analog zum so beschriebenen Vorgehen eines menschlichen Rechners arbeitet. Eine Maschine, „die nur über eine endliche Zahl von Zuständen q_1, q_2, \dots, q_R verfügt“ und die durch ein beidseitig beliebig verlängerbares „Band“ versorgt wird, „das (analog zum Papier) durch sie hindurchläuft und in Sektionen (‚Felder‘ genannt) aufgeteilt ist, von denen jedes ein ‚Symbol‘ tragen kann“ (aus einer endlichen Menge). Die Maschine verfügt über einen Lese- und Schreibkopf, durch den jeweils ein Feld bearbeitet und das dort vorhandene Symbol erkannt und durch ein anderes (oder gleiches) ersetzt werden kann. „Die Maschine kann auch das Feld ändern, das sie abtastet, aber nur durch Verschieben um eine Stelle nach rechts oder links. Zuzüglich zu jeder dieser Operationen kann der [Maschinen-] Zustand geändert werden.“¹³⁵

Ein typischer Arbeitsschritt einer solchen „Turingmaschine“ besteht also im Drucken eines Symbols S_i , einer Bewegung nach links oder rechts und dem Übergang in einen neuen Zustand q_j (ein möglicher Zustand ist, daß die Maschine ihre Arbeit beendet). Wenn nun in einer sogenannten „Maschinentafel“ für jede mögliche Kombination von Zustand der Maschine und auf dem Band vorgefundenem Symbol ein solcher Arbeitsschritt festgelegt wird (etwa $q_4 S_2 : S_5 L q_3$ dafür, daß die Maschine, wenn sie im Zustand q_4 auf dem aktuellen Feld das Symbol S_2 vorfindet, dort das Symbol S_5 druckt, mit dem Lesekopf ein Feld nach links rückt und in den Zustand q_3 übergeht), ist eine spezielle Turingmaschine und damit ein eindeutiger automatischer Ablauf definiert, durch den die ursprüngliche Bandinschrift verändert wird und der möglicherweise nach endlich vielen Schritten stoppt. Wenn das der Fall ist, so TURINGs Vorschlag, soll die dann durch die Bandinschrift repräsentierte Zahl oder Formel „berechenbar“ heißen, und er zeigte, als negative Lösung des Entscheidungsproblems, daß es genau beschreibbare Probleme gibt, die in diesem Sinne nicht berechenbar sind. Weiter konnte er nachweisen, daß einige Turingmaschinen bei geeigneter „Programmierung“ durch eine entsprechende Bandbeschriftung die Arbeit jeder

¹³⁴TURING (1987a), S. 40–43. Über die Beschaffenheit der hier angenommenen „Geisteszustände“ gibt TURING keine weiteren Auskünfte (außer daß die zu berücksichtigende Anzahl endlich ist). An anderer Stelle schreibt er: „Für den Rechnenden ist es immer möglich, seine Arbeit abzubrechen, fortzugehen und alles zu vergessen, um später wiederzukommen und die Arbeit fortzusetzen. Wenn er dies tut, muß er einen Zettel mit Anweisungen (die in irgendeiner standardisierten Form abgefaßt sind) zurücklassen, aus denen hervorgeht, wie die Arbeit fortgesetzt werden soll. Diese Notiz ist das Gegenstück zum ‚Geisteszustand‘. [...] Auf diese Weise ist der Entwicklungsstand der Rechnung in jedem Stadium vollständig abhängig von dem Zettel mit den Anweisungen und den Symbolen, die auf dem Band stehen“ (a. a. O. S. 46). Der hier deutlich werdende Widerspruch zwischen Willensfreiheit und Determinismus spielt in TURINGs Leben und Werk immer wieder eine Rolle (vgl. HODGES (1989), S. 126 ff.).

¹³⁵TURING (1987a), S. 20 f.

beliebigen Turing-Maschine übernehmen können („universelle Turingmaschinen“). Seine These:¹³⁶ Alles, was man vernünftigerweise als „berechenbar“ (*computable*) bezeichnen kann, ist durch eine (und damit wirklich durch *eine*) Turing-Maschine berechenbar.

TURING hat eine Maschine völlig neuer Art ersonnen. Eine Maschine, die nicht die *physischen* Kräfte der Menschen steigert, sondern eine *informationsverarbeitende* Maschine, mit diskreten Zuständen und für diskrete Symbole, tauglich für jede denkbare Berechnung, kurz: den digitalen Universalcomputer. TURING mag bei der Konzeption seiner Maschine (Papierband, Schreib- und Lesekopf) von Fernschreibern und Schreibmaschinen (auch hier wird ja je nach durch Betätigen der „Umschalter-“ bzw. „Shift-Taste“ hergestellten inneren „Zustand“ ein jeweils anderes Zeichen gedruckt) inspiriert gewesen sein, aber er hat nie den Versuch gemacht, seine Maschine in dieser Form auch wirklich zu bauen. Sie diente ihm nur zur präzisen Darstellung einer logischen *Funktion*, die durch ganz unterschiedliche Strukturen physikalisch realisiert werden kann (sei es mechanisch, hydraulisch oder elektronisch, und auch ein „Mensch, ausgestattet mit Papier, Bleistift und Radiergummi sowie strikter Disziplin unterworfen, ist in der Tat eine Universalmaschine“¹³⁷).

4.2 Putnam: Die funktionale Ebene

Den Gedanken, die logisch-*funktionale* Ebene von der physikalisch-strukturellen zu unterscheiden, hat gut zwei Jahrzehnte später Hilary PUTNAM in einem Beitrag zum Leib-Seele-Problem aufgegriffen. Um dafür zu argumentieren, daß die „verschiedenen Streitpunkte und Rätsel“ bei diesem Problem „ihrem Wesen nach rein sprachlich und logisch“, somit „ein leerer Streit um Worte“ sind, entwirft er das Gedankenexperiment einer Turingmaschine, die „elektronische Sinnesorgane“ hat und „Theorien über ihre eigene Struktur formulieren“ und überprüfen kann. Beispielsweise läßt sich dann, so PUTNAM, zu der Frage, ob „zulässigerweise geistige Ereignisse mit körperlichen Ereignissen identifiziert werden können oder nicht“, für die Maschine ein „logisches Analogon konstruieren“, wenn die „geistigen“ Zustände eines Menschen mit den funktionalen einer Maschine und die „körperlichen“ mit den physikalischen in Analogie gesetzt werden. Die Aussage „Ich habe genau dann Schmerzen, wenn ich eine C-Faser-Reizung habe“ entspricht dann der Aussage „Ich befinde mich genau dann im Zustand A, wenn die Flip-Flop-Schaltung 36 an ist“ (und letztere stellt für *uns* überhaupt kein philosophisches Problem dar, wohl aber für die Maschine, so wie die erste für uns). „Kurz, jede philosophische Argumentation, die jemals im Zusammenhang mit dem Körper/Geist-Problem benutzt worden ist . . . hat ihr genaues Gegenstück beim ‚Problem‘ der logischen und strukturellen Zustände von Turingmaschinen.“¹³⁸

Neben diesem Beitrag zum Leib-Seele-Problem entwirft PUTNAM – und das ist für unseren Zusammenhang wichtiger – in dem Aufsatz ein neues psychologisches Forschungsprogramm:

¹³⁶Da hier eine intuitive Vorstellung präzisiert wird, ist diese These selbstverständlich nicht beweisbar (sondern eine Definition). Allerdings hat sich gezeigt, daß alle anderen Vorschläge zur Definition von „effektiver Berechenbarkeit“ bzw. „Algorithmus“ (λ -Kalkül von CHURCH, der an den einfachsten Operationen eines Fließbandarbeiters orientierte Vorschlag von Emil POST, allgemein-rekursive Funktionen) äquivalent sind, so daß die These allgemein akzeptiert wird.

¹³⁷TURING (1987d), S. 91.

¹³⁸PUTNAM (1985a), S. 364, 397, 377, 366, 396.

Es ist interessant zu bemerken, daß es entsprechend den beiden möglichen Beschreibungen des Verhaltens einer Turingmaschine – dem strukturellen Entwurf des Technikers und der „Maschinentafel“ des Logikers – auch zwei mögliche Beschreibungen der Psychologie des Menschen gibt. Der „behavioristische“ Ansatz ... zielt darauf ab, schließlich eine vollständige physikalistische Beschreibung des Menschen zu liefern, die mit der Chemie und Physik verknüpft ist. Dies entspricht der Beschreibung, die ein Techniker oder Physiker von einer physikalisch realisierten Turingmaschine gibt. Aber es wäre auch möglich, nach einer abstrakteren Beschreibung der geistigen Vorgänge von Menschen Ausschau zu halten, in der auf „geistige“ Zustände (ihre physikalische Realisierung, falls sie überhaupt eine haben, wird nicht angegeben) und „Eindrücke“ (sie spielen die Rolle der Symbole auf dem Band der Maschine) zurückgegriffen wird – eine Beschreibung, in der die Gesetze, welche die Aufeinanderfolge der Zustände steuern, angegeben würden, sowie deren Beziehung zur Verbalisierung (oder jedenfalls zum verbalisierten Denken). Diese Beschreibung wäre das Analogon zu einer „Maschinentafel“¹³⁹

Gegen die Identitätstheorie¹⁴⁰ und gegen den Behaviorismus (vgl. den Exkurs im Unterrichtsmodul 1) stellte PUTNAM damit die Hypothese auf, daß „Schmerz zu haben“ und auch andere geistige Zustände (z. B. intentionale wie Überzeugungen und Wünsche) „ein funktionaler Zustand des Organismus“ sind und daß somit für ein angemessenes und gesichertes Verständnis dieser „Zustände“ auch Wissen über die „funktionale Organisation“ eines Lebewesens auf einer eigenen, nicht reduzierbaren Ebene zwischen der physikalischen und der des beobachtbaren Verhaltens nötig ist.¹⁴¹ Für diese These führte er an, daß sie „trotz ihrer zugestandenen Vagheit weit *weniger* vage ist als die ‚physikalisch-chemische Zustandshypothese‘ heute ist und daß sie der Forschung sowohl mathematischer als auch empirischer Art weitaus zugänglicher ist. Tatsächlich heißt, diese Hypothese zu untersuchen, zu versuchen, ‚mechanische‘ Modelle von Organismen zu produzieren – und ist das nicht in einem gewissen Sinn das, wovon die Psychologie handelt?“¹⁴²

Wenn und insofern bei diesem Forschungsansatz von der speziellen physikalischen Realisierung abgesehen wird, gibt es kaum noch einen guten Grund, nichtorganischen Realisierungen dieser „mechanischen Modelle“ Bewußtsein und andere geistige Attribute abzusprechen,¹⁴³ die „funktionale Organisation (Problemlösen, Denken) des Menschen oder der Maschine läßt sich durch die Abfolgen der geistigen bzw. logischen Zustände (und die dazu gehörigen Verbalisierungen)“ in gleicher Weise beschreiben, beim „rationalen Denken (oder Rechnen) ist das ‚Programm‘, das die Reihenfolge der Zustände usw. festlegt,“ in gleicher Weise „rationaler Kritik zugänglich.“¹⁴⁴

¹³⁹A. a. O., S. 379 f.

¹⁴⁰Sie geht (in ihrer starken Version als Typ-Identität) grob gesagt davon aus, daß jeweils bestimmte Typen geistiger Zustände und Prozesse mit spezifizierbaren neurophysiologischen Zuständen und Prozessen identisch sind (in der Art „Ich habe genau dann Schmerzen, wenn ich eine C-Faser-Reizung habe“), somit, in der Terminologie FREGES, einen anderen Sinn, aber dieselbe Bedeutung haben, und daß damit letztlich alles Geistige auf der neurophysiologischen Ebene erklärt werden kann.

¹⁴¹PUTNAM (1981), S. 128 und (1975b), S. 424.

¹⁴²PUTNAM (1981), S. 129.

¹⁴³Nach PUTNAM (1975a, S. 407) ist die Antwort auf die Frage, ob entsprechend hochentwickelte Roboter Bewußtsein haben, keine der *Entdeckung* sondern eine der *Entscheidung*. Und er plädiert dafür, unseren Begriff von Bewußtsein im Sinne einer positiven Antwort auszudehnen, denn „discrimination‘ based on the ‚softness‘ or ‚hardness‘ of the body parts of a synthetic ‚organism‘ seems as silly as discriminatory treatment of humans on the basis of skin color.“ Dieser Vorwurf des „menschlichen Chauvinismus“ wird noch heute gerne von Seiten der KI gegen ihre Gegner erhoben.

¹⁴⁴PUTNAM (1985a), S. 380. Wieweit er hier mehr als nur eine *Analogie* im Sinne hatte, ist nicht ganz

Parallel zu den programmatischen Aufsätzen PUTNAMs hat der Sprachwissenschaftler Noam CHOMSKY auch der Linguistik eine neue Sichtweise und Richtung gegeben. In brillanten Polemiken kritisierte er den Anspruch von SKINNER und anderen Behavioristen, „verbales Verhalten“ vollständig und adäquat im Rahmen einer Stimulus-Response-Theorie behandeln zu wollen. Dagegen argumentierte er überzeugend, daß es mit „Verhaltensanalyse“ und „Analyse der Wechselwirkung physikalischer Körper“ nicht möglich ist, den „kreativen Aspekt des Sprachgebrauchs“ zu erklären, „nämlich die prinzipielle menschliche Fähigkeit, neue Gedanken formulieren und völlig neue Formulierungen von Gedanken verstehen zu können, und zwar im Rahmen einer ‚instituierten Sprache‘, einer Sprache, die als Produkt einer Kultur Gesetzen und Prinzipien unterliegt, die teilweise für sie allein gelten, teilweise aber auch generelle Eigenschaften des Geistes reflektieren.“ Nach CHOMSKY müssen wir, „wenn wir jemals verstehen wollen, wie Sprache gebraucht oder erlernt wird,“ ein „kognitives System voraussetzen, ein System von Wissen und Glauben, das sich in frühester Kindheit entwickelt“. Und er plädiert – unter ausdrücklichem Rückgriff auf DESCARTES – für eine eigene Untersuchungsebene unter der Oberfläche des Verhaltens und über der physiologischen Ebene, dafür, „ein abstraktes theoretisches Modell zu entwickeln, das soweit wie möglich diese Phänomene erklärt und die Prinzipien ihrer Organisation und Funktionsweise aufdeckt“.¹⁴⁵

4.3 Fodor: Die Sprache des Geistes

Vor allem Jerry A. FODOR hat die Gedanken von PUTNAM und CHOMSKY (u. a. den, daß dasselbe äußerlich beobachtbare Verhalten Ausdruck verschiedener (Sprach-) Handlungen sein und dieselbe Handlung durch unterschiedliches Verhalten ausgedrückt werden kann) aufgenommen und zu einer heute als *Funktionalismus* bezeichneten Theorie des Geistes weiterentwickelt. Gegen die Identitätstheorie und das Programm der Einheitswissenschaft des Wiener Kreises¹⁴⁶ einerseits und den Behaviorismus andererseits forderte er das Forschungsprogramm einer *kognitiven* Psychologie:

Zur Konstruktion von psychologischen Theorien veranlaßt uns unter anderem das Vorhandensein der Entsprechungen von Synonymie und Ambiguität in nicht-verbalem Verhalten. Genauso wie – ein Beispiel aus der Linguistik – nicht jede Äußerung der Phonemfolge „Bank“ eine Äußerung desselben Wortes ist, genausowenig ist jedes Auftreten einer bestimmten Bewegung oder Muskelkontraktion für die Psychologie ein Fall desselben Verhaltens. Umgekehrt können in der Linguistik zwei phonetisch distinkte Äußerungen („Junggeselle“, „unverheirateter Mann“) in signifikanten Hinsichten äquivalent sein. [...] Für beide Wissenschaften hat das dieselbe Konsequenz. Wenn wir

klar. In diesem zuerst 1960 erschienenen programmatischen Aufsatz betont er zwar, daß der Verdacht, er „verträte die Ansicht, daß Maschinen denken oder – anders herum – daß Menschen Maschinen sind,“ falsch sei (S. 367). In einem späteren Aufsatz merkt er jedoch an, daß es keine logischen Gründe dafür gebe, daß der menschliche Geist bzw. sein Hirn keine Turingmaschine sei und daß diese Frage nur empirisch entschieden werden könne. Wahrscheinlich sei aber, daß der menschliche Geist (bzw. sein Hirn) zwar nur eine endliche Menge von Zuständen habe, daß aber anders als bei einer Turingmaschine auch probabilistische Gesetze und Zeitverzögerungen eine wichtige Rolle spielen, es sich bei uns also um eine „komplexere Art von Automat“ handelt. Andererseits lassen sich diese Art von Maschinen beliebig gut durch Turingmaschinen annähern, eine empirische Entscheidung sei dadurch schwierig (1975b, S. 412f., vgl. auch die vorherige Anmerkung). In den letzten Jahren ist PUTNAM von dieser Sichtweise und vom durch ihn initiierten Funktionalismus abgerückt (s. u.).

¹⁴⁵CHOMSKY (1973), S. 18, 15 und 30. Mehr zu CHOMSKYS Ansatz im Unterrichtsmodul 2.

¹⁴⁶Vgl. FODOR (1992).

zu den relevanten Generalisierungen kommen wollen, müssen Gleichheit und Verschiedenheit der Ereignisse, mit denen die Wissenschaft es zu tun hat, oft durch Verwendung von Eigenschaften, die nicht direkt beobachtet werden können, bestimmt werden. [...] Psychologische Theorien ... charakterisieren die inneren Zustände von Organismen nur in Hinblick auf die Art und Weise, wie sie zur Erzeugung von Verhalten funktionieren. [...] Das heißt, Theorien werden gebildet unter Verwendung solcher funktional charakterisierten Begriffe wie Erinnerungen, Motive, Bedürfnisse, Triebe, Wünsche, Strategien, Überzeugungen usw., ohne daß dabei auf die physiologischen Strukturen Bezug genommen wird, die diesen Begriffen in irgendeinem Sinn möglicherweise entsprechen.¹⁴⁷

Der von PUTNAM und FODOR propagierte Funktionalismus geht davon aus, „daß der psychologische Typ, dem ein einzelnes mentales Phänomen zugeordnet wird, durch die kausale Rolle bestimmt ist, die diesem Phänomen im mentalen Leben eines Organismus zukommt.“ Anders als bei den Behavioristen hängt dabei die funktionale Rolle eines Zustandes „sowohl von der Beziehung dieses Zustandes zu anderen Zuständen als auch vom Input und Output ab.“ Schmerzen werden beispielsweise nicht nur durch Reize und Reaktionen sondern auch durch den Wunsch, die Schmerzen loszuwerden, charakterisiert. Die von der speziellen physikalischen Realisierung abstrahierende funktionale Beschreibungsebene (so wie „Ventilöffner“ statt „Nockenwelle“) sichert zwar eine eigene, für mentale Phänomene angemessenere Beschreibungsebene, birgt aber immer die Gefahr, daß hier theoretische Entitäten postuliert werden, die keinen empirischen Gehalt haben. Der Verweis auf die zwar bloß logisch beschriebene, aber tatsächlich physikalisch realisierbare universelle Turingmaschine sollte dieser Gefahr vorbeugen: „Die Turingmaschine stellt immer dann, wenn es um die Verarbeitung von Symbolen geht, eine Verbindung zwischen der funktionalistischen und der mechanistischen Erklärung her. Indem man eine psychologische Theorie auf das Programm einer Turingmaschine reduziert, treibt man gleichsam den Ungeist des Homunkulus aus. Es werden keine Operationen gefordert, die nicht durch einen bekannten Mechanismus zustandegebracht werden können.“¹⁴⁸

FODOR entwickelte die von PUTNAM vorgeschlagene funktionale Sichtweise zu einer gegenwärtig einflußreichen und viel diskutierten Theorie des Geistes weiter, indem er – gestützt auf eigene Arbeiten in der Linguistik und auf Entwicklungen in der Informatik (Von-Neumann-Maschinen, höhere Programmiersprachen) – vom Modell der Turingmaschine abrückte und statt dessen die Annahme einer internen „Sprache des Geistes“ – analog zu der bei Computern üblichen zweiten Sprache: „an input/output language in which they communicate with their environment and a machine language in which they talk to themselves (i. e., in which they run their computations)“¹⁴⁹ – zur Beschreibung und Erklärung des Denkens und intelligenten Verhaltens in den Mittelpunkt stellte. FODORS Hauptinteresse richtet sich nicht (wie bei PUTNAM) auf phänomenale oder qualitative mentale Zustände wie z. B. Schmerzen oder eine Rotempfindung, sondern auf *intentionale*, auf sogenannte *propositionale Einstellungen* der Art „Person X glaubt (hofft, befürchtet etc.),

¹⁴⁷FODOR (1985), S. 422, 423, 426 und 427.

¹⁴⁸FODOR (1981), S. 31, 33 und 34.

¹⁴⁹FODOR (1975), S. 65 f. „Roughly, the machine language differs from the input/output language in that its formulae correspond directly to computationally relevant physical states and operations of the machine: The physics of the machine thus guarantees that the sequences of states and operations it runs through in the course of its computations respect the semantic constraints on formulae in its internal language. What takes the place of a truth definition for the machine language is simply the engineering principles which guarantee this correspondence.“ (S. 66)

daß p “ (wobei p für die Repräsentation eines Sachverhaltes, den *Inhalt*, steht).¹⁵⁰ Propositionale Einstellungen sind charakterisiert als eine bestimmte Relation (man kann etwas glauben, wünschen, bezweifeln) zwischen einer Person (Ich, Anna, Bernd) und einem speziellen Inhalt (daß Tante Frieda zu Besuch kommt, daß der HSV den Europapokal gewinnt), wobei die Art der Relation durch ihre funktionale (und, als Äquivalenzklasse irgendwelcher physikalischer Realisierungen, auch kausale) Rolle (zu Reizen, zu anderen mentalen Zuständen und zu Verhaltensreaktionen) bestimmt wird und der Inhalt symbolisch (als „formulae in the internal code“) im Geist repräsentiert wird.¹⁵¹

Darüber hinaus – und das ist der entscheidende Schritt von einer *repräsentationalen* zu einer *komputationalen*¹⁵² Theorie des Geistes – nimmt FODOR an, daß Kognitions- und Denkprozesse nur Zugang zu den *formalen* (syntaktischen), nicht jedoch zu den inhaltlichen (semantischen) Eigenschaften der in der „Sprache des Geistes“ geschriebenen Repräsentationen haben und daß solche Prozesse nichts anderes sind als das regelgeleitete, formale Operieren mit den Symbolen. Genau so, wie die Verarbeitungsprozesse eines Computers durch ausschließliches Betrachten der in der Maschinen- oder Programmiersprache formulierten Daten und Algorithmen verstanden werden können, ohne auf die spezielle Hardware (Vakuumröhren, Siliziumchips oder was auch immer) zu achten, genau so, wie die syntaktischen Operationen der Programme gemäß dem FORMALISTENMOTTO (vgl. oben S. 32) die von der Benutzergemeinde intendierten semantischen Inhalte gleichsam mittransportieren, ohne etwas von ihnen zu „wissen“, genau so lassen sich auch psychologische Prozesse beschreiben und verstehen, ohne auf die physiologischen Prozesse des Gehirns und auf die semantischen Eigenschaften der Repräsentationen Bezug zu nehmen. Geist verhält sich zum Gehirn wie Software zur Hardware, *Denken ist* (wie ein Digitalcomputer) *Rechnen*, nicht nur als Metapher, sondern im Wortsinne: „Insofar as we think of mental processes as computational [...] it will be natural to take the mind to be, inter alia, a kind of computer.“¹⁵³

Die Symbole der internen „Sprache des Geistes“ führen also ein Doppelleben (als dritte Ebene kommt die physische hinzu): sie haben Bedeutung, insofern sie Repräsentationen von etwas sind (und die Bedeutung komplexer Formeln ergibt sich, wie beim logischen Atomismus, kontextfrei ausschließlich aus der Komposition der elementaren Symbole),

¹⁵⁰Sätze dieser Art haben anders als beschreibende Sätze der physischen Welt einige für Logiker unbefriedigende Eigenschaften. Der Wahrheitswert so zusammengesetzter Sätze ist nicht ausschließlich eine Funktion der Teilsätze („Er glaubt, daß Iokaste aus Theben stammt“ kann wahr oder falsch sein, unabhängig davon, ob sie aus Theben stammt oder ob es sie überhaupt gab) und er kann sich ändern, wenn ein Ausdruck durch einen anderen mit demselben Bezug ersetzt wird („... , daß seine Mutter aus Theben stammt“). Für eine detaillierte Untersuchung vergleiche Rainer HEGSELMANN (1980). Diese „Opakheit“ propositionaler Einstellungen stellt auch für FODOR (1980, S. 66) ein Problem dar.

¹⁵¹A. a. O., S. 75. Durch die Repräsentationen in der internen Sprache des Geistes unterscheiden sich kognitive Prozesse von anderen in der Natur vorkommenden (wie z. B. Planetenbewegungen), selbst wenn sie ebenso Gesetzen folgen. Der Geist ist bei FODOR nicht in gleicher Weise „mechanisch“ wie NEWTONS Physik: „What distinguishes what organisms do from what the planets do is that a *representation of the rules they follow constitutes one of the causal determinants of their behavior*“ (S. 74 Anm. 15). Dafür ist, anders als in der Assoziationstheorie und bei den reinen Empiristen, vor jeder Erfahrungsverarbeitung ein gewisser „Apparat“, eine „funktionale Architektur“ nötig: „The moral is: give up the idea of dynamic relations among psychological elements in favor of the computational picture and you thereby give up a lot of what distinguishes Hume’s picture of the mind from, say, Kant’s.“ (FODOR (1983), S. 31 f.)

¹⁵²Die Übersetzung von *computational* ist in der deutschen Literatur nicht einheitlich, so wählt z. B. Joachim SCHULTE bei der Übersetzung von PUTNAM (1991) „kalkülmäßig“ (vgl. auch die Anmerkung 23 bei KRÄMER (1991) und die Anmerkung des Übersetzers von SEARLE (1993), S. 275).

¹⁵³FODOR (1980), S. 65.

aber beim Ablauf kognitiver Prozesse wird die Bedeutung völlig ignoriert und nur die syntaktische Form spielt eine kausale Rolle (daraus folgt, daß sich – bei gegebener Person und Relation – zwei Gedanken nur unterscheiden können, wenn sich die syntaktische Form der Repräsentationen dieser Gedanken unterscheidet, umgekehrt gilt das nicht). Nur so wird der „Ungeist des Homunkulus“ ausgetrieben (und jede Textverarbeitung auf einem PC zeigt, wie es geht), und nur so wird dennoch erlaubt, das Verhalten auch aufgrund des Inhalts einer propositionalen Einstellung zu erklären.¹⁵⁴

Der Funktionalismus in seiner Verbindung von mentalen Repräsentationen in der Tradition der neuzeitlichen Bewußtseinsphilosophie von DESCARTES über LOCKE bis KANT und der idealsprachlichen Repräsentation mit bloß formaler Symbolmanipulation als *calculus ratiocinator* zu der *empirischen* These (und nicht bloß als *methodologische* Forderung für eine intersubjektiv besser gesicherte Erkenntnis und Argumentation wie bei LEIBNIZ über FREGE bis CARNAP), daß Denken Komputieren ist (bzw. daß diese Annahme empirisch am fruchtbarsten ist), ist die Philosophie des Geistes des Computer-Zeitalters, „in der amerikanischen Philosophie . . . die Mehrheitsposition“ und wohl „die gegenwärtig dominierende Interpretation mentalistischer Begrifflichkeit“.¹⁵⁵ Seine Attraktivität hat er erlangt

- in der Philosophie durch die Hoffnung auf die Lösung einiger traditionellen Rätsel der Bewußtseinsphilosophie (Leib-Seele-Problem, Intentionalität, Willensfreiheit), durch die Vereinbarkeit der Annahme einer kausal abgeschlossenen physikalischen Welt mit der einer eigenen mentalen Wirkungsebene, ohne Reduktionismus und ohne Homunkulus;
- in der mit dem Behaviorismus sehr verarmten Psychologie (und Linguistik) durch die Rehabilitierung eines genuinen, gegenüber dem einheitswissenschaftlichen Physikalismus eigenständigen kognitiven Forschungsansatzes (als Informations- bzw. Symbolverarbeitung), der dennoch den natur- (und somit nach diesem Selbstverständnis überhaupt erst) wissenschaftlichen methodischen Ansprüchen genügt;
- in der Computerwissenschaft durch die Vision der Möglichkeit, *wirklich* denkende und intelligente (und nicht jeweils bestimmte Prozesse bloß simulierende bzw. die Ergebnisse dieser Prozesse bloß leistungsanalog erbringende) Computersysteme zu entwickeln;
- und vor allem in einer (mindestens) die genannten Disziplinen vereinenden *Cognitive Science* durch die Erwartung einer nicht anthropozentrischen, einheitlichen Theorie „intelligenter Systeme“, zu denen Menschen, Tiere, Computer und Besucher von

¹⁵⁴„Erlaubt“ heißt nicht unbedingt auch „tatsächlich möglich“, denn dazu müßte genau angegeben werden, in welchem Verhältnis semantische Eigenschaften zu syntaktischen stehen (vgl. FODOR (1987); aber „of the semanticity of mental representations we have, as things now stand, no adequate account“ (Fodor (1990), S. 28). Denn: „Formal operations are the ones that are specified without reference to such semantic properties of representations as, for example, truth, reference and meaning. Since we don’t know how to complete this list (since, that is, we don’t know what semantic properties there are), I see no responsible way of saying what, in general, formality amounts to“ (FODOR (1980). S. 64). Andreas KEMMERLING (1988, S. 39) kommentiert diese Sätze so: „Das ist dankenswert klar gesagt. ‚S spricht Deutsch‘, das heißt für unseren Kognitivisten gegenwärtig also erst einmal nur: ‚S steht dank uns unbekannter Prozesse in seinem Hirn in einer uns unklaren Beziehung zu uns unbekanntem Entitäten‘.“

¹⁵⁵HASTEDT (1988), S. 143, CARRIER & MITTELSTRASS (1989), S. 61. Und FODOR ist – so DENNETT nach PUTNAM (1985b), S. 143 – „the Pope of the High Computationalist Church“.

außerhalb der Erde gleichermaßen gehören können, und durch die Erfahrung einer fruchtbaren Zusammenarbeit, in der Computerprogramme empirische Tests und Gedankenexperimente für philosophische, psychologische und linguistische Theorien seien können und die Bemühungen der KI Antworten auf, zumindest aber ein tieferes Verständnis der Fragen dieser Disziplinen ermöglichen.

Die gegenwärtige Vorherrschaft des Funktionalismus darf nicht darüber hinwegtäuschen, daß mit ihm keineswegs alle oder auch nur die meisten Probleme des Geistes und des Denkens gelöst sind. FODOR selbst sieht das ganz deutlich, wenn er die komputationale Theorie des Geistes in erster Linie bloß als Forschungsstrategie für die (kognitive) Psychologie vorschlägt. Wenn kognitive Leistungen als Ergebnisse bloß formaler Prozesse über mentale Repräsentationen angesehen werden, haben letztlich nur die Operationsregeln und die formale Gestalt der Symbole in der „Sprache des Geistes“ des einzelnen Individuums Erklärungswert im Rahmen dieser Theorie; und das gilt z. B. auch für die Wahrnehmung der Außenwelt: nicht wie die Welt wirklich beschaffen ist, sondern welche Symbole auf dem Weg durch die Sinnesorgane und in der Interaktion des Nervensystems in der internen Sprache am Ende erzeugt werden, zählt für diese Theorie. FODOR nimmt hier – anknüpfend an DESCARTES – bewußt den Standpunkt des „methodologischen Solipsismus“¹⁵⁶ ein, denn für die Erklärung des Verhaltens einer Person im Rahmen der Psychologie sei nur wichtig, was diese „im Kopf“ habe, ein Vergleich mit der Realität und damit semantische Kategorien wie Wahrheit, Sinn, Bedeutung (Bezug) spielen dabei keine Rolle (Ödipus wollte Iokaste heiraten, nicht seine Mutter). Konsequenterweise sieht er auch nicht die Möglichkeit einer Psychologie des *Wissens* (nur eine des *Glaubens* und genaugenommen auch nicht eine der Wahrnehmung, denn „Sehen“ ist ja – wie RYLE schon bemerkte (vgl. oben S. 29) – ein „Erfolgszeitwort“).¹⁵⁷

FODOR will mit seinem Plädoyer für eine „rationalistische“ Psychologie eine „naturalistische“ – eine, die auch die Einbettung des Individuums in die wirkliche Welt zum Thema hat – nicht ausschließen, ist aber äußerst skeptisch, daß es je möglich sein wird, die dafür nötigen Fakten und Kausalverbindungen zwischen allen möglichen Dingen und Personen in wissenschaftlich befriedigender Weise zu spezifizieren. „My point then, is *of course* not that solipsism is true; it's just that truth, reference and the rest of the semantic notions aren't psychological categories. What they are is: they're modes of *Dasein*. I don't know what *Dasein* is, but I'm sure that there's lots of it around, and I'm sure that you and I and Cincinnati have all got it. What more do you want?“¹⁵⁸

Nicht alle sind bei der Einschätzung der Reichweite der funktionalistischen Theorie des Geistes so vorsichtig wie FODOR. Auch gegen seinen Ansatz, erst recht aber gegen den in

¹⁵⁶FODOR (1980), im Titel und S. 65. Diesen Standpunkt hat schon CARNAP vertreten (und auch diesen Begriff verwendet, vgl. oben die Fußnote 125 auf S. 47), FODOR bezieht sich hier aber auf PUTNAM (1979, S. 28), der ihn ablehnt, weil er mit einer zufriedenstellenden Theorie sprachlicher Bedeutung (und damit von Verstehen) nicht zu vereinbaren sei.

¹⁵⁷FODOR (1980), S. 64. Daß der so verstandene und von FODOR propagierte Funktionalismus keine Theorie des *Wissens* liefern kann (und will), wird in der KI häufig von denen übersehen, die einerseits das „Computermodell des Geistes“ übernehmen, andererseits hemmungslos von „wissensbasierten Systemen“ reden (das gilt gelegentlich sogar für Kritiker dieser Sicht, vgl. Barbara BECKER (1992), S. 205 f.; für eine kritische Würdigung des Problems aus der Sicht der Informatik vgl. Herbert STOYAN (1988)). „People who do machine simulation, in particular, very often advertise themselves as working on the question how thought (or language) is related to the world. My present point is that, whatever else they're doing, they certainly aren't doing *that*.“ (FODOR (1980), S. 65)

¹⁵⁸A. a. O., S. 71.

der KI und in der kognitiven Psychologie häufig vertretenen Anspruch, mit dem Funktionalismus und dem Computermodell eine umfassende und befriedigende Theorie des Geistes und des Denkens zu entwickeln, sind eine Reihe gewichtiger Einwände erhoben worden:

- Der Funktionalismus ist eine zu „liberale“ Theorie des Geistes, es können mit ihm auch Systemen mentale Eigenschaften zugesprochen werden, bei denen das offensichtlich zu Unrecht geschieht. Wenn von der speziellen physikalischen Struktur abgesehen wird und nur die funktionale Organisation und *irgendwelche* Inputs und Outputs in Betracht kommen, können auch eine Milliarde Chinesen, die das Verhalten je eines Neurons simulieren und deren Zusammenspiel über Satellitenkommunikation realisiert wird, oder das von einem Tycoon entsprechend manipulierte Finanzsystem eines kleinen Landes als „mentale Systeme“ bezeichnet werden.¹⁵⁹
- Andererseits ist der Funktionalismus eine zu enge, zu reduzierte Theorie des Geistes. Die Konzentration auf propositionale Einstellungen und kognitive Prozesse vernachlässigt andere Bereiche des Mentalen wie phänomenale Bewußtseinszustände („Qualia“), implizites Wissen und Können (*knowing how* statt *knowing that*). Das gilt insbesondere für solche Bereiche, die in enger Verbindung zum Körper stehen wie Empfindungen, Gefühle, kinästhetische Erfahrungen.
- Es bestehen durchaus berechtigte Zweifel daran, daß sich Denken und andere mentale Operationen allein im Medium einer satzartigen Repräsentation in einer „Sprache des Geistes“, durch Komposition und serielle Transformation digitaler Symbole erklären lassen. Zumindest für bildhaftes, assoziatives Denken, für Wahrnehmungen und Mustererkennung sind möglicherweise analoge und verteilte Repräsentationen sowie parallele Prozesse angemessener.
- Der methodische Solipsismus wird der zunehmend allgemein akzeptierten Einsicht nicht gerecht, daß Gebrauch und Verstehen von Sprache (und somit auch wichtige Aspekte des Denkens) ein soziales Gut, Teil der Lebenspraxis in einer Sprachgemeinschaft sind.
- Schließlich hilft die möglicherweise tatsächlich einmal aufweisbare funktionale Äquivalenz zu einer Turingmaschine oder zu einem Computerprogramm überhaupt nicht zur Erklärung der unmittelbaren Erfahrung des Bewußtseins und der Subjektivität, dazu, wie sich der Geist „von innen anfühlt“, wie es „ist“, ein bestimmtes, bewußtes Wesen zu sein.¹⁶⁰

¹⁵⁹Die Beispiele stammen von Ned BLOCK (1992), S. 175 f. und 208 f. Versucht man hingegen, die In- und Outputs genauer zu spezifizieren, landet man wieder beim alten Dilemma von „Liberalismus“ und „Chauvinismus“ (wo genau endet z. B. bei einem Nervensystem die bloße Weiterleitung, wo genau beginnt die *Verarbeitung* eines Reizes?).

¹⁶⁰Dieser Einwand bedeutet nicht unbedingt eine Renaissance der vor-analytischen Bewußtseinsphilosophie. Nachdem „das Thema ‚Erleben‘ in der neuen Ära der Philosophie des Geistes auf merkwürdige Weise ausgespart oder unterdrückt worden war“ – in einer „Art Berührungsangst“, im Glauben, „daß die Rede von der ‚phänomenologischen Innenwelt‘ einen Rückfall hinter Ryle und hinter den wissenschaftlichen Realismus bedeuten würde“ (Peter BIERI (1993), S. 207) –, ist nun, vor allem als Folge des Aufsatzes von Thomas NAGEL (1981), dieser Bann gebrochen (vgl. z.B. Manfred FRANK (1991), Sybille KRÄMER (1996) und Thomas METZINGER (1993, 1994, 1996) dafür, daß Subjektivität auch in der analytischen und naturalistischen Philosophie wieder ein respektables Thema ist).

Ray JACKENDOFF (1987) sieht im Funktionalismus lediglich die Skizze einer möglichen Lösung eines *Teils* des klassischen Leib-Seele-Problems und ersetzt dieses Problem durch *zwei* neue: „The upshot is that psy-

Zu den dezidiertesten Kritikern des Funktionalismus als dem Verlangen, der Überzeugungs- und Wunschpsychologie „zur ‚Wissenschaftlichkeit‘ zu verhelfen, indem man sie unmittelbar mit der Computerpsychologie gleichsetzt“, gehört heute interessanterweise sein Begründer Hilary PUTNAM. In einem Buch, dessen „Hauptzielscheibe“ eines seiner „früheren Ichs“ ist, beantwortet er ausführlich die Frage „Warum hat der Funktionalismus nicht funktioniert?“ Seinen damaligen Ausgangspunkt, die Kritik an der Identitätstheorie, hält er dabei aufrecht: „Der Hinweis darauf, daß den verschiedenen physischen Zuständen, in denen man sich befinden kann, während man glaubt, daß eine Katze auf der Matte liegt, nichts physikalisch/chemisch Spezifizierbares ‚gemeinsam‘ zu sein braucht, war eine einsichtsvolle Erkenntnis des Funktionalismus.“ Nun aber kommt er zu dem Ergebnis, „daß auch den verschiedenen *kalkülmäßigen* [computational] Zuständen, in denen man sich befinden könnte, während man glaubt, daß eine Katze auf der Matte liegt, nichts *kalkülmäßig* Spezifizierbares ‚gemeinsam‘ zu sein braucht“, sie also allein auf dieser funktionalen Ebene nicht eindeutig identifiziert werden können. Schlimmer noch, die Behauptung, daß die menschliche Rationalität durch eine „funktionale Ordnung“ charakterisiert werden könne, verliert dadurch jede nachvollziehbare Bedeutung, daß er in Erweiterung bestimmter modelltheoretischer Theoreme von LÖWENHEIM und SKOLEM (vgl. oben S. 50) beweist: „*Jedes Ding besitzt jede funktionale Ordnung.*“ Daraus folgt, daß zur Erklärung von Intentionalität, von Denken und Verstehen mehr nötig ist als Äquivalenzbeziehungen zwischen komputationalen Zuständen, „daß kalkülmäßige Modelle des Gehirn/Geists nicht ausreichend sind für die kognitive Psychologie. Ohne Bezugnahme auf die jeweilige *Umgebung* können wir keine Begriffe und keine Überzeugungen individuieren. Bedeutungen sind nicht ‚im Kopf.‘“¹⁶¹

5 Intermezzo: Großtantes Alien-Fax

Stellen Sie sich vor, ihre Lieblingsgroßtante sei gestorben und habe Sie zum Testamentsvollstrecker bestimmt. Vom Notar erhalten Sie auch einen verschlossenen Umschlag und den Schlüssel zu einem etwas versteckten Zimmer in ihrem Haus. Sie öffnen beides und stellen zu Ihrer Überraschung fest, daß dort eine Art Faxgerät steht, mit dessen Hilfe Ihre Großtante nach ihrem schriftlichen Bekunden mit geheimnisvollen Unbekannten (die sie die *Aliens* nennt) kommuniziert haben will. Ihr letzter Wille ist, daß Sie diese Kommunikation fortsetzen, und eine umfangreiche Sammlung von Zetteln, Heften und Karteikästen soll Ihnen das ermöglichen. Eine erste Inspektion ergibt, daß die Kommunikation offensichtlich in einer Ihnen unbekanntem Sprache mit unbekanntem Zeichen erfolgt. Sie finden eine Bedienungsanleitung für das Gerät, aber keinerlei Hinweise auf die Beschaffenheit und den Aufenthaltsort der *Aliens*. Auch gibt es keine Wörterbücher und Grammatiken oder sonstige Erklärungen und Übersetzungen der geheimnisvollen Sprache (ob Ihre Tante die Sprache für zu schwierig hielt oder ob sie bestimmte Geheimnisse nicht lüften wollte,

chology now has not two domains to worry about, brain and mind, but three: the brain, the computational mind, and the phenomenological mind. Consequently, Descartes’s formulation of the mind-body problem is split into two issues. The ‚phenomenological mind-body problem‘ . . . is, How can a brain have experiences? The ‚computational mind-body problem‘ is, How can a brain accomplish reasoning? In addition, we have the mind-*mind* problem, namely, What is the relationship between computational states and experience?“ (S. 20)

¹⁶¹PUTNAM (1991), S. 32, 15, 138, 18 und 137. Vgl. auch Unterrichtsmodul 2, insbesondere im Anhang den Abschnitt zur formalen Repräsentation der Semantik.

bleibt unklar, sie war ja schon immer sehr liebenswert, aber auch ein wenig eigenartig), alles, was Sie vorfinden sind Regeln der folgenden Art:

Wenn Du die Zeichenreihe $\mp \square \otimes \square \ominus \angle$ empfangst, sende die Zeichenreihe $\diamond \asymp \triangle \boxtimes \sqsubseteq \asymp$, falls aber in der empfangenen Sendung auch $\square \angle \uplus \uplus \square$ vorkommt, sende solange $\odot, \odot \odot, \odot \odot \odot$ usw., bis $\ominus \mp \exists$ kommt.

Sie haben zwar Zweifel daran, ob man *so* mit Menschen (d. h. mit Ihnen) und mit *Aliens* umgehen sollte, aber sie ringen sich dazu durch, den letzten Willen Ihrer Großtante zu erfüllen und das Gerät – nennen wir es mal GAF (Großtantes *Alien-Fax*) – gemäß ihren Anweisungen zu bedienen. Nach einer Weile wird Ihnen das aber lästig: die Sendungen kommen unregelmäßig, Sie müssen häufig präsent sein und – unter uns – diese Art Tätigkeit ist doch intellektuell ein wenig unter Ihrer Würde. Da Sie ohnehin gerade Ihren alten PC gegen einen neuen, besseren, schnelleren, billigeren austauschen wollten, beschließen Sie, das Altgerät hier günstig zu entsorgen, es mit dem GAF zu verbinden und so zu programmieren, daß die Regeln Ihrer Großtante ausgeführt werden (das ist zwar etwas mühsam, aber prinzipiell kein Problem, und man hofft ja immer, daß sich die viele am Computer investierte Zeit irgendwann mal auszahlt). Kurz: die Kommunikation wird automatisiert durch die Installation von CUGAF (Computer unterstütztes Großtantes *Alien-Fax*).

Bei der systematischen Aufarbeitung aller Karteikästen und Hefte stoßen Sie auf einige etwas abseits liegende und irgendwie verloren wirkende Zettel und finden dort eine Regel, die Sie im „Handbetrieb“ bisher nie benutzen mußten:

Wenn Du $\angle \angle \square \square \boxtimes \otimes \otimes$ empfangst, schaue aus dem Fenster; falls das Wetter gut ist, sende $\sqsupseteq \neg$, sonst $\parallel \oplus \frown$.

Diese Regel bietet zwar Anlaß, auch über die Bedeutung der Zeichen nachzudenken, aber Sie haben keine weiteren Kriterien dafür, ob hier etwa „Hast Du denn bei diesem ganzen Ärger wenigstens gutes Wetter?“ oder „Werden denn bei unserem Besuch auch unsere Solarzellen genügend Energie liefern?“ gefragt wird, die Hinweise für das Erlernen von *Alienisch* sind dabei zu gering. Da Sie auch noch einige weitere Regeln dieser Art finden, ist ein anderes Problem dringender: wie lassen sich diese Regeln in das Projekt CUGAF einbinden? Offensichtlich gibt es zwei recht unterschiedliche Strategien. Erstens können Sie sich sagen, daß die Regel anscheinend nur selten verwendet wird und daß wohl eine noch seltenere falsche Antwort nicht so schlimm ist. Sie können sich dann je nach den Klimaverhältnissen am Wohnort Ihrer Großtante immer für eine der Antworten entscheiden oder die Antwort durch einen entsprechend gewichteten Zufallsgenerator auswählen lassen und die Regel entsprechend modifizieren. Damit bleibt es bei CUGAF. Der zweite Weg ist etwas aufwendiger. Sie bauen mit Photozellen und anderen Instrumenten eine kleine mit Ihrem PC verbundene Wetterstation (das geht recht gut, Schülerinnen und Schüler haben das in der letzten Projektwoche bei uns getan), oder Sie sorgen für eine Verbindung mit dem zuständigen meteorologischen Institut, oder Sie lassen das Programm jeweils an dieser Stelle stoppen und ihren Pieper auslösen, so daß Sie die erforderliche Entscheidung selbst treffen können. Wie auch immer, Sie statten damit Ihr System mit einem sensorischen Input, mit einer Sinneswahrnehmung der Außenwelt aus. Dann haben Sie SUCUGAF (Sensor und Computer unterstütztes ...) gebaut. In dem Bewußtsein, daß CUGAF Sie

recht gut bzw. daß SUCUGAF Sie perfekt dabei vertreten kann, den letzten Willen Ihrer Großtante zu erfüllen, gönnen Sie sich erstmal eine Pause und von dem ebenfalls ererbten Geld eine längere Weltreise.

Auf der Reise haben Sie gelegentlich genügend Muße, Distanz und Lust, über Ihr Werk genauer nachzudenken. Was ist beispielsweise, wenn es Ihnen irgendwo so gut gefällt, daß Sie nicht mehr zurückkehren oder wenn Sie auf der Reise versterben? Wird dann – Energieversorgung und keine technischen Defekte vorausgesetzt – Ihre Anlage sich selbst überlassen (Sie haben selbstverständlich dem Wunsch Ihrer Großtante nach absoluter Verschwiegenheit entsprochen) immer und immer weiter den Anschein erwecken, daß Sie den Anschein erwecken, daß Ihre Großtante nach wie vor mit den *Aliens* kommuniziert? Ihnen kommen Zweifel. Was ist, wenn sich die Umgebung Ihrer Anlage und des Hauses verändert, wenn z. B. ein wachsender Baum oder ein Neubau Schatten auf die Photozelle Ihrer Wetterstation wirft und so der ausgeklügelte Test für „gutes Wetter“ ungültig wird (weil nach wie vor hohe Temperaturen und kein Regen gemessen wird und die eingebaute Uhr nicht „Dämmerung“ anzeigt)? *Sie* würden in diesem Falle aus einem anderen Fenster nach dem Wetter gucken (vor dem Bau von SUCUGAF) oder (nach dem Bau) die Wetterstation an anderer Stelle plazieren. Müßte, um sicher zu gehen, nicht auch Ihre Anlage die Fähigkeit haben, sich zu bewegen und auf die Umwelt einzuwirken, zu prüfen und zu korrigieren: müßte sie neben Sensoren nicht auch über Effektoren verfügen, somit ein weitgehend autonomer Roboter, ein ESUCUGAF (Effektor, Sensor und Computer unterstütztes . . .) sein? Reicht aber Ihr Wissen (nicht Ihr Wissen über Programmieren, sondern Ihr Wissen über die Welt), um alle diese Situationen vorausszusehen und ESUCUGAF entsprechend zu bauen? Müßte nicht Ihre Anlage, damit Sie ohne Gewissensbisse Ihre Reise fortsetzen können, über die Fähigkeit verfügen, dieses Wissen aus gegebenen Anlaß *selbst* zu erwerben und das Programm entsprechend zu modifizieren? Hätten Sie nicht, um das Erbe auf einer Weltreise zu genießen und dennoch verantwortungsvoll und ohne schlechtes Gewissen die Wünsche Ihrer Großtante erfüllen zu können (andererseits: ist das noch sicher, wenn das System selbst die Regeln modifizieren kann?), ein LESUCUGAF (Lernendes, Effektor, . . .) bauen müssen? Der Gedanke beunruhigt Sie einigermaßen, denn Ihre Kenntnisse über KI reichen soweit, daß Sie genau wissen, daß nicht nur Sie das nicht können, sondern daß auch sonst niemand weiß, wie, wann und ob überhaupt das möglich sein wird.

Vielleicht war Ihre gesamte Strategie dafür, Respekt, Verantwortung, Arbeitsentlastung und Genuß unter einen Hut zu bringen, einfach falsch. Vielleicht hätten Sie den Wunsch Ihrer Großtante nicht wörtlich erfüllen sollen, sondern dem Sinn und der Absicht nach; statt durch das sture Befolgen der hinterlassenen Regeln das Surrogat einer Kommunikation zu inszenieren, hätten Sie Ihren Arbeitsaufwand dafür verwenden sollen, aus den Regeln, versuchsweisen Antworten und empfangenen Reaktionen *Alienisch* als Fremdsprache wirklich zu erlernen und zu verstehen (geht das überhaupt, wenn man so wenig über deren Welt und Leben weiß?), um dann in eine „echte“ Kommunikation einzutreten. Aber hätten Sie dabei Ihre Großtante vertreten können, wüßten Sie, was sie den *Aliens* zu sagen wünschte und wie? Was wäre dann aus Ihrer Weltreise geworden, hätten Sie auch dafür eine Sie stellvertretende Anlage bauen können? Beim genaueren Nachdenken sehen Sie keine Möglichkeit, dafür mit weniger als den Fähigkeiten vom Typ LESUCUGAF auszukommen; das Problem, menschlich zu sein und zu wirken, und dennoch zwar verantwortungsvolle, aber lästige Arbeit einer Maschine zu übertragen, bleibt für Sie ungelöst.

Andererseits: Ihre Großtante war nicht nur liebenswert und eigenartig, sondern auch hochintelligent und lebensklug (und sie kannte Sie). Vielleicht sind Ihre Prämissen falsch. Sie haben bisher immer angenommen, die geheimnisvollen Zeichenreihen seien Ausdrücke in der Sprache *Alienisch*. Möglicherweise stimmt das gar nicht, vielleicht handelt es sich um ein spezielles Symbolsystem, um eine Art Privatsprache Ihrer Großtante (und sie, deutlich älter als Sie und anders sozialisiert, hat Ihnen nicht den Gefallen getan, eine Ihnen bekannte formale Sprache oder Programmiersprache zu verwenden). Sie haben auch nie geprüft (prüfen können), wohin genau das GAF sendet und was und woher es empfängt (vielleicht sind die *Aliens* ja auch mikroskopisch klein und *im* Gerät). Damit ergeben sich völlig neue Aspekte. Möglicherweise hat Ihre Großtante alle Ihre Probleme (sowohl Ihre hedonistischen als auch Ihre mit den Regeln zweiter Art, also mit sensorischem Input) vorausgesehen. Sie hat deshalb ihr GAF keineswegs direkt mit den *Aliens* verbunden, sondern mit einem ausgeklügelten, an einem Ihnen unbekanntem Ort installierten System von Sende- und Empfangsteilen für das wirkliche *Alienisch* (aber vielleicht läuft die Kommunikation ja auch auf Deutsch) sowie Sensoren und Effektoren. In diesem Falle wäre GAF nur die Kommando- und Steuerzentrale für das ganze System, und Ihrer Großtante ist es gelungen, ihr gesamtes eigenes Wissen zur Bedienung dieses Systems in rein formale, „idiotensichere“ Regeln zu fassen: dann würde CUGAF völlig reichen! Einige Zweifel und offene Fragen bleiben allerdings. Was sollten dann die Zettel mit den Regeln zweiter Art? Nun, sie lagen ja etwas abseits, vielleicht waren es nur Notizen für die Arbeit an den Regeln erster Art, und Ihre Großtante hat vergessen, sie wegzuerwerfen, oder sie ist mit ihrer Arbeit nicht ganz fertig geworden. So oder so wäre allerdings Ihre Großtante besser als alle KI-Forscher zusammen, aber warum eigentlich nicht? Sie hatte ja auch Verbindung mit den *Aliens* und konnte sich Tips holen.

Da kommt Ihnen ein ganz anderer Gedanke. Könnte es nicht sein, daß auf der anderen Seite, bei den *Aliens*, eine ganz ähnliche Geschichte passiert ist? Vielleicht war der ursprüngliche Faxpartner Ihrer Großtante der Großonkel eines *Aliens*, der oder die (welche und wieviele Geschlechter mag es da geben?) nach seinem Tod ein CUGMF installiert hat, und Ihre Großtante hat das gar nicht bemerkt. Könnten dann nicht CUGMF und CUGAF Jahrzehnte und Jahrhunderte miteinander angeregt plaudern, z. B. über Spaziergänge im schönen Sachsenwald, der dann längst nicht mehr existiert? Wäre das überhaupt noch eine echte *Kommunikation*, und, wenn nein, worin genau liegt der Unterschied zur Paarung CUGMF/Großtante bzw. *Alien*-Großonkel/Großtante?¹⁶² Sie fühlen sich allmählich ein wenig überfordert, und zwar mit Recht, denn warum sollten Sie als fiktive Figur meiner Geschichte Fragen beantworten, auf die auch ich keine guten und endgültigen Antworten weiß? Also Schluß mit der Geschichte und nur noch zwei ganz einfache Fragen:

1. Könnte CUGAF (bzw. SUCUGAF) bei einem beliebigen anderen *Alien*, der oder die Ihre Großtante nicht kannte, den Eindruck hervorrufen, mit einem Menschen verbunden zu sein? Mit anderen Worten: Könnte CUGAF (bzw. SUCUGAF) den Taring-Test¹⁶³ bestehen?
2. Versteht CUGAF (bzw. SUCUGAF) *Alienisch*?

¹⁶²Vgl. dazu das Kapitel I von PUTNAM (1982).

¹⁶³Bei diesem nach seinem Erfinder, dem *Alien*-Logiker Taring, benannten Test geht es etwas verkürzt darum, ob über Fax kommunizierende *Aliens* feststellen können, ob sie mit einem denkenden Lebewesen oder mit einem Computer verbunden sind. Der Test ähnelt damit sehr dem seines irdischen Kollegen Alan TURING (vgl. Unterrichtsmodul 1, Abschnitt 3).

6 Verstehen, Intentionalität, Bewußtsein

Die Antworten auf die am Schluß des letzten Abschnitts gestellten Fragen scheinen auf der Hand zu liegen:

1. Ja, das wäre möglich. Es hängt natürlich vom Geschick Ihrer Großtante ab, aber es gibt keinen guten Grund, es von vornherein auszuschließen.
2. Nein, offensichtlich nicht. Denn CUGAF (bzw. SUCUGAF) „weiß“ und „tut“ höchstens das, was Sie anfangs mit GAF taten, und *Sie* verstehen ja auch nicht *Alienisch*.

6.1 Searle: Ursprüngliche Intentionalität

So hat auch John R. SEARLE geantwortet auf die entsprechenden Fragen im Rahmen seines mittlerweile berühmt gewordenen Gedankenexperiments (*Chinese room*): er sitzt in einem Raum, in den durch einen Schlitz Karten mit irgendwelchen für ihn bedeutungslosen Schnörkeln (chinesische Schriftzeichen, aber das weiß er nicht, er versteht kein Chinesisch) hineingereicht werden. Er hat ein Buch zur Verfügung, geschrieben in seiner Muttersprache, mit Regeln der Art: wenn Dir die und die Schnörkel (die er nur an ihrer Form identifizieren kann) hineingereicht werden, reiche die und die Schnörkel heraus. Wenn das Regelbuch gut genug ist, so SEARLE, kann für einen außenstehenden chinesischen Muttersprachler der Eindruck einer sinnvollen Kommunikation auf Chinesisch entstehen, obwohl er kein Wort Chinesisch versteht und nur nach rein formalen Regeln Symbole manipuliert hat.¹⁶⁴

SEARLE hat mit seinem Gedankenexperiment aus der *Innenperspektive* des Turing-Testes den Nerv der KI getroffen (und so wurde auch reagiert), denn die beiden Antworten führen zu dem Schluß, daß eine Maschine den Turing-Test bestehen kann, ohne wirklich zu verstehen und somit auch ohne wirklich zu denken. Er richtet sich nicht gegen die Strömung innerhalb der KI, die er „schwache KI“ nennt (nach ihr „liegt die Bedeutung des Computers für die Erforschung des Geistes vor allem darin, daß er ein außerordentlich leistungsfähiges Werkzeug ist“), sein Angriff zielt auf die These der „starken KI“, daß „der entsprechend programmierte Computer tatsächlich ein Geist in dem Sinne [ist], daß von Computern mit den richtigen Programmen im wörtlichen Sinne gesagt werden kann, daß sie *verstehen* und andere kognitive Zustände haben.“ Denn solange „das Programm mit Hilfe computationaler Operationen über rein formal definierte Elemente definiert ist, legt das Gedankenexperiment nahe, daß dies an sich in keiner interessanten Beziehung zum Verstehen steht. Sie sind sicherlich keine hinreichenden Bedingungen, und es wurde nicht der geringste Grund für die Annahme genannt, daß sie notwendige Bedingungen seien oder sogar einen entscheidenden Beitrag zum Verstehen leisten sollten.“ Denn das „Computerprogramm ist rein syntaktisch definiert. Aber Denken ist mehr als bloß ein

¹⁶⁴SEARLE (1980 (deutsch (1992), auch in HOFSTADTER & DENNETT (1986)), 1986 (S. 30 ff.), 1990). In der deutschsprachigen Literatur wird als Übersetzung meist „Chinesisches Zimmer“ und nicht „Chinesisch-Zimmer“ gewählt. In der ursprünglichen Version ist die Geschichte etwas komplizierter, neben dem Regelbuch gibt es noch zwei weitere Packen mit chinesischen Symbolen („Skript“ und „Geschichte“ genannt), weil SEARLE hier auf *script*-basierte, angeblich „Geschichten verstehende“ Programme anspielt, die Roger C. SCHANK zusammen mit Robert ABELSON entwickelte (vgl. SCHANK & CHILDERS (1986), S. 163 ff.; mehr dazu im Unterrichtsmodul 2).

Manipulieren bedeutungsloser Symbole; zum Denken gehört semantischer Gehalt mit einer Bedeutung. [...] *Syntax reicht nicht für Semantik aus.*¹⁶⁵

Man kann seine Schlußfolgerung bestreiten (und das wird vor allen von KI-Seite heftig getan), indem man SEARLES Gedankenexperiment als zu stark vereinfacht ablehnt und mindestens eine der oben genannten beiden Fragen anders beantwortet: Dadurch, daß der Turing-Test beliebige Gesprächsthemen erlaubt und insofern von einer sehr großen Allgemeinheit ist, ist es so gut wie ausgeschlossen, daß ein in jedem Detail vorprogrammierter Computer, eine *Ready-made-Intelligenz* vom Typ CUGAF, den Test bestehen kann (schon so einfache Fragen wie z.B. „Wie war denn das Wetter heute morgen, als du aufgestanden bist?“ bereiten dann große Probleme; vgl. den Kommentar zum Abschnitt 3 vom Unterrichtsmodul 1). Nötig wäre ein durch Wahrnehmung und Handeln lernender Computer vom Typ LESUCUGAF, und der wäre dann durch eigene kausale Verbindungen zur Umwelt nicht mehr bloß syntaktisch, sondern auch semantisch („Robotereinwand“). Zweitens, auch wenn man bei dem Konzept keiner weiteren In- und Outputs außer Sprache bleibt, wird nicht behauptet, daß der isolierte Symbolmanipulateur Chinesisch versteht (sowenig, wie man das von der CPU behaupten würde), sondern nur zusammen mit den Regeln; und dabei ist die Vorstellung von ein paar Zetteln und einem überschaubaren Regelbuch viel zu simpel. In Wirklichkeit bräuchte man sehr viele Regeln und eine sehr große Zahl von Sprachebenen (auch Sie mußten ja beim CUGAF die Regeln in Ihrer Sprache lesen, sie in eine Programmiersprache übersetzen, die vom Compiler in Prozessorbefehle umgesetzt wurde, die das Verteilungsmuster in den Speicherplätzen änderten ...), Subroutinen und rekursiven Prozeduren, hierarchisch geordnet und dennoch mit der Möglichkeit der wechselseitigen Modifizierung.¹⁶⁶ Kurz, es wäre ein System von sehr hoher Komplexität nötig; wenn das gelingt und es kommen bei der Eingabe vernünftiger Sätze als Antwort wieder vernünftige Sätze, dann könne man zu Recht sagen, das System als Ganzes versteht die Sprache („Systemeinwand“).

SEARLE läßt diese Einwände nicht gelten. Auch bei sehr komplexen Programmen und bei Robotern werden „letztlich“ bloß nach rein formalen, syntaktischen Regeln Symbole manipuliert (vgl. oben auf S. 67 das GAF als „Kommandozentrale“), und daran ändert sich auch bei künftigen Entwicklungen nichts, solange gemäß der Analogie „der Geist verhält sich zum Gehirn wie das Programm zur Hardware“ von der physikalischen Realisierung des Programms abgesehen wird (auch sogenannte neuronale Netze werden meist auf Digitalcomputern simuliert). Für ihn wird Geist und Denken vom Gehirn verursacht, als biologisches Phänomen wie die Photosynthese.¹⁶⁷ Was unterscheidet uns dann aber einerseits von Schachtelhalmen und andererseits von programmierten Computern? Seine Antwort: ursprüngliche (*intrinsic*) Intentionalität.

¹⁶⁵SEARLE (1992), S. 225 und 230, (1986), S. 31 und 38. Er widerspricht damit insbesondere der für die KI lange Jahre grundlegenden „*Hypothese der physikalischen Symbolsysteme*“ (vgl. oben S. 2). Das erinnert an HUME und „führt uns zum Paradoxon des mechanischen Verstandes zurück: entweder spielen Bedeutungen eine Rolle für die Manipulationen, und in diesem Fall sind die Prozesse nicht wirklich mechanisch (sie setzen einen Homunkulus voraus); oder aber die Bedeutungen spielen keine Rolle, und in diesem Fall sind die Prozesse nicht wirklich rational“ (HAUGELAND (1987), S. 37).

¹⁶⁶Das ist ja der große Vorteil der alten Lieblings-KI-Sprache LISP (und begründete große Hoffnungen), daß modulare und rekursive Konzepte sehr gut unterstützt werden und daß es keine Trennung von Daten und Programm gibt (und somit im laufenden Programm auch Programmteile verändert werden können).

¹⁶⁷Damit ist nach SEARLE keineswegs die Möglichkeit ausgeschlossen, denkende Maschinen zu bauen, „da unsere Körper mit unseren Gehirnen in einem wichtigen Sinn genau solche Maschinen sind“, allerdings nur Maschinen, „die *dieselbe kausale Kraft* wie das Gehirn haben“ (1992, S. 243 u. 250).

Intentionalität (Gerichtetsein) – neben Bewußtsein, Subjektivität und geistiger Verursachung eines der Merkmale geistiger Phänomene, die „das Körper/Geist-Problem wirklich schwierig“ machen¹⁶⁸ – ist ein auf BRENTANO zurückgehendes Konzept, Mentales, Psychisches vor dem bloß Physischen auszuzeichnen (und steht damit in Konkurrenz zum Bewußtseinskonzept nach DESCARTES)¹⁶⁹. Mentale Akte und Zustände wie Wahrnehmen, Erinnern, Glauben, Wünschen und Befürchten sind – anders als Planetenbewegungen und natürliche Kristallstrukturen – *auf* etwas (Objekte oder Sachverhalte) gerichtet, handeln *von* etwas: das Denken kann „nach außen greifen und das Äußere erfassen“.¹⁷⁰

Intentionalität und Verstehen stehen im engen Zusammenhang und sind aufeinander bezogen. Verstehen (im Sinne von: im Gehörten und Geschriebenen das Gemeinte erkennen) zielt auf die Bedeutung und die kommunikative Absicht. Andererseits kann es auch der Inhalt eines Wunsches sein, jemandem etwas (z. B. auch die eigenen intentionalen Zustände) zu verstehen geben zu wollen, und dafür ist auch beim anderen Intentionalität nötig (und mein Wissen davon). Zum Verstehen gehört auch Wahrheit und Erfolg, Verstehen und Verstandenwerden setzt ein Mindestmaß an Kooperation, Klarheit und Aufrichtigkeit voraus. Kurz: Intentionalität und Verstehen sind untrennbar miteinander verbunden und mit einem ganzen Bündel von semantischen, pragmatischen und normativen Aspekten („die wirklichen Tatsachen der Intentionalität enthalten normative Bestandteile; bei funktionalen Erklärungen gibt es hingegen keine *Tatsachen* außer den nackten physischen Fakten; und Normen gibt es nur in uns und nur von unserem Standpunkt aus“¹⁷¹).

Die üblichen Mittel für das Zuverstehengeben sind sprachliche und außersprachliche Zeichen und Symbole. Dazu müssen sie *als* Zeichen und Symbole erkannt und richtig interpretiert, gemäß ihrer *Bedeutung* verstanden werden. Der Lage von Steinen am Wegesrand messen wir normalerweise keine Bedeutung zu, ganz anders ist das bei einer Schnitzeljagd. Aber, und genau so ist es auch bei Schallwellen und Pigmentverteilungen auf dem Papier, die Steine *haben* dann nicht *von sich aus* eine Bedeutung, *wir* haben sie ihnen *verliehen*; und das ist bei unseren intentionalen Zuständen ganz anders: „Weil Sätze – die Geräusche, die aus dem Mund kommen, oder die Schriftzeichen, die auf dem Papier stehen – in einer Hinsicht bloß Gegenstände in der Welt sind wie alle anderen Gegenstände, ist ihr Repräsentationsvermögen nicht intrinsisch, sondern von der Intentionalität des Geistes hergeleitet. Die Intentionalität von Geisteszuständen hingegen ist von keinen vorgängigen Formen der Intentionalität hergeleitet, vielmehr ist sie eine intrinsische Eigenschaft dieser

¹⁶⁸SEARLE (1986), S. 14.

¹⁶⁹Die beiden Konzepte für „Geist“, „Seele“ oder „Mentales“ sind nicht deckungsgleich. Eine allgemeine Nervosität oder Depression mag bewußt, aber auf nichts gerichtet, ein Wunsch oder eine Einstellung gerichtet, aber nicht bewußt sein (vgl. RORTY (1981), S. 33–50). Andere Autoren bestreiten hier einen relevanten Unterschied. „If we strip away the jargon, I think Searle’s argument can be seen as a displacement of a more fundamental and gut-level claim: that running a computer program does not produce consciousness“ (JACKENDOFF (1987), S. 19).

¹⁷⁰PUTNAM (1982), S. 16. Nicht immer, wenn von „Intentionalität“ die Rede ist, sagen die Autoren auch, „was das Wort eigentlich bedeutet.“ Es ist zu einem Ausdruck geworden, der „nicht für ein einziges bestimmtes Thema steht, sondern für einen ganzen Bereich von Themen und Problemen. Speziell die folgenden Fakten werden gewöhnlich als Beispiele für ‚Intentionalität‘ angeführt: (1) Wörter, Sätze und sonstige ‚Repräsentationen‘ haben *Bedeutung*. (2) Repräsentationen können sich auf ein tatsächlich existierendes Ding ... *beziehen* (d. h. auf sie zutreffen). (3) Repräsentationen können von etwas, was *nicht* existiert, handeln. (4) Einem Bewußtseinszustand kann ein ‚Sachverhalt‘ entsprechen. [...] Als die Computerrevolution über die Welt hereinbrach, wurde vielfach damit gerechnet (und es wurde behauptet), Computermodelle würden das Wesen dieser verschiedenen Phänomene erklären“ (PUTNAM (1991), S. 21).

¹⁷¹SEARLE (1993), S. 262.

Zustände selbst. Jemand benutzt einen Satz, um eine Feststellung zu machen oder eine Frage zu stellen; in dieser Weise *benutzt* er seine Überzeugungen und Wünsche allerdings nicht: er hat sie einfach.“¹⁷²

Diesen Unterschied zwischen ursprünglicher und verliehener Intentionalität zu ignorieren, ist nach SEARLE der Hauptfehler des Funktionalismus¹⁷³ und der „starken KI“. „Bei den Gehirnfunktionen sind nicht die formalen Schatten relevant, die durch die Synapsenfolge geworfen werden, sondern die wirklichen Eigenschaften der Folgen. All die Argumente für die starke Version der KI, die ich gehört habe, bestehen darauf, die Umrisse der Schatten nachzuzeichnen, die von den Kognitionen geworfen werden, und danach zu behaupten, diese Schatten seien das wirkliche Ding. [...] Intentionalität, wie sie Computer zu haben scheinen, gibt es nur im Geiste derer, die sie programmieren, und derer, die sie gebrauchen, derer, die ihnen die Inputs geben und die die Outputs interpretieren.“¹⁷⁴

Inzwischen geht SEARLE noch einen Schritt weiter. Er bestreitet nicht nur, daß Computer im Wortsinne Intentionalität haben und verstehen können (These der „starken KI“), er argumentiert auch dafür, daß die im Rahmen der Kognitionswissenschaften gestellte Frage „Ist das Hirn ein digitaler Computer?“ schlecht definiert ist und keinen klaren Sinn hat:

Legt man die Standard-Lehrbuchdefinition von Computation zugrunde, so fällt es schwer, zu sehen, wie sich die folgenden Resultate vermeiden lassen:

1. Für jeden Gegenstand gibt es irgendeine Beschreibung der Art, daß der Gegenstand unter dieser Beschreibung ein digitaler Computer ist.
2. Für jedes Programm und jeden hinreichend komplexen Gegenstand gibt es irgendeine Beschreibung, unter welcher der Gegenstand das Programm implementiert. So implementiert jetzt in diesem Moment zum Beispiel die Wand hinter meinem Rücken das „Wordstar“-Programm, weil es innerhalb der besagten Wand irgendein Muster von Molekülbewegungen gibt, das mit der formalen Struktur des „Wordstar“-Programms isomorph ist. Wenn aber die Wand diese Programm implementiert, dann implementiert sie – vorausgesetzt, sie ist nicht zu klein – jedes beliebige Programm (auch jedes im Hirn implementierte Programm).

[...] *Die Vielfach-Realisierbarkeit computational äquivalenter Vorgänge in unterschiedlichen physischen Medien ist nicht bloß ein Zeichen dafür, daß diese Vorgänge abstrakt sind, sondern auch dafür, daß sie gar nicht dem System an sich angehören: sie sind dem System nicht intrinsisch; sie hängen von einer Interpretation ab, die von außen an sie herangetragen wird.* [...] Computationale Zustände werden nicht *in* der physischen Beschaffenheit entdeckt, sie werden ihr *zugeordnet*. Das ist eine andere Argumentation als die mit dem Chinesisch-Zimmer, und ich hätte das schon vor zehn Jahren sehen sollen, ich habe es aber nicht. Die Argumentation mit dem Chinesisch-Zimmer hat gezeigt, daß die Semantik der Syntax nicht intrinsisch ist. Hier weise ich nun darauf hin – und das ist etwas Anderes und Separates –, daß die Syntax der Physik nicht intrinsisch

¹⁷²SEARLE (1987), S. 9f.

¹⁷³„Functionalism, by the way, is an entire system erected on the failure to see this distinction“ (SEARLE (1980), S. 452). „Wer sich zum Funktionalismus hingezogen fühlt, braucht keine Widerlegung, er braucht Hilfe“ (SEARLE (1993), S. 24).

¹⁷⁴SEARLE (1992), S. 244 und 245. Zu dieser Anleihe bei PLATO vgl. auch WITTGENSTEINS spätere Distanzierung von seiner eigenen Position im *Tractatus*: „Irreführende Parallele: Der Schrei, ein Ausdruck des Schmerzes – der Satz, ein Ausdruck des Gedankens! [...] Denken ist kein unkörperlicher Vorgang, der dem Reden Leben und Sinn leiht, und den man vom Reden ablösen könnte, gleichsam wie der Böse den Schatten Schlemiehls vom Boden abnimmt.“ (*Philosophische Untersuchungen*, §§ 317 und 339)

ist. Zum Zwecke der ursprünglichen Argumentation hatte ich einfach angenommen, die syntaktische Charakterisierung des Computers sei unproblematisch. Doch das ist ein Fehler. Es gibt keine Möglichkeit zu entdecken, daß ein System intrinsischermaßen ein digitaler Computer ist, weil die Charakterisierung des Systems als einem digitalen Computer immer relativ zu einem Beobachter ist, der den rein physischen Merkmalen des Systems eine syntaktische Interpretation zuordnet. Wendet man dies auf die Hypothese von der Sprache des Geistes an, dann hat es zur Konsequenz, daß diese Hypothese inkohärent ist. Es gibt keine Möglichkeit zu entdecken, daß unbekannte Sätze intrinsischermaßen im Kopf sind, weil etwas ein Satz nur relativ dazu ist, daß irgendeine Person oder irgendein Benutzer existiert, die oder der es als Satz verwendet.¹⁷⁵

Als Gründe für diese Fehler – unzureichende Unterscheidung zwischen intrinsischer und abgeleiteter Intentionalität (und damit Semantik), Übersehen der Beobachterrelativität von Syntax – sieht er vor allem eine irreführende Verwendung des Begriffes der „Informationsverarbeitung“¹⁷⁶ und trotz gegenteiliger Beteuerungen einen Rest von Behaviorismus sowie eine Form des Dualismus, der cartesianisch ist in dem Sinne, „daß das, was das spezifisch Mentale des Geistes ist, keine innere Verbindung mit den wirklichen Eigenschaften des Gehirns hat.“ Wie die Beispiele KOPERNIKUS, GALILEI und DARWIN zeigten, „wurden wissenschaftliche Errungenschaften im Westen, die zeigten, daß die Menschheit lediglich ein Teil der gewöhnlichen physikalischen und biologischen Ordnung ist, oft von der Nachhut einer veralteten Weltanschauung bekämpft. [...] Am besten versteht man die starke KI als eine der letzten Ausprägungen dieser antiwissenschaftlichen Tradition, da sie leugnet, daß dem menschlichen Geist irgend etwas essentiell Physikalisches und Biologisches anhaftet“.¹⁷⁷

Niemand wird annehmen, daß die Computersimulation eines Feueralarms die ganze Nachbarschaft niederbrennen oder daß die Computersimulation eines Regenschauers uns völlig durchnäßt zurücklassen wird. Warum sollte man dann aber annehmen, daß eine Computersimulation des Verstehens wirklich etwas versteht? [...] Was Intentionalität auch immer sein mag, sie ist ein biologisches Phänomen, und es ist wahrscheinlich, daß sie von der spezifischen Biochemie ihres Ursprungsstoffes genauso kausal abhängig ist, wie die Milcherzeugung, die Photosynthese oder irgendwelche anderen biologischen Phänomene. Niemand wird annehmen, daß man Milch oder Zucker erzeugen könnte, indem man eine Computersimulation der formalen Abläufe durchführt, die sich bei der Milcherzeugung und der Photosynthese finden. Wenn es aber um den Geist geht, dann wollen viele Leute, motiviert durch ihren tief- und festsitzenden Dualismus, an ein solches Mysterium glauben: sie nehmen an, der Geist sei eine Sache formaler Prozesse und er sei, anders als Milch und Zucker, unabhängig von bestimmten stofflichen Ursachen.¹⁷⁸

¹⁷⁵SEARLE (1993), S. 230 ff. Vgl. dazu oben S. 64 PUTNAMs schon vorher erhobene Kritik am Funktionalismus und seinen Beweis des Theorems „Jedes gewöhnliche offene System ist eine Realisierung jedes abstrakten finiten Automaten“ (PUTNAM (1991), S. 213 ff.).

¹⁷⁶Vgl. SEARLE (1986), S. 47 ff., (1993), S. 245 ff.

¹⁷⁷SEARLE (1992), S. 250, (1990), S. 47. Selbstverständlich fehlt es dann auch nicht an Polemiken von Seiten der so Angegriffenen (und selbsterklärten Antidualisten): DENNETT spricht von „his updated version of *élan vital*“, HOFSTADTER bezeichnet seinen Begriff der Intentionalität als „his name for soul“ und sieht seinen Artikel (1980) als „religious diatribe against AI, ... one of the wrongest, most infuriating articles I have ever read in my life“, RORTY fühlt sich an die Eucharistie erinnert und schreibt, SEARLE argumentiere wie „a fundamentalist Catholic defending transubstantiation“, (SEARLE (1980), S. 430, 433, 445).

¹⁷⁸SEARLE (1992), S. 247 und 250 f.

Allerdings hat SEARLE hier ein ernstes Problem: *wie* der „Ursprungsstoff“ das macht, weiß er selbstverständlich ebensowenig zu sagen wie sonst jemand: „Wenn es uns gelingt, die Rolle des Hirns bei der Hervorbringung von Intentionalität zu verstehen, so wird uns dies wahrscheinlich mit Hilfe von Prinzipien gelingen, die sich vollkommen von den zur Zeit benutzten unterscheiden; der Unterschied wird so groß sein, wie der zwischen den Prinzipien der Quantenmechanik und denen der Newtonschen Mechanik“.¹⁷⁹

SEARLE plädiert ungeachtet dieser Schwierigkeiten offensiv für einen „biologischen Naturalismus“: „Geistige Ereignisse und Vorgänge gehören genauso zu unserer biologischen Naturgeschichte wie Verdauung, Mitose, Meiose oder Enzymsekretion.“ Er wendet sich damit nicht nur gegen die „starke KI“, sondern auch gegen den „Kognitivismus“, gegen eine am „Computermodell des Geistes“ orientierte Kognitionswissenschaft, Linguistik und Psychologie mit einer eigenständigen, nicht reduzierbaren Untersuchungsebene zwischen der Neurophysiologie und unserem bewußten Erleben. Für eine solche Ebene, auf der „irgendwelche Rechnungen unbewußt angestellt werden“, gibt es für ihn weder einen guten Grund noch einen empirisch überprüfaren Sinn; es braucht keinen „Lückenfüller“, denn „es gibt da gar keine Lücke, die zu füllen wäre“: „Das Gehirn macht es einfach so. Wir sind neurophysiologisch so gebaut, daß der Ansturm von Photonen auf unsere Photorezeptoren uns sehen läßt; und wir sind neurophysiologisch so gebaut, daß, wenn wir Leute reden hören und mit ihnen interagieren, wir dadurch die Sprache lernen.“ „Es gibt einfach nur einen Hardware-Mechanismus, der so-und-so funktioniert, aber nicht anders.“¹⁸⁰

In diesem Zusammenhang widerspricht SEARLE auch dem in den Kognitionswissenschaften vorherrschenden „Bemühen um eine Abtrennung des Bewußtseins von der Intentionalität.“ Das Motiv für diese Abtrennung sieht er darin, „daß wir nicht wissen, wie das Bewußtsein zu erklären ist, und daß wir gerne eine Theorie des Geistes hätten, die nicht dadurch in Mißkredit gebracht wird, daß ihr eine Theorie des Bewußtseins fehlt. Die Idee ist, Intentionalität ‚objektiv‘ zu behandeln, sie so zu behandeln, als seien die subjektiven Bewußtseinsmerkmale eigentlich nicht wichtig für Intentionalität.“ Gegen die von den am „Computermodell des Geistes“ orientierten Strömungen in der Philosophie, Linguistik und Psychologie postulierten – dem Bewußtsein prinzipiell nicht zugänglichen –

¹⁷⁹SEARLE (1987), S. 337. Hier trifft er sich mit Roger PENROSE: „Mir scheint, daß weder die klassische Physik noch die Quantenmechanik ... jemals zu erklären vermag, wie wir *denken*“ (PENROSE (1991), S. 394). Die Arbeit des Gehirns, im Niemandsland zwischen klassischer und Quantenmechanik, würde sich nur durch eine bisher nicht vorhandene Theorie, eine von ihm „korrekte Quantengravitation“ genannte Synthese aus Relativitäts- und Quantentheorie, erklären lassen. Ich denke, auch dann wird es einem so gehen wie in LEIBNIZ' Mühle.

¹⁸⁰SEARLE (1993), S. 15, 269, (1986), S. 51 f. Eine unerwartete Konsequenz ist, daß SEARLE „ganz unbeabsichtigt zu einer Verteidigung ... des Konnektionismus gelangt“ ist, wobei die konnektionistischen Modelle vielleicht allesamt falsch sind, aber „nicht in der Weise offenkundig falsch bzw. inkohärent ... wie dies bei den traditionellen kognitivistischen Modellen der Fall ist“ (1993, S. 270). Angesichts des Standes der Hirnforschung, angesichts des schon angedeuteten Fehlens jeder Theorie, die die nötigen kausalen Kräfte des Hirns bei der Hervorbringung von „Geist“ erklären könnte (hier helfen auch die suggestiven Analogien nicht weiter: Intentionalität und Bewußtsein sind eben keine stofflichen Substanzen wie Milch und Enzyme) und auch erklären könnte, warum andere Nervensysteme darüber *nicht* verfügen, und angesichts der Tatsache, daß auch konnektionistische Modelle solche Erklärungen nicht liefern können, ist diese Position recht mutig und ziemlich unbefriedigend: „Das Gehirn hat Searle zufolge also intrinsisch intentionale Zustände, ohne syntaktische oder semantische Zustände zu haben, denn diese seien ja immer nur attribuiert. Außer Searle gibt es allerdings niemanden, der diese Konstruktion plausibel findet“ (KEIL (1993a), S. 346). Vertreter einer „Sprache des Geistes“ fragen bei solchen Verweisen auf „Unknown Neurological Mechanisms“ dann auch an, „whether it's not sort of unreasonable to prefer no psychology of thought to a computational theory of thought“ (FODOR (1987), S. 147).

informationsverarbeitenden, syntaktischen „Rechenprozesse“ setzt er die These: „Nur ein Lebewesen, das bewußte intentionale Zustände haben könnte, könnte überhaupt intentionale Zustände haben, und jeder unbewußte intentionale Zustand ist zumindest potentiell bewußt.“ Bewußtsein (und Subjektivität) ist damit für ihn das zentrale geistige Phänomen: „Die Vernachlässigung des Bewußtseins ist mehr als alles andere schuld an der Dürre und Sterilität in der Psychologie, in der Philosophie des Geistes und in der Kognitionswissenschaft. Die Erforschung des Geistes ist die Erforschung des Bewußtseins, und zwar in ziemlich demselben Sinn, in dem die Biologie die Erforschung des Lebens ist.“¹⁸¹

6.2 Dennett: Intentionale Systeme

Gegen die Ichperspektive SEARLES¹⁸² plädiert Daniel C. DENNETT für eine reine Beobachterperspektive (und bekennt sich damit in der Tat mehr oder weniger offen zu einem „Rest von Behaviorismus“ in der Tradition von RYLE). Sein Grundbegriff ist das *intentionale System*, ein System „dessen Verhalten ... erklärt und vorausgesagt werden kann, indem man ihm Meinungen und Wünsche (und Hoffnungen, Befürchtungen, Absichten, Ahnungen, ...) zuschreibt.“ Bei solchen Systemen sagt man ihr „Verhalten voraus, indem man dem System den *Besitz gewisser Informationen* zuschreibt, von ihm annimmt, daß es von *gewissen Zielen* geleitet wird, und sich dann auf der Grundlage dieser Zuschreibungen und Annahmen die vernünftigste und angemessenste Handlung überlegt.“ Seine Definition intentionaler Systeme sagt wohlgerne nicht, „daß intentionale Systeme *wirklich* Meinungen und Wünsche haben,“ sondern er will nur behaupten, „daß ein rein physikalisches System gelegentlich so komplex und trotzdem so organisiert sein kann, daß wir es bequem, für Erklärungen ergiebig und für Voraussagen nützlich finden, es so zu behandeln, als hätte es Meinungen und Wünsche und wäre rational.“¹⁸³

Diese pragmatisch begründete *intentionale Einstellung* unterscheidet er von der bei der Voraussage über das Verhalten mechanischer Gegenstände üblichen und häufig auch bei Naturgegenständen verwendeten *funktionalen Einstellung* (die bedeutet, „daß wir unsere Voraussagen nur auf der Basis von Wissen oder Annahmen über die funktionale Konstruktion des Systems, unabhängig von der physikalischen Beschaffenheit oder dem inneren Zustand des einzelnen Gegenstandes, machen“; Beispiele: „Sowie der Wagen der Schreibmaschine sich dem Rand nähert, wird ein Klingelzeichen ertönen“, „Ein starker Beschnitt wird ein dichteres Blattwerk und stärkere Äste stimulieren“) und diese wiederum von der *physikalischen Einstellung* (bei der sich unsere Voraussagen „auf den tatsächlichen physikalischen Zustand des einzelnen Gegenstandes“ stützen). Ebenso, wie wir bei Funktionsstörungen gezwungen sind, die funktionale Einstellung zu verlassen und Voraussagen aus der physikalischen Einstellung vorzunehmen („Wenn du den Schalter einstellst, wirst du einen gefährlichen Schlag erhalten“, „Wenn die Schneefälle kommen, wird dieser Zweig abbrechen“), kann es nach Einnahme der intentionalen Einstellung auch passieren, „daß wir die Annahme der Rationalität fallenlassen“ müssen und wir „schließlich bei Voraussagen in der funktionalen Einstellung“ enden. Aber in vielen Fällen kann „sich die Taktik, die intentionale Einstellung einzunehmen,“ auszahlen, nämlich immer „dann, wenn wir Grund

¹⁸¹SEARLE (1993), S. 173, 152 und 250.

¹⁸²„Remember, in these discussions, always insist on the first person point of view“ (SEARLE (1980), S. 451).

¹⁸³DENNETT (1981), S. 162, 165 f. und 167.

zu der Annahme haben, daß die Voraussetzung einer optimalen Konstruktion gerechtfertigt ist, und bezweifeln, daß Voraussagen in der funktionalen oder der physikalischen Einstellung praktikabel sind.“¹⁸⁴

DENNETT erläutert das am Beispiel eines Schachcomputers:

Man betrachte einen Schachcomputer und die verschiedenen Strategien oder Einstellungen, die man als sein Gegner einschlagen bzw. einnehmen kann, wenn man versucht, seine Gegenzüge vorauszusagen. [...] Den Versuch zu unternehmen, eine physikalische Erklärung oder Voraussage des Schachcomputers zu geben, wäre eine witzlose Herkulesarbeit, doch im Prinzip würde es gehen. [...] Die besten heutigen Schachcomputer sind für Voraussagen in funktionaler oder physikalischer Einstellung praktisch unzugänglich. Sie sind sogar für ihre eigenen Konstrukteure zu komplex geworden, um sie in funktionaler Einstellung zu betrachten. Die beste Aussicht für einen Menschen, eine solche Maschine in einem Schachspiel zu schlagen, besteht darin, ihre Gegenzüge dadurch vorauszusagen, daß er sich, so gut er kann, überlegt, welches unter den gegebenen Regeln und Zielen des Schachspiels der beste und vernünftigste Zug wäre [...], wenn man sie eher wie einen intelligenten menschlichen Gegner behandelt. [...] Entwickelt man eine Voraussage dieser Art, so stellt man folgende Frage: Was ist das Vernünftigste, was der Computer machen kann, wenn die Ziele x, y, z, \dots , die Beschränkungen a, b, c, \dots , und die Information (gegebenenfalls einschließlich Fehlinformation) p, q, r, \dots , über die momentane Sachlage gegeben sind? [...] Verbleibende Zweifel, ob der Schachcomputer *wirklich* Meinungen und Wünsche hat, sind unangebracht. Denn die Definition intentionaler Systeme, die ich gegeben habe, sagt nicht, daß intentionale Systeme *wirklich* Meinungen und Wünsche haben, sondern daß man ihr Verhalten erklären und voraussagen kann, indem man ihnen Meinungen und Wünsche *zuschreibt*; und ob man das, was man dem Computer zuschreibt, Meinungen, Meinungsanaloge, Informationskomplexe oder intentionale Sonstwas nennt, macht für die Art der Überlegung, die man auf der Grundlage der Zuschreibungen anstellt, keinen Unterschied. [...] Wir behandeln diese Computer durchaus erfolgreich als intentionale Systeme, und wir tun dies unabhängig von allen Überlegungen über die Materie, aus der sie bestehen, über ihre Herkunft, über die Stellung, die sie in der Gemeinschaft moralisch Handelnder haben oder nicht haben, über ihr Bewußtsein oder Selbstbewußtsein oder über die Determiniertheit oder Indeterminiertheit ihrer Operationen. Die Entscheidung, diese Strategie einzuschlagen, ist pragmatisch und nicht in sich richtig oder falsch. Man kann sich immer weigern, die intentionale Einstellung gegenüber dem Computer einzunehmen, und seine Schachmatts akzeptieren.¹⁸⁵

DENNETT teilt mit den Behavioristen die Ansicht, daß eine „intentionale Theorie als Psychologie leer [ist], weil sie Rationalität oder Intelligenz voraussetzt und nicht erklärt“, daß somit „Intentionalität keine *Grundlage* der Psychologie sein kann“ und man statt dessen „nach rein mechanistischen Regularitäten in den Aktivitäten [der] Subjekte Ausschau“ halten muß. Doch gibt es für ihn „wenig Grund für die Annahme, daß die Regularitäten auf der Oberfläche, im groben Verhalten liegen werden – es sei denn, wir legen den intentionalen Regularitäten eine künstliche Zwangsjacke an.“ Anders als für die Behavioristen

¹⁸⁴A. a. O., S. 163 f., 171 und 167.

¹⁸⁵A. a. O., S. 162, 164–166. „Diese Lobpreisung unseres Schachcomputers soll nicht implizieren, daß er ein völlig adäquates Modell oder eine völlig adäquate Simulation des Geistes oder intelligenter menschlicher oder tierischer Tätigkeit ist. Ich sage auch nicht, daß die Haltung, die wir gegenüber dem Computer einnehmen, genau dieselbe ist, die wir gegenüber einem Lebewesen einnehmen, das wir für bewußt und rational halten. [...] Ein Schachcomputer ist ... eine Maschine zum Schachspielen, was kein Mensch und kein Tier ist; und deshalb ist seine ‚Rationalität‘ beschränkt und künstlich.“ (a. a. O., S. 166 f.)

kann für ihn bei intentionalen Systemen – seien es nun Tiere, Menschen, Computer oder fremde Wesen¹⁸⁶ – die Annahme intentionaler Zustände und verborgener kleiner Männer durchaus zu gehaltvollen Voraussage- und Erklärungsweisen des Verhaltens führen, wenn man darauf achtet, „wo sich die Theorie bezüglich der Aufgabe, Intelligenz zu erklären, *in den roten Zahlen* befindet“, und wenn man bereit ist, diese Schuld durch die funktionale und physikalische Einstellung zu tilgen.¹⁸⁷

Dieser Übergang von intentionalen Erklärungen und Voraussagen zu zuverlässigeren Erklärungen und Voraussagen in der funktionalen Einstellung ... ist ... die Richtung, die der Theoretiker einschlagen sollte, wann immer es möglich ist. Schließlich wollen wir in der Lage sein, die Intelligenz von Menschen oder Tieren durch die Konstruktion dieser Wesen zu erklären und diese wiederum durch die natürliche Auslese dieser Konstruktion. Wir haben also, wann immer wir bei unseren Erklärungen auf der intentionalen Ebene stehen bleiben, einen ungeklärten Fall von Intelligenz oder Rationalität übriggelassen. [...] Jedesmal, wenn ein Theoretiker ein Ereignis, einen Zustand oder eine Struktur usw. in einem System (z. B. im Gehirn des Organismus) ein *Signal*, eine *Botschaft* oder einen *Befehl* nennt oder auf eine ander[e] Weise mit einem Inhalt ausstattet, *nimmt er eine Intelligenz-Anleihe auf*. Mit seinen Signalen, Botschaften oder Befehlen postuliert er implizit *etwas*, das die Signale *lesen*, das die Botschaften *verstehen* oder das *befehlen* kann. Andernfalls werden seine Signale nutzlos sein, sie werden unvernommen und unverstanden verhallen. Diese Anleihe muß man schließlich zurückzahlen, indem man herausfindet, was liest oder versteht und es analysiert. Denn gelingt das nicht, wird die Theorie unanalysierte Menschen-Analoga unter ihren Elementen haben, die mit genügend Intelligenz ausgestattet sind, um die Signale zu lesen usw., und sie wird somit die Beantwortung der Hauptfrage verschieben: *Wodurch kann man Intelligenz erklären?* [...] Unsere vielversprechendste Taktik bei der Suche nach Wissen über die innere Konstruktion besteht darin, Intelligenz-Anleihen aufzunehmen, periphere und innere Ereignisse mit Inhalt auszustatten und dann nach den Mechanismen Ausschau zu halten, die mit solchen ‚Botschaften‘ angemessen funktionieren, so daß wir die Anleihen zurückzahlen können. Diese Taktik ist eigentlich auch nicht unerprobt. Die Forschung auf dem Gebiet der künstlichen Intelligenz, welche unter anderem den Schachcomputer hervorbrachte, geht so vor, daß sie für ein intentional charakterisiertes Problem (wie bringt man den Computer dazu, die richtige Art von Information zu prüfen, die richtige Entscheidung zu treffen) eine Lösung in der funktionalen Einstellung erarbeitet – eine Annäherung an eine optimale Konstruktion.¹⁸⁸

¹⁸⁶Das bedeutet nicht, daß es hier für DENNETT keine Unterschiede gibt; intentionale Systeme können sich auch in ihrer „Ordnung“ unterscheiden, darin, ob es auch Meinungen und Wünsche *bezüglich* Meinungen und Wünschen gibt usw., wobei „eigentliche *Kommunikation*, Sprechakte im strengen, menschlichen Sinne des Wortes, *zumindest* von drei Ordnungen der Intentionalität [x *will*, daß y *meint*, daß x *meint*, daß p] abhängen, und zwar sowohl beim Sprecher wie beim Hörer“ (DENNETT (1992), S. 350f.). Ob dann z. B. Meerkatzen die Fähigkeit zur Intentionalität zweiter oder dritter Ordnung haben, ist für ihn – unter der Voraussetzung der Selektion von Vorteilen durch die Evolution – eine Frage der empirischen Verhaltensforschung (auch sonst ähnelt seine Position der neuerdings viel diskutierten „Evolutionären Erkenntnistheorie“).

¹⁸⁷DENNETT (1981), S. 174f. und 172. Für ihn sind SKINNERS und QUINES „unerbittliche[n] Verbote intentionaler Redeweisen auf allen Ebenen einer Theorie ... das Analogon des steinharten Neuengland-Konservativismus: Keine Verschuldung durch Aufnahme von Anleihen beim Aufbau einer Theorie!“

¹⁸⁸A. a. O., S. 171 und 175. „Ähnlich leihen auch Psychophysiker und Neurophysiologen intentionales Kapital, wenn sie Ereignisse routinemäßig als Informationsübertragung innerhalb des Nervensystems beschreiben; nur neigen sie oft dazu, ihre Schulden zu ignorieren oder zu leugnen“ (a. a. O.). Allerdings muß bezweifelt werden, daß man die Schulden je wirklich los werden kann: „Es gibt keinen neutralen Standpunkt, von dem aus sich die Beziehungen zwischen intentionalen Zuständen und der Welt überblicken und

Für DENNETT dient der Begriff eines intentionalen Systems dazu, „eine Brücke zu bilden, die den Bereich des Intentionalen (welcher die Welt unseres gesunden Menschenverstandes, die Welt von Personen und Handlungen, die Spieltheorie und die ‚neuralen Zeichen‘ des Biologen umfaßt) mit dem nicht-intentionalen Bereich der Naturwissenschaften verbindet.“¹⁸⁹ DENNETT ist nicht so sehr interessiert an einer „philosophischen“ Theorie des Geistes, das Denken ist für ihn eine Spielart intelligenten *Verhaltens*, und seine Theorie intentionaler Systeme ist zu verstehen als eine Strategie zur besseren Erklärung, Voraussage und Kontrolle dieses Verhaltens, als Strategie für eine naturwissenschaftliche Psychologie, als eine Fortentwicklung des Behaviorismus mit neuen, besseren Mitteln. Dabei hat für ihn die KI eine ausgezeichnete Stellung in dem Projekt, diese Brücke wirklich zu bauen und tragfähig zu machen, indem der Homunkulus, der „kleine Mann“ erst durch viele noch kleinere Männer und dann durch einfache Schalter ersetzt wird. An anderer Stelle schreibt er:

Kriegt man ein Team oder ein Komitee *vergleichsweise* unwissender, engstirniger, blinder Homunculi dazu, das intelligente Verhalten des Ganzen zu erzeugen, ist dies Fortschritt ... Schließlich ... hat man dann Homunculi, die so dumm sind (sie brauchen sich nur noch zu erinnern, ob sie ja oder nein sagen müssen, wenn sie gefragt werden), daß sie, wie man sagt, „durch eine Maschine ersetzt“ werden können. Man *entläßt* pffiffige Homunculi aus seinem Schema, indem man zur Erledigung der Arbeit solche Idiotenarmeen organisiert.¹⁹⁰

Das Problem, ob Computer „wirklich“ intentionale Zustände wie Glauben und Wünsche haben, ob sie „wirklich“ denken und verstehen können, stellt sich dabei für DENNETT gar nicht. Seine Position ist *instrumentalistisch*, die Zuschreibung intentionaler Zustände impliziert nicht, daß sie real existieren, und zwar auch nicht – und das ist seine Pointe (und das unterscheidet ihn von SEARLE und FODOR) – bei *uns Menschen*.¹⁹¹ Er stimmt SEARLE zu, daß Computer bloß syntaktische und nicht semantische Maschinen sind, aber, so DENNETT, wir – als Produkte eines blinden und „mechanischen“ Evolutionsprozesses, als „Überlebensmaschinen unserer Gene“ (Richard DAWKINS) – sind es (auf der subpersonalen Ebene) auch; eine Unterscheidung zwischen ursprünglicher und abgeleiteter Intentionalität macht für ihn keinen guten Sinn:

dann nicht-intentional beschreiben ließen. Jede Erklärung der Intentionalität spielt sich demzufolge im Zirkel der intentionalen Begriffe ab“ (SEARLE (1987), S. 46).

Zum „Weganalysieren“ vgl. auch oben die Anmerkung auf S. 39 und das von RUSSELL übernommene Motto von CARNAPS Frühwerk (vgl. S. 47). Statt OCKHAMS Messer nun der elektrische Rasierapparat.

¹⁸⁹A. a. O., S. 181 f.

¹⁹⁰DENNETT, *Brainstorms*, S. 123 f. (deutsch zitiert nach GARDNER (1989), S. 93). Für eine Skizze dafür, wie durch diese „Homunculus-Rekursion“, dem „Paradigma der Kognitionswissenschaft“, das „*Paradoxon des mechanischen Verstandes*“ (bzw. „... *der mechanischen Vernunft*“ [?? R. G.]), „das begriffstheoretische philosophische Problem“ gelöst („indem es zu einem empirischen und wissenschaftlichen gemacht wurde“) und das „*Mysterium der ursprünglichen Bedeutung*“ zwar „nicht ergründet“, aber „durchsichtiger gemacht“ werden soll, vergleiche HAUGELAND (1987), S. 92–106 (dieser Abschnitt ist auch „zum Teil eine Antwort“ auf das Gedankenexperiment von SEARLE). Eine wesentlich ausgearbeitetere Version dieses Grundgedankens hat Marvin MINSKY (1990) mit seiner Theorie der „Gesellschaft des Geistes“ jeweils sehr unbedarfter „Agenten“ vorgelegt. Zur Kritik an diesem Konzept der „Homunculus-Rekursion“ vgl. SEARLE (1993), S. 234 ff. und Geert KEIL (1993a), S. 166 ff. u. 338 ff.

¹⁹¹FODOR (1990, S. 6) kommentiert diese Position so: „The great virtue of instrumentalism – here as elsewhere – is that you get all the goodness and suffer none of the pain: you get to use propositional-attitude psychology to make behavioral predictions; you get to ‚accept‘ all the intentional explanations that it is convenient to accept; but you don’t have to answer hard questions about what the attitudes *are*.“ Als Replik vgl. DENNETT (1987), S. 71 ff.

As a late and specialized product, a triumph of Mother Nature's high tech, our intentionality is highly derived, and in just the same way that the intentionality of our robots (and even our books and maps) is derived. A shopping list in the head has no more intrinsic intentionality than a shopping list on a piece of paper. What the items on the list mean (if anything) is fixed by the role they play in the larger scheme of purposes. We may call our own intentionality real, but we must recognize that it is derived from the intentionality of natural selection, which is just as real – but just less easily discerned because of the vast difference in time scale and size. So if there is to be any original intentionality – original just in the sense of being derived from no other, ulterior source – the intentionality of natural selection deserves the honor.¹⁹²

Wie SEARLE beklagt auch DENNETT, daß in der zeitgenössischen Philosophie des Geistes so wenig das *Bewußtsein* thematisiert wird. Aber anders als für SEARLE ist es für ihn nicht das grundlegende Phänomen; seine Strategie ist vielmehr, „zuerst einen Ansatz für den Inhalt zu entwickeln, der *unabhängig* vom Bewußtsein und *grundlegender* als dieses ist – wobei in gleicher Weise alle unbewußten Inhaltsfixierungen (in Gehirnen, in Computern, in der ‚Erkenntnis‘ von Eigenschaften selektierter Baupläne durch die Evolution) behandelt werden – und dann das Bewußtsein auf dieser Grundlage zu erklären.“ Diese Strategie ist eine Fortsetzung der Behandlung intentionaler Inhalte: das Bewußtsein zuerst von einem intentionalen Dritte-Person-Standpunkt („heterophänomenologisch“) untersuchen, dann funktional rekonstruieren und schließlich (irgendwann) neurobiologisch erklären. Nach seinem Modell „werden alle Arten von Wahrnehmung – tatsächlich auch alle Formen des Denkens und der mentalen Aktivität – im Gehirn durch parallele, vielstufige Prozesse der Interpretation und Elaboration von Sinnesdaten vollzogen“, es „existieren vielfältige Kanäle, in denen spezialisierte Kreise, in paralleler Pandämonie, unterschiedliche Dinge zu tun versuchen, vielfältige Konzepte erzeugen“, „Konzepte von Erzählfragmenten in unterschiedlichen Stadien der Bearbeitung und an unterschiedlichen Orten des Gehirns.“ Der uns durch Introspektion zugängliche kontinuierliche, serielle Bewußtseinsstrom (die „Joyce-Maschine“) ist ein spätes, hauptsächlich durch kulturelle Evolution (Sprache, kulturelle „Meme“ und Bilder) entstandenes Produkt, ein „Komplex von Operationen einer virtuellen Maschine“, eine Art Software: „Ein bewußter menschlicher Geist ist mehr oder weniger eine seriale virtuelle Maschine, die – ineffizient – an der parallelen Hardware montiert ist, die die Evolution uns geliefert hat“ (etwa so, wie mein Textverarbeitungsprogramm eine virtuelle Maschine bezüglich meines Computers ist; die Von-Neumann-Maschine als Prototyp des programmierbaren Computers wurde nach DENNETT in Analogie zu der virtuellen Maschine unseres Bewußtseins entwickelt). Das „Ich“ ist dann eine nützliche Abstraktion, ein „narratives Gravitationszentrum“, aber keinesfalls eine zentrale Steuerungseinheit („Pandämonium“ statt „cartesianisches Theater“). Eingeräumt wird dabei von DENNETT das „Fehlen eines Rechenmodells, das zeigen könnte, was die Maschine tut und wie. Wir befinden uns hier nach wie vor in einem metaphorischen Stadium“.¹⁹³

Nach meinem Eindruck trifft die pragmatisch und naturalistisch begründete intentionale (und dann funktionale und physikalische) *Einstellung* von DENNETT viel besser das Selbstverständnis und die „Philosophie“ der KI als FODORS Funktionalismus mit seinen

¹⁹²DENNETT (1987), S. 318. Vgl. auch dazu KEIL (1993a), S. 340 ff.

¹⁹³DENNETT (1994), S. 575, 151, 333, 182, 557, 288 und 360 (der anspruchsvolle Originaltitel lautet *Consciousness Explained*). SEARLE (1993, S. 22), hält das für eine Verfehlung des Themas, für eine Strategie, „den Begriff des Bewußtseins so umzudefinieren, daß er sich nicht länger auf wirklich bewußte Zustände – d. h. auf innere, subjektive, qualitative, erstpönliche Geisteszustände – bezieht“. Auch John HAUGELAND (1987, S. 216 f.) weist darauf hin, daß bei DENNETTS Ansatz wesentliche Fragen unbeantwortet sind.

recht anspruchsvollen Implikationen und einzulösenden theoretischen Lasten (eine gewisse theoretische Beliebigkeit ist nicht nur Kennzeichen der technischen, sondern auch der kognitiven KI; was zählt, ist der Erfolg, und dabei sind auch andere Ansätze recht als die Theorie der „Sprache des Geistes“¹⁹⁴).

Philosophisch, aber auch aus der Sicht eines an der Lebenswelt orientierten *Common sense*, bleiben die in diesem Abschnitt angebotenen Bemühungen um eine „Lösung“ der Rätsel und Probleme von Verstehen, Intentionalität und Bewußtsein allerdings unbefriedigend. Zu deren Verstehen und Erklären gehört neben der zwischenzeitlich verpönten Ichperspektive der Bewußtseinsphilosophie (Verhältnis *Ich – Es*) und der in der KI bevorzugten Objektivierung des Subjektiven (Verhältnis *Du – Es*) auch die Perspektive des Intersubjektiven (Verhältnis *Ich – Du*), die (auch *leibhaftige*) Teilnahme an einer gemeinsamen Lebenspraxis und Kommunikation mit der gegenseitigen Anerkennung als bedürftige und moralisch verantwortliche *Personen* (und das nicht nur, um im Dreieck *Ich, Du, Es* eine Lücke zu schließen, sondern auch als übergeordnete Klammer für die ganze Figur¹⁹⁵). In diesem Rahmen erfordert die Antwort auf die Frage, ob wir Computern die Fähigkeit zum Denken und Verstehen, zu Intentionalität und Bewußtsein, zusprechen können und sollen, eine *Entscheidung*, die durch den empirischen Aufweis der Möglichkeiten oder des Scheiterns technischer Konstrukte zwar beeinflusst, nicht aber ersetzt werden kann (das zeigte schon KANT bei seiner Behandlung der Antinomie zwischen der „Kausalität durch Freiheit“ und der „Kausalität nach Gesetzen der Natur“¹⁹⁶).

7 Zwischenbilanz: Können Computer denken?

„Können – zumindest prinzipiell und irgendwann – Maschinen denken?“ Zur Beantwortung dieser Frage, darauf hat Alan TURING schon vor mehr als vierzig Jahren hingewiesen,¹⁹⁷ ist eine genauere Bestimmung der Begriffe „Denken“ und „Maschine“ nötig. Um mit letzterem anzufangen: TURING ging bei seiner Prognose von *Digitalcomputern* aus. Zwar lassen sich auch andere Architekturen (massive Parallelverarbeitung, Neuronale Netze) äquivalent auf Von-Neumann-Rechnern nachbilden, aber nicht unbedingt auch deren *Geschwindigkeit*. Wenn es stimmt, daß das Gehirn – mit seinen schätzungsweise 100 000 000 000 (10^{11}) Neuronen mit im Durchschnitt je mehr als 1000 Verbindungen zu anderen Neuronen (die Zahl der möglichen Kombinationen übertrifft die Anzahl der Atome im bekannten Universum) – ca. 10^{15} Operationen pro Sekunde ausführt, so übertrifft das die Leistungsfähigkeit der gegenwärtig vorhandenen Rechner mit massiver Parallelverarbeitung um mehrere Zehnerpotenzen. Auch wenn sich die Technik wie bisher rasant weiterentwickelt, stößt man hier möglicherweise auf Grenzen, die sich mit der vorhandenen und absehbaren Technik (z. B. optische statt elektronische Systeme) nicht überwinden

¹⁹⁴Das ist zwar kritisch, aber nicht unbedingt abwertend gemeint (vgl. auch oben auf S. 4 die Anmerkung mit GALTUNGS Charakterisierung der unterschiedlichen Denkstile); diese Einstellung ist auch Ausdruck eines spezifisch nordamerikanischen Pragmatismus, den ich bei DEWEY und PEIRCE durchaus schätze.

¹⁹⁵Ähnlich plädiert Heiner HASTEDT (1988, S. 235 ff.) für den „Vorrang der Teilnehmerperspektive“ vor der Beobachterperspektive.

¹⁹⁶*Kritik der reinen Vernunft*, B 472 ff. Vgl. dazu auch Ekkehard MARTENS (1993), S. 153. Für zeitgenössische Nachweise, daß keine noch so entwickelte empirische Wissenschaft und Technik die einer solchen Entscheidung notwendigerweise zugrundeliegende Selbstdeutung ersetzen kann, vgl. so unterschiedliche Schriften wie die von Geert KEIL (1993a), Hilary PUTNAM (1982) und Robert SPAEMANN & Reinhard Löw (1985).

¹⁹⁷Vgl. TURING (1967, 1987b) und Unterrichtsmodul 1, Abschnitt 3.

lassen. Die Arbeitsweise von Neuronen auf subzellulärer Ebene, ihre Biochemie und ihr Zusammenspiel mit anderen organischen Substanzen wie Enzymen ist (wie insgesamt die Arbeitsweise des Gehirns) wenig geklärt. Es kann sich herausstellen, daß die Leistung (Arbeit pro Zeit) des Gehirns tatsächlich organische, molekulare Substanzen und Prozesse erfordert und daß diese Leistung mit vertretbarem Aufwand vielleicht nur auf die durch die Evolution entwickelte Weise zu haben ist. Sind Systeme, die auf dieser Basis arbeiten, dann noch „Maschinen“?¹⁹⁸

Noch größere Probleme hat man mit dem „Denken“, mit der Frage danach, was *genau* man denn eigentlich von Computern erwartet, um ihnen diese Fähigkeit zuzusprechen. FREUD sah beispielsweise im Denken eine „Aufhaltung der motorischen Abfuhr (des Handelns)“, ein inneres „Probegedanken“, das sich „aus dem Vorstellen herausbildete“ und darüber „erhob und sich den Relationen der Objekteindrücke zuwendete“; es „war wahrscheinlich ursprünglich unbewußt, ... und erhielt weitere für das Bewußtsein wahrnehmbare Qualitäten erst durch die Bindung an die Wortreste.“¹⁹⁹ Für DESCARTES war dagegen Denken noch untrennbar mit *Bewußtsein* verbunden. Unter Denken verstand er „alles, was derart in uns geschieht, daß wir uns seiner unmittelbar aus uns selbst bewußt sind“, und alle Bewußtseinsarten führte er auf den Verstand einerseits und den Willen andererseits zurück (vgl. oben S. 8). Auch für LOCKE war das Bewußtsein zentral, und zwar als fortlaufende Identität eines *Ichs*, die es erst ermöglicht, von einer vernunftbegabten und verantwortungsvollen *Person* zu sprechen (vgl. oben S. 12). Ähnlich mußte für KANT das „*Ich denke*“ alle „Vorstellungen begleiten können;“ das Denken war für ihn, als Vermögen des *Verstandes*, die „Handlung, die Synthesis des Mannigfaltigen“, welches durch die Anschauung gegeben wird, „zur Einheit der Apperzeption zu bringen“. Darüber stand für ihn die *Vernunft*, eine Fähigkeit die letztlich „aufs Moralische gestellt“ ist (vgl. oben S. 23 f.). Weiter wird seit KANT bei Denkprozessen (als Operationen mit Vorstellungen) deutlicher und in methodischer Absicht zwischen Begriffsbildung, Urteilen und Schließen unterschieden. Mit der Entwicklung der formalen Logik durch FREGE und andere wurden diese Operationen objektiviert, vom Bewußtsein und den subjektiven Vorstellungen abgelöst; Gedanken werden in diesem Verständnis durch Denken nicht erzeugt, sondern *gefaßt* (vgl. oben S. 34). Damit rückte beim Nachdenken über das Denken immer mehr die *Sprache* in den Mittelpunkt des Interesses, zunächst als Konzept einer idealen Sprache (für den frühen WITTGENSTEIN war Denken ein Abbilden der logischen Form der Welt in Zeichen, vgl. oben S. 43), dann die ganz unterschiedlichen „Sprachspiele“ der Alltagssprache mit ihren vielfältigen pragmatischen und sozialen Kontexten. Auch wenn der genaue Zusammenhang von Denken und Sprechen (nach PLATON sind Denken und Rede dasselbe, Denken ist „das innere Gespräch der Seele mit sich selbst“²⁰⁰) bis heute strittig ist, gibt es kaum einen Zweifel daran, daß Sprache das wichtigste Medium des Denkens und Sprachverstehen die zentrale, spezifisch menschliche kognitive Leistung ist (vgl. Unterrichtsmodul 2). Mit dieser Wende zur Sprache rückte *Intentionalität* in das Zentrum der Diskussion (vgl. oben S. 70): Das Denken handelt *von* etwas, es hat *Bedeutung*, es ist auf etwas *gerichtet* (da sich das Denken aber nicht nur auf äußere Gegenstände und Sachverhalte richten kann, sondern auch auf sich selbst, wird man damit allerdings die Probleme von Subjektivität und Bewußtsein nicht los). Versucht man eine Zusammenfassung, so könnte man sagen: Denken ist ein in der Regel bewußter, subjektiv gefärbter und durch das Erleben eines Ichs

¹⁹⁸ Vgl. zu diesen Spekulationen den Aufsatz *Fast Thinking* in DENNETT (1987), S. 323–337.

¹⁹⁹ FREUD (1911), S. 20.

²⁰⁰ *Sophistes*, 263 e; vgl. auch *Theaitetos*, 189 e.

begleiteter geistiger Vorgang, der sich in theoretischer und praktischer Absicht (z. B. zur Vorbereitung von Handlungen, wenn instinktives oder erlerntes Verhalten nicht ausreicht) auf äußere und innere Sachverhalte richtet und sich dabei sozial und kulturell vermittelter Zeichensysteme bedient.

Werden in der Philosophie eher die Grundprinzipien und Grenzen des Denkens thematisiert, untersucht die Psychologie – im Anschluß an die Assoziationstheorie HUMES (vgl. oben S. 15) – die „Mechanik“, die „Funktionsweise“ des Denkens. Die Ich-Perspektive der Bewußtseinsphilosophie wird durch die Beobachterperspektive ersetzt, Denken wird als „objektives Verhalten“ untersucht (vorzugsweise als „Problemlösen“²⁰¹). Nach der Assoziationstheorie, der Gestaltpsychologie und der „Eliminierung“ des Mentalen durch den Behaviorismus (vgl. den Exkurs im Unterrichtsmodul 1) stehen nun Ansätze im Vordergrund, die sich am Paradigma der *Informationsverarbeitung* orientieren, z. B. als weitgehend unbewußte, regelgeleitete Manipulation „Mentaler Modelle“ als interne Repräsentationen der Welt.²⁰² Insgesamt hat sich dabei das Interesse weg vom recht unscharfen Begriff des Denkens hin zu dem der *Intelligenz* (als hinter dem Vollzug des Denkens angenommene, allgemeine Fähigkeit; vgl. dazu den Abschnitt 1 im Teil 1) einerseits und enger umgrenzten einzelnen Leistungen und Fähigkeiten wie Wahrnehmen, Erinnern, Lernen, Schlußfolgern und Sprachgebrauch andererseits verlagert, und als Oberbegriff für diese psychischen Prozesse hat sich der Begriff der *Kognition* durchgesetzt. Die meisten Kognitionspsychologen verstehen sich mittlerweile als Teilnehmer einer *Cognitive science*, die sich als am „Computervermodell des Geistes“ orientierte Zusammenarbeit mit Linguisten, Informatikern und Neurobiologen etabliert hat und der sich auch viele KI-ler zugehörig fühlen (auch immer mehr Philosophen – insbesondere die, die nach dem *Linguistic turn* nun den *Naturalistic turn* propagieren – fühlen sich von der Verheißung (natur-)wissenschaftlicher²⁰³ Erklärungen für die Rätsel des Geistes angezogen). Durch diese Entwicklung werden zwei Trends unterstützt, die das im letzten Absatz vorgestellte, aus der philosophischen Tradition stammende Verständnis vom Denken verändern. Wenn Denken als „objektives Verhalten“ verstanden und in erster Linie im Hinblick auf die erbrachte Leistung untersucht (und vom Bewußtsein abgekoppelt) wird, verändert sich erstens auch der Sprachgebrauch. Der Ausdruck „Denken“ wird dann – wie schon bei „Fliegen“ oder „Schwimmen“ – auch auf Artefakte übertragen, die möglicherweise auf andere Art, aber im Ergebnis dasselbe leisten (vgl. dazu den Text von Walter SCHULZ im Unterrichtsmodul 1, Abschnitt 4). War zweitens der von DESCARTES, LOCKE und KANT entwickelte Begriff des Geistes noch „gekennzeichnet durch die Synthese von intellektuellem und moralischem Tun“, erleben wir nun zunehmend eine „Dissoziation von kognitiven und moralischen Kompetenzen, von Intellekt und Urteilskraft.“²⁰⁴

²⁰¹ „Das Zweigespann Bewußtsein-Gehirn (wie immer es zusammenarbeitet) ist . . . eine Kontextmaschine, welche sich auf geheimnisvolle Weise die Übersicht im Meer der Kontexte bewahrt. *Das* ist die Zentralleistung der Intelligenz, nicht die Problemlösung“ (Heinz ZEMANEK (1993), S. 93).

²⁰² Vgl. JOHNSON-LAIRD (1983) und den Anhang von Unterrichtsmodul 2.

²⁰³ Wie attraktiv ein solcher „Aufstieg“ insbesondere im angelsächsischen Kulturraum ist, mag folgende Anekdote illustrieren. In einer großen Buchhandlung in Cambridge befinden sich die Abteilungen für Philosophie, Geschichte, Literatur (die *Humanities*) im zweiten, die für Mathematik, Physik, Chemie, Biologie, Informatik usw. im obersten, dritten Stock. Ich suchte ein Buch von einem Philosophen über Künstliche Intelligenz und war nicht sicher, in welcher Abteilung ich es finden könnte. Als ich einer Verkäuferin sagte, daß man es möglicherweise auch der *Cognitive science* zuordnen könnte (eine solche Abteilung war nicht eigens ausgewiesen), war sie sicher: Dritter Stock! „Because *every* science is on the third floor.“

²⁰⁴ Sybille KRÄMER (1992), S. 340 f. Vgl. dazu auch meine in der Einleitung geschilderte Erstbegegnung mit einem Vertreter der KI.

Eine klare Antwort auf die zu Beginn dieses Abschnitts gestellte Frage wird von Seiten des Funktionalismus, des „Computermodell des Geistes“, der „starken KI“ gegeben: Ja! Entsprechend programmierte Digitalcomputer können deshalb im Wortsinne denken, weil wir selbst in einem relevanten Sinne solche Computer sind. Gegenüber der über dreihundert Jahre alten These „Denken ist Rechnen“ von HOBBS kann die zeitgenössische These „Denken ist Komputieren“ auf ihrer Habenseite verbuchen.²⁰⁵

- Das in der Bewußtseinsphilosophie in der Nachfolge von DESCARTES über LOCKE und HUME bis KANT entstandene Konzept, daß Denken das regelgeleitete Operieren mit mentalen Repräsentationen der Außenwelt, mit „Ideen“ oder „Vorstellungen“, sei.
- Das von LEIBNIZ entworfene, von FREGE, RUSSELL, WITTGENSTEIN und CARNAP ausgebaute Konzept einer eindeutigen, kontextfreien, idealen Sprache als exaktes Abbild der Welt, als rationale Nachkonstruktion der gesamten Erkenntnis der Wirklichkeit, als narrensichere Methode, als „symbolische Maschine“ zum Erfinden und Prüfen durch bloß formale, syntaktische Transformationsregeln, die automatisch der intendierten Semantik Rechnung tragen (wobei Sinn und Bedeutung komplexer Symbole ausschließlich eine Funktion der elementaren Symbole und ihrer Kombination sind).
- TURINGS These, daß jedes nicht willkürliche, regelgeleitete Verhalten ergebnisgleich durch Input, Maschinentafel und Output einer Turingmaschine beschrieben und simuliert werden kann, und sein Nachweis, daß dafür eine universelle Maschine genügt (und äquivalent damit auch jeder uns vertraute, programmierbare Computer vom Von-Neumann-Typ).

Die Computer-Philosophie des Denkens, mit ihrer These, daß Computer gleichsam *per definitionem* denken können, teils Ergebnis der Reflexion über die KI, teils deren theoretische Basis, vereint somit zwei traditionelle Stränge der abendländischen Philosophie mit der Erfindung der informationsverarbeitenden Universalmaschine zu einer komputationalen Theorie des Geistes, nach der Denkprozesse formale Operationen über eine unsere Überzeugungen über die Außenwelt symbolisch repräsentierenden, biologisch-physikalisch wie eine Maschinsprache realisierten „Sprache des Geistes“ sind (ironischerweise wird damit das Programm einer formalen Logik und einer idealen Sprache, mit dem sich FREGE ausdrücklich gegen eine psychologische Begründung der Logik und des Denkens wandte, zu einem Fundament der Psychologie, und Sybille KRÄMER²⁰⁶ weist mit Recht darauf hin, daß hier eine *Kulturtechnik* (nämlich Rechnen als symbolisches Operieren) als Vorbild für eine empirische These über einen biologischen Vorgang dient). Mit der von PUTNAM vorgeschlagenen Unterscheidung einer physikalisch-strukturellen und einer logisch-funktionalen Ebene bei der Beschreibung dieser Prozesse will man dabei sowohl den einheitswissenschaftlichen Reduktionismus als auch die Schwächen früherer repräsentationaler Theorien des Geistes vermeiden, den „Vorhang der Ideen“ zerreißen, den „Ungeist des Homunkulus“ und das von RYLE verspottete „Gespenst in der Maschine“ vertreiben (und manche

²⁰⁵Neben den nachfolgend genannten Entwicklungen und Ergebnissen aus Philosophie und Logik gibt es noch weitere aus den Bereichen kognitive Psychologie, Linguistik, Anthropologie, Neurobiologie und aus der KI selbst. Für eine umfassende Darstellung der Geschichte und Ergebnisse dieser Disziplinen vergleiche Howard GARDNER (1989).

²⁰⁶KRÄMER (1991), S. 6 ff., (1994a), S. 99 f. und 114.

möchten dann auch noch Begriffe wie „Selbst“, „Wissen“, „Bedeutung“ und „Verstehen“ – „prescientific idea germs“ wie MINSKY gerne sagt²⁰⁷ – „weganalysieren“). Dadurch wird Denken unabhängig von der durch die Evolution hervorgebrachten spezifisch menschlichen Lösung: Geist verhält sich zum Gehirn wie das Programm zur Hardware.

Das „Computermodell des Geistes“ hat, wie angedeutet, entscheidende Schwächen und Lücken: eine befriedigende Erklärung der Semantik und Pragmatik der „Sprache des Geistes“, der Verbindungen zur Außenwelt und zu anderen Menschen, von Intentionalität und Verstehen, von der Erfahrung des Bewußtseins und des Ichs, von der Realisierung der funktionalen Zustände durch neurophysiologische Prozesse, ist nicht in Sicht; aber dieser Einwand ist nicht ganz fair, konkurrierende Theorien stehen hierbei nicht besser da. Aus der Sicht der KI ist dieses „Paradigma der Kognitionswissenschaft“ kein Glaubenssatz, sondern eine empirisch überprüfbare These (wie wir aus den Bemühungen der Wissenschaftstheorie wissen, ist ein solcher Trennstrich allerdings kaum zu ziehen). Auch wenn nicht ganz klar ist, auf Grund welcher empirischer Ergebnisse die Vertreter dieser Theorie bereit sind, sie als gescheitert anzusehen, haben empirische Befunde somit sicher einige Bedeutung.

Zwei Dinge fallen dabei auf. Erstens haben sich die großen Versprechungen der Anfangszeit nicht erfüllt. Zwar gibt es gewisse Erfolge bei jeweils einzelnen kognitiven Leistungen, aber der große Durchbruch, die Entwicklung von Computerprogrammen oder Robotern mit einem wirklich umfassenden Denkvermögen, wurde nicht erreicht und ist auch nicht in Sicht. Sprachverstehen (vgl. Unterrichtsmodul 2), Repräsentation von und Umgang mit Alltagswissen, maschinelles Lernen sind weitgehend ungelöst (vgl. den Abschnitt 3.2 im Teil 1), und wie die Überlegungen zum Turing-Test (Unterrichtsmodul 1, Abschnitt 3) und zu CUGAF (oben Abschnitt 5) zeigen, sind Lösungen nur zu erwarten, wenn die Systeme durch Sensoren und Effektoren „eigene“ Erfahrungen mit der Wirklichkeit machen können; dabei werden jeweils entscheidende Fortschritte in den genannten Teilbereichen erst erwartet, wenn die Probleme der anderen Bereiche gelöst sind (vgl. die Zwischenbilanz am Ende vom Teil 1). Die einst optimistische und vom Paradigma (im Sinne KUHNS) der „Hypothese der physikalischen Symbolsysteme“ inspirierte Forschergemeinde, die „*Good Old Fashioned Artificial Intelligence*“ (John HAUGELAND), zerfällt in zunehmend unübersichtliche Teilgruppen und Forschungsbereiche mit jeweils eigenen Themen und spezifischen, häufig *ad hoc* gewählten Lösungsansätzen, innerhalb der KI sinkt der Stern des „Computermodells des Geistes“ zugunsten eines „Gehirnmodells des Computers“ (dabei ist wieder, wie schon bei der Metapher „Elektronenhirn“ in den 50er Jahren, der Mensch das Vorbild; auch das ist ein Abrücken vom ursprünglichen, nämlich symbolverarbeitenden Ansatz). Kurz: die „starke KI“ zeigt deutliche Anzeichen eines degenerierenden Forschungsprogramms (im Sinne von LAKATOS).²⁰⁸

²⁰⁷In seiner Antwort zu SEARLE (1980), S. 439.

²⁰⁸Vgl. KUHN (1973), S. 28 ff. und LAKATOS (1974), S. 129 ff. Hier kann man sich kaum einen besseren Zeugen wünschen als Jerry FODOR: „... the psychology of thought has proved quite intractable. In particular, on my view, the attempt to develop general models of intelligent problem-solving – which one associates most closely with work in artificial intelligence by such figures as Schank, Minsky, Newell, Winograd, and others – has produced surprisingly little insight despite the ingenuity and seriousness with which it has often been pursued. I have the impression that it is becoming rather generally conceded that this early, as one might say Wagnerian, phase of AI research has led to a dead end [...] Intellectual capacities were divided into what seem, in retrospect, to be quite arbitrary subdepartments (proving theorems of elementary logic; pushing blocks around; ordering hamburgers), and the attempted simulations proceeded by supplying machines with very large amounts of more or less disorganized, highly topic-specific facts and

Zweitens, und auch das ist ein empirischer Befund: Bedeutende Philosophen (und einige Theoretiker und Praktiker der KI und der Kognitionswissenschaft), die zu dieser Sicht entscheidend beigetragen haben, indem sie, meist mit „jugendlichem“ Elan, den Versuch unternahmen, das Denken mit formalen Mitteln einzufangen, haben sich davon später selbst distanziert. Gewisse Revisionen im Werk von RUSSELL und CARNAP wurden oben schon angedeutet, besonders illustrativ allerdings – neben dem bekannten Umschwung bei WITTGENSTEIN – ist die Entwicklung von Hilary PUTNAM, einem der Väter des funktionalistischen Gedankens:

Wir alle haben mitbekommen, wie eine Wissenschaft von der Gesellschaft oder vom Menschen nach der anderen – Psychologie, Soziologie und Ökonomie – unter den Einfluß dieser oder jener Modemasche geraten ist. In den Vereinigten Staaten waren solche Maschen besonders häufig das Ergebnis einer reduktionistischen Vorstellung dessen, was „Wissenschaftlichkeit“ bedeutet. Die Idee, etwas könne nur dann zur „Kognitionswissenschaft“ beitragen, wenn es in der Terminologie der „mentalenen Repräsentationen“ dargestellt wird (und wenn dies Repräsentationen „kalkülmäßig“ beschrieben werden), ist lediglich ein weiteres Beispiel für diese beklagenswerte Tendenz. [...] Mir scheint, daß wir die Forderung aufgeben müssen, alle Begriffe, die wir ernst nehmen, müßten sich auf das Vokabular und den Begriffsapparat der exakten Wissenschaften zurückführen lassen. Nach meiner Überzeugung ist es nicht die Intentionalität selbst, die in Schwierigkeiten steckt, sondern der Reduktionismus.²⁰⁹

Vielleicht muß man – wie der Professor für Moderne Mathematik und Mathematische Logik an der Harvard Universität – wirklich erst die Leiter hinaufsteigen, bevor man sie wegwerfen kann (und barfuß gehen schickt sich nur für die, die Schuhe haben): „Daß die Rationalität durch ein ideales Computerprogramm definiert wird, ist eine szientistische Theorie, die von den exakten Wissenschaften inspiriert ist; [...] Ich bestreite nicht, daß die Logik wichtig ist oder daß formale Untersuchungen der Bestätigungstheorie, der Semantik natürlicher Sprachen usw. wichtig sind. Ich neige jedoch zu der Auffassung, daß sie eher Randgebiete der Philosophie“ und ein Versuch sind, „dem wirklichen Problem aus dem Weg zu gehen: eine verständige und menschliche Beschreibung der Reichweite der Vernunft zu geben.“²¹⁰

Auch Terry WINOGRAD – ein Star der KI-Szene, sein in einer virtuellen „Mikrowelt“ von geometrischen Körpern operierendes Programm SHRDLU²¹¹ galt als lang ersehnter Durchbruch beim Problem des Sprachverstehens – hat eine radikale Wandlung vollzogen und sich von der „rationalistischen Tradition“, in der er mit seinen eigenen Arbeiten stand, losgesagt:

Der rationalistische Zugang zum Problem der Bedeutung, der Systemen wie SHRDLU anhaftet, gründet auf der Voraussetzung, daß sich die Bedeutung von Worten und der

heuristics. [...] What emerged was a picture of the mind that looked rather embarrassingly like a Sears catalogue“ (1983, S. 126 f.). Nun propagiert er „Fodor’s First Law of the Nonexistence of Cognitive Science“: „the more global ... a cognitive process is, the less anybody understands it. *Very* global processes, like analogical reasoning, aren’t understood at all“ (a. a. O. S. 107).

²⁰⁹PUTNAM (1991), S. 112 f. und 136.

²¹⁰PUTNAM (1982), S. 171 f.

²¹¹„it’s a mere computer which dreams that it’s a robot“, „precisely the situation that Descartes dreads“ (FODOR, (1980), S. 65).

damit gebildeten Sätze und Redewendungen unabhängig von der situationsabhängigen und je individuellen Interpretation ermitteln lassen. [...] Hintergrundbezug und Interpretation durchziehen unser gesamtes Alltagsleben. Jede Bedeutung leitet sich ab aus einer in einer Situation wurzelnden Interpretation.²¹²

Auf die Bedeutung des situativen Kontextes und des „Immer-schon-in-der-Welt-sein“ für das Verstehen und Handeln sowie auf die Schwierigkeiten seiner formalen Beschreibung als von der KI unzulässigerweise ignoriertes Problem hat mit Nachdruck immer wieder Hubert L. DREYFUS²¹³ auf Grund seiner Kenntnis der europäischen Sprachphilosophie und Phänomenologie hingewiesen. Nun folgt WINOGRAD seinen Ratschlägen, rezipiert HEIDEGGER und zitiert mit den einleitenden Sätzen „Das Wesen unseres Seins ist durch kulturellen Hintergrund festgelegt; und da dieser bereits durch die Art und Weise unserer Spracherefahrung und unseres Lebens in der Sprache selbst geprägt ist, kann er nicht vollständig in dieser Sprache erklärt werden“²¹⁴ den modernen Klassiker der Hermeneutik, Hans-Georg GADAMER:

Die Gewinnung des Bewußtseins einer Situation ist aber in jedem Falle eine Aufgabe von einiger Schwierigkeit. Der Begriff der Situation ist ja dadurch charakterisiert, daß man sich nicht ihr gegenüber befindet und daher kein gegenständliches Wissen von ihr haben kann. Man steht in ihr, findet sich immer schon in einer Situation vor, deren Erhellung die nie ganz zu vollendende Aufgabe ist.²¹⁵

In der Tat: „Die Welt gibt es nur als Idee der Einheit aller Folgen von Situationen und niemals als ‚fertigen‘ Gegenstand“ und die Selbstbestimmung der Menschen „ist der großangelegte Versuch, die Folge von Situationen, also die Handlungs- und Redezusammenhänge der Menschen über Orte und Zeiten hinweg in *eine* zusammenhängende Situation, *die Welt*, zu verwandeln, was natürlich selbst nur als offener Prozeß mit zahllosen Verzweigungen“ möglich ist.²¹⁶ Dagegen ist der von der Vorstellung der „Welt als der Gesamtheit der Tatsachen im logischen Raum“ des frühen WITTGENSTEIN inspirierte Versuch, Denken, Intentionalität und Verstehen, umfassendes Wissen (*knowing that*) und Können (*knowing how*) in einer Art Momentaufnahme, als „Modellierung von Weltausschnitten“, von diesen Handlungs- und Redezusammenhängen abzulösen (so wie FREGE und RUSSELL – nach dem Vorbild einer „platonisch“, als „ewige Wahrheiten“ verstandenen Mathematik – eine

²¹²WINOGRAD & FLORES (1989), S. 185 ff. Der Ausdruck „rationalistische Tradition“ (vgl. a. a. O., Kapitel 2) ist ein wenig mißverständlich, zu ihr haben, wie wir gesehen haben, auch verschiedene Empiristen wie LOCKE, RUSSELL und CARNAP beigetragen (im übrigen bin ich an *diesem* Punkt mit CARNAP und der evolutionären Erkenntnistheorie einig, daß der Unterschied letztlich nicht so wichtig ist).

²¹³Vgl. DREYFUS (1985). Er hatte entdeckt, daß vieles von dem, was die KI versucht, schon Jahrzehnte vorher HUSSERL versucht hat. Dessen Scheitern und HEIDEGGERS Kritik an dem Unternehmen ist – neben der Spätphilosophie WITTGENSTEINS – einer der Grundpfeiler von DREYFUS' KI-Kritik. Die KI konnte – vor der Arbeit von Terry WINOGRAD & Fernando FLORES – mit diesen Hinweisen lange Zeit wenig anfangen. Inzwischen gibt es aber, wegen mangelnder Erfolge, auch in der technischen KI einige, die das Problem der „*situatedness*“ sehen und ein „Situating Adaptive Design“ versuchen (vgl. z. B. Rolf PFEIFER & Philip RADEMAKERS (1991)). Zu den Beschränkungen neuerer Ansätze wie LENATS CYC und Neuronale Netze vgl. DREYFUS (1993).

²¹⁴WINOGRAD & FLORES (1989), S. 58.

²¹⁵GADAMER, *Wahrheit und Methode*, Tübingen ⁶1990, S. 307.

²¹⁶Kuno LORENZ (1990), S. 90 und 84 f. Und deutlich wird, „daß Selbstbestimmung, so verstanden, Selbsterkenntnis nur vereint mit Welterkenntnis nach sich zieht: Anthropologie und Kosmologie sind siamesische Zwillinge, deren Trennung beide sterben ließe.“ (a. a. O.)

formale, kontextbefreite Logik entwickelten) und in die Zeilen eines Computerprogramms zu gießen, ein sehr heroisches oder ein sehr naives Unternehmen.

Theoretische Lücken, ausbleibende Erfolge, ein degenerierendes Forschungsprogramm, das Aufgeben von Jugendträumen zugunsten einer reiferen Sicht mit Blick auch auf andere, vorher ignorierte philosophische Traditionen: also bei der Frage „Können Computer denken?“ ein Sieg der „Humanisten“ über die „Reduktionisten“? Ja, ich glaube schon; zwar kein K. O., aber einer nach Punkten, vielleicht nur ein Etappensieg. Aber ich bin mit dieser Antwort nicht ganz zufrieden.

Sowenig ich dafür plädiere, unkritisch und ohne Widerspruch die An-Sprüche und das Menschenbild der KI sich selbst zu überlassen, sowenig möchte ich die KI „entlarven“ und ein bequemes „Na also, ich habe es ja immer gewußt“ stützen. Ich glaube nicht, daß es mit dem Ansatz der physikalischen Symbolsysteme gelingen wird, Computerprogramme oder Roboter zu entwickeln, die unseren Denk- und Intelligenzleistungen nahe kommen, und ich glaube auch, daß die neuerdings wieder propagierten subsymbolischen, konnektionistischen Ansätze – neuronale Netze, massive Parallelverarbeitung – nur bei Problemen Erfolg haben werden, die im weitesten Sinne Mustererkennung erfordern.²¹⁷ Aber die Beschäftigung mit der KI ist dennoch nicht einfach nur eine gefährliche Zeit- und Ressourcenverschwendung, nicht nur eine aufwendige Sackgasse (eher schon die Windung auf einem Wendel). Ich denke, daß die KI auch den Anlaß und die Chance bietet, über vieles neu nachzudenken, unser Selbstverständnis, unsere Vorstellung vom „guten Leben“ und unser Konzept von Rationalität zu überprüfen (dazu gehören nicht nur Antworten darauf, ob intelligente Maschinen möglich sind, sondern auch darauf, ob und zu welchen Zwecken wir solche Maschinen *wollen*). Das Aufspüren der ideengeschichtlichen Wurzeln der KI, die Reflexion über ihre Vorhaben und Resultate (z. B. daß so hoch geschätzte geistige Fähigkeiten wie Schachspielen oder das Lösen von Differentialgleichungen viel leichter durch Computer realisiert werden können als das visuelle Wahrnehmungssystem einer Katze oder das Sprachvermögen eines sechsjährigen Kindes) samt der gesellschaftlichen und kulturellen Veränderungen, und auch Gedankenexperimente wie der Turing-Test und das Chinesisch-Zimmer, bereichern dabei den philosophischen und den interdisziplinären Diskurs. Die KI kann hier auch von der Philosophie lernen, wie umgekehrt auch Erfolge und Mißerfolge der KI als „empirische Tests“ philosophischer Thesen gelten können.²¹⁸

²¹⁷Für eine kritische Gegenüberstellung symbolischer und konnektionistischer „Modelle der menschlichen Informationsverarbeitung“ vgl. Gerhard HELM (1991). Abgesehen davon, welches Modell beim *Menschen* angemessener ist, wird zunehmend auch die zentrale Grundannahme der „klassischen“ KI bestritten, daß Computer „wirklich“, d. h. unabhängig von der Zuschreibung durch ihre Nutzer, *symbolverarbeitende*, syntaktische Maschinen sind. Vgl. dazu oben S. 71 die Einwände von SEARLE. Auch Winfried D’AVIS (1994, S. 327) kommt „zu dem Schluß, daß schon der Ausdruck ‚symbolverarbeitende Maschine‘ eine unzulässige Kennzeichnung ist, denn die Maschine verarbeitet keine Symbole, sondern letztlich nur durch *physikalische Informationsträger* ausgelöste elektrische Signale, so daß der Ausdruck ‚symbolvernichtende Maschine‘ dem, was sie über die verschiedenen Repräsentationsebenen wirklich tut, ein Stück näher kommt.“

²¹⁸Nur drei Beispiele:

„Das Ergebnis lautet, daß die klassische Formulierung der Einheitswissenschaft dem Erfolg der Computersimulation preisgegeben ist. . . . Die Einheitswissenschaft war als eine empirische Hypothese intendiert, die durch mögliche wissenschaftliche Befunde widerlegbar ist. Niemand dachte jedoch daran, daß sie durch Newell, Shaw und Simon widerlegt werden sollte.“ (FODOR (1992), S. 144 f.)

„Was das Problem der Bedeutung angeht, so konnten wir in seinem Kontext das Verhältnis von Philosophie und KI dahingehend klären, daß sich die KI – sei es im positiven Sinne des Aufgreifens, sei es im negativen Sinne des Vermeidens von Fehlern – von philosophischen Lösungsversuchen profitieren kann wie sie auch umgekehrt einen empirischen Prüfstein für philosophische Bedeutungstheorien darzustellen vermag.“

Die „starke KI“ sagt, daß auch wir Computer sind (und wer das leugnet, denkt vorwissenschaftlich). Wir „Humanisten“ lehnen das natürlich ab. Warum? Weil wir evidentermaßen anders sind als Computer: wir sind nicht „mechanisch“, wir haben einen freien Willen, setzen uns Zwecke und übernehmen Verantwortung, wir sind kreativ, haben Überzeugungen, Wünsche und Emotionen. Wenn das „ComputermodeLL des Geistes“ Recht hat, wird dies alles durch eine „Idiotenarmee“ erledigt, durch ein komplexes Ensemble von im einzelnen ausgesprochen anspruchslosen Prozessen, zu denen wir introspektiv keinen Zugang haben. Was ist eigentlich so schrecklich an dieser Vorstellung, was wäre so schlimm daran, wenn sich herausstellen würde, daß Wahrnehmungen, Schlußfolgerungen, das Erfinden einer neuen Melodie, das bewußte Ablehnen einer Zigarette und anderes, auf das wir so stolz sind, auf diese Weise erfolgt? Wir haben uns an den Gedanken einer engen naturgeschichtlichen Verwandtschaft mit den Tieren gewöhnt und daran, daß unserem Denken physiologische Prozesse zugrundeliegen; einige „Humanisten“ sind ausgesprochen verliebt in den Gedanken, daß wir eine durch frühkindliche Traumata mehr oder weniger deformierte Hydraulikanlage zum Ausgleich von *Eros* und *Thanatos* sind; wir haben uns daran gewöhnt, bei verhältnismäßig einfachen Rechnungen mehr einem Gerät für zwanzig Mark als unserer langjährigen Schulbildung zu vertrauen (und jene, die das als Kulturverlust beklagen, lassen sich von ihren Kindern zeigen, wie man ein modernes Autoradio bedient); und dennoch streiten wir mit Recht und dessen ungeachtet darüber, ob eine neue Asylregelung die Menschenwürde verletzt und gegen den Geist des Grundgesetzes verstößt. Wie gesagt, was *genau* ist so schlimm an dieser Vorstellung? Auf der anderen Seite sehen wir Menschen, die jahrelang mit Menschen und an Computern gearbeitet haben, Kognitions- und Intelligenzprozesse erforschen und im vollen Ernst vorschlagen, die Entscheidung, ob eine Interkontinentalrakete mit atomaren Mehrfachsprengköpfen abgefeuert werden soll, einem Computersystem zu übertragen, weil der Mensch zu langsam und zu fehlbar, zu emotional und zu unberechenbar, also der schlechtere Computer sei.

Wie kommt es zu diesem tiefgreifenden Aneinandervorbeireden, zu der Vorstellung einerseits, daß unser Verständnis von Freiheit und Würde die Aufklärung mentaler Prozesse in dieser Richtung von vornherein verbietet, und zu der Vorstellung andererseits, daß das Abweichen von selbst gesetzten formalen Standards unser Abtreten zugunsten von Maschinen gebietet? Wie kommt es zu dieser heillosen Verwirrung von Erklärung und Rechtfertigung und warum spielt gerade der Computer, ein Gerät, an dem ich diese Zeilen schreibe und das dafür sehr nützlich ist, dabei eine so große Rolle?²¹⁹

Die eingangs genannten zwei Fraktionen (vgl. oben S. 3) haben eine je eigene Sichtweise

(KESE (1988), S. 64.)

„So wie das Scheitern der maschinellen Übersetzung bewies, daß Sprachkompetenz Verstehen voraussetzt, verdanken wir dem Mikrowelt-Ansatz vielleicht den Nachweis, daß die Welt nicht in unabhängige Fragmente zerlegt werden kann. [...] Forscher, die sich mit Mikro-Welten beschäftigten, waren jedoch nicht die ersten, die darauf hinwiesen. Philosophen wie Husserl, Heidegger, Dewey und dem späteren Wittgenstein war dieses Problem wohlbekannt, auch wenn es nicht empirisch bewiesen war.“ (HAUGELAND (1987), S. 170 und 227.)

²¹⁹ „Diese vierhundert Jahre alte Vorstellung, wir könnten mehr darüber erfahren, was wir glauben sollen, indem wir eingehender erforschen, wie wir funktionieren, kann als ein Gedanke eingesehen werden, der so irrig ist wie die Vorstellung, wir könnten in Erfahrung bringen, ob wir Robotern bürgerliche Rechte gewähren sollen, indem wir ihre Funktionsweise besser erforschten. Man wird dann die Analogie zwischen Mensch und Maschine nicht bloß als Quelle brauchbarer Modelle für Organismen ansehen, sondern auch als eine Eselsbrücke, die uns den Unterschied zwischen dem Menschen als einem Gegenstand der Erklärung und dem Menschen als einem moralischen Agenten, der seine Handlungen und Meinungen rechtfertigt, nicht vergessen läßt.“ (RORTY (1981), S. 281.)

mit jeweils gewisser Plausibilität (und je eigenen blinden Flecken). Für die einen ist der Mensch, auch seine Kultur und Technik, Produkt und Teil der Natur. Das Denken (als Hirntätigkeit und objektives Verhalten) ist wohl komplexer, aber nicht grundsätzlich anders als andere Naturvorgänge, es hat nichts Magisches oder Geheimnisvolles, sondern läßt sich mit (natur-) wissenschaftlichen Mitteln erforschen und erklären (und was ich verstanden habe, kann ich auch nachbauen). Für die anderen sind die Gesetze und Theorien der Naturwissenschaften kulturelle Leistungen, sprachlich vermittelt und abhängig von unseren Zwecken und Möglichkeiten, „Natur“ also ein Teilbereich der menschlichen Kultur und Technik; und Computer gibt es entsprechend nur deshalb, weil es Nutzer, Konstrukteure und Programmierer gibt, sie sind das Mittelstück und Medium eines Prozesses, der beim Menschen beginnt und wieder endet. Somit wäre es einfach absurd, ihnen eine „eigene“ Denkfähigkeit zuzuschreiben (vgl. den Text von Walter SCHULZ im Unterrichtsmodul 1, Abschnitt 4).

Ich denke, daß beide Seiten, wir alle, noch nicht wirklich verstanden haben, worum es hier geht. Um nocheinmal ganz von vorne anzufangen: DESCARTES hat die Menschen vom Rest der Welt, von den Tieren, Pflanzen und Dingen, durch das Innen und Außen von Bewußtsein, durch zwei völlig verschiedene Substanzen geschieden. BACON hat die Naturbetrachtung für gänzlich unfruchtbar erklärt, wenn sie nicht dem Wohl der Menschen dient, und diese Macht und Herrschaft über die Natur sei nur möglich, wenn man ihre Gesetze erforscht und sich ihnen unterwirft. KANT hat gesagt, daß die Moral, als Reich der Freiheit, dort beginnt, wo die Natur, als Reich der Notwendigkeit, aufhört. Wir Menschen haben lange mit dieser strikten Trennung in Subjekt und Objekt, in Geist und Materie, in Kultur und Natur, in Kunst und Mechanik, in Ethik und Wissenschaft, in Gründe und Zwecke und in Ursachen und Wirkungen, in Verstehen und Erklären, in Intentionalität und Kausalität, in Handeln und Verhalten, gut gelebt. Aber seit einiger Zeit beginnen die Ränder und Grenzen zu zerfransen, die scharfen Trennungslinien zu verschwimmen.²²⁰

Die Aufklärung hat uns die Gewißheit unserer Verantwortung, die Menschenrechte und den Abbau der Furcht vor den unheimlichen Kräften der Natur gebracht. Aber, auch das gehört zur Dialektik der Aufklärung, die zur Naturbeherrschung entwickelten Mittel – Mathematik, Naturwissenschaften, Logik – lassen sich auch auf den Menschen anwenden, die häufig szientistisch verkürzte, instrumentelle Vernunft sieht zunehmend weniger Grund, sich an die von DESCARTES, BACON und KANT gezogenen Grenzen zu halten. Auch hat uns die Aufklärung nicht davor bewahrt, in vielen Bereichen – z. B. bei der Fließbandarbeit²²¹ – selbst immer methodischer, regulierter und „maschinenhafter“ zu werden. Schließlich werden auch die Grenzen zwischen den „Zwei Kulturen“ der Natur- und Geisteswissenschaften immer durchlässiger: von der Evolution bis zur Synergetik, von der Quantentheorie bis zur Autopoiese, werden Begriffe und Konzepte aus dem einen Bereich in den anderen übernommen und häufig auch wieder, empirisch und begrifflich angereichert, zurückgegeben.

²²⁰Nach Hilary PUTNAM (1982, S. 9) geht es heute darum, „den Würgegriff“ solcher „Dichotomien zu durchbrechen,“ und Donald DAVIDSON (1993, S. 91) schreibt: „nach meinem Eindruck läuft der am meisten versprechende und interessanteste Wandel, der sich heute in der Philosophie abspielt, darauf hinaus, daß diese Dualismen auf neue Weise in Frage gestellt oder gründlich umgestaltet werden. [...] Was uns bevorsteht, ist der Blick auf das Auftauchen einer radikal umgemodelten Auffassung der Beziehung zwischen Geist und Welt.“

²²¹„Ein anderer klassischer Ort ‚maschinenhaften‘ Handeln ist das Militär (und die Tanzrevue).“ (Bettina HEINTZ (1993), S. 295.)

Prominentestes und epochentypisches Beispiel für diese Entwicklung ist allerdings der Computer als zur Zeit vielversprechendstes Mittel zur Naturbeherrschung und zur Verwaltung der Menschen. Der programmierbare Computer ist eine Universalmaschine, aber nicht wirklich „mechanisch“ (jedenfalls nicht in der Hinsicht, die ihn für uns interessant macht). Hier geht es nicht um Kategorien wie Kraft, Druck oder Impuls, es geht um *Symbolmanipulation*, verarbeitet und erzeugt wird *Information* (und keiner weiß genau, was dieser „Stoff“ – weder Materie noch Energie, wie Norbert WIENER sagte – eigentlich ist: „eine Art modernen Phlogistons“²²²), diese Maschine liegt quer zu den vertrauten Dualismen von Geist und Körper und von Form und Inhalt.

Durch diese Verobjektivierung des Subjektiven, durch die maschinelle Realisierung mentaler Prozesse, ist eine neue Herausforderung für die Selbstbesinnung darüber entstanden, was den Menschen zum Menschen macht. Der Versuch, intelligente Maschinen zu entwickeln, zwingt zum genauen Nachdenken über den Unterschied von Daten, Information, Wissen und Sinn, von Syntax, Semantik, Pragmatik, vom Folgen, Finden, Wählen, Ignorieren, Erfinden einer Regel, von Verstand, Urteilskraft, Vernunft, zum genauen Nachdenken über die Rolle und Bedeutung unseres biologischen und kulturellen Erbes, der Sprachgemeinschaft und der Lebenspraxis (und Erfolg und Mißerfolg bei dem Versuch, intelligente Maschinen zu entwickeln, ist auch ein Test dafür, wie gut das Nachdenken ist; und so wie ab der Antike der Animismus überwunden und begonnen wurde, eine strikte Trennung der subjektiven und der objektiven Welt zu etablieren, so ist allein die Tatsache, daß dieser Versuch jetzt unternommen wird, ein Vorschein davon, daß diese strikte Trennung nun „aufgehoben“ wird und ein „neues Zeitalter“ beginnt). Intentionalität und Verstehen, Denken, Bewußtsein, Rationalität und Willensfreiheit, Sinn, Bedeutung und Wahrheit sind genuin philosophische Themen. Aber, auch hier gilt der Satz von KANT, „Gedanken ohne Inhalt sind leer, Anschauungen ohne Begriffe blind“. Ein unverkrampfter, interessierter, aber auch kritischer Blick auf die empirischen Wissenschaften, die sich mit diesen Themen beschäftigen, dient – wie schon bei KOPERNIKUS, DARWIN und FREUD – auch den philosophischen Interessen und der Selbstaufklärung der Menschen.

SEARLES Gedankenexperiment vom „Chinesisch-Zimmer“ aus der Innenperspektive des Turing-Tests lenkt die Aufmerksamkeit auf die ursprüngliche Intentionalität und damit auf das Sprachverstehen, das Erfassen der *Bedeutung* sprachlicher Ausdrücke als wesentlichen Prüfstein für die Frage, ob Computer denken können. Mit dem Sprachverstehen beim *Menschen* befassen sich neben der Philosophie und KI z.B. auch die Linguistik und Psychologie, also Wissenschaften mit ganz unterschiedlichen Ansätzen und Methoden, Fachsichten und Grenzen. Das Projekt des maschinellen Sprachverstehens erfordert die strikte Formalisierung wesentlicher Teile der Semantik – und die interdisziplinäre Zusammenschau und -arbeit der genannten Disziplinen. Dazu mehr im folgenden Kapitel.

²²²So Francisco J. VARELA (1990), S. 18.

Literaturhinweise

Darstellungen der **Geschichte der Philosophie** gibt es viele. Über eine bloße Aufzählung von Daten und Werkinhalten hinaus zeichnet Christoph HELFERICH (1985) ein lebendiges und spannend zu lesendes (auch durchaus persönlich gefärbtes) Bild unter Einbeziehung des jeweiligen politischen, sozialen und wissenschaftlichen Umfeldes. All das gilt auch für das Buch von Bertrand RUSSELL (1975), darüber hinaus bietet es eine interessante Erweiterung unserer häufig ein wenig eingeschränkten kontinentalen und deutschen Perspektive. Über die neuere angelsächsische Philosophie des Geistes von RYLE bis zur Gegenwart gibt Peter LANZ (1993) einen kompakten Überblick (in dem Sammelband sind eine Reihe weiterer einschlägiger Aufsätze enthalten, z. B. zu FREGE, RUSSELL, WITTGENSTEIN und CARNAP). Der von Peter BIERI (1993) herausgegebene Band versammelt zu den Themen Materialismus, Intentionalität, Selbstbewußtsein und Person Aufsätze, die für die neuere Diskussion in der analytischen Philosophie des Geistes grundlegend sind (mit ausführlichen Einleitungen und Literaturhinweisen).

Einen Bogen vom aristotelischen Begriff der Seele über den cartesischen Dualismus bis zur KI als epochale Auffassungen vom „Geist“ schlägt Ansgar BECKERMANN (1989, 1994). Einen knappen und doch umfassenden Überblick über die **Bewußtseinsphilosophie** und ihren Zusammenhang zu anderen inner- und außerphilosophischen Disziplinen gibt Heiner HASTEDT (1991b). Ein „Prolegomenon zu einer Geschichte der erkenntnistheoretisch orientierten Philosophie als einer Episode der europäischen Kulturgeschichte“ bietet Richard RORTY (1981) mit seinem provozierenden und viel diskutierten Buch an. Er plädiert für eine Verabschiedung der auf mentale Repräsentationen, auf unsere „gläserne Natur“ aufbauenden Erkenntnistheorie (samt ihrer Nachfolger in der analytischen Philosophie) als einer auf besondere Gewißheit gerichtete, anderen Wissenschaften vor- oder übergeordnete spezielle Disziplin, für eine vom Begriff des Bewußtseins abgekoppelte Thematisierung von Vernunft und Personalität im Rahmen einer „bildenden Philosophie“: „Erst wenn wir Bewußtsein und Vernunft auf diese Weise trennen, können wir die Personalität als das sehen, was sie meiner Meinung nach ist — eine Funktion unserer Entscheidungen, nicht unserer Erkenntnis, die Anerkennung eines anderen Wesens als einer Gemeinschaft zugehörig, nicht das Bewußtsein einer gemeinsamen Natur“ (S. 50).

Wir alle gebrauchen dieses Wort ständig, bei näherem Nachdenken erweist es sich aber als recht „flüchtiger“ Begriff: die vielen Facetten und Verwicklungen des Nachdenkens über das **Ich** und das Selbstbewußtsein sind Thema des sehr unterhaltsamen, von Douglas R. HOFSTADTER & Daniel C. DENNETT (1986) herausgegebenen und kommentierten Sammelbandes. Zum gleichen Thema, aber mit einem ganz anderen Ansatz plädieren Francisco VARELA & Evan THOMPSON (1992) für einen „mittleren Weg der Erkenntnis“ im Dialog zwischen Kognitionswissenschaft und Meditationserfahrungen. Warum trotz einiger empirischer Fortschritte das **Bewußtsein** für uns ein Rätsel bleibt, verdeutlicht Peter BIERI (1992); daß sich daran so bald wohl auch nicht allzuviel ändern wird, zeigt der Bericht von John HORGAN (1994) über eine Tagung zum Thema „Towards a Scientific Basis for Consciousness“. Julian JAYNES (1993) hat eine hochinteressante und vieldiskutierte These zur Phylogenese des Bewußtseins. Sehr ausführlich und im besten Sinne populär behandelt Tor NØRRETANDERS (1994) die wissenschaftlichen Grundlagen für die Erforschung des menschlichen Bewußtseins, mit Rekurs nicht nur auf die Psychologie, sondern auch auf die Informationstheorie und die Thermodynamik. Längere Zeit eher verpönt, ist nun, wie die Beiträge in den von Sybille KRÄMER (1996) und Thomas METZINGER (1996) herausge-

gebenen Bänden zeigen, ist nun das Bewußtsein auch wieder ein wichtiges Thema für die Philosophie. Die beiden ausgezeichneten Untersuchungen zum **Leib-Seele-Problem** von Martin CARRIER & Jürgen MITTELSTRASS (1989) und Heiner HASTEDT (1988) rezipieren den neuesten Stand der (meist angelsächsischen) Diskussion und berücksichtigen auch die aktuellen Ergebnisse der beteiligten Fachwissenschaften (Psychologie und Neurobiologie), ohne einem eindimensionalen, szientistischen Reduktionismus nachzugeben.

Die ideengeschichtliche Bedeutung mechanischer **Automaten** als Abbild des Lebendigen untersucht Alex SUTTER (1988), J. David BOLTER (1990) ergänzt diese Betrachtung zeitprägender Techniken und ihrer symbolischen Bedeutung mit einem Überblick von der Antike (kurz) bis zum **Computer** (ausführlich). Das Maschinenhafte in uns und die Rolle des Computers bei der Vermenschlichung der Maschinenwelt sowie die Rückwirkungen auf die Selbstdefinition des Menschen stehen auch im Mittelpunkt der Studien von Arno BAMMÉ, Günter FEUERSTEIN, Renate GENTH, Eggert HOLLING, Renate KAHLE & Peter KEMPIN (1983), HOLLING & KEMPIN (1989) und Uwe HERITH (1989) (und sie gehören zu den wenigen Arbeiten, die dabei das Werk von Gotthard GÜNTHER zur Kenntnis nehmen). Eine hervorragende Darstellung der geschichtlichen Entwicklung der Idee der **Formalisierung**, der „symbolischen Maschine“ im Sinne des LEIBNIZ-Programmes, von den frühen Zahlzeichen bis zu den Präzisionen des Algorithmusbegriffes gibt Sybille KRÄMER (1988). Weiter zeichnet sie die Genese solcher Hilfsmittel in methodologischer Absicht hin zur empirischen These der „computational theory of mind“ nach (1991, 1994a, 1994b) und zeigt die mit der Reduktion von Geist auf Kognition und Komputation verbundene „Dissoziation von kognitiven und moralischen Kompetenzen“ auf (1992, 1993, 1994a).

Die historische Kontinuität und die Übernahme grundlegender Annahmen des **Logischen Empirismus** in der KI werden öfter kritisch vermerkt (so z. B. von Wolfgang COY (1988a), Gerhard HEYER (1988), Alfred Lothar LUFT (1988), Heinz ZEMANEK (1993)), Ralf KESE (1988) zeigt diese Verbindung (namentlich zu CARNAP) im Detail auf und plädiert ausdrücklich für eine Fortsetzung dieses Programmes mit den Mitteln der KI.

Der klassische Aufsatz zur **Turingmaschine** stammt selbstverständlich von ihrem Erfinder Alan TURING (1987a) selbst, populärere Darstellungen gibt es sehr viele, gut verständlich und informativ sind neben anderen die von John HAUGELAND (1987, S. 116 ff.) und Roger PENROSE (1991, S. 32 ff.). Bettina HEINTZ (1993) ordnet TURINGs Arbeit in das formalistische Programm HILBERTS und beides in den weiteren Rahmen der sozialen und kulturellen Aspekte von moderner Mathematik, Mechanisierung und Rationalisierung ein. John L. CASTI (1998) läßt in einem fiktiven *Dinner*-Gespräch TURING mit C. P. SNOW, J. B. S. HALDANE, Erwin SCHRÖDINGER und Ludwig WITTEGENSTEIN über seine Maschine und Thesen diskutieren.

Als Gründungsdokument des **Funktionalismus** gilt der zuerst 1960 erschienene Aufsatz von Hilary PUTNAM (1985a). Eine allgemeinverständliche Einführung in den Funktionalismus und einen Vergleich mit anderen Theorien des Geistes (Behaviorismus, Identitätstheorie) bietet Jerry A. FODOR (1981). Der Beitrag des Funktionalismus zur Lösung des Leib-Seele-Problems wird kritisch von Heiner HASTEDT (1988) und Martin CARRIER & Jürgen MITTELSTRASS (1989) gewürdigt, zur analytischen Handlungstheorie von Ansgar BECKERMANN (1985), zur Kognitionswissenschaft von Howard GARDNER (1989) und Dieter MÜNCH (1992), zur „Neuroepistemologie“ von Erhard OESER & Franz SEITELBERGER (1988), zur Psychologie des Sprachverstehens von Jürgen SCHRÖDER (1992) und zur Repräsentation sprachlichen Wissens von Andreas KEMMERLING (1988). Geert

KEIL (1993b) rekonstruiert die Grundannahmen des Funktionalismus in zwei unabhängige Teilthesen, nämlich daß erstens mentale Entitäten wie propositionale Einstellungen „abstrakte Gegenstände“ sind (wie sie in jeder wissenschaftlichen Theorie, auch in der Physik, gebraucht werden) und somit einen anderen ontologischen Status haben als konkrete physische Einzeldinge und daß zweitens „mentale Zustände und Prozesse auf vielfache Art und Weise und in verschiedener Hardware physisch realisiert sein können“, unterstützt die erste und verwirft die zweite. Eine grundlegende Kritik daran, daß sich der Funktionalismus tatsächlich zu einer umfassenden Theorie des Geistes entwickeln kann, haben Ned BLOCK (1992) und, besonders ausführlich und unter Berücksichtigung auch semantischer Probleme, Hilary PUTNAM (1991) vorgelegt.

SEARLES (zuerst (1980, deutsch 1992), eingebettet in weitere Überlegungen zur Philosophie des Geistes (1986), erweitert zur „Turnhalle“ (1990) als Reaktion auf neuere Entwicklungen in der KI) Gedankenexperiment vom **Chinesisch-Zimmer** hat heftige Diskussionen ausgelöst (vgl. die 27 Stellungnahmen in (1980), u. a. von BLOCK, DENNETT, FODOR, HAUGELAND, HOFSTADTER, MINSKY, PYLYSHYN und RORTY) und viele Einwände provoziert, meist Variationen des Roboter- (vgl. auch Ansgar BECKERMANN (1988, 1990), der die These vertritt, daß mit Wahrnehmungskomponenten zur Überprüfung der Wahrheitsbedingungen von Sätzen ausgestattete Computer tatsächlich im Wortsinne eine Sprache verstehen können) bzw. Systemeinwandes (vgl. auch Douglas R. HOFSTADTER & Daniel C. DENNETT (1986), S. 357–366, Dieter MÜNCH (1990), S. 371 ff.). Manfred BIERWISCH untersucht und kritisiert die Grundannahmen und Konsequenzen von SEARLES Argumentation aus der Sicht einer kognitiven (an CHOMSKY anknüpfenden) Linguistik. Margaret A. BODEN (1988, S. 238 ff., 1990, S. 89–104, 1992, S. 331 ff.) bestreitet, daß Computerprogramme nur eine Syntax und keine Semantik haben (denn das Programm bewirkt auch physisch etwas in der Maschine und an der Peripherie). Rosemarie RHEINWALD (1991) diskutiert diese Einwände und die dabei verwendeten Begriffe, betont die Relevanz von *Handlungsmöglichkeiten* und plädiert für eine Verwendung von „Denken“, „Sprache“ und „Verstehen“ als graduelle Begriffe. Paul M. CHURCHLAND & Patricia SMITH CHURCHLAND (1990) stimmen in Bezug auf symbolverarbeitende Maschinen in der herkömmlichen Architektur SEARLE zu, bestreiten aber die Gültigkeit seiner Argumente für subsymbolische Prozesse in massiv parallelverarbeitenden, der Architektur von Gehirnen nachempfundenen neuronalen Netzwerken (vgl. dazu ausführlicher Patricia S. CHURCHLAND & Terrence J. SEJNOWSKI (1997), Paul M. CHURCHLAND (1997) und Manfred SPITZER (1996)). Elmar HOLENSTEIN (1987) untersucht den Unterschied zwischen Mensch und Maschine am Beispiel der Intuition und kommt ähnlich wie SEARLE zu dem Schluß, daß nur dann Computer die gleichen Leistungen wie Menschen erbringen werden, „wenn sie aus derselben Bio- oder Neuromasse gemacht sind.“ Erich SCHÄFER (1994) würdigt ausführlich, mit einer Darstellung seiner Philosophie der Sprache und Intentionalität, SEARLES Argumentation und die dagegen erhobenen Einwände. Matthias TICHY (1987) weist auf die aus der Wissenschaftstheorie bekannte Tatsache hin, daß Gedankenexperimente (wie auch empirische Experimente) nicht schlicht für sich sprechen, sondern nur im größeren Rahmen der theoretischen Hintergrundsannahmen Thesen stützen oder widerlegen können; nach einer Thematisierung auch der unausgesprochenen Voraussetzungen von TURING und SEARLE sieht er neben allen verbleibenden Schwierigkeiten auch ein positives Ergebnis: die Kontroverse „hat sich als die Suche nach einem tragfähigen Intentionalitätsbegriff entpuppt.“

Sein Konzept der intrinsischen Intentionalität als nicht weiter reduzierbare Grundlage von Bedeutung und Verstehen hat SEARLE (1987) in einer ausführlichen Abhandlung vorge-

stellt. „**Intentionalität** und **Verstehen**“ ist auch der Titel eines vom Forum für Philosophie Bad Homburg (1990) herausgegebenen Sammelbandes, in dem es unter anderen um die Kontroverse der (genetischen oder logischen) Priorität von Intentionalität oder Kommunikation (und einer entsprechend je anderen Begründung der Semantik) geht (Karl-Otto APEL verteidigt hier SEARLE als Autor der *Sprechakte* gegen SEARLE als Autor von *Intentionalität*, dem er eine Rückkehr zur Bewußtseinsphilosophie — statt einer Weiterentwicklung des *linguistic turn* zu einem *pragmatic turn* — vorwirft). SEARLE (1993) ist diesen Weg weitergegangen mit einer Streitschrift gegen jede komputationale Theorie des Geistes – sei es in der symbolischen (Zenon W. PYLYSHYN (1980)), konnektionistischen (Paul M. CHURCHLAND (1988)) oder beides verbindenden (Daniel C. DENNETT (1994)) Spielart – und für die Untersuchung von **Bewußtsein** als dem zentralen geistigen Phänomen. Der letzte Satz des Buches lautet: „Wir müssen den sozialen Charakter des Geistes wiederentdecken.“ Daran knüpft SEARLE (1997) an: Wie konstruieren wir eine objektive gesellschaftliche Wirklichkeit?

Mit seinem „biologischen **Naturalismus**“ ist er allerdings wie die anderen zuletzt genannten Autoren Teilnehmer eines *Naturalistic turn* in der Philosophie des Geistes (der schon von CARNAP und QUINE eingeleitet wurde und die sprachanalytischen Bemühungen von WITTGENSTEIN und RYLE zunehmend ablöst), „naturalistische“ Theorien des Geistes, des Bewußtseins und der Subjektivität haben Konjunktur (vgl. z. B. Ray JACKENDOFF (1987), Thomas METZINGER (1993, 1994), Holm TETENS (1990, 1993, 1994); vgl. auch den Beitrag von Ansgar BECKERMANN (1992), der bestreitet, daß es bei intentionalen Zuständen überhaupt Naturalisierungsbedarf gibt, und die Kritiken dazu, die u. a. zeigen, wie schwer es hier ist, zwischen der *Sache* und der zu ihrer Beschreibung benutzten *Sprache* zu unterscheiden). Geert KEIL (1993a) unterzieht diese „naturalistische Revolution unseres Selbstverständnisses“ einer gründlichen Kritik und erinnert daran, daß es sich auch hierbei um *Diskurse* handelt, die ohne intentionales, Handlungen beschreibendes Vokabular nicht auskommen. Bernhard IRRGANG (1993) stellt diese Naturalisierungstendenz in der Philosophie in den weiteren Zusammenhang von Evolutionärer Erkenntnistheorie, Soziobiologie und Radikalem Konstruktivismus. Oswald SCHWEMMER (1994) skizziert die Konzeption einer „symbolischen Existenz des Geistes“, mit der, in „Verbindung mit den kulturellen Symbolismen“, „die isolationistischen Auffassungen naturalistischer“ und „idealistischer“ Auffassungen vom Geist überwunden und „als Teilaspekte einer übergreifenden, ‚interaktiven‘ Wirklichkeit des Geistes“ der philosophischen Reflexion zugänglich gemacht werden sollen.

Ekkehard MARTENS (1992) untersucht, wieweit die sokratisch-platonische Tradition in der KI wirklich „aufgehoben“ wurde, und verteidigt diese Tradition — entgegen einer häufig üblichen „rationalistischen“ Rezeption — gerade als kritisches Gegenstück zu einer Verkürzung der Vernunft durch den Anspruch der „Expertensysteme“. Hubert L. DREYFUS (1985, mit Stuart E. DREYFUS (1989), dort auch weitere Literatur) geht den Parallelen zwischen den Grundannahmen der KI und den (gescheiterten) Bemühungen von Edmund HUSSERL nach. Karl LEIDLMAIR (1991) konfrontiert die Grundannahmen der KI (sowohl die These der physischen Symbolsysteme als auch den Konnektionismus) mit HEIDEGGERS Technikkritik. Martin KURTHEN (1994) lotet die Möglichkeiten und Grenzen einer — mit einer „Vernutzung HEIDEGGERS“ auf die Krise des orthodoxen Symbolverarbeitungsansatzes reagierenden — „hermeneutischen Kognitionswissenschaft“ aus. Weitere und andere Plädoyers dafür, daß sich die Philosophie (und allgemeiner die Geistes- und Sozialwissenschaften) den **Herausforderungen der KI** stellt, haben Winfried D’AVIS (1988 und — in

einer interessanten Verbindung von Bedeutung und Zeitlichkeit als grundlegende Kriterien für „Denken“ — 1994), Ansgar BECKERMANN (1986), Bernhard IRRGANG (1992), Karl LEIDLMAIR (1989), Käte MEYER-DRAWE (1993), Klaus-Dieter SCHELLSCHMIDT (1988), Friedhelm SCHNEIDER & Joachim VAHLAND (1991), Otto ULRICH (1990), Hartmut von VOIGT (1988), Walter Ch. ZIMMERLI (1989) und (in einer Einleitung zu einer Sammlung „klassischer“ Texte) ZIMMERLI & Stefan WOLF (1994) vorgelegt.

Die vorliegende Schrift konzentriert sich mehr auf das Können der KI als auf das Sollen (in der schlichten Einsicht, daß Sollen Können voraussetzt). Allerdings — darauf hat als einer der ersten Joseph WEIZENBAUM (1977) hingewiesen — kann auch ein bloßer Glaube an die Leistungskraft der KI einschneidende Veränderungen bewirken. **Ethische Probleme** der Informatik und der KI werden im Teil 3 (Abschnitt „KI als interdisziplinäre Herausforderung“) genauer behandelt. Hier zunächst als Vorgriff der Hinweis auf einige Beiträge zu diesem Thema: Walter BAUMER (1992), Rafael CAPURRO (1987, 1992, 1993) Christiane FLOYD (1985, 1994), Heiner HASTEDT (1991a, S. 67 ff.), Sybille KRÄMER (1992), Hans LENK (1989, 1990), Bernd LUTTERBECK & Reinhard STRANSFELD (1992), Bernd MAHR (1992), Klaus MAINZER (1995, Abschn. 4.46), Ekkehard MARTENS (1988b, 1989 1993), Karl-Heinz RÖDIGER u. a. (1989), Peter SCHEFE (1992), Reinhard STRANSFELD (1992), Otto ULLRICH (1992), Otto ULRICH (1992), Godela UNSELD (1992), Rüdiger VALK (1987) und H. WEDEKIND (1987).

Literaturverzeichnis*

- AYER, A. J. (1972). *Sprache, Wahrheit und Logik*. Stuttgart: Reclam (Original: Language, Truth and Logic, London 1936)
- BAMMÉ, A., G. FEUERSTEIN, R. GENTH, E. HOLLING, R. KAHLE & P. KEMPIN (1983). *Maschinen-Menschen, Mensch-Maschinen: Grundrisse einer sozialen Beziehung*. Reinbek bei Hamburg: Rowohlt
- BAUMER, W. (1992). Sozialethische Folgenabschätzung der KI-Forschung und -Anwendung. In: J. Schmidt (Hrsg.), *Denken und denken lassen: Künstliche Intelligenz — Möglichkeiten, Folgen, Herausforderung* — (S. 237–247). Neuwied: Luchterhand
- BECKER, B. (1992). *Künstliche Intelligenz: Konzepte, Systeme, Verheißungen*. Frankfurt/Main, New York: Campus
- BECKERMANN, A. (1985). Handeln und Handlungserklärungen. In: ders. (Hrsg.), *Analytische Handlungstheorie, Band 2. Handlungserklärungen* (S. 7–84). Frankfurt am Main: Suhrkamp
- BECKERMANN, A. (1986). Kann die Künstliche Intelligenz-Forschung Fragen der Philosophie beantworten? In: H. Stoyan (Hrsg.), *GWAI – 85* (S. 2–25). Berlin/Heidelberg: Springer (Informatik-Fachberichte 118)
- BECKERMANN, A. (1988). Sprachverstehende Maschinen: Überlegungen zu John Searle's Thesen zur Künstlichen Intelligenz. *Erkenntnis* 28, S. 65–85
- BECKERMANN, A. (1989). Aristoteles, Descartes und die Beziehungen zwischen Philosophischer Psychologie und Künstlicher-Intelligenz-Forschung. In: E. Pöppel (Hrsg.), *Gehirn und Bewußtsein* (S. 105–123). Weinheim: VCH
- BECKERMANN, A. (1990). Semantische Maschinen. In: Forum für Philosophie Bad Homburg (Hrsg.), *Intentionalität und Verstehen* (S. 196–211). Frankfurt am Main: Suhrkamp
- BECKERMANN, A. U. A. (1992). Das Problem der Intentionalität — Naturalistische Lösung oder meßtheoretische Auflösung? *Ethik und Sozialwissenschaften* 3, Heft 4, S. 433–522 (mit Kritiken von 28 Autorinnen und Autoren, einer Replik und einer Metakritik von Holm Tetens)
- BECKERMANN, A. (1994). Der Computer — ein Modell des Geistes? In: S. Krämer (Hrsg.), *Geist — Gehirn — künstliche Intelligenz: zeitgenössische Modelle des Denkens, Ringvorlesung an der FU Berlin* (S. 71–87). Berlin, New York: de Gruyter
- BIERI, P. (1992). Was macht Bewußtsein zu einem Rätsel? *Spektrum der Wissenschaft*, 10/1992, S. 48–56
- BIERI, P. [HRSG.] (1993). *Analytische Philosophie des Geistes*. Bodenheim: Athenäum Hain Hanstein (2., verbesserte Auflage)
- BIERWISCH, M. (1990). Perspectives on mind, brain, and language: Linguistics as a cognitive science or Touring the Chinese room again. In: A. Burkhardt (Hrsg.), *Speech Acts, Meaning and Intentions: Critical Approaches to the Philosophy of John R. Searle* (S. 391–428). Berlin, New York: de Gruyter
- BLOCK, N. (1992). Schwierigkeiten mit dem Funktionalismus. In: D. Münch (Hrsg.), *Kognitionswissenschaft: Grundlagen, Probleme, Perspektiven* (S. 159–224). Frankfurt am Main: Suhrkamp (Original: Troubles with Funktionalism, in: C. W. Savage (Hrsg.), *Perception and Cognition*, University of Minnesota Press 1978, S. 261–325)

* „Klassiker“ wie z. B. HOBBS, DESCARTES, LOCKE, LEIBNIZ, HUME, KANT, FREGE, RUSSELL und WITTGENSTEIN werden im Text mit den üblichen Angaben zitiert und sind in diesem Literaturverzeichnis nicht eigens angeführt.

- BODEN, M. A. (1988). *Computer Models of Mind: Computational Approaches in Theoretical Psychology*. Cambridge University Press
- BODEN, M. A. [HRSG.] (1990). *The Philosophy of Artificial Intelligence*. Oxford: Oxford University Press
- BODEN, M. A. (1992). *Die Flügel des Geistes: Kreativität und künstliche Intelligenz*. München: Artemis & Winkler (Original: *The Creative Mind. Myths & Mechanisms*, London 1990: Weidenfeld and Nicolsen)
- BOLTER, J. D. (1990). *Der digitale Faust*. Stuttgart, München: Oktogon (Original: *Turing's Man. Western Culture In The Computer Age, 1984*, The University of North Carolina Press: Chapel Hill)
- BRÜGGEN, M. (1974). Wissen. In: H. Krings, H. M. Baumgartner & C. Wild (Hrsg.), *Handbuch philosophischer Grundbegriffe*, Band 6 (S. 1723–1739). München: Kösel
- CAPURRO, R. (1987). Zur Computerethik. Ethische Fragen der Informationsgesellschaft. In: H. Lenk & G. Ropohl (Hrsg.), *Technik und Ethik* (S. 259–273). Stuttgart: Reclam
- CAPURRO, R. (1992). Die Herausforderung der Informatik für die praktische Philosophie. In: W. Coy u. a. (Hrsg.), *Sichtweisen der Informatik* (S. 343–353). Braunschweig, Wiesbaden: Vieweg (Theorie der Informatik)
- CAPURRO, R. (1993). Zur Frage der professionellen Ethik. In: P. Schefe, H. Hastedt, Y. Dittrich, G. Keil (Hrsg.), *Informatik und Philosophie* (S. 121–140). Mannheim, Wien, Zürich: BI-Wissenschaftsverlag
- CARNAP, R. (1993). *Mein Weg in die Philosophie*. Stuttgart: Reclam (Original: *Rudolf Carnap: Intellectual Autobiography*. In: *The Philosophy of R. Carnap*. Ed. by P. A. Schilpp, La Salle/London 1963)
- CARRIER, M. & J. MITTELSTRASS (1989). *Geist, Gehirn, Verhalten: das Leib-Seele-Problem und die Philosophie der Psychologie*. Berlin, New York: de Gruyter
- CASTI, J. L. (1998). *Das Cambridge Quintett. Eine wissenschaftliche Spekulation*. Berlin: Berlin Verlag (Original: *The Cambridge Quintett. A Work of Scientific Speculation*, Reading, MA, 1998: Addison Wesley)
- CHOMSKY, N. (1973). *Sprache und Geist*. Frankfurt am Main: Suhrkamp (Original: *Language and Mind*, New York u.a. 1968)
- CHURCHLAND, P. M. (1988). *Matter and Consciousness: a Contemporary Introduction to the Philosophy of Mind*. Cambridge, Mass., London: MIT Press, Bradford (Revised Edition)
- CHURCHLAND, P. M. (1997). *Die Seelenmaschine: eine philosophische Reise ins Gehirn*. Heidelberg u. a.: Spektrum (Original: *The Engine of Reason, the Seat of the Soul*, Cambridge, Mass., 1995: MIT Press)
- CHURCHLAND, P. M. & P. SMITH CHURCHLAND (1990). Ist eine denkende Maschine möglich? *Spektrum der Wissenschaft*, 3/1990, S. 47–54
- CHURCHLAND, P. S. & T. J. SEJNOWSKI (1997). *Grundlagen zur Neuroinformatik und Neurobiologie*. Braunschweig, Wiesbaden: Vieweg (Original: *The Computational Brain*, Cambridge, Mass., 1992: MIT Press)
- COY, W. (1988a). ... und was jenseits der Grenze liegt, wird einfach Unsinn sein. *Sprache im technischen Zeitalter*, 105, S. 34–41
- COY, W. & L. BONSIEPEN (1989). *Erfahrung und Berechnung: Kritik der Expertensystemtechnik*. Berlin, Heidelberg, New York: Springer (Informatik-Fachberichte 229)
- DAVIDSON, D. (1993). *Der Mythos des Subjektiven. Philosophische Essays*. Stuttgart: Reclam

- D'AVIS, W. (1988). KI, ein Angriff der Computer auf den menschlichen Geist? *Ästhetik und Kommunikation*, Heft 69, S. 75–83
- D'AVIS, W. (1994). *Können Computer denken? Eine bedeutungs- und zeittheoretische Analyse von KI-Maschinen*. Frankfurt/Main, New York: Campus
- DENNETT, D. C. (1981). Intentionale Systeme. In: P. Bieri (Hrsg.), *Analytische Philosophie des Geistes* (S. 162–183). Königstein/Taunus: Hain (Original: *Intentional Systems*, *The Journal of Philosophy* 68 (1971), 87–106, Nachdruck in: Dennett, *Brainstorms*, Hassocks 1978)
- DENNETT, D. C. (1987). *The Intentional Stance*. Cambridge, Massachusetts: MIT Press
- DENNETT, D. C. (1992). Intentionale Systeme in der kognitiven Verhaltensforschung. In: D. Münch (Hrsg.), *Kognitionswissenschaft: Grundlagen, Probleme, Perspektiven* (S. 343–386). Frankfurt am Main: Suhrkamp (Original: *Intentional Systems in Cognitive Ethology. The Behavioral and Brain Sciences* 6 (1983), 343–355)
- DENNETT, D. C. (1994). *Philosophie des menschlichen Bewußtseins*. Hamburg: Hoffmann und Campe (Original: *Consciousness Explained*, Boston u. a. 1991: Little, Brown & Co.)
- DIEDERICH, J. (1991). Neue Trends im Konnektionismus. *KI*, 2/1991, S. 6–11
- DOTZLER, B. J. (1987). Die Revolution der Denkart und das Denken der Maschine: Kant und Turing. In: F. A. Kittler, M. Schneider & S. Weber (Hrsg.), *Diskursanalysen 1: Medien* (S. 150–163). Opladen: Westdeutscher Verlag
- DREYFUS, H. L. (1985). *Die Grenzen künstlicher Intelligenz: Was Computer nicht können*. Königstein/Ts.: Athenäum (Original: *What Computers can't do — The Limits of Artificial Intelligence*, New York 1979: Harper & Row)
- DREYFUS, H. L. (1993). Was Computer noch immer nicht können. *Deutsche Zeitschrift für Philosophie* (41), Heft 4, S. 653–680
- DREYFUS, H. L. & S. E. DREYFUS (1989). *Schöpfung des Geistes oder Modellierung des Gehirns? Künstliche Intelligenz am Scheideweg*. Klagenfurter Beiträge zur Technikdiskussion, Heft 23 (Original: *Making a Mind Versus Modeling the Brain: Artificial Intelligence Back at a Branchpoint*, *Artificial Intelligence* 117, Nr. 1 (Winter 1988). Nachdruck auch in: M. A. Boden (1990), S. 309–333)
- DUMMETT, M. (1988). *Ursprünge der analytischen Philosophie*. Frankfurt am Main: Suhrkamp
- FLOYD, C. (1985). Wo sind die Grenzen des verantwortbaren Computereinsatzes? *Informatik-Spektrum* 8, S. 3–6
- FLOYD, C. (1994). Künstliche Intelligenz — Verantwortungsvolles Handeln. In: S. Krämer (Hrsg.), *Geist — Gehirn — künstliche Intelligenz: zeitgenössische Modelle des Denkens, Ringvorlesung an der FU Berlin* (S. 256–278). Berlin, New York: de Gruyter
- FODOR, J. A. (1975). *The Language of Thought*. Cambridge, Mass.: Harvard University Press
- FODOR, J. A. (1980). Methodological solipsism considered as a research strategy in cognitive psychology. *The Behavioral and Brain Sciences* 3, S. 63–109 (mit Open Peer Commentary und Author's Response)
- FODOR, J. A. (1981). Das Leib-Seele-Problem. *Spektrum der Wissenschaft*, 3/1981, S. 27–37
- FODOR, J. A. (1983). *The Modularity of Mind*. Cambridge, Mass.: MIT Press (Seventh printing, 1991)
- FODOR, J. A. (1985). Erklärungen in der Psychologie. In: A. Beckermann (Hrsg.), *Analytische Handlungstheorie, Band 2. Handlungserklärungen* (S. 412–434). Frankfurt am Main: Suhrkamp (Original: *Explanations in Psychology*, in: M. Black (Hrsg.), *Philosophy in America*, Ithaca, NY, 1964, S. 161–179)

- FODOR, J. A. (1987). *Psychosemantics: the Problem of Meaning in the Philosophy of Mind*. Cambridge, Mass.: MIT Press
- FODOR, J. A. (1990). *A Theory of Content and other essays*. Cambridge, Mass.: MIT Press
- FODOR, J. A. (1992). Einzelwissenschaften. Oder: Eine Alternative zur Einheitswissenschaft als Arbeitshypothese. In: D. Münch (Hrsg.), *Kognitionswissenschaft: Grundlagen, Probleme, Perspektiven* (S. 134–158). Frankfurt am Main: Suhrkamp (Original: *Special Sciences (or The Disunity of Science as a Working Hypothesis)*, *Synthese*, 28 (1974), S. 97–115)
- FORUM FÜR PHILOSOPHIE BAD HOMBURG [HRSG.] (1990). *Intentionalität und Verstehen*. Frankfurt am Main: Suhrkamp
- FRANK, M. (1991). *Selbstbewußtsein und Selbsterkenntnis: Essays zur analytischen Philosophie der Subjektivität*. Stuttgart: Reclam
- FREUD, S. (1911). Formulierungen über die zwei Prinzipien des psychischen Geschehens. In: ders., *Studienausgabe Band III: Psychologie des Unbewußten* (S. 17–24). Frankfurt am Main 1982: Fischer
- GALTUNG, J. (1983). Struktur, Kultur und intellektueller Stil: Ein vergleichender Essay über sachsenische, teutonische, gallische und nipponische Wissenschaft. *Leviathan*, Nr. 3/1983, S. 303–338
- GARDNER, H. (1989). *Dem Denken auf der Spur: der Weg der Kognitionswissenschaft*. Stuttgart: Klett-Cotta (Original: *The Mind's New Science. A History of the Cognitive Revolution*, New York 1985: Basic Books)
- GOLECKI, R. (1991). *Naturverständnis und Umgang mit der Natur*. (Handreichung für das Fach Philosophie, hrsg. von der Freien und Hansestadt Hamburg, Amt für Schule, Referat S 13/33)
- HASTEDT, H. (1988). *Das Leib-Seele-Problem: zwischen Naturwissenschaft des Geistes und kultureller Eindimensionalität*. Frankfurt am Main: Suhrkamp
- HASTEDT, H. (1991a). *Aufklärung und Technik. Grundprobleme einer Ethik der Technik*. Frankfurt am Main: Suhrkamp
- HASTEDT, H. (1991b). Bewußtsein. In: E. Martens & H. Schnädelbach (Hrsg.), *Philosophie: ein Grundkurs*, Band 2 (S. 642–682). Reinbek bei Hamburg: Rowohlt (Überarbeitete und erweiterte Neuausgabe in zwei Bänden)
- HAUGELAND, J. (1987). *Künstliche Intelligenz — Programmierte Vernunft?* Hamburg: McGraw-Hill (Original: *Artificial Intelligence: The Very Idea*, 1985, MIT Press)
- HEGSELMANN, R. (1980). Zur Logik propositionaler Einstellungen. In: C. F. Gethmann (Hrsg.), *Theorie des wissenschaftlichen Argumentierens* (S. 213–237). Frankfurt am Main: Suhrkamp
- HEINTZ, B. (1993). *Die Herrschaft der Regel: zur Grundlagengeschichte des Computers*. Frankfurt/Main, New York: Campus
- HELFERICH, C. (1985). *Geschichte der Philosophie: von den Anfängen bis zur Gegenwart und Östliches Denken*. Stuttgart: Metzler
- HELM, G. (1991). *Symbolische und konnektionistische Modelle der menschlichen Informationsverarbeitung: eine kritische Gegenüberstellung*. Berlin, Heidelberg: Springer
- HERITH, U. (1989). *Hinter dem Rücken der Maschinen: über die Synkrisis von Autonomie und Kommando als Voraussetzung bereinigten Erkennens*. Gießen: Focus
- HEYER, G. (1988). Geist, Verstehen und Verantwortung. Philosophische Grundlagen der künstlichen Intelligenz. *KI*, 1/88, S. 36–40, 2/88, S. 24–27
- HODGES, A. (1989). *Alan Turing, Enigma*. Berlin: Kammerer & Unverzagt (Original: *Alan Turing: The Enigma*, London 1983: Burnett Books)

- HOFSTADTER, D. R. & D. C. DENNETT [HRSG.] (1986). *Einsicht ins Ich: Fantasien und Reflexionen über Selbst und Seele*. Stuttgart: Klett-Cotta (Zweite, unv. Auflage, Original: *The Mind's I*, New York 1981: Basic Books)
- HOLENSTEIN, E. (1987). Maschinelles Wissen und menschliches Bewußtsein. *Studia Philosophica* 46, S. 145–163
- HOLENSTEIN, E. (1992). Mentale Gebilde. In: D. Münch (Hrsg.), *Kognitionswissenschaft: Grundlagen, Probleme, Perspektiven* (S. 319–342). Frankfurt am Main: Suhrkamp
- HOLLING, E. & P. KEMPIN (1989). *Identität, Geist und Maschine: auf dem Weg zur technologischen Zivilisation*. Reinbek bei Hamburg: Rowohlt
- HORGAN, J. (1994). Ist das Bewußtsein erklärbar? *Spektrum der Wissenschaft*, 9/1994, S. 74–80
- IPTS [HRSG.] (1991). *Unterrichtsfach: Philosophie. Der Philosophieunterricht und die Computertechnologie. Hinweise und Materialien*. (Hrsg.: Landesinstitut Schleswig-Holstein für Praxis und Theorie der Schule (IPTS), 2575/90)
- IRRGANG, B. (1992). Die Maschinisierung des Subjektes und die rationale Konstruktion der Gesellschaft. In: J. Schmidt (Hrsg.), *Denken und denken lassen: Künstliche Intelligenz — Möglichkeiten, Folgen, Herausforderung* — (S. 115–154). Neuwied: Luchterhand
- IRRGANG, B. (1993). *Lehrbuch der Evolutionären Erkenntnistheorie: Evolution, Selbstorganisation, Kognition*. München, Basel: Reinhardt, UTB
- JACKENDOFF, R. (1987). *Consciousness and the Computational Mind*. Cambridge, MA und London: MIT Press, Bradford (Second printing, 1989)
- JAYNES, J. (1993). *Der Ursprung des Bewußtseins*. Reinbek bei Hamburg: Rowohlt (Original: *The Origin of Consciousness in the Breakdown of the Bicameral Mind*, Boston 1976: Houghton Mifflin Co.)
- JOHNSON-LAIRD, P. N. (1983). *Mental Models: towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge: Cambridge University Press (3. Auflage 1990)
- KEIL, G. (1993a). *Kritik des Naturalismus*. Berlin, New York: de Gruyter (Quellen und Studien zur Philosophie, Band 34)
- KEIL, G. (1993b). Die zwei Teilthesen des Funktionalismus. In: P. Scheffe, H. Hastedt, Y. Dittrich, G. Keil (Hrsg.), *Informatik und Philosophie* (S. 195–209). Mannheim u. a.: BI-Wissenschaftsverlag
- KEMKE, C. (1988). Der Neuere Konnektionismus. Ein Überblick. *Informatik Spektrum*, 11, S. 143–162
- KEMMERLING, A. (1988). Philosophischer Kognitivismus und die Repräsentation sprachlichen Wissens. In: G. Heyer, J. Krems & G. Görz (Hrsg.), *Wissensarten und ihre Darstellung: Beiträge aus Philosophie, Psychologie, Informatik und Linguistik* (S. 21–46). Berlin, Heidelberg, New York: Springer (Informatik-Fachberichte 169)
- KESE, R. (1988). Wissensrepräsentation, Bedeutung und Reduktionismus — Einige neopositivistische Wurzeln der KI. In: G. Heyer, J. Krems & G. Görz (Hrsg.), *Wissensarten und ihre Darstellung: Beiträge aus Philosophie, Psychologie, Informatik und Linguistik* (S. 47–66). Berlin, Heidelberg, New York: Springer (Informatik-Fachberichte 169)
- KRÄMER, S. (1988). *Symbolische Maschinen: Die Idee der Formalisierung in geschichtlichem Abriß*. Darmstadt: Wissenschaftliche Buchgesellschaft
- KRÄMER, S. (1991). Denken als Rechenprozedur: Zur Genese eines kognitionswissenschaftlichen Paradigmas. *Kognitionswissenschaft* 2, S. 1–10

- KRÄMER, S. (1992). Symbolische Maschinen, Computer und der Verlust des Ethischen im geistigen Tun: sieben Thesen. In: W. Coy u. a. (Hrsg.), *Sichtweisen der Informatik* (S. 335–341). Braunschweig, Wiesbaden: Vieweg (Theorie der Informatik)
- KRÄMER, S. (1993). Operative Schriften als Geistestechnik. Zur Vorgeschichte der Informatik. In: P. Scheffe, H. Hastedt, Y. Dittrich, G. Keil (Hrsg.), *Informatik und Philosophie* (S. 69–83). Mannheim u. a.: BI-Wissenschaftsverlag
- KRÄMER, S. (1994a). Geist ohne Bewußtsein? Über einen Wandel in den Theorien vom Geist. In: S. Krämer (Hrsg.), *Geist — Gehirn — künstliche Intelligenz: zeitgenössische Modelle des Denkens, Ringvorlesung an der FU Berlin* (S. 88–110). Berlin, New York: de Gruyter
- KRÄMER, S. (1994b). Künstliche Intelligenz — Der Aufstieg des Computers zum Modellbaukasten des Geistes. *Information Philosophie* 2/1994, S. 20–28
- KRÄMER, S. [HRSG.] (1996). *Bewußtsein: philosophische Beiträge*. Frankfurt/Main: Suhrkamp
- KUHN, T. S. (1973). *Die Struktur wissenschaftlicher Revolutionen*. Frankfurt am Main: Suhrkamp (Original: *The Structure of Scientific Revolutions*, University of Chicago Press, 1962)
- KURTHEN, M. (1994). *Hermeneutische Kognitionswissenschaft: Die Krise der Orthodoxie*. Bonn: Dje
- LAKATOS, I. (1974). Falsifikation und die Methodologie wissenschaftlicher Forschungsprogramme. In: I. Lakatos & A. Musgrave (Hrsg.), *Kritik und Erkenntnisfortschritt* (S. 89–189). Braunschweig: Vieweg (Original: *Falsification and the Methodology of Scientific Research Programmes*, in: dies., *Criticism and the Growth of Knowledge*, Cambridge University Press, London 1970, 91–195)
- LANZ, P. (1993). Vom Begriff des Geistes zur Neurophilosophie: das Leib/Seele-Problem in der angelsächsischen Philosophie des Geistes von 1949 bis 1987. In: A. Hügli & P. Lübcke (Hrsg.), *Philosophie im 20. Jahrhundert. Band 2: Wissenschaftstheorie und Analytische Philosophie* (S. 270–314). Reinbek bei Hamburg: Rowohlt
- LEIDLMAIR, K. (1989). KI - Eine Teildisziplin der Philosophie. In: J. Retti & K. Leidmaier (Hrsg.), *5. Österreichische Artificial-Intelligence-Tagung. Proceedings* (S. 313–324). Berlin, Heidelberg, New York: Springer (Informatik-Fachberichte 208)
- LEIDLMAIR, K. (1991). *Künstliche Intelligenz und Heidegger: Über den Zwiespalt von Natur und Geist*. München: W. Fink
- LENK, H. (1989). Können Informationssysteme moralisch verantwortlich sein? *Informatik-Spektrum* 12, S. 248–255
- LENK, H. (1990). Information zwischen Verantwortung und humaner Eigentätigkeit: Chancen und Probleme der neuen Informationstechnologien für Gesellschaft, Moral und Ausbildung. In: B. Irrgang & J. Klawitter (Hrsg.), *Künstliche Intelligenz* (S. 99–115). Stuttgart: Hirzel, Wissenschaftliche Verlagsgesellschaft
- LORENZ, K. (1990). *Einführung in die philosophische Anthropologie*. Darmstadt: Wissenschaftliche Buchgesellschaft (2., unveränderte Auflage 1992)
- LUFT, A. L. (1988). Was wissen Wissensingenieure? *Computer Magazin*, 4/1988, S. 12–15
- LUTTERBECK, B. & R. STRANSFELD (1992). Ethik in der Informatik — Vom Appell zum Handeln. In: W. Coy u. a. (Hrsg.), *Sichtweisen der Informatik* (S. 367–378). Braunschweig, Wiesbaden: Vieweg (Theorie der Informatik)
- MAHR, B. (1992). Zur Diskussion um die Verantwortung in der Informationstechnik. In: W. Coy u. a. (Hrsg.), *Sichtweisen der Informatik* (S. 355–360). Braunschweig, Wiesbaden: Vieweg (Theorie der Informatik)

- MAINZER, K. (1995). *Computer – Neue Flügel des Geistes? Die Evolution computergestützter Technik, Wissenschaft, Kultur und Philosophie*. Berlin, New York: de Gruyter
- MARTENS, E. (1988b). Rehabilitierung der Angewandten Philosophie? Zum Beispiel Computer-Ethik. *Zeitschrift für Didaktik der Philosophie* 4/1988, S. 204–210
- MARTENS, E. (1989). Das Subjekt der Computer-Ethik. In: M. Gatzemeier (Hrsg.), *Verantwortung in Wissenschaft und Technik* (S. 239–255). Mannheim u. a.: BI-Wissenschaftsverlag
- MARTENS, E. (1992). Sokratisch-platonische Tradition im „Expertensystem“. *Zeitschrift für philosophische Forschung*, Band 46, I, S. 57–75
- MARTENS, E. (1993). Computereethik. In: P. Schefe, H. Hastedt, Y. Dittrich, G. Keil (Hrsg.), *Informatik und Philosophie* (S. 141–154). Mannheim u. a.: BI-Wissenschaftsverlag
- METZINGER, T. (1993). *Subjekt und Selbstmodell: die Perspektivität phänomenalen Bewußtseins vor dem Hintergrund einer naturalistischen Theorie mentaler Repräsentation*. Paderborn u. a.: Schöningh
- METZINGER, T. (1994). Schimpansen, Spiegelbilder, Selbstmodelle und Subjekte: drei Hypothesen über den Zusammenhang zwischen mentaler Repräsentation und phänomenalem Bewußtsein. In: S. Krämer (Hrsg.), *Geist — Gehirn — künstliche Intelligenz: zeitgenössische Modelle des Denkens, Ringvorlesung an der FU Berlin* (S. 41–70). Berlin, New York: de Gruyter
- METZINGER, T. [HRSG.] (1996). *Bewußtsein: Beiträge aus der Gegenwartsphilosophie*. Paderborn u. a.: Schöningh (2., durchgesehene Auflage)
- MEYER-DRAWE, K. (1993). Und noch einmal: „Ob Computer denken können?“ *Deutsche Zeitschrift für Philosophie*, (41) H. 4, S. 681–691
- MINSKY, M. (1990). *Mentopolis*. Stuttgart: Klett-Cotta (Original: *The Society of Mind*, New York 1985: Simon & Schuster)
- MOHNHAUPT, M. & K. REHKÄMPER (1990). Gedanken zu einer neuen Theorie der Kognition. *Kognitionswissenschaft* 1, S. 36–45
- MÜNCH, D. (1990). Minds, brains and cognitive science. In: A. Burkhardt (Hrsg.), *Speech Acts, Meaning and Intentions: Critical Approaches to the Philosophy of John R. Searle* (S. 367–390). Berlin, New York: de Gruyter
- MÜNCH, D. (1992). Computermodelle des Geistes. In: D. Münch (Hrsg.), *Kognitionswissenschaft: Grundlagen, Probleme, Perspektiven* (S. 7–53). Frankfurt am Main: Suhrkamp
- NAGEL, T. (1981). Wie ist es, eine Fledermaus zu sein? In: P. Bieri (Hrsg.), *Analytische Philosophie des Geistes* (S. 261–275). Königstein/Taunus: Hain (Original: *What is it like to be a bat?* *Philosophical Review* 83 (1974), 435–50, Nachdruck in T. Nagel, *Mortal Questions*, Cambridge 1979)
- NEWELL, A. & H. SIMON (1992). Computerwissenschaft als empirische Forschung: Symbole und Lösungssuche. In: D. Münch (Hrsg.), *Kognitionswissenschaft: Grundlagen, Probleme, Perspektiven* (S. 54–91). Frankfurt am Main: Suhrkamp (Original: *Computer Science as Empirical Inquiry: Symbols and Search*, *Communications of the Association for Computing Machinery* 19 (1976), S. 113–126; Nachdruck auch in M. A. Boden (1990), S. 105–132)
- NØRRETRANDERS, T. (1994). *Spüre die Welt: die Wissenschaft des Bewußtseins*. Reinbek: Rowohlt (Original: *Mærk verden*, Kopenhagen 1991: Gyldendal)
- OESER, E. & F. SEITELBERGER (1988). *Gehirn, Bewußtsein und Erkenntnis*. Darmstadt: Wissenschaftliche Buchgesellschaft (Dimensionen der modernen Biologie, Bd. 2)

- PENROSE, R. (1991). *Computerdenken: Die Debatte um Künstliche Intelligenz, Bewußtsein und die Gesetze der Physik*. Heidelberg: Spektrum der Wissenschaft (Original: *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*, 1989, Oxford University Press)
- PFEIFER, R. & P. RADEMAKERS (1991). Situated Adaptive Design: Toward a New Methodology for Knowledge Systems Development. In: W. Brauer & D. Hernández (Hrsg.), *Verteilte Künstliche Intelligenz und kooperatives Arbeiten. 4. Int. GI-Kongreß Wissensbasierte Systeme. Proceedings* (S. 53–64). Berlin u.a.: Springer
- PUTNAM, H. (1975). *Mind, Language and Reality. Philosophical Papers, Volume 2*. Cambridge: Cambridge University Press
- PUTNAM, H. (1975a). Robots: machines or artificially created life? In: ders., *Mind, Language and Reality. Philosophical Papers, Volume 2* (S. 386–407). Cambridge: Cambridge University Press (Erstveröffentlichung in: *The Journal of Philosophy*, LXI (November 1964), S. 668–691)
- PUTNAM, H. (1975b). The mental life of some machines. In: ders., *Mind, Language and Reality. Philosophical Papers, Volume 2* (S. 408–428). Cambridge: Cambridge University Press (Erstveröffentlichung in: H. Castaneda (Hrsg.), *Intentionality, Minds and Perception*, Detroit 1967)
- PUTNAM, H. (1979). *Die Bedeutung von „Bedeutung“*. Frankfurt am Main: Klostermann (2., durchges. Aufl., Original: *The Meaning of „Meaning“*, in: K. Gunderson (Hrsg.), *Language, Mind, and Knowledge*, University of Minnesota Press, 1975)
- PUTNAM, H. (1981). Die Natur mentaler Zustände. In: P. Bieri (Hrsg.), *Analytische Philosophie des Geistes* (S. 123–135). Königstein/Taunus: Hain (Original: *Psychological predicates*, in: Capitan & Merrill (Hrsg.), *Art, Mind and Religion*, Pittsburgh 1967, Nachdruck als: *The nature of mental states*, in: Putnam (1975), S. 429–440)
- PUTNAM, H. (1982). *Vernunft, Wahrheit und Geschichte*. Frankfurt am Main: Suhrkamp (Original: *Reason, Truth and History*, Cambridge University Press 1981)
- PUTNAM, H. (1985a). Geist und Maschine. In: A. Beckermann (Hrsg.), *Analytische Handlungstheorie, Band 2. Handlungserklärungen* (S. 364–396). Frankfurt am Main: Suhrkamp (Original: *Minds and Machines*, in: S. Hook (Hrsg.), *Dimensions of Mind*, New York University Press, 1960, S. 138–164, Nachdruck in: Putnam (1975), S. 362–385)
- PUTNAM, H. (1985b). Reflexive Reflections. *Erkenntnis*, 22, S. 143–153
- PUTNAM, H. (1991). *Repräsentation und Realität*. Frankfurt am Main: Suhrkamp (Original: *Representation and Reality*, Massachusetts Institute of Technology, 1988)
- PYLYSHYN, Z. W. (1980). Computation and cognition: issues in the foundations of cognitive science. *The Behavioral and Brain Sciences* 3, S. 111–169 (mit Open Peer Commentary und Author's Response)
- RHEINWALD, R. (1991). Können Maschinen eine Sprache sprechen? Sind Computerprogramme syntaktisch oder semantisch? *Kognitionswissenschaft* 2, S. 37–49
- RÖDIGER, K.-H. U. A. (1989). Informatik und Verantwortung. *Informatik-Spektrum*, 12, S. 281–289 (auch in: W. Coy u. a. (Hrsg.), *Sichtweisen der Informatik*, Braunschweig 1992: Vieweg, S. 311–325)
- RORTY, R. (1981). *Der Spiegel der Natur: eine Kritik der Philosophie*. Frankfurt am Main: Suhrkamp (Original: *Philosophy and the Mirror of Nature*, 1979, Princeton University Press)
- RUSSELL, B. (1975). *Philosophie des Abendlandes: ihr Zusammenhang mit der politischen und der sozialen Entwicklung*. Wien: Europa (Lizenzausgabe der deutschen Übersetzung, Zürich 1950, Original: *A History of Western Philosophy*, London 1946)

- SCHÄFER, E. (1994). *Grenzen der Künstlichen Intelligenz: John R. Searles Philosophie des Geistes*. Stuttgart, Berlin, Köln: Kohlhammer (Münchener philosophische Studien, Neue Folge, Band 9)
- SCHANK, R. C. & P. G. CHILDERS (1986). *Die Zukunft der Künstlichen Intelligenz: Chancen und Risiken*. Köln: DuMont (Original: *The Cognitive Computer: on Language, Learning, and Artificial Intelligence*. Reading, MA 1984: Addison Wesley)
- SCHEFE, P. (1992). Theorie oder Aufklärung? Zum Problem einer ethischen Fundierung informatischen Handelns. In: W. Coy u. a. (Hrsg.), *Sichtweisen der Informatik* (S. 311–334). Braunschweig, Wiesbaden: Vieweg (Theorie der Informatik)
- SHELLENSCHMIDT, K.-D. (1988). Von existierenden Menschen und laufenden Programmen: ein philosophischer Blick auf die „Künstliche Intelligenz“. *Philosophie. Beiträge zur Unterrichtspraxis*, Heft 22, S. 58–69
- SCHNÄDELBACH, H. (1991). Philosophie. In: E. Martens & H. Schnädelbach (Hrsg.), *Philosophie: Ein Grundkurs*, Band 1 (S. 37–76). Reinbek bei Hamburg: Rowohlt (Überarbeitete und erweiterte Neuauflage in zwei Bänden)
- SCHNEIDER, F. & J. VAHLAND (1991). Die Herausforderung der Philosophie durch die neuen Kommunikationstechniken. In: IPTS (Hrsg.), *Unterrichtsfach: Philosophie. Der Philosophieunterricht und die Computertechnologie. Hinweise und Materialien* (S. 58–72). (Hrsg.: Landesinstitut Schleswig-Holstein für Praxis und Theorie der Schule (IPTS), 2575/90)
- SCHRÖDER, J. (1992). *Das Computermodell des Geistes in der analytischen Philosophie und in der kognitiven Psychologie des Sprachverstehens*. Würzburg: Königshausen & Neumann
- SCHULTE, J. & B. MCGUINNESS [HRSG.] (1992). *Einheitswissenschaft*. Frankfurt am Main: Suhrkamp (Mit einer Einleitung von Rainer Hegselmann: Einheitswissenschaft — das positive Paradigma des Logischen Empirismus)
- SCHWEMMER, O. (1994). Die symbolische Existenz des Geistes. In: S. Krämer (Hrsg.), *Geist — Gehirn — künstliche Intelligenz: zeitgenössische Modelle des Denkens, Ringvorlesung an der FU Berlin* (S. 3–40). Berlin, New York: de Gruyter
- SEARLE, J. R. (1980). Minds, brains, and programs. *The Behavioral and Brain Sciences* 3, S. 417–457 (mit Open Peer Commentary und Author's Response)
- SEARLE, J. R. (1986). *Geist, Hirn und Wissenschaft*. Frankfurt am Main: Suhrkamp (Original: *Minds, Brains and Science. The 1984 Reith Lectures*, BBC 1984)
- SEARLE, J. R. (1987). *Intentionalität: eine Abhandlung zur Philosophie des Geistes*. Frankfurt am Main: Suhrkamp (Original: *Intentionality. An essay in the philosophy of mind*, Cambridge University Press 1983)
- SEARLE, J. R. (1990). Ist der menschliche Geist ein Computerprogramm? *Spektrum der Wissenschaft*, 3/1990, S. 40–47
- SEARLE, J. R. (1992). Geist, Gehirn und Programme. In: D. Münch (Hrsg.), *Kognitionswissenschaft: Grundlagen, Probleme, Perspektiven* (S. 225–252). Frankfurt am Main: Suhrkamp (Original: *Minds, brains, and programs. The Behavioral and Brain Sciences* 3 (1980), 417–424)
- SEARLE, J. R. (1993). *Die Wiederentdeckung des Geistes*. München: Artemis & Winkler (Original: *The Rediscovery of the Mind*, Cambridge, Mass. 1992: MIT Press)
- SEARLE, J. R. (1997). *Die Konstruktion der gesellschaftlichen Wirklichkeit: zur Ontologie sozialer Tatsachen*. Reinbek: Rowohlt (Original: *The Construction of Social Reality*, New York 1995: The Free Press)
- SEIFFERT, H. (1992). *Einführung in die Wissenschaftstheorie*, Dritter Band. München: Beck (zweite, überarbeitete Auflage)

- SIMON, H. A. (1990). *Die Wissenschaften vom Künstlichen*. Berlin: Kammerer & Unverzagt (Original: *The Sciences of the Artificial*, Cambridge, MA, London 1981: MIT Press)
- SPAEMANN, R. & R. LÖW (1985). *Die Frage Wozu? Geschichte und Wiederentdeckung des teleologischen Denkens*. München: Piper
- SPITZER, M. (1996). *Geist im Netz: Modelle für Lernen, Denken und Handeln*. Heidelberg u. a.: Spektrum
- STEGMÜLLER, W. (1978). *Hauptströmungen der Gegenwartsphilosophie: eine kritische Einführung*, Band 1. Stuttgart: Kröner (Sechste Auflage)
- STEINVORTH, U. (1985). Ludwig Wittgenstein: Sprache und Denken. In: J. Speck (Hrsg.), *Grundprobleme der großen Philosophen, Philosophie der Gegenwart Band 1* (S. 100–154). Göttingen: Vandenhoeck & Ruprecht (dritte, teilw. überarbeitete Auflage)
- STENIUS, E. (1969). *Wittgensteins Traktat: eine kritische Darlegung seiner Hauptgedanken*. Frankfurt am Main: Suhrkamp (Original: *Wittgenstein's Tractatus*, Oxford 1960: Blackwell and Moss)
- STOYAN, H. (1988). Wissen wissensbasierte Programme etwas? In: G. Heyer, J. Krems & G. Görz (Hrsg.), *Wissensarten und ihre Darstellung: Beiträge aus Philosophie, Psychologie, Informatik und Linguistik* (S. 250–261). Berlin, Heidelberg, New York: Springer (Informatik-Fachberichte 169)
- STRANSFELD, R. (1992). Verantwortungslosigkeit. In: W. Coy u. a. (Hrsg.), *Sichtweisen der Informatik* (S. 361–366). Braunschweig, Wiesbaden: Vieweg (Theorie der Informatik)
- SUTTER, A. (1988). *Göttliche Maschinen: die Automaten für Lebendiges bei Descartes, Leibniz, La Mettrie und Kant*. Frankfurt am Main: Athenäum
- TETENS, H. (1990). Geist, Gehirn, Maschine: ein philosophischer Versuch über ihren Zusammenhang. *Wechselwirkung* Nr. 45/46, S. 34–39
- TETENS, H. (1993). Informatik und die Philosophie des Geistes. In: P. Scheffe, H. Hastedt, Y. Dittrich, G. Keil (Hrsg.), *Informatik und Philosophie* (S. 175–193). Mannheim u. a.: BI-Wissenschaftsverlag
- TETENS, H. (1994). *Geist, Gehirn, Maschine: philosophische Versuche über ihren Zusammenhang*. Stuttgart: Reclam
- TICHY, M. (1987). Auf der Suche nach dem Subjekt in der Maschine: zu John R. Searles Kritik am Computermodell des Denkens. *Zeitschrift für philosophische Forschung*, 3/1987, S. 429–443
- TURING, A. M. (1967). Kann eine Maschine denken? *Kursbuch* 8, S. 106–138 (Original: *Computing Machinery and Intelligence*, *Mind* 59, 1950)
- TURING, A. M. (1987a). Über berechenbare Zahlen mit einer Anwendung auf das Entscheidungsproblem. In: B. Dotzler & F. Kittler (Hrsg.), *Alan Turing: Intelligence Service* (S. 17–60). Berlin: Brinkmann & Bose (Original: *On Computable Numbers with an Application to the Entscheidungsproblem*, *Proceedings of the London Mathematical Society* (2) 42, 1937)
- TURING, A. M. (1987b). Rechenmaschinen und Intelligenz. In: B. Dotzler & F. Kittler (Hrsg.), *Alan Turing: Intelligence Service* (S. 147–182). Berlin: Brinkmann & Bose (Original: *Computing Machinery and Intelligence*, *Mind* 59, 1950)
- TURING, A. M. (1987d). Intelligente Maschinen. In: B. Dotzler & F. Kittler (Hrsg.), *Alan Turing: Intelligence Service* (S. 81–113). Berlin: Brinkmann & Bose (Original: *Intelligent Machinery*, in: B. Melzer/D. Michie (Hrsg.), *Machine Intelligence* 5, Edinburgh 1969)
- ULLRICH, O. (1992). Was spricht gegen Plastikhirne? Ursachen und Folgen der Maschinisierung des Lebendigen. In: J. Schmidt (Hrsg.), *Denken und denken lassen: Künstliche Intelligenz — Möglichkeiten, Folgen, Herausforderung* — (S. 167–202). Neuwied: Luchterhand

- ULRICH, O. (1990). Über Denken nachdenken! „Künstliche Intelligenz“ aus naturphilosophischer Sicht. In: B. Irrgang & J. Klawitter (Hrsg.), *Künstliche Intelligenz* (S. 131–144). Stuttgart: Hirzel, Wissenschaftliche Verlagsgesellschaft
- ULRICH, O. (1992). Entscheidungen aus der Retorte oder Orientierung am Lebendigen? „Künstliche Intelligenz“ — ein notwendiger „Umweg“ zu einem verantwortbaren Fortschritt. In: J. Schmidt (Hrsg.), *Denken und denken lassen: Künstliche Intelligenz — Möglichkeiten, Folgen, Herausforderung* — (S. 203–221). Neuwied: Luchterhand
- UNSELD, G. (1992). *Maschinenintelligenz oder Menschenphantasie? Ein Plädoyer für den Ausstieg aus unserer technisch-wissenschaftlichen Kultur*. Frankfurt am Main: Suhrkamp
- VALK, R. (1987). Der Computer als Herausforderung an die menschliche Rationalität. *Informatik-Spektrum* 10, S. 57–66
- VARELA, F. J. (1990). *Kognitionswissenschaft — Kognitionstechnik: eine Skizze aktueller Perspektiven*. Frankfurt am Main: Suhrkamp (Original: *Cognitive Science. A Cartography of Current Ideas*, 1988)
- VARELA, F. J. & E. THOMPSON (1992). *Der Mittlere Weg der Erkenntnis: die Beziehung von Ich und der Welt in der Kognitionswissenschaft — der Brückenschlag zwischen wissenschaftlicher Theorie und menschlicher Erfahrung*. Bern, München, Wien: Scherz (Original: *The Embodied Mind*, 1991, Massachusetts Institute of Technology)
- VOIGT, H. v. (1988). Die philosophischen Grundlagen Künstlicher Intelligenz. *Chip Professional: Programmieren* 3, S. 19–31 (Nachdruck in: IPTS (1991), S. 266–294)
- WAHLSTER, W. (1982). Natürlichsprachliche Systeme: eine Einführung in die sprachorientierte KI-Forschung. In: W. Bibel & J. H. Siekmann (Hrsg.), *Künstliche Intelligenz* (S. 203–283). Berlin/Heidelberg: Springer (Frühjahrsschule Teisendorf März 1982, Informatik-Fachberichte 59)
- WEDEKIND, H. (1987). Gibt es eine Ethik der Informatik? Zur Verantwortung des Informatikers. *Informatik-Spektrum* 10, S. 324–328
- WEIZENBAUM, J. (1977). *Die Macht der Computer und die Ohnmacht der Vernunft*. Frankfurt am Main: Suhrkamp (Original: *Computer Power and Human Reason. From Judgement to Calculation*, 1976, Freeman a. Co.)
- WINOGRAD, T. & F. FLORES (1989). *Erkenntnis Maschinen Verstehen: zur Neugestaltung von Computersystemen*. Berlin: Rotbuch (Original: *Understanding Computers and Cognition: A New Foundation for Design*, Norwood, NJ 1986: Ablex)
- ZEMANEK, H. (1993). Philosophische Wurzeln der Informatik im Wiener Kreis. In: P. Schefe, H. Hastedt, Y. Dittrich, G. Keil (Hrsg.), *Informatik und Philosophie* (S. 85–117). Mannheim u. a.: BI-Wissenschaftsverlag
- ZIMMERLI, W. C. (1989). Künstliche Intelligenz: die Herausforderung der Philosophie durch den Computer. *Themen aus der Wissenschaft* 1, S. 7–22, Universität Kaiserslautern (auch in: J. Schmidt (Hrsg.), *Denken und denken lassen: Künstliche Intelligenz — Möglichkeiten, Folgen, Herausforderung* —, Neuwied 1992: Luchterhand, S. 101–114)
- ZIMMERLI, W. C. & S. WOLF [HRSG.] (1994). *Künstliche Intelligenz: Philosophische Probleme*. Stuttgart: Reclam