

The Blind's World I

Gerd Döben-Henisch

Ein philosophisches Experiment auf dem Weg zum digitalen Bewußtsein

1 EINLEITUNG

Wer die Installation "The BLIND's WORLD I" (= BW1) betritt, dessen Blick wird zuallererst auf die bunten lustigen Pictogramme fallen, die die virtuelle Welt von BW1 repräsentieren. Die Assoziation zu einem Computerspiel drängt sich unweigerlich auf. Doch wäre der weitere Schluß, daß es sich bei BW1 auch "nur" um ein Spiel handelt, ein Fehl-Schluß.

BW1 versteht sich in erster Linie als ein philosophisches Experiment. BW1 soll entscheidend mit dazu beitragen, zentrale Fragen der modernen Philosophie zu beantworten. z. B.:

- Was ist das Bewußtsein?
- Welche Funktionen lassen sich im Bewußtsein unterscheiden?
- Wie entsteht ein "Wissen von der Welt"?
- Wie entwickelt sich Sprache?
- Welchen Einfluß haben die Gefühle?

Die Auffassung von BW1 als einem "philosophischem Experiment" steht in einiger Spannung zu einer Auffassung von Philosophie, die philosophische Erkenntnis ansieht als "in sich stehend", "nur auf sich beruhend", als "autark", "keiner zusätzlichen Hilfsmittel bedürftend", mit dem Ziel, das "Wesen der Dinge" zu erfassen, "ewige Wahrheiten" bzw. "allgemeingültige Prinzipien allen Erkennens" aufzudecken.

Dazu kommt, daß das Experiment als solches von Philosophen in der Regel dem empirischen Erkenntnisparadigma zugeschlagen wird, war es in der Vergangenheit doch gerade die "Kanonisierung des Experimentes" als Gültigkeitskriterium wissenschaftlicher Aussagen, die zur Abspaltung der modernen Naturwissenschaften von der klassischen Philosophie geführt haben. Für viele Philosophen — und wohl für die Mehrheit der Naturwissenschaftler — gilt diese "Spaltung" bis heute als unüberbrückbar.

BW1 hält noch eine weitere Provokation bereit: die Rede vom Bewußtsein. In den modernen experimentellen Wissenschaften ist kein Platz für den Begriff eines "sich selbst gewissen Bewußtseins", nicht einmal im Rahmen der Psychologie. Selbst in der modernen Philosophie ist dieser Begriff in bestimmten Richtungen — wie z.B. in der analytischen Philosophie — stark diskreditiert.

Um also die philosophische und wissenschaftliche Bedeutung von BW1 verstehen zu können, bedarf es einiger Erläuterungen, die über eine bloße Deskription des Programms von BW1 hinausgehen.

Die Geschichte von BW1 beginnt mit Alan Matthew TURING.

2 VISIONÄR TURING

1936 erschien die berühmte Arbeit *On computable numbers with an Application to the Entscheidungsproblem* von Alan Matthew TURING. Mit dieser Arbeit beantwortete er nicht nur die Frage nach der Entscheidbarkeit der Mathematik, die HILBERT 1928 explizit gestellt hatte, negativ, sondern er stellte in ihr auch eine Definition von berechenbaren Prozessen vor, die schon zu seinen Lebzeiten unter der Bezeichnung Turing Maschine [TM] zum Allgemeingut der Mathematiker und Logiker wurde. Für GÖDEL war die TM die befriedigendste aller vorgeschlagenen Definitionen eines "mechanischen Verfahrens" (DAVIS 1965:72 Anmk).

Mit dem Konzept der Turing Maschine, insbesondere mit ihrer verallgemeinerten Form, der Universellen Turing Maschine [UTM], verbindet sich die Hypothese, daß sich alle berechenbaren Prozesse durch sie darstellen lassen.

TURING war mit seiner idealen Maschine den technischen Möglichkeiten seiner Zeit weit voraus, desgleichen mit seinen praktischen und philosophischen Visionen, zu denen er durch sein neu gefundenes Konzept angeregt wurde.

Seine von vielen als provozierend empfundenen Überlegungen lassen sich vor allem zwei Themenkreisen zuordnen: (1) Bau eines elektronischen Gehirns und (2) Wie können Computer lernen?

Trotz seines starken Interesses für konkrete Physiologie und speziell auch Neurophysiologie — schrieb er doch 1952 eine der ersten bahnbrechenden mathematischen Arbeiten zur Chemie der Morphogenese — lehnte er es ab, die physiologischen Strukturen des Gehirns in der Hardware zu imitieren. Er war vielmehr an der Analyse der logischen Struktur des Gehirns interessiert und er ging davon aus, daß sich jedes stetige System mit beliebiger Genauigkeit durch ein diskretes System annähern läßt. Diese Annahmen eröffneten die Perspektive, ein durch diskrete Zustände approximiertes Gehirn durch eine UTM zu simulieren.

Auf der Basis einer solchen strukturellen In-Beziehung-Setzung von menschlichem Gehirn und UTM lagen natürlich auch Überlegungen zu einer möglichen Imitation der menschlichen Intelligenz durch eine UTM nahe. Schon 1941 hatte sich TURING mit der Frage von schachspielenden Maschinen beschäftigt und diese Fragen dann dahingehend ausgedehnt, wieweit eine Maschine, d.h. eine UTM, im allgemeinen Sinne lernen könne.

Bei der Verfolgung dieser Frage wirkt TURING ein wenig zwiespältig (siehe dazu [TURING 1948] und [TURING 1950]).

Einerseits ist er sich offensichtlich bewußt, daß zu einer allgemeinen Lernfähigkeit, so, wie sie der Mensch besitzt, ein entsprechend elaborierter Weltbezug gehört. Man müsse eine solche Maschine mit Fernsehkameras, Mikrofonen, Lautsprechern etc. ausstatten, damit sie möglichst umfassend zu Interaktionen mit der Außenwelt fähig sei, ebenso müßte sie "über Land streifen können", sie müsse "Eigeninitiative" besitzen, sie müsse "trainiert" und "unterrichtet" werden, kurzum alles, was ein menschliches Kind zur Verfügung hat, um zu lernen, das müsse man auch einer Maschine zur Verfügung stellen.

An anderen Stellen sprach er sich allerdings dagegen aus, den Menschen in seinen natürlichen Eigenschaften übermäßig imitieren zu wollen. Auch Maschinen mit einem reduzierten Körper waren für ihn interessante Kandidaten.

Das Erkennen des Vorliegens von "künstlicher Intelligenz" sollte mittels des Kriteriums der Imitation gewährleistet werden. Immer dann, wenn ein Mensch in einem ausschließlich durch einen über ein Terminal geführten Dialog zur "Meinung" käme, daß der Gesprächspartner ein Mensch sein könnte, dann könnte man auf die "künstliche Intelligenz" auch die Beschreibungsprädikate anwenden, die man sonst nur auf den Menschen anwenden würde.

3 BEDEUTUNG, WELTWISSEN UND BEWUSSTSEIN

Die durch TURING ausgelöste Diskussion, ob Maschinen nun eine Intelligenz entwickeln können, die der menschlichen vergleichbar ist, ob sie gar so etwas wie "Bewußtsein" entwickeln können, hält bis heute an und kann eigentlich nicht als definitiv entschieden gelten. Trotz aller faktischen Gewöhnung an Computer verbleibt der Fragestellung eine hohe philosophische Brisanz (DÖBEN-HENISCH 1993). Ein Schlüssel zur weiteren Klärung dieser Fragestellung scheint bei der Sprache zu liegen: dies sah auch schon TURING sehr klar. Im Rahmen der Taxierung möglicher Einsatzbereiche für eine intelligente Maschine stellt er z.B. fest: "Das Lernen von Sprachen wäre unter den oben genannten Anwendungen die beeindruckendste, weil es die menschlichste dieser Tätigkeiten ist" (TURING 1948; dt 1987:98).

TURING hat auch schon, wenngleich halbherzig, vorausgesehen, daß eine UTM, die ähnlich wie ein Mensch in der Lage sein soll, eine beliebige natürliche Sprache zu lernen und dann auch situationsgemäß anzuwenden, dazu auch die entsprechenden Interaktionen mit der Welt benötigt, um jenes Weltwissen erwerben zu können, das notwendig ist, um eine Sprache angemessen verwenden zu können (z.B. TURING 1948: dt 1987:98).

Doch aus der bloßen Feststellung, daß man Weltwissen benötigt, folgt noch in keiner Weise, (1) welches Wissen das genau ist, und (2) wie dieses Wissen erworben werden kann und auch nicht, (3) wie dieses Wissen intern repräsentiert werden muß, damit es mit dem Sprachsystem in Wechselwirkung treten kann.

Ohne Festlegung des theoretischen Rahmens, innerhalb dessen die obigen Fragen beantwortet werden sollen, ist die Beantwortung von (1) bis (3) beliebig, da ortlos. Dies gilt auch für Fragen, wie: (4) wie entsteht ein Sprachsystem in einem Sprecher-Hörer und (5) wie kann das Sprachsystem mit dem Weltwissen in Beziehung treten.

Im Gegensatz zu den empirischen Wissenschaften spielten die Fragen (2), (4) und (5) in der Philosophie bis heute — wenn überhaupt — nur eine marginale Rolle. Dies hatte zur Folge, daß die Fragen (1) und (3) sowie spezielle Teilaspekte von (5), wie z.B. die Frage der Bedeutung sprachlicher Ausdrücke, nur als isolierte statische Aspekte eines an sich dynamischen Lerngeschehens behandelt wurden. Nur vor diesem Hintergrund kann man ansatzweise verstehen, warum es gerade auch in der Philosophie zu der Auffassung kommen konnte, daß sich die Bedeutung sprachlicher Ausdrücke auch ohne Bezugnahme auf Bewußtseinstatbestände rekonstruieren läßt. (Zu verweisen ist hier z.B. auf [FREGE 1892], [WITTGENSTEIN 1921], [DAVIDSON/ HARMAN 1972], [BARWISE/ PERRY 1983], um nur einige zu nennen.)

Nimmt man hingegen den Standpunkt des Lernens bzw. des Lernenden — hier als "Agent" bezeichnet — ein und akzeptiert man die Fragen (2), (4) und (5) als Leitfragen, dann wird man gezwungen, nicht nur ein Agenten-Welt-System als minimalen Rahmen anzunehmen, sondern man muß sich auch über das "Welt-Interface" des Agenten Gedanken machen, wie

auch über all jene "internen Zustände" des Agenten, die notwendig sind um die durch (2), (4) und (5) geforderten Prozesse realisieren zu können.

Turing selbst hat diese Fragen nur ansatzweise erörtert.

Um die interne Funktion f eines Agenten konstruieren zu können, muß man eine Grundsatzentscheidung treffen: Sollen die Annahmen über f (a) beliebig sein oder (b) will man sich an bestimmten Vorgaben orientieren?

Eine Entscheidung für (a) würde das Projekt der Konstruktion von f dem Bereich einer allgemeinen Strukturwissenschaft wie der Mathematik zuordnen, eventuell eingeschränkt auf die berechenbaren Funktionen.

Eine Entscheidung für (b) fordert zwei weitere Entscheidungen heraus: Will man die Konstruktion von f (b.1) abhängig machen von Verhaltensdaten menschlicher Versuchspersonen oder (b.2) will man das Wissen einbeziehen, über das jeder Mensch in der Perspektive seines Selbstbewußtseins verfügt.

Eine Entscheidung für (b.1) führt zum Paradigma der Kognitionswissenschaften, die auf der Basis von Verhaltensdaten (Phonetik, Sprachpsychologie, Physiologie, ...) versuchen, Computermodelle zu erarbeiten, die in ihren Funktionen möglichst weitgehend mit diesen Verhaltensdaten übereinstimmen sollen. Das Vorgehen nach (b.1) besitzt jedoch mindestens einen gravierenden Mangel: eine Modellbildung auf der Basis von Verhaltensdaten ist, bezogen auf die formal möglichen Strukturen interner Prozesse, hochgradig unterbestimmt.

Es verbleibt die Frage, was mit Variante (b.2) ist. Zunächst steht dieser Variante das heute weit verbreitete "Vor-Urteil" entgegen, daß Bewußtseinsdaten für eine ernsthafte Rekonstruktion unbrauchbar sind.

Für eine Einbeziehung von Bewußtseinsdaten sprechen jedoch ein Reihe von philosophischen Arbeiten, die zeigen, daß eine Vielzahl von Erkenntnisaspekten nur durch Bezugnahme auf das eigene Bewußtsein zugänglich sind (z.B. [MACH 1922], [HUSSERL 1913], [MERLEAU-PONTY 1945]). Eine indirekte Bestätigung der philosophischen Argumente findet sich auch in experimentellen psychologischen Arbeiten zur Wahrnehmung und zum Gedächtnis (z.B. in [MURCH/ WOODWORTH 1978], [KLIX 1980], [HOFFMANN 1982], [SHIFF 1980]).

Dieser hier grob als phänomenologisch zu bezeichnende Erkenntnisansatz zeichnete sich in der Vergangenheit allerdings durch das Fehlen einer hinreichenden Sprachkritik aus, was seine Akzeptanz im Rahmen der Philosophie bis heute deutlich beeinträchtigte. Durch Einbeziehung der neuzeitlichen Sprachkritik in die Phänomenologie sowie der Entwicklungen in der modernen Wissenschaftstheorie nach 1970 ((LUDWIG 1978). (BALZER/ MOULINES/ SNEED 1987)) läßt sich jedoch ein Erkenntnisparadigma formulieren, das es ermöglicht, auf der Basis einer phänomenologischen Analyse formale Theorien von Bewußtseinsstrukturen zu erarbeiten, die eine formale Theoriebildung erlaubt, die "nahezu alle" (DÖBEN-HENISCH 1994c) Aspekte des Bewußtseins darzustellen erlaubt und diese Darstellung kontrollierbar macht.

Der soeben unter (b.2) skizzierte Lösungsansatz ist so angelegt, daß er die Anwendung der Theorie auf sich selbst erlaubt (eine Antwort auf (NAGEL 1986)). Durch die Wahl der Mittel ist eine solche philosophische Theorie ferner "kompatibel" mit jeder empirischen Theorie, da

sie jede mögliche empirische Theorie auf formaler Ebene als Teiltheorie enthält. Die Unüberbrückbarkeit zwischen Naturwissenschaften und Geisteswissenschaften erweist sich aus dieser Sicht als bloßes Artefakt falsch gewählter Paradimgrenzen.

Entscheidend ist nun, daß sich auf der Basis einer solchen formalen philosophischen Theorie des Selbstbewußtseins Simulationsmodelle definieren lassen. Eine spezielle Teilklasse von Simulationsmodellen sind solche, die sich auf den Einsatz von UTM's beschränken.

Simulationsmodelle sind jedoch keinesfalls bloße "Abfallprodukte" einer in sich abgerundeten formalen Theorie des Bewußtseins. Aufgrund der schwindelerregenden Komplexität der Bewußtseinsprozesse muß man vielmehr annehmen, daß die Simulationsmodelle ein notwendiges Hilfsmittel sind, um überhaupt Theorien des Selbstbewußtseins mit einem gewissen Komplexitätsgrad ausarbeiten zu können. In diesem Sinne sind Simulationsmodelle Instrumente für philosophische Experimente.

Diese hier nur kurz skizzierte Vorgehensweise kann eine neue Disziplin begründen, die man, je nach Standpunkt, entweder als Computergestützte Philosophie [CGP] (ComputerAided Philosophy [CAP] — oder besser Computational Philosophy [CP] -) bezeichnen könnte, wenn man sich mehr dem philosophischen Erkenntnisinteresse zugehörig fühlt, oder als Künstliches Bewußtsein [KB] (Artificial Consciousness [AC]), will man sie als neuer Sparte neben der bisherigen Künstlichen Intelligenz [KI] (Artificiat Intelligence) im Rahmen einer allgemeinen Informatik betreiben.

4 KNOWBOTIC INTERFACE

Die Umsetzung des Konzeptes einer computergestützten Philosophie setzt voraus, daß man simultan sowohl eine formale Theorie des Selbstbewußtseins entwickelt wie auch ein zu dieser Theorie passendes Simulationsmodell.

BW1 soll der erste Prototyp eines solchen theoriegeleiteten Simulationsmodells des Selbstbewußtseins sein, das im Rahmen eines Knowbotic Interface [KInt] realisiert wird.

Im Frühjahr 1994 zum ersten Mal formuliert (DÖBEN-HENISCH 1994a) und in einem Forschungsseminar 1994 zusammen mit Prof. HOCHÉ in Bochum theoretisch weiter ausgearbeitet (siehe DÖBEN-HENISCH 1994b), greift die Idee des KInt neben der Bewußtseinsproblematik einen Gedanken TURINGS wieder auf, der seine Vision von einer lernenden Maschine in die Gestalt einer Kind-Maschine gekleidet hatte: die Kind-Maschine — die natürlich eine UTM ist — sollte so beschaffen sein, daß sie wie ein Kind einem Erziehungs- und Lernprozeß unterworfen werden könnte (TURING 1950:1987 pp.177f).

Mit dem KInt wird ein Softwarebaukasten zur Verfügung gestellt, der es erlaubt, genau solche UTM-Kind-Maschinen zu definieren wie sie TURING vorschwebten, dazu beliebige UTM-Welten, in die auch "normale Menschen" "eintreten" können, letztere allerdings unter dem Gewande von UTM-Kind-Maschinen. Diese UTM-Kind-Maschinen nennt der Verfasser Knowbots (to know + robot = knowbot) und die UTM-Kind-Maschinen, die Menschen "verkleiden", nennt er Pseudo-Knowbots. (Die Anregung zum Begriff Knowbot bekam der Verfasser aus intensiven Gesprächen mit Christian Hübler von der Gruppe knowbotic research. Die Gruppe kr+cf verwendet auch die Bezeichnung knowbots, allerdings in einem anderen Sinne. Ferner wird der Begriff "knowbot" im Internet gelegentlich für intelligente Agenten benutzt, die im Netz verschiedenste Informationen sammeln. Diese "gewöhnlichen" Knowbots haben mit den Knowbots vom KInt nur den Namen gemein.)

Das KInt bietet somit die Möglichkeit, Knowbots von Menschen trainieren und unterrichten zu lassen, ohne daß die Knowbots die Pseudo-Knowbots notwendigerweise als etwas von ihnen Verschiedenes erkennen müssen. Damit erschließt das KInt eine neue Variante des TURINGschen Imitationstests.

Gegenüber TURING wird in KInt entscheidendes Gewicht darauf gelegt, daß sowohl die innere Struktur der Knowbots, d.h. ihr Bewußtsein, wie auch die Struktur der KInt-Welt mit der uns Menschen bekannten Welt soweit übereinstimmt, daß im Prinzip alle bedeutungsrelevanten Sachverhalte simuliert werden können, die im Kontext der dem Menschen bekannten natürlichen Sprachen auftreten können.

Aufgrund der Flexibilität des KInt kann man natürlich auch andere Welt- und Bewußtseinsstrukturen als die aus der menschlichen Welt bekannten simulieren. Prinzipiell könnte man im Rahmen des KInt daher auch Themen behandeln, die sonst nur im Rahmen von Science-fiction-Romanen vorkommen: künstliche Bewußtseinsstrukturen, die sich von den menschlichen substantiell unterscheiden, die ihre eigene Sprache sprechen, die eventuell unsere menschliche Sprachen verstehen, wir aber nicht ihre. Im Rahmen des KInt kann man diese Experimente direkt ausführen.

5 THE BLINDS WORLD I

Die Grundstruktur von BW1 ist in Bild 1 zu sehen.

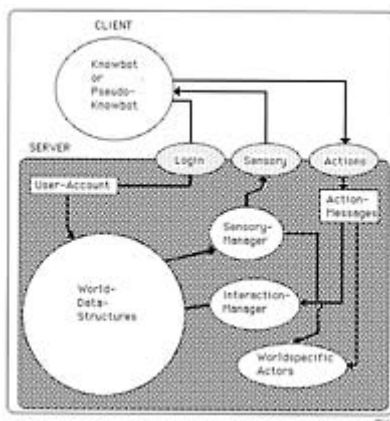


Bild 1

Zunächst wird zwischen einem Server-Programm unterschieden, das eine bestimmte Welt verwaltet, und einem Client-Programm. Letzteres kann ein Knowbot oder ein Pseudoknowbot sein. Client- und Serverprogramme können entweder auf dem gleichen Rechner oder auf verschiedenen Rechnern gestartet werden, d.h. diese Client-Server-Architektur ist voll netzwerkfähig. Es können auch mehrere Server-Programme gleichzeitig nebeneinander auf dem gleichen Rechner oder auf verschiedenen Rechnern existieren.

Die Welt, die durch das Server-Programm realisiert werden kann, wird zu Beginn aus einer Textdatei geladen und intern vom Server als Weltdatenstruktur aufgebaut. Die Textdatei kann wie ein normales Textdokument editiert werden, d.h. der Benutzer kann sich nach Belieben eine Welt definieren: das Innere eines Hauses, eine Stadt, eine Behörde, ein bestimmtes Land oder einen ganzen Planeten. In der ersten Version sind die möglichen Welten eingeschränkt auf 2-dimensionale Darstellungen, in denen 5 Schichten unterschieden werden. Die Welten von BW1 ähneln daher alle bunten Landkarten, in denen sich farbige Pictogramme hin und her bewegen.

Sämtliche Objekte von BW1 sind multisensorielle Objekte. Jedes Objekt hat nicht nur in der visuellen Dimension eine Gestalt mit innerer Struktur und farblichen Mustern, sondern es kann in der akustischen Dimension mit Geräuschen verbunden sein, in der olfaktorischen mit einem Geruch, in der gustatorischen mit einem Geschmack und in der taktilen mit diversen Berührungswerten. Diese Werte können sich außerdem in Abhängigkeit von inneren Zuständen oder von äußeren Einflüssen beständig ändern. Dieser multisensorielle Charakter aller Objekte ist eine notwendige Konsequenz der Forderung, daß die (Pseudo)Knowbots mit einer BW1-Welt ausschließlich über Sinnesorgane kommunizieren sollen.

Die prototypische Welt von BW1 ist insgesamt eine sehr einfache Welt: eine Welt wird als eine große Fläche angesehen, deren Enden z.Zt. nicht festgelegt sind. In Anlehnung an den biblischen Schöpfungsbericht besteht eine Welt grundsätzlich aus Wassermassen, in die man nach Belieben isolierte oder verbundene Landmassen einfügen kann. Jedes Land kann mittels einer Liste verfügbarer Landschaftstypen gestaltet werden. Es stehen z.B. zur Verfügung "Wüste", "Gras", "Wald", "Felsen", "Fluß", "See" und "Weg". In eine solcherart charakterisierte Landmasse kann man dann beliebig viele einzelne Objekte einfügen. Als Objekttypen stehen zur Verfügung die Objektklassen Pflanzen und Lebewesen. Bei den Pflanzen wird unterschieden zwischen "Kleinpflanzen", "Büschen" und "Bäumen". Bei den Lebewesen zwischen "Kleintieren", "Raubtieren" und "Großwild", wobei hier als besondere Objektart noch (Pseudo-)Knowbots auftreten können, falls sich solche als Clients bei dem Weltprozeß anmelden.

Neben den sensorischen Eigenschaften besitzen sowohl die Pflanzen wie die Lebewesen zusätzliche Eigenschaften. Pflanzen haben z.B. bestimmte Nährwerte, die sich auf die Lebewesen unterschiedlich auswirken können. Die Lebewesen sind alle zumindest mit den Grundbedürfnissen "Hunger", "Durst", "Müdigkeit" und "Fortpflanzungsbedürfnis" ausgestattet. Ferner können sie "Schmerzen" und "Angst"* empfinden.

Nachdem sich Clients durch ein Login bei einem Server als "Weltbürger" angemeldet haben, verläuft der Weltprozeß als eine Folge von Weltzyklen. Ein Weltzyklus ist ein Verarbeitungszyklus im Server, der eine bestimmte physikalische Zeit benötigt (vgl. Bild 2).

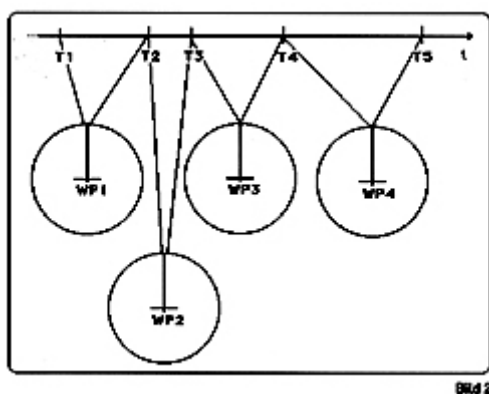


Bild 2

In Abhängigkeit von der anfallenden Datenmenge und der Leistungsfähigkeit der Hardware kann die physikalische Zeit für einen einzigen Weltzyklus stark schwanken. Innerhalb des Weltprozesses definiert ein Durchgang aber genau eine minimale Weltzeiteinheit, d.h. der Weltprozeß realisiert eine logische Zeit.

Innerhalb eines Weltzyklus gibt es im wesentlichen zwei Aktivitäten: (1) Der Interaction Manager [IM] liest sämtliche Aktions-Mitteilungen, die seit der letzten Auswertung im Briefkasten für Aktions-Mitteilungen eingegangen sind. Er überprüft dann jede Mitteilung daraufhin, inwieweit dadurch irgendwelche Kollisionen eintreten können. Stellt er eine Kollision fest, dann berechnet er eine korrigierte Endposition für die Bewegung und gibt diese weiter. Nachdem die Wirkungen der Aktions-Mitteilungen erfaßt sind, tritt (2) der Sensory-Manager [SM] auf den Plan. Für jedes Lebewesen, insbesondere natürlich für die (Pseudo-)Knowbots, berechnet er sämtliche mögliche sensorische Reize, die an der Stelle, wo sich das Lebewesen gerade befindet, "nach den Weltgesetzen" möglich sind. Die Gesamtheit der Werte, die der SM für ein bestimmtes Lebewesen findet, werden dann in eine sensorische Mitteilung verpackt und ihm zugeschickt. Wenn der SM seine Arbeit beendet hat, ist ein Weltzyklus beendet. Die Aktions-Mitteilungen wie auch die sensorischen Mitteilungen sind Zeichenketten (ASCII-Strings), deren syntaktischer Aufbau durch kontextfreie Grammatiken (Contextfree Grammars) bestimmt sind.

6 KNOWBOTS

Das vorstehend skizzierte KInt ist notwendig, um minimale Umgebungsbedingungen für die Simulation eines strukturellen Bewußtseinsäquivalentes bereit zu stellen.

Es seien hier nun jene Rahmenbedingungen geschildert, die in BW1 für erste Experimente mit dem Phänomen Selbstbewußtsein vorgesehen sind.

6.1 Minimales reaktives System

In BW1 wird angenommen, daß die Knowbots zumindest über eine Sensorik und eine Effektorik verfügen können.

Die Sensorik ist beschränkt auf die Sinnesarten "Hören", "Riechen", "Tasten" und "Schmecken" (Das Sehen wurde für BW1 ausgeklammert). Die Wertverläufe, die die Sensoren in das "Innere" des Knowbot übermitteln, werden dort zunächst in einem 3dimensionalen sensorischen Plan unter Bewahrung ihrer Topologie aufbereitet. Von hier aus stehen sie zur weiteren Verwendungen zur Verfügung.

Die Effektorik basiert auf einer endlichen Menge von Elementar-Handlungen (Gehen, sich drehen, die Hände bewegen, etwas fassen, etwas aus der Hand geben, essen, trinken, schlafen, Paarungsverhalten, spielen), die parametrisiert werden können. Mehrere solcher Elementarhandlungen können zu komplexen Handlungen zusammengefaßt werden, wobei die Parametrisierungsbedingungen erhalten bleiben. In BW1 wird angenommen, daß jeder Knowbot "von Geburt an" über eine Reihe von vorgefertigten Verhaltensschemata verfügt, die basale Verhaltensweisen wie Nahrungssuche, Nahrungsaufnahme, Fortpflanzung und Flucht sicherstellen. Ein Verhaltensschema ist eine Menge von Verhaltensregeln, wobei jede Verhaltensregel ein Bi-Konditional der Art darstellt: ZIEL Z kann erreicht werden genau dann, wenn BEDINGUNG B1 ist erfüllt & ... & BEDINGUNG Bn ist erfüllt. Der linke Teil der Regel heißt ihr "Kopf" und der rechte Teil ihre "Bedingung". Im allgemeinen Fall kann im Kopf mehr als ein Ziel genannt werden und in der Bedingung können auch Disjunktionen (oder) auftreten. Der "Schema"-Charakter von Verhaltensregeln entsteht dadurch, daß als Argumente der Ziele und Bedingungen auch Variablen auftreten können. Ein solcherart gegebenes Verhaltensschema steht dann für eine ganze Menge möglicher konkreter Verhaltensregeln, je nachdem, welche konkreten Werte für die Variablen eingesetzt werden dürfen.

Zur Effektorik gehört auch eine einfache motorische Planung, die die Aktivierung der jeweiligen Verhaltensschemata vornimmt und die Koordinierung der auszuführenden Handlungen überwacht. Dazu gehört sowohl die Verwertung der Rückmeldungen von der Sensorik wie auch die Umsetzung der "Tendenzvorgaben" der emotionalen Steuerungseinheit.

Die Emotionen bilden zusammen mit diversen Körperzuständen eine dritte Basiskomponente. Zu den Körperzuständen gehören "Energiebilanz", "Flüssigkeitsbilanz", "Wach/Schlafzustand". Zu den Emotionen werden Triebe gerechnet (Hunger, Durst, Müdigkeit, sexuelles Verlangen, Spielverhalten), Schmerzen, Angst sowie ein nicht weiter definierter "emotionaler Gesamtzustand". Diese Komponenten beeinflussen sich gegenseitig, wobei der emotionale Gesamtzustand die "Führungsrolle" hat und die Triebe wie auch die Schmerzen ihn je nach aktueller Stärke "übertönen" können. Für jede Emotion gibt es Aktivatoren oder Deaktivatoren. Hunger wird z.B. durch eine negative Energiebilanz aktiviert, und durch Nahrungsaufnahme — in Abhängigkeit von der Beschaffenheit der Nahrung — deaktiviert.

Zu jedem Zeitpunkt gibt es immer eine Emotion, die "dominant" ist. Diese beeinflusst den Planer bei der Gestaltung seiner Handlungspläne. Wechseln die Emotionen, dann führt dies u.U. auch zur Unterbrechung laufender Handlungen bzw. zur Wiederaufnahme eines vorher unterbrochenen Planes.

Die motorische Planungseinheit wird somit vollständig von dem aktuellen emotionalen Zustand dominiert, falls es nicht gerade einen Verhaltensplan gibt, der eine noch höhere Priorität hat.

Die Körperzustände werden sowohl von Handlungen wie auch von sensorischen Ereignissen beeinflusst und sie wiederum beeinflussen die Emotionen.

Die Gesamtheit dieser Basiskomponenten bildet das, was in BW1 ein reaktives System genannt wird (vgl. Bild 3), d.h. diese Komponenten bilden für sich genommen schon eine funktionsfähige Einheit, die in der Lage ist, sich in der Welt zu bewegen und sich notdürftig mit Nahrung zu versorgen. Zu diesem reaktiven System können jetzt weitere Komponenten hinzutreten, um es zu modifizieren.

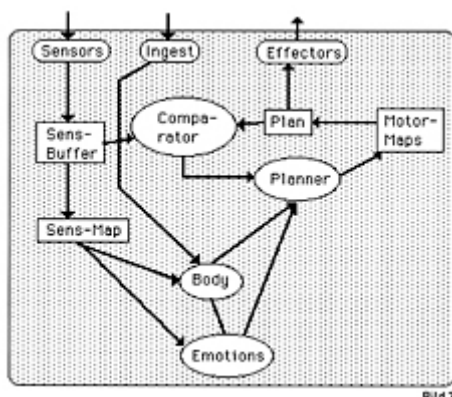


Bild 3

6.2 Die Notwendigkeit des Lernens

Ein großer Nachteil des reaktiven Systems besteht in seiner Starrheit: es kann, sieht man von einfachen Variationen innerhalb vorgegebener Grenzen einmal ab, nichts dazu lernen. Es setzt

"auf eine bestimmte Weise" Einwirkungen der Umwelt in bestimmte Reaktionsweisen um, ohne daß sich diese Reaktionsweise f für die Dauer der Existenz dieses Systems nennenswert ändert. Die "Reaktionsweise f " kann man als eine Funktion auffassen, die Elemente aus der Menge S der Stimuli in Elemente der Menge R der Reaktionen abbildet.

Damit aus einem nicht lernfähigen reaktiven System ein lernfähiges System wird, bedarf es der Möglichkeit, die Reaktionsweise f substantiell ändern zu können, d.h. man muß eine aktuelle Version von f durch eine andere Version f' aus der Menge F der möglichen Reaktionsweisen ersetzen können.

In solch einem System gibt es eine zusätzliche Lernfunktion L , die die aktuelle Version einer Reaktionsweise f durch Rückgriff auf die Menge F der möglichen Reaktionsweisen gegen eine neue Version f' austauscht.

Ein blindes Lernen, das eine aktuelle Reaktionsweise f durch "irgendeine andere" Reaktionsweise f' ersetzt, ist jedoch nicht sehr effektiv. Bei großer Populationsdichte und hohen Vermehrungsraten läßt sich über "lange" Zeiträume möglicherweise auch auf diese Weise eine Verbesserung erzielen. Verhaltensbezogenes Lernen beschränkt sich jedoch auf den Bereich eines Individuums und ereignet sich in einem vergleichsweise "kurzen" Zeitraum. Dies scheint nur möglich zu sein, wenn man annimmt, daß verhaltensbezogenes Lernen als informiertes Lernen realisiert ist. Das Mindeste, was man dann von einem lernfähigen System fordern müßte, wäre, daß es über Bewertungskriterien verfügen kann, die über die Ausgestaltung der Reaktionsweise f mitentscheiden können. Eine solche Evaluation könnte z.B. Umwelteinwirkungen zum Ausgangspunkt von Bewertungen zu machen.

Neben der reinen Verhaltensfunktion f gibt es dann sowohl eine Bewertungsfunktion ev , die Umwelteinwirkungen mittels positiver oder negativer Bewertungszahlen indiziert, und eine modifizierte Lernfunktion L , die bei der Auswahl von neuen Verhaltensweisen f' aus F auch Bewertungszahlen berücksichtigt.

- (1) $F: S \rightarrow R$
- (2) $ev: S \rightarrow E$
- (3) $L: F \times E \rightarrow F$

6.3 Lernen in BW1

Die Umsetzung des abstrakten Postulates, Knowbots als Systeme mit einer informierten Lernfunktion zu konzipieren, geschieht in BW1 an mehreren Stellen zugleich.

Die Bewertungsfunktion ev wird in einem Knowbot in erster Linie durch die Auswirkungen von Ingestionen und Perceptionen auf körperliche und emotionale Zustände realisiert. Wenn z.B. das Essen eines bestimmten Nahrungsmittels zur Linderung des Hungers führt, dann korreliert solch ein Essen samt den zugehörigen Geschmacks- und Geruchsempfindungen mit einem positiven Wert. Dieser kann benutzt werden, um sowohl ein bestimmtes Objekt wie auch eine bestimmte Handlung im Kontext einer bestimmten Bedürfnisbefriedigung zu bewerten.

Damit solche Zusammenhänge von Handlungen, Perceptionen, Ingestionen sowie deren Auswirkungen, die sich in einem bestimmten Zeitintervall t_i ereignen, auch zu einem späteren Zeitpunkt t_j+c verfügbar sein können, bedarf es der Möglichkeit, diese in geeigneter Weise zu "erinnern", d.h. sie in "erinnerbarer Weise zu speichern". Man benötigt also eine

Gedächtnisstruktur, die in der Lage ist, Objekte unter Berücksichtigung ihrer unterschiedlichen sensorischen Qualitäten sowie ihrer positiven oder negativen Auswirkungen auf emotionale Zustände und unter Berücksichtigung beteiligter Handlungen und relevanter räumlicher Strukturen "abzulegen" und "nach Bedarf" wieder zu "finden".

Schließlich muß auch die Möglichkeit bestehen, solcherart bewertete erinnerbare Zusammenhänge für die konkrete Verhaltensplanung nutzbar zu machen. D.h. wenn die Handlungs-Planungseinheit eine bestimmte Handlungsfolge beschlossen hat, um ein bestimmtes Ziel zu erreichen, dann muß es in einem lernfähigen System möglich sein, daß es die "vorgegebenen Verhaltensschemata" aufgrund von "Erfahrungen" "abändert". Solche Änderungen können auf unterschiedlichster Ebene stattfinden. Insgesamt stellt die Notwendigkeit der möglichen Einbeziehung von "erinnerbaren Zusammenhängen" in aktuelle Handlungsplanungen eine zusätzliche Anforderung an die Art und Weise dar, wie diese Zusammenhänge "gespeichert" werden.

Aufgrund der sehr kurzen Realisierungszeit konnten in BW1 diese Gedächtnisstrukturen vorerst nur sehr rudimentär implementiert werden.

6.4 Sprache

Zu einer natürlichen Sprache gehört neben dem Ausdrucksmaterial (Laute, Zeichen) auch das, was die sprachlichen Ausdrücke "bezeichnen", was sie "sagen", und das, was sie dadurch "bewirken", daß man sie benutzt. Diese über das reine Ausdrucksmaterial hinausgehenden Aspekte eines sprachlichen Ausdrucks sollen hier vereinfachend die "Bedeutung" eines sprachlichen Ausdrucks genannt werden.

Für uns Menschen hängt die sprachliche Bedeutung durchgängig mit unserem Selbst- und Weltbezug zusammen.

Für die Knowbots von BW1 steht ein Selbst- und Weltbezug nur in dem Umfang zur Verfügung, wie ihn das zuvor geschilderte informierte lernfähige reaktive System bereitstellt. Wie sich im weiteren Verlaufe zeigen wird, ist diese Struktur für die Realisierung differenzierter Bedeutungsstrukturen, wie wir sie aus der menschlichen Alltagssprache kennen, noch bei weitem zu einfach. Nur primitivste Benennungen und einfachste 1-, 2-, 3-Wort-Sätze sind möglich.

Das Sprachmodul verfügt über mehrere Teilmodule.

Eine wichtige Basiseinheit bildet das Lexikon, das unterschiedliche Teilfunktionen umfaßt. Ein Wortgenerierer versucht aus einer Menge elementarer Laute und Intonationsmustern Phonemsequenzen zu bilden, die im Lexikon gesammelt werden. Im Lexikon können sich aber nur solche Phonemsequenzen halten, die innerhalb eines bestimmten Zeitintervalls von außen "bestätigt" werden. Ein Wort-Objekt-Generierer bildet Wort-Objekt-Hypothesen, d.h. er versucht Worte mit Zuständen oder Handlungen zu verknüpfen. Die aktuelle Situation besitzt für solche Hypothesenbildungen dabei Priorität. Auch die Wort-Objekt-Hypothesen können sich nur halten, wenn sie "von außen" bestätigt werden.

Um die Grammatik zu verstehen, muß man sich klar machen, welche Aufgabenstellungen innerhalb einer sprachlichen Kommunikation zu bewältigen sind. Grundsätzlich wird in BW1 zwischen Sprechen und Hören unterschieden. Bild 4 demonstriert den Fall des Sprechens:

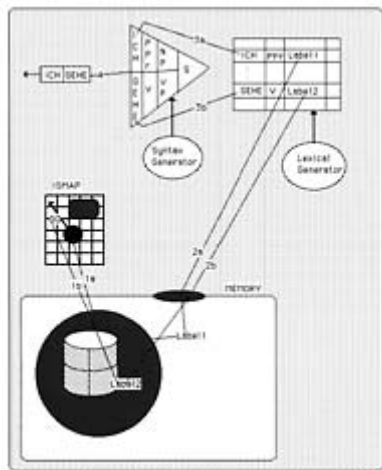


Bild 4

Ausgehend von dem interpretierten sensorischen Plan (interpreted sens-map [ISMAP]), werden von einem Pragmatik-Modul — genannt language map [LMAP] — die Unterscheidungen zwischen Sprecher und potentiellm Hörer erarbeitet sowie verschiedene Beziehungen, die aktuell zwischen dem Sprecher, seiner Umgebung und dem Hörer bestehen. Als Ergebnis gibt es Vorschläge für verschiedene Sprechakte. Nachdem ein Sprechakt ausgewählt worden ist, werden dann im Lexikon jene Objekte aktiviert, die in den Sprechakt eingehen sollen. Die Wortformen des Lexikons einschließlich diverser lexikalischer Kategorien verweisen dann auf die Grammatik. In dieser gibt es grammatische Regeln, die festlegen, wie Worte einer bestimmten lexikalischen Kategorie unter Berücksichtigung ihrer Objektindizierung und unter Berücksichtigung eines bestimmten Sprechaktes miteinander kombiniert werden können. Das Ergebnis ist dann eine konkrete Folge von Wortformen, die dann als konkrete Äußerung "zu hören" ist.

Das Bild 5 demonstriert den umgekehrten Fall, das Hören einer Äußerung. Dieser Fall stellt erheblich höhere Anforderungen als das Reden.

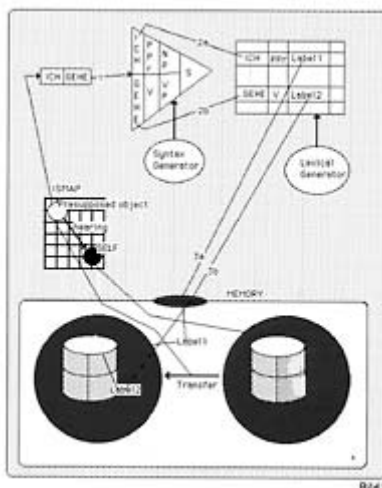


Bild 5

Ein vom Hörer verschiedener Sprecher sagt "Ich gehe". Die Regelschemata der Grammatik liefern zunächst eine erste Zerlegung der Wortfolge in einzelne Elemente, die über das Lexikon auch wieder auf entsprechende Objekte verweisen können. Mit Hilfe der LMAP versucht der Hörer gleichzeitig, den potentiellen Sprecher und die daraus sich ergebenden

Relationen zu ihm selbst zu erfassen, um damit mögliche Sprechaktformen zu ermitteln. In diesem Zusammenhang muß er eine ziemlich komplizierte Transferleistung erbringen, insofern er nämlich das dem Wort "Ich" zugeordnet Ich-Objekt als Modell für ein anderes Objekt "ansehen" muß. Entsprechend sind die an der Äußerung "Ich gehe" zugänglichen Aspekte auf das Fremd-Ich-Objekt und dessen Beziehungen zur Umgebung zu übertragen.

Es sei darauf hingewiesen, daß der alltagsprachliche Verstehensbegriff erheblicher umfangreicher ist als der oben skizzierte sprachliche Dekodierungsvorgang während der grammatischen Analyse einer Äußerung. Im Alltag führt das Ergebnis der sprachlichen Erkennung in der Regel zu weiteren Erkenntnisprozessen wie z.B. Schlußfolgerungen, Analogieschlüssen und bildhaften Assoziationen. Von diesem allgemeinen Verstehensbegriff ist das rein sprachliche Verstehen als ein — wenngleich notwendiger — Teilprozeß abzugrenzen. In den Knowbots von BW1 konnte aufgrund der sehr knappen Zeit nur das sprachliche Verstehen realisiert werden und auch dies nur in ersten Ansätzen.

Dadurch, daß in BW1 alle Beteiligten blind sind, kann das Erlernen von Zusammenhängen zwischen Sprachzeichen und beliebigen Wahrnehmungsereignissen nur insoweit "von außen" beeinflußt werden, als es möglich ist, durch das Moment der Gleichzeitigkeit einen potentiellen Zusammenhang zwischen Teilen einer sprachlichen Äußerung und Teilen der Wahrnehmung bzw. Teilen des aktuellen Situationsmodells herzustellen. Durch den Ausfall des Sehsinns muß die mangelnde Eindeutigkeit dieses Vorgehens häufig durch ergänzende Tastwerte ausgeglichen werden, was nicht immer ganz einfach ist.

7 HARDWARE, SOFTWARE, MANPOWER

Die hohen Anforderungen an Interprozeß- und Interobjektkommunikation seitens der Installation BW1 konnte zum Zeitpunkt der Erstellung nur von dem Betriebssystem NEXTSTEP vollständig erfüllt werden. Ferner stellten die Systeme von Hewlett-Packard die einzige leistungsfähige Hardware dar, auf der im Entwicklungszeitraum NEXTSTEP portiert war. Wir entschieden uns daher, HP-Workstations mit dem Betriebssystem NEXTSTEP Version 3.2 als Entwicklungsplattform zu wählen: eine HP 712/80 und eine HP 712/60 mit jeweils 64 MB RAM, 1GB Platte und mit NEXTSTEP 3.2. Zusätzlich benutzten wir einen Aquarius-PC 486/66 mit 32 MB RAM und 1 GB Platte mit NEXTSTEP 3.3. Für die Entwicklung der XWindow-Motif-Version des Normal-User-Clients stand ein ESCOM-PC 486/66 mit 16MB RAM 1,3GB Platte und dem Betriebssystem Unifix 1.5 inclusive Motif 2.0 zur Verfügung. Zusätzlich wurde eine Version des XClients unter IRIX auf Silicon-Graphics-Rechnern compiliert. Alle Rechner waren mittels EtherNet verbunden.

Die Server-Software sowie die NEXTSTEP-Version des Super-User-Clients wurden von Leo POS und Thore SWINDALL programmiert. Das Perzeptions- und das Memory-Modul im Knowbot erstellten Thore SWINDALL und Raoul SCHOLZ. Das Action-Modul sowie Teile des Language-Moduls programmierte Leo POS. Michael KLÖCKNER zeichnet verantwortlich für das Language-Module.

Die XWindow-Motif-Version des Normal-User-Clients wurde von Joachim RASCH und Sonja SCHELLENBERG im Rahmen ihrer Diplomarbeiten programmiert.

Die Vorlage für die Programmierung der Software unter NEXTSTEP bildet das Arbeitspapier Agents with Consciousness. The Theoretical framework for the Knowbotic-Interface Project. Phase 1: The BLINDS WORLD 1 vom Verfasser. Ende März umfaßte es 130 Seiten, ergänzt durch zahlreiche Schautafeln.

Sämtliche Arbeiten erfolgten im Institut für Neue Medien, dessen Direktor Michael KLEIN uns für dieses Projekt in großzügiger Weise auch einen eigenen Projektraum zur Verfügung gestellt hatte. Zu nennen sind auch Markus FIX, Peter FRANK sowie Sewo STILLE, deren konstruktive Kritik das Projekt in allen Phasen begleitet hat.

Der Künstler. Dr. Gerd DÖBEN-HENISCH, ist von seiner Ausbildung her Kognitionswissenschaftler und Philosoph. Auf eine Initiative von Peter WEIBEL hin kam er 1992 an das Institut für Neue Medien, um seine bisherige wissenschaftliche Forschungsarbeit zum Thema sprachliche Bedeutung und selbstlernende Systeme unter den stimulierenden Bedingungen des Instituts für Neue Medien fortzusetzen. Inspiriert durch Multi-User Dungeons (MUDs) und den Knowbots von Christian Hübler entwickelte sich im Frühjahr 1994 die erste Vision vom Knowbotic Interface als einem Multi-Agenten-Betriebssystem für migrationsfähige adaptive Agenten, die in ihrer virtuellen Welt wie in einer realen Welt leben. Die Basis für eine neue Form digitaler Interaktion ist damit gelegt.

8 LITERATUR

- W. BALZER/ C.U. MOULINES/ J.D. SNEED [1987]. An Architectonic for Science. The Structuralist Program. Dordrecht.
- J. BARWISE/ J. PERRY [1983]. Situation and Attitudes. MIT Press. Cambridge (Mass.).
- D. DAVIDSON/ G. HARMAN (eds.) (1972). Semantics of Natural Language. D.Reidel. Dordrecht.
- M. DAVIS (ed) [1965]. The Undecidable. Basic Papers On Undecidable Propositions. Unsolvble Problems And Computable Functions. Raven Press. Hewlett (NY).
- G. DÖBEN-HENISCH [1993]. Philosophie des Subjekts und der Symbolautomat. MS. Institut für Neue Medien. Frankfurt (Erscheint in der Zeitschrift MINUS 5°, Sommer 1995. minus-Verlag 10791 Berlin, Postfach 620 340).
- [1994a]. Knowbots und Interaktives Fernsehen. Das Knowbotic Interface Projekt als Herausforderung an die KI., Gekürzte Fassung eines Vortrages im Rahmen des Workshops Interaktives Fernsehen im Institut für Neue Medien. 7. Juni 1994. Frankfurt. Engl. Fassung in Tightrope. HBK Saar -Forschungsprojekt Internet. Febr. 1995. SS.10-14: Engl. und Deutsch unter <http://www.phil.uni-sb.de/tightrope.html>.
- [1994b]. Fremdbeschreibung und Selbstbeschreibung. Ein Problem der Philosophie des Subjekts, Arbeitsgrundlagen für ein gleichnamiges Forschungsseminar mit Prof. H.-U. HOCHÉ an der Ruhruniversität Bochum. Fakultät für Philosophie. MS Institut für Neue Medien. Frankfurt.
- [1994c]. Warum bin ich keine Maschine? Thesen zur philosophischen Bestimmung des Mensch-Maschine Verhältnisses. MS. Institut für Neue Medien e.V.. erscheint in: Jahrbuch der Humboldt-Gesellschaft, 1995.
- G. FREGE [1892]. Über Sinn und Bedeutung. in: Ztschr. f. Philos. u. philos. Kritik, NF 100. pp. 25 — 50. Abgedruckt in: G.PATZIG (Hrsg.). Funktion, Begriff und Bedeutungen. Göttingen. 1975. pp.40-65.
- K. GÖDEL [1931]. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. in: Monatshefte f. Mathem. u. Physik. Vol. 38. pp. 173 — 198.
- A. HODGES [1994 (engl.: 1983). 2.Aufl.] Alan Turing. Enigma. transl. R.HERKEN/ E.LACK. Springer-Verlag. Wien.
- J. HOFFMANN [1982) Das aktive Gedächtnis. VEB Deutscher Verlag der Wissenschaften. Berlin.
- E. HUSSERL [1913], Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie. Edited by E. STRÖKER. Felix Meiner Verlag. Hamburg.

- F. KLIX [1980. 5th ed.]. Information und Verhalten, VEB Deutscher Verlag der Wissenschaften. Berlin.
- G. LUDWIG [1978]. Die Grundstrukturen einer physikalischen Theorie. Springer. Berlin — Heidelberg — New York.
- E. MACH [1886. 6. korr. Aufl. 1911. 1985 repr. von 9. Aufl. 1922] Die Analyse der Empfindungen und das Verhältnis des Physischen zum Psychischen. Wiss. Buchgesellschaft. Darmstadt.
- M. MERLEAU-PONTY [1945. dt. 1965]. Phänomenologie der Wahrnehmung. Übersetzt von R. Boehm, Walter de Gruyter. Berlin.
- G.M. MURCH/ G.L. WOODWORTH [1978]. Wahrnehmung, Kohlhammer. Stuttgart -Berlin — Köln — Mainz.
- Th. NAGEL [1986]. The View from Nowhere. Oxford University Press. New York. Oxford.
- W. SHIFF [1980]. Perception: An Applied Approach. Houghton Mifflin Company. Boston — Dallas — London et al.
- A.M. TURING [1936-7]. On Computable Numbers with an Application to the Entscheidungsproblem. In: Proc. London Math. Soc., Ser. 2, vol. 42. pp. 230-265; corr. vol. 43. pp. 544-546 (Reprint in M. DAVIS 1965. pp. 116-151; corr. ibid. pp. 151-154).
- [1987 (engl.: 1948)]. Intelligente Maschinen. in: Intelligent Service — Schriften. Hrsg, v. B. DOTZLER u. F. KITTLER, pp. 81-113. (Engl. Titel: Intelligent Machinery).
- [1987 (engl.: 1950)]. Rechenmaschinen und Intelligenz. in: Intelligent Service — Schriften. Hrsg, v. B. DOTZLER u. F. KITTLER). pp. 147-280 (Engl. Titel: Computing machinery and intelligence. in: Mind. vol. 59. pp.433-460).
- [1987]. Intelligence Service. Schriften. Hrsg, v. B.DOTZLER u. F. KITTLER. Verlag Brinkmann & Bose. Berlin.
- L. WITTGENSTEIN [1921. dt. 9. Aufl. 1973], Tractatus logico-philosophicus. Logisch-Philosophische Abhandlung. Suhrkamp. Frankfurt.