

# Mosaiksteine einer Wissenschaft vom Geiste

W. Bibel

TU Darmstadt, Germany

Bibel@Informatik.TU-Darmstadt.De

9. April 2003

## Zusammenfassung

In diesem Aufsatz wird die Intellektik, dh. Künstliche Intelligenz und Kognitionswissenschaft, charakterisiert sowie ihre Zielsetzungen, Erfolge und Perspektiven beschrieben. Unter ihren Perspektiven findet sich auch das Heranwachsen einer Wissenschaft vom Geiste im strengen Sinne, von der inzwischen erste Mosaiksteine identifiziert werden können. Zudem wird gegen wissenschaftlich bemäntelte und fruchtlose Versuche argumentiert, die die Unmöglichkeit künstlicher Intelligenz „beweisen“ möchten.

## 1 Geist, Intellektik und Philosophie

Kein anderes Thema hat die Denker aller Zeiten seit Jahrtausenden mehr bewegt als die Fragen nach dem Geist und der Seele des Menschen. Was sind das, Geist und Seele? Wie hängen Leib und Seele zusammen? Wie denken und fühlen wir? Das sind nur einige der unzähligen Fragen, die sich Menschen zu allen Zeiten gestellt haben – und noch immer stellen – und die alle um die Grundfrage „*Was ist Geist?*“ kreisen.

Diese Grundfrage gehört dementsprechend auch zu den fundamentalsten Fragestellungen der Wissenschaften. Sie hat mindestens den gleichen Bedeutungsrang wie „*Was ist Materie?*“ oder „*Was ist Leben?*“. All diese drei Fundamentalfragen sind bis heute nicht letztgültig beantwortet, am wenigsten jedoch die nach dem Geist.

Während nämlich die Naturwissenschaften Physik und Biologie mit Experimenten untermauerte, begrifflich präzise und mathematisch formulierte Theoriegebäude erstellt haben, die die Begriffe Materie und Leben recht weitgehend eingrenzen, ist es der sich für den Geistesbegriff zuständig erklärenden Disziplin Philosophie bis heute nicht gelungen, wirklich erkennbare Fortschritte in dieser Richtung zu erzielen. Dabei ist das Thema immerhin seit Platon und Aristoteles über Descartes, Kant und Hegel bis heute intensivst studiert worden und füllt mit den darüber verfassten Schriften ganze Bibliotheken. Nur Klassiker darunter – wie beispielsweise der Phaidon von Platon [Pla02] oder das Seelenbuch

von Aristoteles [Ari95] – erscheinen einem wegen ihrer bis heute spürbaren intellektuellen Frische überhaupt noch als lesbar.

Mitte letzten Jahrhunderts setzte jedoch mit der Erfindung des universellen Computers eine Entwicklung ein, die mit der Disziplin Philosophie absolut nichts zu tun hatte und deren Tragweite sich die Philosophie erst so langsam bewusst wird. Die Computerwissenschaftler der ersten Stunde wie Zuse, Turing, von Neumann und viele andere hatten sofort die Relevanz des dem Computer zugrundeliegenden Denkparadigmas für jene Fundamentalfolge nach dem Geist erkannt. Von diesen war es dann Alan Turing, der diese Relevanz in einem bis heute grundlegenden Artikel [Tur50] im Einzelnen ausgeführt hat. Um diese Gedanken herum hat sich schließlich im Jahre 1956 eine neue Disziplin konstituiert.

Als Name für diese neue Disziplin hat sich in den angelsächsischen Ländern *Artificial Intelligence* eingebürgert. Weil die wörtliche Übersetzung *Künstliche Intelligenz* (KI) im Deutschen eine andere Bedeutung als der englische Terminus hat und aus linguistischen Gründen für eine Gebietsbezeichnung ohnehin nicht wirklich taugt, wurde stattdessen *Intellektik* vorgeschlagen [Bib80, Bib92]. Etwa gleichzeitig formierte sich aus der KI heraus die Kognitionswissenschaft und die KI fokussierte immer stärker auf ihren informatisch-technischen Teil, so dass Intellektik heute als die Vereinigung von KI und Kognitionswissenschaft(en) angesehen wird. Intellektiker ist danach jeder Wissenschaftler, dessen Forschungsgegenstände intelligente (technische oder biologische) Systeme unter den Gesichtspunkten des Computationalismus sind, auf den im nächsten Abschnitt noch genauer eingegangen wird.

Die Intellektik hat nicht nur eine in der Geschichte der technologischen Entwicklung nie dagewesene Erfolgsbilanz aufzuweisen, sondern auch in den vormals philosophischen Fragestellungen rund um den Geistesbegriff erstmals echte Fortschritte erzielen können. Auf beides werden wir im nächsten Abschnitt ausführlich zu sprechen kommen.

Dieser Erfolg eines jungen, respektlosen, aber eben unübersehbar erfolgreichen Konkurrenten im ureigensten Terrain musste die altehrwürdige Philosophie wie ein Schock treffen. So schreibt ein Philosoph über die KI: „*Besonders einige Philosophen versetzt sie in Aufregung.*“ [Tet94, S.106]. LaForte et al. beschreiben die Reaktionen auf das Projekt KI deutlich drastischer und damit nach meiner persönlichen Erfahrung zutreffender [LHF98, S.1] (in der im Web stehenden Version).

*This project produces in some a reaction of revulsion, even horror. Indeed, the spectre of artificial intelligence has inspired many of the world's scholars to engage in rituals of exorcism – their goal being the banishing of this spirit to the mechanistic netherworld from which it came.*

Dieses Zitat beschränkt sich zurecht nicht auf die Philosophie allein, da die vom Wesen her interdisziplinär angelegte Intellektik das Potenzial zur radikalen Veränderung der Forschungsmethoden in vielen anderen Disziplinen hat, worauf wir im vierten Abschnitt noch zu sprechen kommen werden.

Der Komputationalismus, auf dem die Intellektik beruht, ist ein revolutionärer Denkansatz von ähnlich großer Tragweite wie der Galileische Ansatz, der die Naturwissenschaften

begründete, oder wie die Evolutionstheorie von Darwin, die bis heute das wissenschaftliche Weltbild entscheidend prägt. Jede derart tiefgreifende wissenschaftliche Revolution hat zu heftigsten Reaktionen geführt, bei denen sich beiläufig immer auch die Kirchen aufgrund ihrer teilweise ideologischen Prägung in die vorderste Front der Kritiker eingereiht hatten. Eine durchgreifende Idee hat sich wohl aber noch nie verhindern lassen. So ist auch die Intellektik in allerersten Ansätzen bereits dabei, mit ihrer Methodik alle Geistes- und Sozialwissenschaften zu revolutionieren, insbesondere zu exakten Wissenschaften im naturwissenschaftlichen Verständnis zu machen.

Aus diesem Grunde wird dieser Aufsatz nur periphär auf die Angriffe gegen die KI eingehen und sich mehr um eine Klarstellung der Vorstellungen von KI bemühen, die außerhalb der Intellektik noch immer recht verworren sind. In diesem Sinne umreißt der nachfolgende Abschnitt die Arbeitshypothese, das Profil und die Inhalte der Intellektik. Vor allem beschreiben wir auf dieser Grundlage, wie sich innerhalb dieses methodologischen Rahmens eine Wissenschaft vom Geiste entwickeln könnte, die diesen Namen verdient. Hierauf kontrastieren wir im Abschnitt 3 diese wissenschaftliche Methodik mit der bis heute in den Geistes- und Sozialwissenschaften, vor allem in der Philosophie vorherrschenden Methodik, die wir anhand der Diskussion über die Frage „*Können Maschinen denken?*“ illustrieren. In Abschnitt 4 entwerfen wir schließlich eine Perspektive, die sich aus der weiteren Entwicklung der Intellektik ergibt. Einige Bemerkungen beschließen diesen Artikel.

Natürlich ist dies ein viel zu umfangreiches Programm für einen Aufsatz, der die Grundlage für einen Vortrag auf einer Konferenz bildet. Es muss dem Autor daher nachgesehen werden, dass er ausgiebig von der Möglichkeit Gebrauch macht, auf Textabschnitte seines jüngst erschienenen Buches [Bib03] zu verweisen.

## 2 Umriss der Intellektik

Jede wissenschaftliche Disziplin ist durch ihr Forschungsziel und ihre Forschungsmethodik charakterisiert. Das *Ziel* der Intellektik ist – kurz gesagt – ein wissenschaftliches Verständnis des Intellekts. Statt vom Intellekt können wir (in diesem Zusammenhang) synonym auch von Geist und Psyche sprechen. Der Gegenstand der wissenschaftlichen Untersuchung in der Intellektik ist also alles, was mit Geist und Psyche im Zusammenhang steht. Dazu gehören dann beispielsweise auch ein Verständnis von Intelligenz, Wahrnehmung, Gedächtnis, Intentionalität, Bewusstsein, freier Wille, Gefühle usw.

Die grundlegende *Forschungsmethodik* der Intellektik ist durch die Arbeitshypothese des (von Newell und Simon formulierten) Komputationalismus [NS72] charakterisiert. Der Komputationalismus besagt, dass alle geistig-psychischen Phänomene auf einer bestimmten Abstraktionsebene als informationsverarbeitende, komputationale Prozesse verstanden werden können. Mit der Erfindung des universellen Computers steht uns ein Werkzeug zur Verfügung, auf dem wir alle informationsverarbeitenden Prozesse modellieren können. Im Sinne des Komputationalismus erfordert die Theoriebildung in der Intellektik zu ihrer Validierung Experimente mit solchen Computermodellen. Insoweit ist ihre Entwicklung eng mit derjenigen der Informationstechnologie (IT) verknüpft.

Die so gegebene Charakterisierung bedarf für den Uneingeweihten noch einer Reihe weiterer Erläuterungen der darin verwendeten Begriffe wie „wissenschaftlich“, „Arbeitshypothese“, „Abstraktionsebene“, „informationsverarbeitender Prozess“, „universeller Computer“, „modellieren“ usw. Wie im letzten Abschnitt angekündigt, verweisen wir diesbezüglich auf [Bib03, S.48ff] und beschränken uns hier auf wenige Anmerkungen dazu. Vor allem sei auf den in der Charakterisierung beschriebenen experimentellen Charakter bei der Theoriebildung in der Intellektik hingewiesen, der hier die analoge Funktion wie in den Naturwissenschaften hat; nur eine experimentell bestätigte Theorie ist letztlich akzeptabel.

Weiter betonen wir auch hier nochmals den Charakter des Komputationalismus als *Arbeitshypothese*, der von vielen philosophischen Autoren offenbar noch immer nicht verstanden ist. So finden sich in [Tet94, S.108,117] Formulierungen wie: *die KI „behauptet“ diese These* oder *„ist die These überhaupt wahr, oder ist sie eine großsprecherische Ankündigung der KI“* oder *ihr Nichtgelten „wäre für die Ingenieure vermutlich eine herbe Enttäuschung“*. (Arbeitshypo-) Thesen werden als Richtlinie für die Forschungsarbeit zugrundegelegt und eben nicht „behauptet“; ob sie „wahr“ oder falsch sind wird völlig emotionslos offengelassen. Sollte ihr Wahrheitsgehalt jemals entschieden werden können, wird dies angesichts der damit täglich erzielten konkreten Erfolge mit Sicherheit zu keinen „Enttäuschungen“ führen.

In diesen beiden Punkten der experimentellen Begründung und des komputationalen Ansatzes unterscheidet sich die Intellektik ganz wesentlich von den Geistes- und Sozialwissenschaften herkömmlicher Prägung, die aus unserer Sicht keine Wissenschaften in dem hier geforderten strengen Sinne sind (siehe Abschnitt 5.4 in [Bib03]). Sie befinden sich aber bereits mitten in der Umbruchphase, so dass eine Reihe ihrer Wissenschaftler schon eindeutig zur Intellektik zu rechnen sind. Dazu gehören beispielsweise führende Psychologen wie J.R. Anderson, W. Edwards, D.E. Rumelhart und J.L. McClelland, Philosophen wie D.C. Dennet, J.A. Fodor und Z.W. Pylyshyn oder komputationale Hirnforscher wie P.S. Churchland und T.J. Sejnowski, um nur einige wenige (Disziplinen und Forscher) zu nennen. Die verbleibenden stehen vor der Entscheidung, ob sie sich einer exakten Wissenschaft verpflichten oder in einem präformalen akademischen Bereich verharren wollen, dessen messbare Erfolge schon immer sehr zweifelhaft waren.

Der experimentelle Teil der Intellektik hat eine ungeheure Eigendynamik entwickelt, weil die entstehenden Computersysteme nicht nur eine Bedeutung im Hinblick auf das Fernziel der Intellektik haben, sondern in den meisten Fällen auch von größter praktischer Bedeutung sind. Es ist daher nur allzu verständlich, dass die Mehrzahl der Intellektiker den Nutzen ihrer Arbeit vorrangig in diesen praktischen Anwendungen sehen und dabei das Gesamtziel der Intellektik gern aus den Augen verlieren. Genau aus diesem Grunde hat es den im letzten Abschnitt bereits erwähnten Formierungsprozess der Kognitionswissenschaft und die Bedeutungseinengung des KI-Begriffs auf diesen praktischen Teil gegeben. Es ist aber in beiden Lagern wohl immer noch Konsens, dass letztlich nur ein vereintes Vorgehen den angestrebten langfristigen Erfolg bringen werden. Methodologisch verwenden ohnehin alle das gleiche Instrumentarium.

Zu diesem Instrumentarium gehören eine Vielzahl formaler Konzepte, die meist aus anderen Disziplinen entlehnt wurden, in der Intellektik dann aber oft eine intensive Weiterentwicklung, vor allem eine Präzisierung erfuhren, die der Modellierung in Form von Computerprogrammen gerecht wird. Es ist natürlich völlig unmöglich, hier in irgendwelche Details dieser Konzepte zu gehen, die etwa in dem Lehrbuch [RN03] auf nahezu eintausend Seiten ausgebreitet sind ([Nil98] verschafft bereits mit 500 Seiten einen guten Überblick). Davon sei hier nur eine grobe Vorstellung vermittelt.

Im Hinblick auf das Ziel der Intellektik ist es natürlich, ihren konzeptuellen Apparat anhand der abstrakten Modellierung eines menschlichen Akteurs in einer Umwelt zu strukturieren, wie es in dem genannten Lehrbuch konsequent durchgeführt ist und weshalb dieses sich als „*The Intelligent Agent Book*“ bezeichnet. *Ein solcher (modellierter) Akteur nimmt über seinen sensorischen Apparat aus der Umwelt (zu der auch die eigene Körperlichkeit gehört) Wahrnehmungen auf, deren Verarbeitung zu einer inneren „Vorstellung“ vom gegenwärtigen Zustand der Welt führt. In dieser Vorstellung sind natürlich auch frühere Wahrnehmungen und Erfahrungen repräsentiert. Auch hat der Akteur Kenntnisse über mögliche Auswirkungen des ihm zur Verfügung stehenden Aktionsrepertoires. So kann er „überlegen“, welche unter diesen Aktionen mit einer hohen Wahrscheinlichkeit zu Veränderungen führen, die den aufgrund seiner Wertvorstellungen eingeschätzten Nutzen für ihn maximieren. Aus den Wahrnehmungen nach der Ausführung der ausgewählten Aktionen kann er auch Schlüsse ziehen, die zu einem Lernverhalten im Hinblick auf künftige Aktionen führen.*

Diese letzten fünf Sätze beschreiben das Konzentrat des Forschungsprogramms der Intellektik. Dahinter verbergen sich die fast tausend Seiten dieses Lehrbuchs und der inzwischen unermesslichen Literatur dieses erfolgreichen Gebietes, in der diese vage Beschreibung in mathematisch präzisen Begriffen so weit formalisiert ist, dass sie sich unmittelbar in Computersystemen modellieren lässt. Deren Verhalten lässt sich dann mit dem Verhalten menschlicher Akteure messen. Je größer die Übereinstimmung, umso überzeugender ist die zugrundeliegende Theorie. Die Forschung besteht darin, noch vorhandene Diskrepanzen durch Verfeinerungen der Theorie zu verringern.

Ein allgemeinverständliches Beispiel ist die vergleichsweise winzige Welt des Schachspiels, in der es nur endlich viele, genau beschreibbare Weltzustände gibt, wenn auch die Anzahl der Zustände schon in dieser Miniwelt astronomisch groß ist. Der Leser möge die obigen fünf Sätze im Lichte dieses geläufigen Beispiels nochmals lesen. Die im ersten Schritt aufgenommenen Wahrnehmungen bestehen in der Information über den gegnerischen Zug, auf den der Akteur nun irgendwie reagieren muss; usw.

Das Verhalten heutiger Schachsysteme unterscheidet sich phänomenologisch praktisch nicht mehr von dem der Menschen. So besiegte 1997 das System Deep Blue den damaligen Schachweltmeister Kasparov. Während für dieses Ereignis noch ein ungeheurer Ressourcenaufwand betrieben werden musste, schaffen dieses Leistungsniveau heute Schachprogramme, die man für etwa 50 Euro im Kaufhaus erwerben kann. Ein solches Programm, genannt Deep Fritz, hat 2002 unter Turnierbedingungen gegen den damaligen Weltmeister Kramnik auf einem handelsüblichen Rechner ein Unentschieden erzwungen, ebenso wie vor kurzem (Februar 2003) das System Deep Junior gegen Kasparov.

Der moderne Mensch ist schon so abgestumpft gegenüber derartigen Ereignissen, dass der unglaubliche technologische Fortschritt hinter diesen Systemerfolgen nur noch mit einem Achselzucken wahrgenommen wird. Dabei sind die Schachsystemerfolge nur die Spitze eines riesigen Eisberges. Viele modernen Geräte des täglichen Gebrauchs wie Autos, Flugzeuge, Buchungssysteme, Waschmaschinen, PCs, Suchmaschinen im Internet usw. verwenden die in der KI entwickelte Technologie, ohne die sie den heute selbstverständlich erwarteten Standard nicht erbringen könnten.

Natürlich ist das Intellektziel selbst im Hinblick auf den winzigen Bereich des Schachspiels damit noch lange nicht erreicht, weil die von der Maschine eingesetzte Methode noch zu weit ab von der menschlichen ist. Daher erklären die heutigen Schachsysteme die Vorgänge im Denken eines menschlichen Weltmeisters erst äußerst rudimentär, weil sie auf einem bestimmten (formal definierten) Akteurstyp beruhen, der beispielsweise kein Wissen explizit verfügbar macht. Viele weitere Theorieverfeinerungen werden daher selbst im Hinblick auf das Schachspielen noch nötig sein, um einer solchen Erklärung näher zu kommen. Solange jedoch die wissenschaftlichen Fortschritte in der IT insgesamt und der Intellektik im besonderen von Jahr zu Jahr so rasant anhalten, tritt die Frage in den Hintergrund, wie weitgehend eine solche Erklärung je gelingen kann. Niemand kann das heute vorhersagen. Allerdings hat derzeit auch niemand eine alternative Forschungsmethode anzubieten, mit der man solche Erklärungen schneller erzielen könnte. Also müsste sich eigentlich jeder an den Fragen des Geistes Interessierte an dem Intellektikprojekt mit Engagement und nach Kräften beteiligen.

Soweit also dieses geläufige Beispiel des Schachspiels, das in der modernen Intellektikforschung allerdings nur noch eine untergeordnete Bedeutung hat. Heute ist man mehr an Modellierungen interessiert, die der Wirklichkeit unserer Welt näher kommen. Ein Beispiel sind fußballspielende Robotermannschaften, welches eine Fülle neuer Herausforderungen im Vergleich mit Schachcomputern stellt. Das Geschehen ist hier dynamisch, die Impräzision der Sensoren und Effektoren ist zu berücksichtigen, Kommunikation unter den Akteuren ist nötig, gemeinsame strategische Planung ist erforderlich und vieles andere mehr, was im Schach überhaupt keine, in der Wirklichkeit aber eine große Rolle spielt. Dies erfordert ein sehr allgemeines Akteurmodell mit vielen Fähigkeiten, die der Mensch beherrscht.

Aber auch Fußballer sind sehr einseitig, brauchen beispielsweise keinen großen Wissensschatz zur Lösung anstehender Probleme. In Bezug auf Wissen und seine exakte Verarbeitung erscheinen uns eher Mathematiker als Vorbilder. Die Automatisierung mathematischer Fähigkeiten ist daher ein weiteres der vielen Anwendungsdomänen in der Intellektik. Mathematiksysteme sind heute in der Lage, Tausende von mathematischen Theoremen vollautomatisch zu beweisen, wobei die Beweise manchmal noch von keinem Menschen vorher gefunden wurden. Das System Otter/EQN hat beispielsweise im Jahre 1996 vollautomatisch den Beweis für eine mathematische Vermutung erbracht, den die Weltbesten der menschlichen Mathematiker 60 Jahre gesucht und nicht gefunden haben. Otter hat sich hier also mit den Besten in der „Königin aller Wissenschaften“ auf ein Niveau gestellt. Nebenbei bemerkt, liegen den hier verwendeten Systemen teilweise auch die grundlegenden theoretischen Beiträge von Gödel, vor allem sein Vollständigkeitssatz für die Prädikatenlogik erster Stufe zugrunde. In Darmstadt haben wir ein aus drei Pro-

logeregeln bestehendes System entwickelt, das das hundertfach größere Otter-System in manchen Aspekten sogar übertrifft.

Komplementär zu diesen Logiksystemen ist CYC als ein riesiges Wissenssystem entwickelt worden, dessen Stärke auf einem beeindruckend umfangreichen Allgemeinwissen beruht und quasi eine inferenzfähige EnZYKlopädie darstellt.

Bei all diesen beispiellosen Erfolgen ist andererseits jedem Experten klar, wie weit die Technik gleichwohl noch von den *umfassenden* Fähigkeiten des Menschen entfernt ist. Bei der Rasanz des Entwicklungstempos sind aber weder negative – „auch in  $n$  Jahren wird ein Computersystem die Fähigkeit  $x$  noch nicht aufweisen“ – noch positive Vorhersagen wirklich möglich. Irgendwelche Grenzen sind derzeit allerdings nicht erkennbar, auch nicht in Bezug auf die Qualität der Fähigkeiten. Um dies zu illustrieren seien hier noch einige typisch menschliche Fähigkeiten und deren Modellierungen kurz angedeutet.

Ist es möglich, dass ein System ebenso wie ein Mensch seine eigenen Fähigkeiten selbst einzuschätzen lernt? Genau dieses Thema der Selbstreflexion wurde an der TU Darmstadt von Herrn Grieser in seiner Dissertation ausführlich behandelt [BG01]. Es zeigt sich kein Grund, weshalb Systeme ihre eigenen Fähigkeiten nicht auch selbst einschätzen können.

Aber wie sollten Bewusstsein und Intentionalität, zwei entscheidende Eigenschaften mentaler Zustände beim Menschen, in einem Computer modelliert werden können? In [McC95] skizziert der „Vater“ der Intellektik, John McCarthy, einen Formalismus, der genau diese Modellierung auf einer sehr hohen Ebene (der reinen Logikebene) leistet. Eine ähnliche Skizze findet sich in [Bib03, S.79ff]. Dort wird im Abschnitt 3.4 sogar ein computationales Modell vom Verlieben und der Liebe skizziert.

Zugegeben, noch sind dies alles erst viele Mosaiksteine, die zu einer kohärenten *Theorie des Geistes* erst noch zusammengefügt werden müssen. Gleichwohl bilden sie eben bereits heute die Fragmente einer künftigen *Wissenschaft vom Geiste* – wohlgermerkt „Wissenschaft“ (im strengen Sinne) und nicht nur „Philosophie“!

Alle bis hier betrachteten Modellierungen müssen beim Leser den Anschein erwecken, als würde die Intellektik ihre Modellierungen nur aus der phänomenologischen Sicht außenstehender Beobachter und ohne Bezugnahme auf die Vorgänge im Gehirn vornehmen. Das ist keineswegs der Fall. Viele Modellierungsversuche dringen in diesem Sinne zu einer wesentlich tieferen Ebene vor. So sind die psychologischen Phänomene beim Farbsehen des Menschen erst durch eine computationale Modellierung von entsprechenden Gehirnmolekülen weitestgehend erklärt worden [AAS<sup>+</sup>00, S.130], um nur eines von unzähligen Beispielen aus der intellektischen Hirnforschung [CS95] zu erwähnen.

### 3 Die Sicht der Philosophie

Mit der Informationstechnologie (IT) ist ein neues Zeitalter angebrochen. Darüber sind sich alle Kulturexperten einig. Immer sind die Übergänge in neue Zeitalter von fundamentalen Krisen begleitet gewesen. Dies ist auch in dem jetzigen Übergang deutlich erkennbar [Bib03, S.277]. Vor allem sind von der jetzigen Krise die Geistes- und Sozi-

alwissenschaften betroffen [Bib03, S.286], zu denen auch die Philosophie gehört. Nicht zuletzt die Intellektik macht ihnen allen schwer zu schaffen. Zusätzlich zu den bereits im Einleitungsabschnitt gegebenen Zitaten hierzu soll als weiterer Beleg an dieser Stelle einer der führenden amerikanischen Psychologen, Ward Edwards von der University of California (UCLA), zu Wort kommen [Edw98, S.416], der in diesem – anlässlich seiner Auszeichnung verfassten – Artikel seine Kollegen mit ersichtlich eindringlichen Worten auf die Werkzeuge der Intellektiker, vor allem die Bayesschen Netze, aufmerksam macht.

*... the decision-theoretic tools are for the most part from new subfields of artificial intelligence .... Psychology as a science must decide whether it is or is not a part of .... If it is, then psychologists have some catching up to do; artificial intelligence professionals know what ....*

In der Philosophie hat sich – vor allem in Deutschland – noch kein „Edwards“ in diesem Sinne zu Wort gemeldet, obwohl es dazu allerhöchste Zeit wäre und das Zitat fast wortwörtlich auch für sie zutrifft. Der in Abschnitt 1 zitierte Tetens verkennt sogar das eigentliche Problem, wenn er sagt: „*In der Philosophie des Geistes kann Entwarnung gegeben werden*“ [Tet94, S.121]. Vielmehr zeigen sich bereits Anzeichen einer tieferen Veränderung, wie beispielsweise in der Vorankündigung einer Tagung in Arnoldshain 2003 mit dem Thema *Bildung durch Naturwissenschaft*, wo es heißt: „*In Deutschland fällt Bildung noch immer vorwiegend in die Zuständigkeit der Geistes- und Kulturwissenschaften. Diese Zuordnung ist von der internationalen Vergleichsstudie PISA radikal in Frage gestellt worden.*“ Im Kern geht es, wie im letzten Abschnitt ausgeführt, um die beiden methodologischen Vorteile der Intellektik, die in der Präzision der Begriffsbildung und der daraus resultierenden mächtigen Werkzeuge bestehen, mit denen Experimente wie in den Naturwissenschaften durchgeführt werden können.

Oft in peinlicher Unkenntnis des Entwicklungsstandes der Intellektik wehrt sich die Philosophie mit konkreten Argumenten, die ihre Eigenständigkeit und Unentbehrlichkeit belegen sollen. So versuchte ein Philosoph in der Diskussion zu einem Vortrag des Autors [Bib02] eine solche Rechtfertigung mit dem Hinweis auf die grundsätzlich erforderliche Unschärfe in Begriffen sowie auf den Kontext, als ob die Intellektik nicht längst das präzise Instrumentarium für den Umgang mit unscharfen Begriffen ebenso wie mit Kontext geschaffen hätte. [Tet94] argumentiert, um ein weiteres Beispiel anzuführen, durchweg ausschließlich auf der Grundlage der Identifikation des Maschinenbegriffs mit einem endlichen Automaten, als ob alle Rechenmodelle in der Intellektik (wie zB. Kalküle der Prädikatenlogik oder Bayessche Netze) auf diesen primitivsten Typ zurückgeführt werden könnten (ungeachtet der Tatsache, dass unsere Rechner im Prinzip – aber nicht für diese Argumentation – auch als äußerst komplexe endliche Automaten aufgefasst werden können). Auf weitere Beispiele aus einer langen Reihe derartiger Peinlichkeiten werden wir weiter unten im Kontext von Gödel noch eingehen.

Natürlich ist es nicht nur legitim, sondern auch verständlich, dass eine Disziplin ihre Existenzberechtigung in dieser Weise zu erbringen versucht. Ein stichhaltiges Argument ist dabei aber noch nicht hervorgebracht worden. Natürlich benötigt jede Wissenschaft auch einen präformalen Bereich, in dem die formalen Vorgehensweisen diskutiert und

dargestellt werden; man muss sich ja auch beim Kaffee beispielsweise über die Schrödingergleichung unterhalten können. Es kann aber beim heutigen Stand der wissenschaftlichen Methodik nicht mehr angehen, dass sich eine Disziplin anmaßt zu behaupten, sie könnte auf den formalen Bereich vollständig verzichten, beispielsweise eine „Philosophie des Geistes“ entwickeln und die mühsame Entwicklung der harten Formalismen anderen überlassen. Wenn sich die Philosophie nicht in gleicher Weise präziser Methoden befließigt wie die Intellektik und damit sich zu einer intellektischen Wissenschaft entwickelt, wird sie ihre Existenzberechtigung als akademische Disziplin vollends einbüßen. Philosophie für den Kaffeepausch können auch die anderen Disziplinen bieten.

Eine intellektische Philosophie könnte hingegen wertvolle Beiträge auch für die Natur- und Intellektikwissenschaften liefern, worauf schon in [BHS93, S.310f] hingewiesen wurde. Denn diese sind sich oftmals der zugrundeliegenden „philosophischen“ Annahmen überhaupt nicht bewusst. Mit induktiven Methoden aus der Intellektik könnten solche Annahmen in einem Theoriebildungsprozess formal präzise erarbeitet werden, was von der Zielsetzung her eine typisch philosophische Aufgabe wäre. Der nächste Schritt wäre dann der Vergleich eines so erarbeiteten Annahmenbündels – was man in der Logik eine Theorie nennt – mit anderen (Theorien) in Bezug auf die logischen Konsequenzen, wobei durchaus auch Inferenzsysteme zum Einsatz gebracht werden könnten; denn die Zeit ist schon längst vergangen, in der man derartige Theorievergleiche nur „im Lehnstuhl“ ausführen konnte. Von der Intension her machen Philosophen dies, nur eben auf der Grundlage nicht präzise definierter Begriffe und ohne die Schärfe mathematischer Beweisführung. Insofern erscheint der Vergleich der Philosophie mit der Mathematik in [Tet94, S.141, Fußnote 14] als weit verfehlt.

Als Zwischenbilanz halten wir fest, dass (i) durch die intellektische und IT Herausforderung eine Krise der Philosophie (und aller Geistes- und Sozialwissenschaften) zu konstatieren ist, (ii) der wiederholte Versuch der Philosophie zu einer Revierverteidigung mit einfachen Argumenten gegenüber den Formalisten bislang immer gescheitert ist und dass es (iii) aber eine Fülle genuin philosophischer Aufgaben gibt, die auf eine Lösung mit intellektischen Methoden warten.

Anstatt diese anstehenden Aufgaben in Angriff zu nehmen, haben einige Philosophen und philosophisch gesinnte Autoren noch einen anderen Ausweg aus der Krise versucht: sie meinten – ua. auf der Grundlage des berühmten (ersten) Gödelschen Unvollständigkeitssatzes (kurz Gödelsatz) – einen präzisen Beweis für die Unmöglichkeit künstlicher Intelligenz erbringen zu können, indem sie einen Unterschied zwischen Mensch und Maschine nachweisen. Auch dieses Unterfangen ist gescheitert und war von Anfang an zum Scheitern verurteilt, was ich im Rest dieses Abschnitts nun noch darlegen werde.

Als prominenteste Wortführer dieses letzten Auswegs haben sich die beiden Engländer J.R. Lucas und Sir R. Penrose einen Namen gemacht. Unzählige andere Wissenschaftler geringerer Prominenz haben Ähnliches wie diese beiden versucht. Umgekehrt gibt es unzählige Arbeiten, die Gegenargumente vertreten. So umfasst die Literatur zu diesem Thema inzwischen wohl viele Tausende von Arbeiten, die bis zu Descartes zurückreichen. In [And64, S.1] werden eintausend Arbeiten allein für den Zeitraum zwischen 1950 und

1964 genannt. Die Argumente auf beiden Seiten wiederholen sich in vielfacher Variation. Ein Fortschritt über die Jahrhunderte ist nicht feststellbar.

Hier ist ein erstes und sehr lehrreiches Beispiel. Anfang der Siebziger Jahre – also vor über drei Jahrzehnten – nahm ich an einer Tagung in den USA teil, auf der der Philosoph Hubert Dreyfus einen Vortrag über die Grenzen der KI hielt. Unter Bezugnahme auf Heideggersche Schlussketten stellte er einen Beweis der folgenden Aussage vor: ein Computer wird aus grundsätzlichen Gründen niemals die besten menschlichen Schachspieler der Welt besiegen können. Wie wir aus dem letzten Abschnitt wissen, ist diese Aussage durch die Fakten widerlegt. Der „Beweis“, den man noch immer in seinem Buch von 1972 nachlesen kann, war falsch und damit kein Beweis. Er erfüllte schon aufgrund seiner Form nicht die Minimalansprüche eines jeden Mathematikers an einen Beweis.

Lucas und Penrose haben bewiesen: aus dem Gödelsatz folgt, dass ein Mensch mehr mathematische Fähigkeiten als jede denkbare Maschine hat. Die entsprechende Arbeit von Lucas [And64, S.43–49] enthält nicht eine einzige Definition oder Formel, erfüllt also schon deswegen nicht die Minimalansprüche an einen Beweis. Penrose hat mehrere und voneinander stark differierende Versionen eines Beweises für die Aussage geliefert, die derzeit vorletzte in [Pen94]. Penrose ist von Hause aus Physiker und weiß damit, was ein Beweis ist. Dennoch sind ihm ua. von dem weltbesten Spezialisten für Gödel, dem Logiker Fefermann, eine Fülle von gravierenden Fehlern im Beweis nachgewiesen worden [Fef96]. Da ein einziger Fehler genügt, einen mathematischen Beweis zum Einsturz bringen zu lassen, gibt es keinen Grund dafür, in dem „Beweis“ noch länger herumzustochern oder gar die behauptete Aussage als wahr zu halten.

Penrose ist für seine hervorragenden wissenschaftlichen Leistungen in der Physik geadelt worden. Sein erstes Buch mit einem falschen Beweis hätte man ihm noch verzeihen können. Es ist aber untragbar, dass er daraus nicht Lehren gezogen hat und das Gleiche nochmals versucht hat, nunmehr insgesamt schon auf über Tausend Buchseiten. Er hat ein stattliche Anzahl von hervorragenden Wissenschaftlern damit beschäftigt, sich durch diese Tausend Seiten durchzuwühlen und seine Fehler aufzudecken. Meiner Vorstellung eines ethisch und wissenschaftlich akzeptablen Verhaltens, noch dazu eines weltbekannten Wissenschaftlers, entspricht das absolut nicht.

*„It is clear that Penrose ... is ... searching for a weapon to help him refute a view he finds offensive. When a club breaks he picks up a new one“* [LHF98, S.266]. Leider ist dieses Verhalten nicht untypisch für viele Reaktionen gegen die KI. Greift ein Argument nicht, wird das nächste versucht. So hat sich die Philosophie bei den Intellektikern inzwischen den Nimbus einer „Wissenschaft der Fehlschlüsse“ erworben, auf die man besser überhaupt nicht mehr eingeht. Sie musste sich schon aus den naturwissenschaftlichen Bereichen völlig zurückziehen und ist jetzt dabei, sich auch den letzten (intellektischen) Bereich zu verspielen und sich einer idiosynkratischen Haltung [Tet94, S.116] hinzugeben. Dabei könnte die Philosophie gerade an dem Bereich der Grundlagen der Mathematik, aus dem der Gödelsatz stammt, lernen, dass dieser Bereich erst dann zum Blühen gebracht werden konnte, als die Logiker im Neunzehnten Jahrhundert begannen, sich mathematisch präziser Begriffe und Beweistechniken zu bedienen.

Was aber ist die Bedeutung des so arg strapazierten Gödelsatzes? Mein Doktorvater Schütte pflegte auf eine solche Frage zu antworten: der Inhalt des Satzes (samt Beweis). Der Gödelsatz bezieht sich auf eine genau definierte Klasse formaler Systeme. Informell: Jede hinreichend starke formale (und konsistente) Theorie der Arithmetik ist unvollständig, dh. es gibt Sätze in der Sprache der Theorie, so dass weder der Satz selbst noch sein Gegenteil in der Theorie liegen [BJ89, S.179]. Gödel gibt sogar ein Verfahren an, mit dem man einen solchen wahren Satz, der nicht zur Theorie gehört, konstruieren kann.

Um ihn auf das menschliche Denken anwenden zu können, müsste man dieses in einen präzisen Bezug mit einem formalen System der Arithmetik setzen können. Da niemand genau weiß, wie menschliches Denken genau funktioniert bzw. wie ein solcher Bezug hergestellt werden könnte, muss ein solches Unterfangen beim gegenwärtigen Kenntnisstand – und mutmaßlich noch für sehr lange Zeit – von vorneherein zum Scheitern verurteilt sein. Ein formales System ist etwas Statisches, Unveränderbares. Der Mensch ändert sich durch sensorische und kommunikative Erfahrungen fortwährend. Das Gleiche gilt für moderne Intellektiksysteme (zB. in der Robotik), die beispielsweise in der Lage sind, ganze Module untereinander auszutauschen. Wie sollte beispielsweise ein solcher kommunikativer Systemeingriff in einem formalen System (oder einer Turingmaschine) angelegt sein? Natürlich haben die Intellektiker längst auch Systeme formalisiert, die das Zusammenwirken vieler Akteure modellieren. Die dabei ins Auge gefassten Akteure sind aber noch immer von recht armer Struktur, die mit der Reichhaltigkeit der menschlichen Phänomene recht wenig gemein haben. Mit „handwaving“ Argumenten, wonach man mit dem „Rest“ dann analog verfahren würde, ist es in präzisen Aussagen wie denen im Gödelsatz aber nicht getan. [LHF98] und eine Reihe weiterer Autoren geben noch viele weitere informelle Argumente dafür, dass der Gödelsatz für diese gesamte Diskussion nichts hergeben kann. Allerdings: beweisen, dass er nichts hergibt, kann man auch nicht; das hat aber – wohl mangels einer entsprechenden emotionalen Motivation – auch noch niemand versucht.

Was tut man, wenn man weder das eine noch das andere beweisen kann: man trifft Annahmen. Niemand würde Penrose die Annahme verübeln, Menschen seien keine Maschinen irgendwelcher Art. Und niemand sollte der Intellektik verübeln, dass sie unter der Annahme forscht, alle menschlichen Fähigkeiten ließen sich komputational modellieren. In [Cha95, Abschn.4.7ff] findet sich beispielsweise ein Ansatz möglicher Annahmekombinationen für unser Thema hier. Aber mit Annahmen emotionsfrei umzugehen scheint selbst mathematisch geschulten Philosophen wie der Autorin von [Rhe91, S.3f] schwer zu fallen, die Turing „Behauptungen“ unterstellt, während Turing korrekt von „conjectures“ und „beliefs“ spricht und sogar ausdrücklich den Unterschied zu (behaupteten) Fakten betont [Tur50, S.443].

Mit Recht weist [Rhe91, S.15] schließlich auf die eigentliche Frage hin: *Wie funktioniert das Denken?* Die philosophischen Debatten, von denen in diesem Abschnitt die Rede war, haben zur Beantwortung dieser Frage so gut wie nichts beigetragen. Im Vergleich damit ist der durch die Intellektikforschung in wenigen Jahrzehnten erzielte Fortschritt sensationell. Absolut gesehen, das heißt gemessen an der Schwierigkeit der Aufgabenstellung, ist er natürlich noch höchst bescheiden, nimmt aber täglich weiter zu und liefert vor allem echte Hilfen für die täglichen Probleme von Mensch und Gesellschaft. Statt sich Glas-

perlenspielen hinzugeben, sollten sich auch Philosophen in einer derart nützlichen Weise betätigen.

## 4 Perspektiven der Intellektik

Nach der im letzten Abschnitt dargelegten kritischen Einschätzung philosophischer Methoden soll hier der Faden aus dem Abschnitt 2 mit einer Darstellung der Perspektiven wieder aufgenommen werden, die sich aus einer Weiterentwicklung ergeben könnten. Es sei ausdrücklich betont, dass jede technologische Entwicklung auch große Gefahren in sich birgt, auf die ich beispielsweise in [Bib89] ausführlich eingegangen bin. Ich beschränke mich hier daher allein auf positiv einzuschätzende Chancen.

Natürlich kann man in jedem einzelnen Fall darüber debattieren, ob eine technologisch herbeigeführte Veränderung von Vorteil ist oder nicht. Unterlegen wir daher dieser Diskussion das folgende Wertungskriterium: wenn die Veränderung eine Tätigkeit verbessert (bequemer, weniger zeitaufwendig usw. macht), die auch vorher schon allgemein als nützlich eingeschätzt wurde, dann ist die Veränderung positiv einzuschätzen.

In dem so definierten Sinne hat die gesamte IT einschließlich der KI bereits eine Fülle positiver Veränderungen erbracht. Ein naheliegendes Beispiel ist die von mir in diesem Moment ausgeübte Tätigkeit des Schreibens eines solchen Aufsatzes. Einen umfassenden Überblick über die zugrundeliegende Literatur holt man sich innerhalb weniger Stunden aus dem Internet, wofür ich als junger Wissenschaftler Monate benötigt hätte. Die Textformatierung, Fehlerkorrektur, Literaturverzeichniserstellung wird nahezu vollständig von zugrundeliegenden Systemen übernommen und liefert ganz nebenbei ein druckreifes Resultat. Mit Grauen denke ich vergleichsweise noch an das Abtippen von (mathematischen) Texten durch Sekretärinnen und das mühevollen Bemühen, diese in endlosen Versionen fehlerfrei zu bekommen, an das Korrekturlesen von Druckfahnen usw. usf. Der Fortschritt ist für diese Tätigkeit unbestritten revolutionär. Und er ist es für Tausende weiterer Tätigkeiten (Kommunikation, Bankwesen, Produktentwicklung und Produktion, Verwaltung, Medien usw.).

Der Transfer wissenschaftlicher Erkenntnis in die Anwendung ist nicht einfach. Am schnellsten gelingt er innerhalb eines Gebietes, weshalb die IT selbst bislang am meisten von ihren eigenen Erkenntnis profitiert hat, während für den IT-Fachmann die Unkenntnis in den anderen Gebieten oft unverständlich ist. Dies gilt auch für Intellektiker, vor allem wenn sie mit den verwandten Disziplinen Philosophie, Psychologie usw. in Kontakt kommen, die ja eigentlich an gleichen Zielsetzungen arbeiten. Hiervon war im letzten Abschnitt ausgiebig die Rede. Vor allem haben wir dort auf wichtige intellektische Fragen hingewiesen, die Philosophen mit intellektischen Methoden endlich anpacken könnten.

Im Abschnitt davor habe ich bereits von den großen Herausforderungen im Hinblick auf die komputationale Modellierung des menschlichen Denkens und der menschlichen Psyche bis hin zu Intentionalität und Bewusstsein gesprochen. Die sich daraus ergebenden – im strengen Sinne – wissenschaftlich gesicherten Erkenntnisse könnten segensreichste Veränderungen mit sich bringen. Denken wir beispielsweise an das große Feld psychischer

Schwächen, Leiden und Erkrankungen, etwa an die weitverbreiteten Ängste von Menschen. Wüßte man genauer, worauf sie beruhen, würde man sicher auch Verfahren finden, sie zu mildern oder sogar zu beseitigen.

Oder denken wir an Erziehung und Schule. Pädagogik ist wie alle Geistes- und Sozial-,wissenschaften“ keine Wissenschaft im strengen Sinne, könnte es aber mit dem Einsatz intellektischer Methoden werden. Denn hätte man Modelle lernender Kinder, könnte man verschiedene Methoden experimentell im Hinblick auf ihre Ergebnisse miteinander vergleichen, was mit Menschen grundsätzlich nicht geht. Pädagogen haben davon leider noch nicht die geringste Ahnung. Sie führen bis heute endlose ideologische Positionskämpfe und zählen Noten, Schüler und deren Merkmale, was zwar erste Anhaltspunkte (wie durch PISA), aber in keinem Fall gesicherte und nachprüfbar Theorien liefert.

Was für die Pädagogik gilt und dem Leser vielleicht selbst in dieser kurzen Andeutung einleuchten könnte, ist in gleicher Weise auf alle Bereiche anwendbar, die von menschlichem Geist und Psyche geformt werden – und welche sind das nicht. Dazu gehören die Mechanismen des Gemeinschaftslebens in Familie, Ortsgemeinschaft, Unternehmen, Parteien usw., die Organisationsformen der Wirtschaft und des Staates und deren Verwaltungen, die Bereiche Politik, Justiz, Gesundheitswesen, Sozialsysteme, Verkehr, innere und äußere Sicherheit usw. Für Einzelheiten zu all diesen Themen verweise ich auf [Bib03, Kap.4]. Das Potenzial für eine intellektische Revolution unserer Gesellschaft und zur Überwindung vieler ihrer gravierenden Defizite, über die täglich in den Medien lamentiert wird, ist schon aufgrund der bereits erforschten Erkenntnisse immens. Nur hat davon außerhalb des IT-Bereichs kaum jemand mehr als einen Schimmer von Ahnung.

Auch innerhalb des – ausschließlich vom menschlichen Geist geprägten – Wissenschaftsbereiches reicht die Potenz der Intellektiktechnologie weit über die IT hinaus, wobei wir den Bereich der Geistes- und Sozialwissenschaften schon ausführlich behandelt haben. Beispielsweise ist im Abschnitt 2 bereits das automatische Beweisen für die Mathematik genannt. Weltweit angelegte Projekte (wie zB. Mizar) haben die maschinelle Korrektheitsprüfung aller bekannten mathematischen Beweise zum Fernziel. Über formale Methoden für den Begriffs- und Theoriebildungsprozess, für logisches Schließen in Diagrammen, für informellere Schlussweisen, für Beweisstrukturierung und -präsentation wird intensiv geforscht. Auch die deutschen Mathematiker verschließen vor dieser Entwicklung die Augen.

Intellektiktechnologie dringt bereits stärker in die technischen Disziplinen ein, da Ingenieure offenbar aufgeschlossener gegenüber neueren Entwicklungen sind. Auch in den Naturwissenschaften, allen voran der Biologie, wurde in den letzten Jahren ein erstes vorsichtiges Interesse geweckt. Genaueres zu diesem Thema in Bezug auf den gesamten Wissenschaftsbereich findet sich in [Bib03, Kap.5].

## 5 Schlussbemerkungen

Dem Auftrage nach sollte dies ein Aufsatz zum Thema *Gödel und KI* aus intellektischer Sicht werden. Das Thema ist eine bescheiden formulierte Variante der provozierenderen Fragen: „Können Maschinen denken?“ oder auch „Haben Menschen von Maschinen prin-

ziell nie erreichbare Fähigkeiten?“ Wie im Abschnitt 3 ausgeführt, sind diese Themen nicht allzu fruchtbar und haben seit Jahrhunderten zu einem Austausch der immer wieder gleichen Argumente geführt. Die Aufgabe konnte daher von einem Intellektiker nur so verstanden werden, den bis heute nicht müde werdenden Diskutanten über dieses fruchtlose Thema den beispiellosen Erfolg und die Perspektiven der Intellektik als Kontrast vor Augen zu führen; dies ist in den Abschnitten 2 und 4 geschehen. Vielleicht führt dieser Kontrast ja dazu, dass der eine oder andere sein bislang fruchtloses Bemühen gegen einen konstruktiven Beitrag zu den faszinierenden Fragen der Intellektik eintauscht.

Natürlich bewegen uns die obigen Fragen, auch wenn sie wissenschaftlich unfruchtbar sind. Wie jede revolutionierende Technologie bereitet auch die der Intellektik vielen Menschen ein großes Unbehagen. Daraus resultierende emotionale Angriffe sind daher psychologisch durchaus verständlich und legitim, solange sie nicht in verwerflicher Absicht mit einem pseudowissenschaftlichen Mäntelchen behängt werden. Wie immer in vergleichbaren Situationen ist Aufklärung das Mittel der Wahl, das ich auch hier einzusetzen versucht habe. In der Kürze eines solchen Aufsatzes kann sie aber über zusammenfassende Beschreibungen und zugehörige Literaturhinweise nicht hinausgehen. Die Entscheidung zugunsten eingehenderen Studien muss nun dem Leser überlassen bleiben.

Im Hinblick auf eine solche Entscheidung soll nochmals auf die Geschichte verwiesen werden. Alle wissenschaftlichen Revolutionen (Galileo, Kopernikus, Darwin usw.) sind in gleicher Weise wie die Intellektik angefeindet worden. Noch im 19. Jahrhundert hat ein bis dahin angesehener Mathematiker einen exakten Beweis dafür geliefert, dass die damals entworfenen Flugmaschinen grundsätzlich nie fliegen werden können. Heute sind ständig Tausende der so konzipierten Flugmaschinen in der Luft, befördern Hunderttausende von Menschen und wir verehren die Pioniere der Flugtechnik. Von dem irrenden Mathematiker und seinem „Beweis“ spricht dagegen niemand mehr.

## Literatur

- [AAS<sup>+</sup>00] Rita L. Atkinson, Richard C. Atkinson, Edward E. Smith, Daryl J. Bem und Susan Nolen-Hoeksema. *Hilgard's Introduction to Psychology*. Harcourt College Publishers, Fort Worth, 13. Auflage, 2000. Inzwischen in deutscher Fassung [AAS<sup>+</sup>01] verfügbar.
- [AAS<sup>+</sup>01] Rita L. Atkinson, Richard C. Atkinson, Edward E. Smith, Daryl J. Bem und Susan Nolen-Hoeksema. *Hilgards Einführung in die Psychologie*. Spektrum Akademischer Verlag, Heidelberg, 2001.
- [And64] Alan Ross Anderson (Hg.). *Minds and Machines*. Contemporary Perspectives in Philosophy. Prentice Hall, Englewood Cliffs NJ, 1964.
- [Ari95] Aristoteles. *Über die Seele*. Meiner, Hamburg, 1995.
- [BG01] Wolfgang Bibel und Gunter Grieser. Maschinelles Lernen und Automatische Reflexion. *thema Forschung*, (1):96–101, 2001.

- [BHS93] Wolfgang Bibel, Steffen Hölldobler und Torsten Schaub. *Wissensrepräsentation und Inferenz*. Vieweg Verlag, Braunschweig, 1993.
- [Bib80] W. Bibel. “Intellektik” statt “KI” — Ein ernstgemeinter Vorschlag. *Rundbrief der Fachgruppe Künstliche Intelligenz in der Gesellschaft für Informatik*, 22:15–16, Dezember 1980.
- [Bib89] W. Bibel. The Technological Change of Reality — Opportunities and Dangers. *AI & Society*, 3:117–132, 1989.
- [Bib92] W. Bibel. Intellectics. In S. C. Shapiro (Hg.), *Encyclopedia of Artificial Intelligence*. John Wiley, New York, 1992.
- [Bib02] Wolfgang Bibel. Knowledge in IT – absolutely crucial, mostly ignored. [www.Intellektik.Informatik.TU-Darmstadt.de/~Bibel](http://www.Intellektik.Informatik.TU-Darmstadt.de/~Bibel), June 2002. Invited Talk on the Conference on Logic and Knowledge, Darmstadt, 14.–1.6.02.
- [Bib03] Wolfgang Bibel. *Lehren vom Leben – Essays über Mensch und Gesellschaft*. Sozialwissenschaft. Deutscher Universitäts-Verlag, Wiesbaden, 2003.
- [BJ89] George S. Boolos und Richard C. Jeffrey. *Computability and Logic*. Cambridge University Press, Cambridge, 3 Auflage, 1989.
- [Cha95] David J. Chalmers. Minds, Machines, and Mathematics. *Psyche*, 2(9), June 1995.
- [CS95] Patricia S. Churchland und Terrence J. Sejnowski. *Das Neurohirn*. Vieweg Verlag, Braunschweig, 1995. Originalausgabe: The computational brain, MIT Press 1992.
- [Edw98] Ward Edwards. Hailfinder – Tool for and Experiences With Bayesian Normative Modeling. *American Psychologist*, 53(4):416–428, 1998.
- [Fef96] Solomon Feferman. Penrose’s Gödelian argument. *Psyche*, 2:21–32, 1996.
- [FF63] E. A. Feigenbaum und J. Feldman (Hg.). *Computers and Thought*. McGraw-Hill, New York NY, 1963.
- [LHF98] Geoffery LaForte, Patrick J. Hayes und Kenneth M. Ford. Why Gödel’s Theorem Cannot Refute Computationalism. *Artificial Intelligence*, 104:265–286, 1998.
- [McC95] John McCarthy. Awareness and Understanding in Computer Programs. *Psyche*, 2(11), 1995.
- [Nil98] N. J. Nilsson. *Artificial Intelligence: A New Synthesis*. Morgan Kaufmann, San Francisco, CA, 1998.
- [NS72] A. Newell und H. A. Simon. *Human Problem Solving*. Prentice-Hall, Englewood Cliffs NJ, 1972.

- [Pen94] Roger Penrose. *Shadows of the Mind*. Oxford University Press, New York, 1994.
- [Pla02] Platon. *Sämtliche Werke*, volume 2. rororo, Reinbek, 29. Auflage, 2002. Übersetzt von Friedrich Schleiermacher.
- [Rhe91] Rosemarie Rheinwald. Menschen, Maschinen und Gödels Theorem. *Erkenntnis*, 34:1–21, 1991.
- [RN03] Stuart J. Russell und Peter Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall Series in Artificial Intelligence. Pearson Education, Upper Saddle River, N.J., 2 Auflage, 2003.
- [Tet94] Holm Tetens. *Geist, Gehirn, Maschine – Philosophische Versuche über ihren Zusammenhang*. Reclam, Stuttgart, 1994.
- [Tur50] Alan M. Turing. Computing Machinery and Intelligence. *Mind*, 59:433–460, 1950. Also in: [FF63], pp. 11–35, and [And64], pp. 4–30.