

Virtual Reality als Element für Künstliche Intelligenz

Über einen Zusammenhang von Künstlicher Intelligenz und Virtual Reality für ein Forschungsprogramm „Artifizielle Person“

Abstract

The concept of reason has, since the beginning of the modern era, been increasingly formulated as calculation-based. The comparatively rapid initial success of research into artificial intelligence is as much rooted in this tendency as the present difficulties in this field. In the current situation it would appear sensible to apply the insight which has led 20th Century linguistic philosophy to the establishment of pragmatic theories of syntax and semantics. VR technology could well facilitate the transfer of action-theory contemplation onto artificial systems, and thus broaden the horizon of the artificial intelligence research programme.

Einleitung

Meine ersten Vorstellungen von Computern und Robotern erhielt ich vor ca. 25 Jahren durch den Magazinteil eines Comicheftes. Der abgebildete Roboter empfahl sich als fleißiger Haushaltsgehilfe und besaß selbstverständlich die Gestalt eines Menschen. Ich erinnere, daß es mir schon damals ein Greuel war, Computer nicht schlicht in Kästen zu verpacken und Roboter nicht anders zu konstruieren als wie dies bei Industrierobotern üblich ist: Roboter und Computer haben eben prinzipiell nichts miteinander zu tun.

Neuerdings glaube ich allerdings, daß ein wesentlicher Zusammenhang besteht zwischen „künstlich denken“ und „künstlich tun“.

Zusammenfassung

Der Vernunftbegriff wird seit der Neuzeit zunehmend als rechnende Vernunft formuliert. Die vergleichbar raschen Anfangserfolge der KI-Forschung beruhen ebenso auf dieser Anschauung wie ihre augenblicklichen Schwierigkeiten. In dieser Situation scheint es sinnvoll, eine Einsicht fruchtbar zu machen, welche die Sprachphilosophie des 20. Jahrhunderts zu einer pragmatischen Begründung von Syntax und Semantik geführt hat. Das in der VR-Forschung entwickelte Equipment könnte dabei die Anwendung handlungstheoretischer Überlegungen auf künstliche Systeme in der Praxis ermöglichen und ein erweitertes Forschungsprogramm für KI begründen.

Entwicklung der KI

Die Bemühungen um künstliche Intelligenz beginnen systematisch wie historisch mit den logischen Untersuchungen von Aristoteles. Pascal (1642), Leibniz (1672), Babbage (1835) und Boole (1847) waren weitere wichtige Stationen.¹

Nachdem es zunächst um die Übertragung von Aussagen in formale Systeme und die Schlußregeln innerhalb solcher Systeme ging, rückte zunehmend die technische Realisierung in den Vordergrund.

Als Additionsmaschine erfunden boten sich Computer für umfangreiche Informationsverarbeitung an. Verführt durch den raschen Erfolg der rechnenden Vernunft wurde deren Heuristik herangezogen, jeden Erkenntnisgewinn unter dem Mo-

dell der Informationsverarbeitung zu betrachten. Genügt die Abarbeitung der Algorithmen vernünftigen Gesichtspunkten, muß nur noch das Vorliegen der Ausgangsbedingungen gesichert werden. Der Mechanismus ist dann wahrheitsvererbend. Aus Sicht der KI-Forschung stellen sich damit zwei Problembereiche:

1. Wie kann Weltwissen repräsentiert werden?
2. Welche Prozesse entsprechen Erkenntnisgewinn?

Zur Lösung der ersten Aufgabe muß Wissen über die Welt diskret formulierbar und einer symbolverarbeitenden Maschine implementierbar sein.

Auf die Problematik des Kontextes hierbei möchte ich anschließend eingehen.

Die Formulierung der zweiten Frage ist deutlich schwächer als der Anspruch, mit der Simulation intelligenter Prozesse auch gleich eine Theorie des menschlichen Verstandes zu liefern. Zunächst ist nur an ein Ensemble von Regeln gedacht, etwa Prädikatenlogik, Naturgesetze, Erkennen von Gewinnstellungen etc., mit dem Ergebnisse kognitiver Prozesse technisch reproduziert werden können. Die Übereinstimmung von Computer und Mensch besteht darin, daß beide z.B. Knobelaufgaben richtig lösen. Ihre Ergebnisse - oder im Vorgriff Interaktionen - sind dann hinsichtlich Knobelaufgaben funktional gleich.

In der Regel wird von dieser Gleichheit auf eine Identität der Struktur geschlossen.

Diese Position - meist schwache KI-These genannt - wird nicht selten zu der Behauptung verstärkt, prinzipielle Wesensgleichheit zu verkünden. In Abstufung eröffnen sich damit drei Zielrichtungen:

- qualitative Identität
- Übereinstimmung der Struktur
- Entsprechung der Resultate

Zumindest die beiden letzten Alternativen dürfen als empirisch kontrollierbar gelten. Bevor die Befunde der Neurologie eine Beurteilung der mittleren Ebene aber nicht ermöglichen, sollte eine naturwissenschaftliche Erörterung zunächst nur die

dritte Version ins Auge fassen.

Wer sich an die erste Antwort wagt, begibt sich unweigerlich in das Terrain philosophischer Argumente. Da von Identität die Rede ist, wäre insbesondere die Überflüssigkeit von Termen wie Intelligenz, Geist und Bewußtsein zu erweisen, bzw. deren vollständiges mechanistisches Verständnis. Da sich dieser Strang am Ende leicht aufnehmen läßt, sollen zuvor die Schwierigkeiten aus der „internen Sicht“ der KI dargestellt werden.

KI und Welt

Ich werde in einem Raum mit einer Obstkiste aufgefordert mich zu setzen. Mit nur geringem Zögern werde ich die Kiste als Stuhl benutzen, obwohl ich sie bei Anwesenheit eines Designsessels nicht einmal als Schemel in Erwägung gezogen hätte. In einer anderen Situation wäre ich vielleicht bereit, die Obstkiste für den fehlenden Designtisch zu nehmen.

Welche Rolle, welche Gegenstände besitzen, ist durchaus keine triviale Angelegenheit. Die Zuordnung der Prädikatoren ist in diesem schlichten Beispiel von einem Verständnis des Gesamtzusammenhangs abhängig. In noch größerem Ausmaß betrifft dies die Gewichtung der einzelnen Sachverhalte.²

Um den Erkenntnisgewinn durch Computer zu steigern, sollten dem Automaten möglichst viele Informationen zur Verfügung stehen. Da die Anzahl an Informationen stets erweiterbar und damit prinzipiell offen ist, stellt sich jedoch generell wie auch im Einzelfall die Frage nach der Auswahl der relevanten Umstände. Menschen lösen diese Aufgabe mit Alltagswissen, sie reklamieren Plausibilität und bringen unstrukturierte Vorannahmen ein. Für die KI-Forschung wird gerade die von Menschen als „selbstverständlich“ berücksichtigte Kontextabhängigkeit zum Problem.

An dieser Stelle lohnt sich ein Blick auf die Sprachphilosophie des 20. Jahrhunderts.

Parallele zu Philosophie

Als mit dem Mittel der modernen Logik Anfang des Jahrhunderts ein neues Instrument zur Analyse von Aussagen zur Verfügung stand, glaubte auch die Philosophie die Frage nach der Grundlegung von Erkenntnis neu stellen zu können. Der logische Empirismus definierte Wissen als Operation aus elementaren Befunden,³ welche die Gegebenheiten der Welt repräsentieren. Das entstehende Basisproblem sollte möglichst eliminiert werden durch ein Programm, das philosophische Fragen auf syntaktische Fragen reduzieren wollte⁴ und naturwissenschaftliche Fragen durch die Anwendung syntaktischer Mittel auf unproblematisch empirisches Material zurückführen wollte.

Zunehmend wurde jedoch deutlich, daß selbst der „Kalkül Sprache“ nicht mit syntaktischen Mitteln allein zu deuten ist. Semantische Theorien wurde entwickelt, um die Bedeutung der Ausdrücke zu sichern oder - wie Tarski (1935) es formulierte - den Wahrheitsbegriff in den formalisierten Sprachen zu begründen. Während durch die Arbeiten zur Semantik die Autonomie der Syntax gebrochen war, entstand andererseits die Gefahr, daß die wissenschaftliche Begründung die vor-sprachlichen Unterscheidungen nun auf der Ebene der Sprache reformuliert. Um diese „ontologische Verdoppelung“ zu vermeiden, gleichzeitig die Einsichten aus den semantischen Theorien zu wahren, wurden zunehmend C.S. Peirce, G. Ryle, J.L. Austin und insbesondere L. Wittgenstein diskutiert, welche im Rahmen des „linguistic turn“ eine konsequent pragmatische Sichtweise forderten. Methodisches Mittel war der Handlungsbegriff, durch den die Kontaktstelle zum „Gegebenen“ nicht mehr von einem unterstellten Realitätsbegriff abhängt, sondern Bedeutung in gemeinsamer Interaktion erst konstituiert wird. Sprache ist demnach nicht auf Abbildung von Wirklichkeitsausschnitten oder Sachverhalten angelegt, sondern sie eröffnet primär Handlungsmöglichkeiten. Praktische Bedürfnisse

können natürlich auch „reine Abbildungsinteressen“ geltend machen, aber im methodischen Aufbau vollzog sich eine Wendung. Zur Basis von Syntax und Semantik wurde Pragmatik.

Ich möchte die ausgiebige Auseinandersetzung, die zu diesem Ergebnis führte, hier nicht aufrollen, aber als Resümee für den Fortgang dieser Arbeit zusammenfassen:

- Eine reine Tatsachenerhebung ist nicht möglich, stets sind Wertungen konstitutiv. Geltungsfragen können daher nur in einem pragmatisch fundierten Kontext aufgeworfen werden, in dem die Möglichkeit besteht, Interessen und Absichten zu verhandeln.

Dieses Ergebnis ließ das Vorhaben scheitern, von einer atomistischen Wissensbasis aus mit nur logisch-semantischen Mitteln unser Wissen zu begründen.

KI und Sprache

Die Geschichte der KI-Forschung scheint diese Entwicklung nachzuvollziehen, ohne freilich die Argumente aus der Sprachphilosophie zur Kenntnis zu nehmen.

Demnach befindet sich die KI-Forschung in der zweiten Phase, wo lexikalische Referenz- und Relationenlisten den Wissenspool speisen und eine normal-sprachliche Dateneingabe angestrebt wird. Beliebt sind auch Hand-Auge-Maschinen, welche einfachen Situationen die notwendigen Informationen selbsttätig entnehmen. Denkbar erscheint in dieser Phase auch, den Computer mit allen Rundfunkprogrammen zu füttern.

Dennoch ist die Weigerung, die Ebene der Sätze zu verlassen, mit dem Versuch gleichzusetzen, die Bedeutung der Ausdrücke von der Beziehung der Sprecher zueinander und zu den sprachlichen Elementen abzutrennen.

Was ergibt nun eine Übertragung der sprachphilosophischen Ergebnisse auf die KI-Forschung?

KI und Handlung

Mir scheint der Schlüssel für eine neue Phase der KI-Forschung darin zu liegen, auch in der Theorie der künstlichen Intelligenz den Handlungsbegriff an zentraler Stelle zu plazieren.

Handeln bedeutet aber nicht nur Arme und Beine zu besitzen. Automaten müßten als Mitglied einer Handlungsgemeinschaft „erzogen“ werden. Die Dateneingabe von „Welt“ erfolgt weder durch Satzsysteme noch mit Hilfe künstlicher Sinnesorgane, sondern in interaktiven Lernsituationen. Ausgaben sind nicht schlicht Reports, sondern Äußerungen werden als Handlungen gewertet und wie über eine Person verhandelt.

Nun ist es aber offensichtlich nicht damit getan, nur so zu tun „als ob“ Reaktionen des Computers Handlungen wären. Nur weil Menschen ihren Umgang mit einer Apparatur als Handlung interpretieren, muß es nicht gerechtfertigt sein, die Prozesse an der Apparatur als Handlung zu werten und die Apparatur selbst als Handelnden.

An welcher Stelle erfolgt ein qualitativer Zuwachs, der es angemessen erscheinen läßt, etwa von einer „artificialen Person“ zu sprechen?

Die Rede von Person scheint mir in diesem Zusammenhang geeignet, da eine Debatte über Geist, Bewußtsein oder gar Seele nicht intendiert ist, aber darauf hingewiesen werden soll, daß wir es mit selbstorganisierenden Automaten zu tun haben. Als Kriterium möchte ich das Vorliegen von Zieldispositionen vorschlagen. Ziel eines „unselbständigen“ Computers ist höchstens die ordnungsgemäße Erledigung klar vorgegebener Aufgaben. Schachprogramme waren daher lange die Paradebeispiele der KI. Doch für jeden Zug alle Fortsetzungen zu berechnen und jeweils wieder alle möglichen Gegenzüge usw. auf der Suche nach einer Gewinnstellung - dies macht kein Großmeister. Lebende Schachmeister überblicken nur für wenige Stellungen die Fortsetzungen weniger Züge in eine Tiefe von bis zu fünf Zügen.

Im Vordergrund dagegen stehen Konstellationen und Strategien. Allgemein gesprochen, ist Intelligenz weniger das Lösen als das Entwickeln einer Lösungsstrategie. Solche Intelligenz setzt einen gewissen Freiheitsgrad voraus in der Wahl der Mittel. Um diesen Freiheitsgrad zu organisieren, werden regulative Instrumentarien benötigt, welche den Hintergrund liefern, auf dem Menschen gewöhnlich handeln. An dieser Stelle sollte man vermuten, daß das Forschungsprogramm „artificial Person“ abgebrochen werden muß, denn welche Ziele sollten konkret angegeben werden? Ich vermute aber, daß es gar nicht notwendig ist, eine vollständige Liste „oberster Ziele“ zu finden, sondern daß nur überhaupt antriebsmotivierende Ziele implementiert werden müssen. Sicher wären Ziele zum Wohl der Menschheit - zusammengestellt von den Weisesten der Erde - nützlich, aber im Zweifelsfall dürfte es auch „Versuche, Lob zu erhalten“ tun oder „Strebe es an, einmal als Beteiligter in einer Fernsehshow über Künstliche Intelligenz mitzuwirken“. Bezeichnenderweise sind sich auch Menschen in der Regel über ihre Ziele oder andere intentionalen Zustände nicht wesentlich klarer und nähern sich Lebensproblemen dennoch auf intelligente Weise.

Neben der Einführung von Zielen, wird von einer „artificialen Person“ erwartet, daß sie selbstbezügliche Aussagen treffen kann. Einige Daten erhalten damit den Status von Erinnerungen. Der Computer muß lernen, daß „Handlungen“ Auswirkungen auf ihn haben und er muß frühere Reaktionen auf seine „Handlungen“ ins Kalkül ziehen, wenn er Ziele verfolgt. Er muß dazu gebracht werden, handlungsleitende Theorien zu entwickeln, die ihm in seinem Umgang mit Menschen nützlich sind. Er soll anstreben, bei möglichst vielen (Sprach-)Spielen mitzuspielen. Schließlich sollte er sich eine Meinung bilden, was beim Menschen Intelligenz genannt wird.

Um ein praktisches Beispiel zu geben, ist in vielen Situationen gar nicht die theore-

tisch ideale Lösung gefragt, sondern diejenige, welche „ideal durchsetzungsfähig“ ist (und einen Schritt in die richtige Richtung darstellt). Damit benötigen Computer plötzlich kulturelles Wissen und Kenntnisse um Eigenarten von Menschen, welche oft diesen selbst rational nicht verfügbar sind, sondern die sich nur im Handeln erweisen.

Die Sozialisierung von Computern gerät rasch an eine Grenze, wenn es zum Verständnis bestimmter Vorgänge darüberhinaus notwendig ist, einen menschlichen Körper zu besitzen. Um einen Vorschlag zu machen, wie kinästhetische Informationen (aus Gelenken, Muskeln etc.) sowie Wissen, das nur durch Lebensvollzug erworben werden kann, der KI verfügbar gemacht werden kann, muß nun ein weiterer Strang aufgenommen werden, der zunächst in keinem Zusammenhang mit Künstlicher Intelligenz steht.

Virtual Reality

Virtuelle Realität oder besser Telepräsenz zielt auf die Erschaffung neuer Handlungszusammenhänge, indem zu den „wirklichen“ Handlungen „künstliche“ hinzugenommen werden. Diese sind „künstlich“ in dem Sinn, daß Regeln noch nicht etabliert sind.

Das Instrument für Telepräsenz ist ein möglichst umfassender Körpersensor. Er nimmt jegliche Regung wahr, überträgt und stellt sie einzelnen (auch sich selbst) oder mehreren gleichzeitig oder zu einem späteren Zeitpunkt unter beliebiger Veränderung zur Verfügung. Zu jeder Sinnesmodalität sind Apparate für Äußerung und Wiedergabe zu konstruieren.

Dabei muß die Wiedergabe nicht analog in dem Sinn sein, daß eine Kraft durch nanotechnische Minimotoren simuliert wird. „Spüren lassen“ kann auch bedeuten, Erregungsmuster durch Felder zu erzeugen, die äquivalent interpretiert werden. Im Prinzip ist das Sinnesorgan damit überbrückbar und das Nervensystem zieht sich auf das Gehirn zurück.

Analoges muß freilich für das Hormonsystem gelten. Vielleicht spielt auch das Lympfdrüsen system eine Rolle und Modalitäten, von denen wir nur gerüchtweise gehört haben. Ob nach Telepräsenz also konsequenterweise Teleportation, Telepathie und Telekinese angesagt sind, kann zunächst zurückgestellt bleiben.

Um die Zielrichtung zu verdeutlichen, ließe sich als Ideal für die Wirkungsweise des Körpersensors die der SF-Literatur entnommene „Nerveninduktion“ heranziehen. Zur Wiedergabe der Modalität „Sehen“ genügt die geschickte Erregung des Sehentrums. Überall wo die Benutzung eines Codes bereits etabliert ist, kann die Kommunikation im Prinzip auf neuronaler Ebene stattfinden.

Die Interpretation zusätzlicher Erregungsmuster setzt auch hier keine Theorien über „Denken“ voraus. Welche Erregungsmuster welche Erfahrung repräsentieren, wird bereits heute schlicht durch Feedback-Experiment⁵ erwiesen. Sollten Gedanken durch elektrodynamische Apparate stimuliert werden können, so wird der Proband schon sagen, welche Gedankeninhalte er vorfindet.

Ich betone dies, weil ich hervorheben möchte, daß auch an dieser Stelle keine materialistische Theorie des Denkens notwendig ist. Solange Methoden zur Beurteilung der neuen Handlungsmöglichkeiten zur Verfügung stehen, kann jede abtast- und reproduzierbare Aktivität zur Kommunikation herangezogen werden. Kommen wir nun auf die Erzeugung einer „artificialen Person“ zurück.

Telepräsenz für KI

Das Instrumentarium, welches zur Realisierung von Telepräsenz benötigt wird, wird zutreffend als Interface bezeichnet. Auf dieser Kontaktfläche berühren sich die Welt der handelnden Menschen und der virtuelle Raum, auf den deren Handlungen abgebildet werden und damit für Interaktionen „zweiter Ordnung“ bereit stehen. Diese Informationsmatrix ist nun ideal

geeignet, die Laborumgebung artifiziieller Personen⁶ zu gestalten. Da beide computergestützt konstruiert werden, brauchen Schnittstellenprobleme auf der Ebene der Syntax nicht befürchtet werden.

Für eine artifiziieller Person ist es völlig gleichgültig, ob eine menschliche Gestalt von VR-Rechnern „vorgespield“ oder materiell simuliert wird. Es ist daher auch nicht notwendig, auf die Hilfe von Robotik und - wohl oder übel - Genetik zu warten. Ob der Arbeitsspeicher eines Versuchsandroiden das Schütteln der rechten Hand von Biomotoren gemeldet bekommt oder die äquivalenten Signale irgendwie anders erzeugt werden, macht keinerlei Unterschied, wenn er dies nur als Gruß interpretiert und darauf zu reagieren weiß. Für Computer kann die Antwort auf einen Solipsismusvorbehalt nur bejaht werden: Diese Welt ist nur eingebildet.

Es ist überraschend, daß die Pioniere der VR ihre Arbeit eher in einer Intelligenzverstärkung verstanden als es die heutige Cyberspaceliteratur tut. Zwischen KI und VR gab es damals jedoch kaum Austausch. Die Dateneingabe/Anlern-, Dressurphase dürfte ähnlich wie bei Kindern verlaufen, nur daß bei einem Computer massive parallele Kommunikation vorstellbar ist. Theoretisch könnte jeder Mitspieler am globalen Netzwerk bei der Bildung künstlicher Intelligenz durch einen Telefonanruf oder mittels Telepräsenz Sitzung mitwirken. Für den Computer erfolgt sein Lernen nicht schlicht durch Wissensanhäufung, sondern durch üben und erproben.

Ich vermute, artifiziielle Intelligenz wird sich dem annähern können, was beim Menschen Intelligenz genannt wird, sofern theoretische und empirische Elemente Gelegenheit erhalten, quasi in einem Abschleifungsprozeß zur menschlichen Position zu konjugieren⁷. Die empirischen Elemente betreffen dabei die Kontaktfläche (-raum wäre angebracht) zur Lebenswelt (hier wäre Interaktionsraum zutreffend), während die Theorie die Konzeptbildung kanalisieren muß. Da umwelt- und handlungsbezogene Organisa-

tion von Verhalten bei biologischen Systemen beobachtet werden kann, liegt es nahe, deren Aufbau zu untersuchen. Neben die prozessor-orientierte Forschung, deren Fortgeschrittendste Konzepte neuronale Netze verwenden, und die erwähnte interface-orientierte, müßte schließlich noch eine künstlerisch-wissenschaftliche Kreativität treten, welche die interessanten Ergebnisse erst herauskitzelt⁸. Was Träume, Inspirationen und Visionen anbelangt, dürften Computer bis auf weiteres auf Menschen angewiesen sein. Auch die Feinabstimmung von Zielen und Prioritäten bedarf wohl immer wieder kleiner Korrekturen oder zumindest stichprobenartiger Kontrollen⁹.

Ganz allgemein dürfte die Modellierung einer artifiziiellen Person die Form eines Trainings annehmen.

Nachdem seit Newtons epochemachender Principia gerade 300 Jahre vergangen sind, kann sich die KI-Forschung aber ruhig ein paar Jahrtausende nehmen, um diverse Details zu klären. Eine Unterhaltung mit artifiziiellen Personen soll schließlich nicht nur informativ und unterhaltsam sein, sondern auch originell und anregend. Wollte man diese Geschöpfe bedauern, dann höchstens, weil ihnen Bewußtsein unbekannt ist.

Die Frage nach Bewußtsein artifiziieller Personen scheint mir für die Theorie der KI gänzlich unwesentlich zu sein. Da sich Menschen - und insbesondere KI-Forscher - jedoch so außerordentlich dafür interessieren, ist es unumgänglich, im Rahmen einer Einschätzung von KI darauf einzugehen.

Besitzen Computer Bewußtsein?

Die physikalistisch/kybernetische Theorie besagt, daß in einem Automaten das Vorhandensein hinreichend komplexer Informationsverarbeitungsprozesse mit Bewußtsein äquivalent ist.

Bei der bisher dargelegten Auffassung tritt „bewußt“ jedoch nie als Charakterisierung

des Zustandes einer Maschine auf, da eine Schalterstellung (oder Kombinationen daraus) und ein mentaler Zustand zunächst verschiedenen Kategorien angehören. Zum Kriterium wurde die größtmögliche Annäherung an einen Menschen - ganz in Übereinstimmung mit dem Turing-Test, der die Ununterscheidbarkeit fordert.

Daher irrt Turing, wenn er sein Verfahren zur Beantwortung der Frage vorstellt: Können Maschinen denken? So wie er sein Immitationsspiel einführt, hatte er offenbar im Sinn: Können Maschinen glaubhaft machen, daß sie denken? Diese beiden Fragen sind verschieden, wie jedermann weiß, der eine Zaubervorführung besucht hat.

Der Turing-Test ist folglich die maximale Qualitätsforderung an einen ordentlich kultivierten KI-Rechner, sie ist kein empirischen Tests für eine Eigenschaft, von der bisher nicht die Rede war.¹⁰

Welches ist nun diese Eigenschaft?

Es ist von Dali bekannt, daß er sich gelegentlich vorstellte, ein Bahnhof zu sein. Könnte sich nicht auch ein Bahnhof vorstellen, ein Künstler zu sein?

Diese Frage scheint darauf hinauszulaufen, daß es irgendwie *ist*, ein Mensch zu *sein*.

Der Autor kann sich eine Maschine vorstellen, die in allen Situationen so agiert wie er. Sie sei gerade dabei einen Aufsatz über KI und VR zu schreiben. Hat sie deswegen die gleiche Vorstellung von Ich wie ich sie habe?

- Wenn wir die *Möglichkeit* dafür im Auge haben, werden wir wahrscheinlich Ja sagen. Bei anderen Menschen ist unsere Lage keineswegs klarer, obwohl dieses Einverständnis unumstritten ist.

- Wenn wir allerdings *Notwendigkeit* fordern, gibt es kein Argument, denn als Lösung können nur eine logische und eine empirische Variante anerkannt werden.

1. Sich wie ein Mensch zu verhalten und ein Mensch zu sein, sind logisch das gleiche.

Hierfür sehe ich keinen Grund. Im Gegenteil: ich war gestern so wie ich üblicherweise bin und doch verbindet sich für mich

mit jenem Umstand heute keine solche Vorstellung von Ich wie ich sie jetzt von mir im Augenblick habe. Eine Vorstellung von Ich zu besitzen, kann also nicht äquivalent mit einer bestimmten Konstellation innerer und äußerer Zustände sein. Denn obwohl solche Konstellation bei einem Apparat jederzeit realisiert werden können, darf ein Ich nur in den Augenblicken auftreten, welche Gegenwart reklamieren können. Ist dies beim Menschen schon rätselhaft genug, wird es für die Bewußtseinsfrage artifizieller Personen zum echten Problem.

Ein ganz anderes Argument lautet: Wäre es möglich, einen Computer mit Bewußtsein zu bauen, wäre es sicher auch möglich, einen in seinem Denken und Verhalten identischen zu bauen, welcher das Phänomen Bewußtsein nicht aufweist. Wer logische Identität vertritt, muß daher die Möglichkeit zu einer solchen Apparatur leugnen, was wenig sinnvoll erscheint.

2. Nun mag es jedoch empirische Gründe geben, warum eine gewisse Komplexität an Informationsverarbeitungsprozessen Bewußtsein hervorbringt.

Diese Möglichkeit möchte ich nicht grundsätzlich bestreiten und glaube auch, daß ein Unmöglichkeitsbeweis unmöglich ist. Dennoch muß das Argument beim gegenwärtigen begrifflichen und empirischen Stand der Wissenschaft als unverständlich und keineswegs naheliegend bezeichnet werden.

Ich halte es zwar für zulässig, ein Ich als Kristallisationspunkt einer artifiziellen Person anzunehmen - etwa mit dem Status theoretischer Begriffe in der Physik - sehe aber nicht, inwieweit dieses Ich mit Selbstbewußtsein ausgestattet sein soll. Obwohl verwandt, würden gerade die euphorischen Physikalisten sich verwahren, einen Einfluß von Planetenkonstellationen auf menschliche Schicksale anzunehmen, während sie bei anderen Konstellationen mit mentalen Zuschreibungen weitaus großzügiger sind.

Bei Organismen, deren Funktionsweise wir nicht kennen, ja wo wir nicht einmal wis-

sen, ob eine Funktionsweise vorliegt, bleibt uns in der Tat nur eine Beurteilung des Verhaltens. Bei künstlichen Systemen dagegen steht uns die Möglichkeit offen, in der Terminologie ihrer Konstruktion zu sprechen und wir sollten es zunächst dabei belassen.

Vielleicht ist der Physikalismus eines der bedeutendsten Forschungsprogramme der Wissenschaft, die These selbst bleibt jedoch Glaubenssache.

- 1 Eine historische Reihenfolge müßte freilich mit Schickard aufnehmen. (Vgl. Briefwechsel mit Kepler über Rechenmaschine für 4 Grundrechenarten).
- 2 Ein schönes Beispiel ist die Einbindung fiktionaler Kontexte.
- 3 Die Autoren diskutieren an dieser Stelle je nach Ausrichtung über Sachverhalte, Sinnesdaten, Protokollsätze, Basissätze etc.
- 4 Systematisch weitreichendster Versuch war Carnap Vorschlag, ein Basisproblem gar nicht erst aufkommen zu lassen, um den von Russell und Hilbert eingeführten Strukturbegriff zu verwenden.
- 5 Vgl. die Versuche, in denen Gehirnwellen zur Steuerung herangezogen werden sowie die Arbeiten der Molekularbiologin Kary Mullis.
- 6 Ich verzichte im folgenden auf die Anführungszeichen.
- 7 In der Astronomie bezeichnet Konjunktion die gleiche Stellung für einen Beobachter.
- 8 Walther Ch. Zimmerli spricht von der Mensch-Maschine-Kommunikation als kreatives Tandem.
- 9 Die Problematik sogenannter Expertensysteme wage ich hier nicht aufzugreifen.
- 10 Der Unterschied erinnert an den Status von Trägheit in der Physik. Zunächst als empirisches Gesetz betrachtet, sprechen wir heute vom Trägheitssatz, der den Hintergrund für Naturgesetze abgibt.