

# Wie menschlich sind Maschinen?

Vortrag Siemens München 9.1.1997

© by Britta Schinzel

## 1. Einleitung:

### Der Computer

Schon mechanische Maschinen wurden als Ergänzungen des menschlichen Körpers angesehen, als Organersatz oder Organ-Erweiterungen in Bereichen, wo der Mensch scheinbar defizitär ist (Flug), später als elektromechanische Organerweiterungen (Motor), und nun ist das Organ *Gehirn* und seine Funktion *Denken* an der Reihe: Computer sind Geistesmaschinen, Verstärker von Intelligenz, indem sie das Gedächtnis, die Rechenleistung, die logischen Denkleistungen, die Problemlösungsfähigkeiten des Menschen unterstützen und zum Teil weit überflügeln.

Aber Ist das Intelligenz? So gesehen ist der Computer ein Intelligenzverstärker, aber er ist auch ebenso willig ein Unintelligenzverstärker.

Der Computer verstärkt unsere logisch-mathematischen Fähigkeiten auf atemberaubende Weise. Wirklich gut ist der Computer bei der automatischen Speicherverwaltung; das ist ein formales Problem, und solche Probleme löst der Computer ideal: Klar, rasch und verlässlich - vorausgesetzt, daß das erforderliche Programm so geschrieben ist, daß die guten Eigenschaften zum Tragen kommen. Und der Computer ist ein hervorragender Prüfer von Abhak-Listen - auch das kann er klar, rasch und verlässlich. Beides sind nicht eben Intelligenz-Leistungen, aber sie unterstützen die menschliche Intelligenz.

Es ist die Schnelligkeit des Computers im Absuchen und schnellen Auffinden, ohne dabei etwas auszulassen, die ihn so wirksam macht, die fehlerlose Benutzung von Schablonen, von einem intelligenten Programmierer zielführend organisiert.

Mit seinen Prozeßschablonen wird er aber auch zum schnellen und billigen Entlastungsinstrument gegen Routinearbeit, aber dies auf Kosten von Kontingenzen und von Ausnahmen. Der Kontrast zwischen algorithmischer Regelmäßigkeit und menschlicher und sozialer Kontingenz macht sich bei allen Anwendungen bemerkbar.

Der Computer ist aber auch ein Intelligenzverstärker in dem Sinne, als er den Menschen zu Intelligenzleistungen anregt, zu denen es ohne Computer nicht

gekommen wäre. So dient beispielsweise die Computerleistung zur Hypothesenbildung in der Mathematik und in der Physik, zur Generierung von Beispielen für solche Hypothesen, oder von Gegenbeispielen. Der Computer als aus integrierten Schaltkreisen bestehendes Gerät, in dem durch äußere Trigger elektronische Zustände verändert werden können und so kaskadenartig Abläufe stattfinden, ist ein rein syntaktisches Gerät: Zeichen (repräsentiert als Zustände spezieller Schaltelemente oder anders ausgedrückt als Inhalte von Registern) werden - ohne Ansehen ihrer Bedeutung - systematisch in andere Zeichen verwandelt, seien es die Vorgänge bei den Grundrechnungsarten oder bei der Kompilierung.

## **Künstliche Intelligenz**

Der Computer ist ein automatisches Werkzeug, und Automaten tun, wie der Name sagt, doch alles von selbst. Die Täuschung besteht darin, den Computer für einen Roboter zu halten, dem nur eine Mutation zum Menschen fehlt. Der Ausdruck Künstliche Intelligenz unterstellt, daß man dem Computer Verständnis, Einsicht und Erkenntnisfähigkeit zuschreiben könne. Das ist ein Denkfehler.

Intelligente Schablonenverwendung ergibt nicht eine intelligente Schablone. Das intelligent mit Programmen ausgestattete Terminal wird nicht ein intelligentes Terminal. Die Kombination von fixer, präziser, rasch beweglicher und rasch systematisch modifizierbarer Schablone mit intelligenter Verwendung ergibt Intelligenzverstärkung, nicht nur bei Programmierern und ihren Professoren, sondern auch bei Schalterbeamten mit einem Bildschirm vor sich, die alle auf diese Weise zu Intelligenzleistungen kommen, zu denen sie allein nicht fähig wären. Ein 'intelligentes Terminal' ist nichts als ein mit Mechanismen gefülltes Terminal, intelligent konzipiert und für intelligente Benutzung bestimmt.

In menschlicher Verstärkung der Unintelligenz wird neuerdings ein Gebäude, das mit Leitungsbündeln und Mikrocomputern ausgestattet ist, als "intelligentes" Gebäude bezeichnet, vorläufig noch mit Anführungszeichen, aber die sind, man weiß es, bald abgestreift. (Zitat Zemanek)

## **Der Mensch**

**Der Mensch** lernt vom Babyalter an, Symbole und Sprache mit erlebter semantischer Realität zu verbinden, ausdrücklich - und noch mehr im Unterbewußtsein. Worte wie weich, glatt, Samt, spitz erhalten ihre Semantik aufgrund sinnlicher Erfahrung. Kinder mit gestörter sensomotorischer Integration haben deshalb auch Schwierigkeiten beim Sprachverständnis, beim Sprechen und beim Lesen und Schreiben. Der Computer ist nicht nur weit von einer Veranlagung für diese hervorragende Eigenschaft entfernt (weder die selbsttätige Ineinanderkoppelung von Sinnesorganen, Verarbeitung und Effektoren steht ihm zur Verfügung, noch die Universalienbildung - beim

Menschen vollautomatisch, auch bei denen, die man mit Recht als dumm bezeichnen dürfte, diese betreiben nur die falschen Verallgemeinerungen - sondern es fehlen dem Computer auch Motivation und Ehrgeiz).

## **Vergleich ZNS - Computer Konnektionismus**

Informationsverarbeitung geschieht in Computersystemen physikalisch als über elektronische Impulsmuster, in biologischen Systemen sowohl durch physikalische Impulse als auch durch chemische Signale, und zwar gemäß der topologischen Kodierung der Verknüpfungen im biologischen, d.h. stark diversifizierten Neuronennetzwerk. Biologische Systeme haben die Fähigkeit zur Ausbildung neuer Strukturen: Die Plastizität des Gehirns wird durch die Möglichkeit zur Bildung neuer Dendriten an einem Neuron und neuer Verbindungen oder Verstärkung / Abschwächung von Verbindungen erreicht. Das Gehirn ist zu autonomer neuronaler Aktivität in der Lage, es ist ein offenes System. Es arbeitet nicht sequentiell, sondern parallel und unsynchronisiert, es hat keinen festgelegten Zustandsraum, es schafft stets neue und eliminiert alte Systeme. Auch die Veränderungen sind zufällig und variabel. Die mentalen Repräsentationen sind nicht festgelegt, weder intra- noch interpersonell. Partielle Stabilisierung erfahren sie durch Bewertungs- und Kontrollprozesse, die dem menschlichen Bewußtsein eigen sind. Mentale Prozesse also lassen sich nicht unabhängig von ihrer physikalischen (+chemischen+biologischen) Realisierung beschreiben und verstehen.

Wer Gehirn und Computer gleichsetzt, muß also entweder das Bewußtsein negieren oder in einem Dualismus das Geistige in eine transzendente Dimension verlagern, einen Leib-Seele-Dualismus vertreten. Sinnvoll ist ein Vergleich, der ein materielles Substrat als Basis für Unterschiede sieht, nur bei einer materialistischen Position, etwa der des emergentistischen Materialismus: Durch das komplexe Zusammenspiel biologischer Prozesse ergibt sich eine neue Qualität. Das Geistige, das auf den Körper bezogen ist (ohne von einem ontologischen Dualismus auszugehen), ist von anderer Art als in Silizium eingebettete Rechenleistungen. Geistige Objekte sind bestimmte neuronale (Teil-)Netze, psychische Zustände und Prozesse des Gehirns. Psychophysische Beziehungen sind Wechselwirkungen zwischen verschiedenen Teilneuronalen Systemen des Organismus. Die mentalen Konzepte sind durch sensomotorische Erfahrungen geprägt. Auch Abstrakta werden aufgrund sinnlicher Erfahrungen geformt. Intentionales Denken operiert mit diesen mentalen Objekten in bewußter ziel- und handlungsorientierter Weise. Die menschliche Art des vorstellungsgeprägten lebendigen Denkens wird einer Maschine kaum möglich sein.

Allerdings liefern die mit den Computern und der Programmierung in Mode gekommenen Computermetaphern über Gehirn, Gedächtnis u.s.w. Begriffe und

Bedeutungen intelligenter Eigenschaften, die z.B. in der Kognitionspsychologie neu interpretiert werden. Sie bieten damit neue Konzepte für "Denken", welche aber im Sinne der Vorstellung des Menschen über sich selbst auch bedenkliche Folgen haben können. Wir beobachten jedenfalls neben der zunehmenden „Vermenschlichung“ von Maschinen auch eine zunehmende Mechanisierung und Machinisierung des Menschen ( etwa durch die Reparationsmedizin und Gentechnologie) und seines Bildes von sich selbst.

## **2. Die symbolische KI**

Die KI hat ihre Wurzeln in der philosophischen Tradition. Die Griechen gingen von der Möglichkeit der Formalisierung und Mathematisierung von Denkleistungen aus und betonten logische Schlüsse und abstrakte Denkvorgänge als Ausdruck von Intelligenz. Mit der Entwicklung autonomer Maschinen wie der Uhr wurde die Vorstellung von einem künstlichen Geist möglich. Auch die Literatur bemächtigt sich dieser Idee: der Talmud berichtet von der Erschaffung eines künstlichen Menschen, dem Urvater von Gustav Meyrink's Golem; Homunculus entsteht in der Retorte im Faust II und mit dem Roboter des tschechischen Autors Karel Capek ist der Name für eine ganze Gattung gefunden, die sowohl in der Automatisierung als auch in der Science Fiction-Literatur in vielen Varianten auftritt.

Die früheste explizite Äußerung zu Künstlicher Intelligenz aus der Mathematik ist der sogenannte Turing-Test: Turing, der ein allgemeines Berechenbarkeitsmodell, die Turing-Maschine, definierte, bezeichnet ein System als intelligent, wenn ein Tester, der es befragt, nicht erkennen kann, ob er einen Menschen oder ein System vor sich hat. Gleichzeitig begann John von Neumann mit Modellen zur maschinellen Realisierung zu experimentieren. Ihn interessierten u.a. Probleme der Selbstreproduktion, der Simulation neuronaler Strukturen mit Netzen und Schwellenfunktionen, aber auch der mathematischen Automatenkonstruktion. Die Konsequenzen der Zusammenführung der technischen und mathematisch-logischen Modellierung in heutigen Computern sind Ihnen alle bekannt. Ebenso die Tatsache, daß sie in der Lage sind, Teile der menschlichen Arbeit zu automatisieren, und zwar gerade jene, die uns lange als dem Menschen vorbehalten erschienen, nämlich Denkvorgänge und organisationale Anteile.

Kein Wunder, daß diese Maschinen Philosophen, Psychologen und Soziologen besonders interessieren. Hier soll [aber] nicht von den sozialen Auswirkungen der Computerisierung gesprochen werden, sondern von den mathematisch-technischen Modellen und deren kognitionspsychologischem und philosophischem Interesse.

## **Paradigmen**

Obgleich also schon in den 40er Jahren verteilte, am biologischen Gehirn angelehnte Modelle untersucht worden waren, feierte zunächst in den 60er und 70er Jahren die symbolische Repräsentation sowohl als Mittel zur Maschinisierung der Kopfarbeit wie als Modell der künstlichen Intelligenz erste Erfolge.

John McCarthy, einer der Begründer der KI, [er]stellte 1956 in Dartmouth eine Liste von Anforderungen an KI-Systeme:

1. Metaphysisches Postulat: Das System muß eine Repräsentation der Welt haben, sie manipulieren und erweitern können.
2. Erkenntnistheoretisches Postulat: Diese Repräsentation geschieht durch einen Formalismus, der Problemlösungen als logische Folgerungen enthält.
3. Heuristisches Postulat: Die Implementierung des Formalismus muß das Auffinden der Problemlösungen bzw. der Folgerung gestatten, d.h. eine Suchstrategie oder einen Beweis ermöglichen.

Voraussetzung für die Erfüllbarkeit dieser Postulate ist die Annahme der Existenz einer realen Welt und ihrer grundsätzlichen Erkennbarkeit. Ist die erste Forderung zwar eine nur von wenigen Philosophen in Frage gestellte, so ist die zweite innerhalb der Erkenntnistheorie zumindest sehr umstritten. Doch gehen wir zunächst von der epistemologischen Annahme der analytischen Philosophie aus, daß [den] Menschen ein objektives Bild der Welt und von sich selbst möglich ist. Und weiter von der Annahme, es gäbe eine maschinelle Intelligenz. Intelligente Systeme bilden mentale Modelle über die Welt. Solche Modelle bestehen aus verfeinerbaren Objekten und Beziehungen zwischen ihnen und sollen so Weltausschnitte repräsentieren. Durch Benennung erhält man natürlich-sprachliche Beschreibungsmodelle, welche durch das mentale Modell, aber auch, wenn das mentale Modell diesen richtig beschreibt, durch den Weltausschnitt interpretiert werden können. Das natürlich-sprachliche Beschreibungsmodell kann nun - eine weitere Annahme - durch Formalisierung in ein hinreichend genaues formal-algorithmisches Modell überführt werden. Letzteres besteht aus Datenstrukturen und Algorithmen, die durch Implementierung in ein Stück Software umgewandelt und damit weiter auf einer physikalischen Maschine realisiert werden können.

Dem symbolischen Repräsentationsparadigma steht das konnektionistische Paradigma gegenüber, das in den letzten Jahren zunehmende Bedeutung erlangt hat. Hier hat man keine expliziten Repräsentationen und Modelle, sondern sieht intelligente Fähigkeiten als Emergenzphänomene in verteilten vernetzten Systemen von Recheneinheiten an. Die Eigenschaften des Netzes bilden sich in Abhängigkeit von der uniformen oder spezialisierten Aktivierungsfunktion der einzelnen Recheneinheiten, nach einer Lernphase - z.B. durch Präsentation einer Beispielmenge - als stabile Gewichte der Kanten des Netzes aus. Im Konnektionismus werden Lernvorgänge und ganzheitliches Erfassen von Bildern und Vorgängen durch evolutive Methoden eher behandelbar als durch symbolische Repräsentationen; diese hingegen bevorzugen Logik, arithmetisches

Rechnen, allgemein diskrete Vorgänge.

## **Realisierte KI-Software**

Konkret hat sich die KI vor allem mit automatischem Beweisen, Problemlösen, insbesondere Spielen, Agenten, Planen, wissensbasierten Systemen, natürlich-sprachlichen Systemen, akustischer Sprachverarbeitung, Mustererkennung und Bildverarbeitung, Robotik und lernenden Systemen befaßt. In dieser Reihenfolge finden absteigend symbolische Repräsentationen und aufsteigend konnektionistische Systeme Verwendung.

Automatische Beweissysteme haben einfache zahlentheoretische oder gruppentheoretische Sätze gefunden, auch solche, die vorher noch nicht gefunden worden waren. Eine solche Verwendung gelingt jedoch nur in einfach und damit voll formalisierbaren Teilen der Mathematik, für die Analysis z.B. sind automatische Beweissysteme schwer denkbar.

Problemlöseaufgaben bestehen typischerweise darin, einen Anfangszustand durch Anwendung von Transformationen aus einem gegebenen endlichen Vorrat von Operationen in einen erwünschten Endzustand zu überführen. Solche Aufgaben können Brettspiele, Puzzles oder das in der KI sehr intensiv bearbeitete Schachproblem sein.

Wissensverarbeitende Systeme und Expertensysteme werden für Probleme verwendet, für die eine geschlossene Lösung nicht möglich ist, oder auf Gebieten, wo dynamische Reaktionen auf Benutzerwünsche notwendig sind, die in einer geschlossenen Lösung zu explodierenden Komplexitäten führen würden. Sie haben typischerweise eine Repräsentation von Wissen und eine Inferenzkomponente, die es erlaubt, neues Wissen aus gegebenem Wissen abzuleiten oder unter Einbeziehung neuer Eingabedaten zu lernen. Es gibt sie auf allen möglichen Gebieten, in der Medizin zur Diagnose und Therapieunterstützung, in der Technik zur Fehleranalyse, in der Biologie zur Kategorisierung, also Bestimmung der Zugehörigkeit eines Objekts zu einer bestimmten Kategorie aus einer Menge von möglichen, in der Architektur zur Entwurfsunterstützung, z.B. bei graphischen Entwürfen, oder im Verkauf zur Beratungsunterstützung.

Natürlichsprachliche Systeme werden zur Sprachgenerierung und Sprachanalyse durch den Rechner gebraucht. Anwendungen finden sie in Dialogkomponenten, Erklärungskomponenten, Frage-Antwort-Systemen, als Übersetzungshilfen und als Hilfen beim Korrekturlesen.

Bei der Bildverarbeitung unterscheidet man zwischen **Bildererkennung**, und der **Verarbeitung**, d.h. also der Manipulation von Bildern mit dem Zweck der Extraktion von Informationen. Typische Probleme bei der Bildverarbeitung sind das Umgehen mit riesigen Datenmengen sowie Probleme der Zentrierung und des 'Matchens', also des Vergleichs mit Normbildern.

In der Robotik werden vorwiegend Industrieroboter mit Lenk- und Greifmechanismen, die u.U. sensorgesteuert sein können und programmierbare Bewegungsfolgen ausführen, entwickelt. Ein wichtiges Problem ist dabei die

Rücktransformation, also die Bestimmung der Parameter der Gelenkstellungen und Drehwinkel, welche nötig sind, damit der Greifarm einen bestimmten Punkt im Raum erreichen kann.

Eine solche Berechnung ist jedoch nicht notwendig bei konnektionistischen Lösungen. Hier existieren bereits Roboter, die in der Lage sind, auf einem Bein zu hüpfen, ohne umzufallen oder solche, die in der Lage sind, einen Stab so zu balancieren, daß er nicht herunterfällt. Oder es gibt Realzeitlösungen für Steuerungsprobleme, auch von Autos in realistischen Verkehrssituationen, die ziemlich gut funktionieren.

Lernende Systeme werden bei allen erwähnten Problemen auch mitbenutzt. Dabei werden über deduktive Schlüsse hinausgehende induktive und analoge Schlüsse verwendet. Logisch orientierte Methoden, die dem Repräsentationsparadigma folgen, stehen konnektionistischen Lernmethoden gegenüber, die eher für holistische Lernverfahren geeignet sind, z.B. zur Mustererkennung/Musterergänzung.

Sie sehen bereits an dieser Liste, daß nur ganz spezifische Ausschnitte menschlicher Fähigkeiten bisher modelliert worden sind, und für symbolische Repräsentation gilt, daß es gerade die von körperlichen Existenz und Bewegung nicht abhängigen abstrakten Leistungen sind. Alles, was mit Handlungen, Sensomotorik, und integrativen Fähigkeiten zusammenhängt, insbesondere natürlich die im Menschen realisierte Integration aller dieser Leistungen, kommt hier nicht vor.

## **Die heutige Bedeutung der KI innerhalb der Informatik**

Heute sieht man die KI und ihre Methoden nicht mehr getrennt von der übrigen Informatik: Selbstlernende Teile sind prinzipiell überall zu finden, Agenten haben Eingang in die [Zugang zur] Netzprogrammierung gefunden, und die KI berücksichtigt die Ergebnisse der Informatik, z.B. der Komplexitätstheorie.

## **Methoden der KI**

### **1. Suchverfahren**

Ein großer Teil der Intelligenzleistungen des Menschen muß im Rechner durch Suchstrategien nachgebildet werden, so z.B. Gedächtnisleistungen und Assoziationen. Auch beim Problemlösen muß die richtige Folge von Transformationen zur Erreichung des Zieles gesucht werden. Gleiches gilt für das Finden von Beweisen, also nicht für die Abfolge der Beweisschritte selbst, sondern den Vorgang, diesen Beweis unter einer Menge von möglichen herausauszufiltern.

Es gibt kaum ein KI-Problem, das ohne Suchvorgänge auskommt. Im Allgemeinen handelt es sich dabei um eine "trial and error"- Suche aus einem Suchraum möglicher Lösungswege oder Lösungen. Da diese Art von Suchen zu

einer kombinatorischen Explosion, also zu exponentiellem Wachstum der Suchzeiten mit der Zahl der notwendigen Transformationen (bereits bei nur jeweils 2 möglichen Folgezuständen) führt, müssen die Möglichkeiten der Lösungen oder Lösungswege eingeschränkt werden. Dies kann einerseits durch Voreinschränkungen des Suchraumes geschehen, wenn die Lösung in einem speziellen Bereich erwartet wird, andererseits durch die Verwendung von heuristischen Verfahren. Dies sind Erfahrungsstrategien, die in einer oft nur statistisch erfaßbaren Anzahl von Problemstellungen zum Erfolg führen.

## 2. Logische Methoden.

Neben dem Suchen in Zustandsräumen bildet den zweiten großen methodischen Komplex die Logikorientierung der KI. Diese drückt sich in der Verwendung logischer (Prolog, OPS 5) und funktionaler (LISP) Programmiersprachen aus, welche wiederum durch die zu Hilfe genommenen Wissensrepräsentationsformen nahegelegt werden. Symbolische Repräsentationen werden wesentlich in Prädikatenlogik erster Stufe (PL 1) formuliert. Auch andere Repräsentationsformen, wie Produktionssysteme (wenn/dann-Regeln), semantische Netze oder Frames können - mit Abstrichen - in PL 1 umformuliert werden.

Man verwendet die PL 1 trotz ihrer beschränkten Ausdrucksfähigkeit deshalb, weil sie besonders zugänglich ist. So gilt für sie der Gödelsche Vollständigkeitsatz, der für eine Ausgewogenheit von Syntax und Semantik sorgt: Alle wahren Sätze (Tautologien) sind beweisbar.

Aber die Beweissuche ist schwierig: Church's Unvollständigkeitsatz (eigentlich Unentscheidbarkeitssatz) stellt fest, daß es kein Entscheidungsverfahren gibt für die Eigenschaft, Tautologie zu sein.

Weiter gelten gewisse Endlichkeitseigenschaften, die Inferenzen erleichtern.

Über PL 1 hinausgehend können erweiterte Ausdrucksmöglichkeiten für unsichere Aussagen, performative Äußerungen, zeitliche Zusammenhänge usw. durch höherstufige Logiken erreicht werden, die Quantifizierung nicht nur über Variablen erlauben, sondern auch über Prädikats- und Mengenvariablen. Sie führen jedoch zu Gödels Unvollständigkeitsresultaten, welche besagen, daß jedes formale System, in dem mindestens die elementare Arithmetik ausdrückbar ist, wahre Sätze enthält, die in diesem System nicht beweisbar sind.

Im Einzelnen werden Modallogiken und andere verwendet, z.B.

1. Für Situationen mit unvollständiger Information oder für dynamische Situationen die Modallogik in engerem Sinne mit den verallgemeinerten Quantoren notwendigerweise und möglicherweise:

2. Für die Ausdrückbarkeit von zeitlichen Zusammenhängen, Zeiträumen und

Zeitpunkten existieren verschiedene Zeitlogiken.

3. Der S4-Kalkül erlaubt es, Wissen und Überzeugungen zu unterscheiden. Ein typisches Axiom ist z.B. das positive (und negative) Introspektionsaxiom, welches besagt, daß wenn Person A eine Tatsache phi (nicht) weiß, folgt, daß A weiß, daß sie phi (nicht) weiß.

4. Für die Darstellung natürlicher Sprachen wurde von Barwise und Perry die sogenannte Situationssemantik konzipiert, welche alle genannten umfaßt.

Weitere wichtige Logiken sind induktive und nicht monotone Logiken. Sie werden benutzt, um sinnvolle Hypothesen über Beispielphänomene aufzustellen. Nicht monoton heißt eine Logik, wenn eine neue Beobachtung oder ein neues Beispiel die vorherige Hypothese falsifizieren kann.

Allgemein kämpfen Modallogiken mit

1. Unvollständigkeit. Diese kann durch die sogenannte Closed-World-Assumption praktisch unschädlich gemacht werden, also durch die Annahme, daß jede nicht ableitbare Aussage falsch ist, also in dem Kalkül nicht gilt. Sie kostet allerdings u.U. die Monotonie.

2. Mit Widersprüchen. Die Widersprüche führen dazu, daß man lokale Konsistenzüberwachung betreiben und lokal Konsistenz herstellen muß. Hierfür gibt es verschiedene Methoden des "Model-checking" und der "Truth-Maintenance".

Ein Problem stellt dabei die Integration aller dieser Logiken dar: Kann man auch jedes einzelne Konzept der Aufweichung harter formaler Methoden logisch formulieren, so gelingt dies nicht ohne Widersprüche und Unvollständigkeiten für die Gesamtintegration.

## **Vergleich zu Denken, Wissen, Expertentum des Menschen**

Gehen wir zurück zur Auffassung von Realität und ihrer Sicht durch den Menschen im Lichte dieser Kalküle: Wissen besteht für die KI aus "kodierbaren" "wahren" Tatsachen, kodierbar im Sinne von Ausdrückbarkeit durch logische Formeln, die in einem mengentheoretischen Modell (einer "möglichen Welt" im Kopf des Menschen) interpretierbar sind, und wahr im Sinne der Existenz in diesem Modell, das zusätzlich ein "reales Welt"-Modell sein muß. Auch den ausgefeiltesten Kalkülen, wie der Situationssemantik, liegt so eine [eine solche] Abbildtheorie der Welt in symbolischen Repräsentationen zugrunde, also die

naiv platonistische Annahme, daß die Außenwelt sich symbolisch im menschlichen Gehirn abbildet. Intelligenz ist für die KI die operationale Verarbeitung von KI-Wissen. Die Leistungen des Gehirns kommen dieser Annahme zufolge durch Transformationen der erwähnten symbolischen Repräsentationen von Weltausschnitten oder deren Transformierten [??] zustande. Sprache ist für die KI (linguistisch) kodierte Wissen und dient der Übermittlung von Information über "gewußte" Tatsachen in möglichen Welten. Die symbolische KI also sieht **Denken** als komplexes System logischer Schlüsse an, "Verstehen, was man tut", oder **Bewußtsein** sieht die KI als Fiktionen an. Viele KI-Wissenschaftler meinen, wenn Computer nur genügend komplexe Intelligenzaufgaben lösen könnten, würde sich die Frage, ob ein solches System Bewußtsein habe, nicht mehr stellen. Es wird weiter argumentiert, daß Schmerz, Freude, **Fühlen** usw. durchaus simuliert werden können. Doch fehlt der Simulation selbstredend die Unmittelbarkeit der körperlichen Erfahrung. Analoges gilt für **ethische** und moralische Verantwortlichkeiten: Man könne ein ethisches Axiomensystem ebenfalls simulieren. Natürlich haben subjektive und situative Abhängigkeiten in einer solchen Modellierung keinen Platz. Die stark operationale Betrachtungsweise läßt überdies die Motive und Begründungen für den Einsatz von Intelligenz außer Acht.

Es gibt also viele Kritikmöglichkeiten an diesen Vorstellungen. Aus philosophischer Sicht zeigt sich zuerst, daß die Modellbildungen der KI immer auch ein epistemologisches Problem darstellen, das der rationalen Rekonstruktion. Umgekehrt kann die Frage "Ist Künstliche Intelligenz möglich?" nicht unabhängig davon beantwortet werden, welche Intelligenzleistungen dabei in den Vordergrund gestellt werden und in welcher Welt diese künstliche Intelligenz sich bewegen soll. Mithin ist sie nicht unabhängig davon, ob wir vorab Maschinen Intelligenz zuzuschreiben geneigt sind oder nicht. Betrachten wir den aktuellen Stand der KI mit Repräsentationsparadigma, so wird der menschliche Problemlöser bei der Ausführung komplexer Algorithmen und logischer Schlüsse, also bei mechanischen Vorgehensweisen, sicher übertroffen. Hingegen sind kreative Fähigkeiten des Menschen, Sprechen, Mustererkennen, die Fähigkeit, Zusammenhänge zu sehen, ganzheitliches Erfassen, aber auch Integrationsleistungen, Verarbeitung von Erfahrungen, u.v.a. durch den Rechner uneinholbar. Vor allem jene Leistungen, bei denen der Mensch nicht bewußt denkt, also das Erkennen von Figuren, das Aufkommen von Konnotationen und Assoziationen, nicht Bewußtes, sind wohl kaum explizit programmierbar, da sie ja schon nicht zu explizieren sind. Auch die Integration verschiedener, im einzelnen durchaus lösbarer Intelligenzleistungen ist schwierig. Gesamtansätze gelingen bisher nur in Mikrowelten. Jedoch sind Erweiterungen oder Verknüpfungen einer Mikrowelt mit verschiedenen anderen Mikrowelten nicht oder nur sehr eingeschränkt möglich. Als Ursache wurde das fehlende "Weltwissen" von KI-Systemen angesehen. Auswege wurden mit sogenannten Scripts versucht (Roger SHANK) Die Annahme, mit etwa 10000 solchen Scripts auszukommen, erwies sich als falsch.

Auch berücksichtigt sie nicht bestehende kulturelle, historische und subjektbezogene Kontingenzen. So sieht man unmittelbar ein, wie wenig ein amerikanisches Script zur Orientierung etwa in China nützen könnte.

Ein weiterer gravierender Einwand ist, daß Computerprogramme und ihre Herstellung nicht reflektieren, wie Intelligenz phylogenetisch und ontogenetisch entstanden ist, daß sie also weder die Evolution noch die kindliche Entwicklung von Intelligenz erklären können. Ein Computersystem, das auch nur eingeschränkte Gehirnkapazitäten und Ausgangsintelligenz hätte (und wie sollten wir wissen, was diese wären), könnte nicht die gleichen Erfahrungen sammeln wie ein Kind, weil die Körperlichkeit von Silizium-basierter und Eiweiß-basierter Existenz sehr verschieden ist.

Diese Überlegungen an sich schon führen zu einer Entethisierung des Problems und zu stark deterministischen Denkweisen. [*ist es das? Einschub S.11*]

### **3. Methodische Probleme der (symbolischen) KI**

(verschieden von menschlichen Intelligenzproblemen):

Sie betreffen vor allem die Machbarkeit und entstehen wesentlich durch drei Hauptfaktoren:

- A) die rationale Rekonstruktion von Wissen, Intelligenz und Sprache
- B) die Abgeschlossenheit formaler Systeme und
- C) Komplexitätsprobleme der Manipulation.

#### **1. Rationale Rekonstruktion**

Das Paradigma von der rationalen Rekonstruktion geht 1. von der Existenz einer objektiven Welt und 2. ihrer grundsätzlichen Erkennbarkeit durch den Menschen aus; 3. liege Wissen in symbolischer Form, gleichsam als kodierte Abbilder von Wirklichkeit im menschlichen Gehirn vor, 4. bestünden die Mechanismen der Intelligenz im Operieren mit diesen Symbolen; 5. folgt der Erkenntnisprozeß der Realität mit Hilfe der "rationalen Methode", die die Existenz kleinster atomarer Einheiten, z.B. von Symbolen, und die Rekonstruierbarkeit etwa von Wissen durch induktiven Aufbau größerer Bestandteile aus kleineren annimmt.

Die moderne Philosophie stellt (und zwar spätestens seit Kant) Existenz und damit auch Erkennbarkeit einer objektiven Welt in Frage. Was Wirklichkeit ist, wird heute vielmehr als durch soziale Übereinkunft als gemeinsame Festlegung von Vorstellungen konstituiert. Verständnis zeigt sich weniger in der Fähigkeit zur Übernahme von Informationen (durch Abbilder) als in den daraus folgenden Handlungen.

Auch die Ergebnisse der Neurologie widersprechen der KI-Vorstellung.

MATURANA und VARELA zeigen die Abhängigkeit der Erkenntnisfähigkeit von neurologischen Gegebenheiten als typisch für die biologische Art. Sie stellen die Verständnis erzeugenden Leistungen der Sinnesorgane als überlebensorientiert und nicht eine objektive Welt abbildend dar. Einen Organismus bezeichnen sie als autopoetisches selbstorganisierendes System, das durch Außenreize mehr getriggert als mit Abbildern von Wirklichkeit ausgestattet wird. Verständnis entsteht nach ihrer Vorstellung über sogenannte konsensuelle Bereiche, in denen gemeinsame Verständnisebenen existieren. Ähnlich stellt sich Jakob von ÜXKÜLL in seiner Bedeutungstheorie Verständnis und Handeln von Lebewesen vor, nur herstellbar mittels der Existenz von Merkzeichen und Wirkzeichen, die einer Art gemeinsam sind. Sie konstituieren sowohl Möglichkeiten für Verstehen als auch für Handeln. Diese Abhängigkeit der Erkenntnisfähigkeit von den Sinnesorganen wird auch beim Menschen deutlich, wenn man sich vor Augen führt, daß die Empfindlichkeit des Tastsinns  $10^{30}$  mal schlechter ist als die der optischen Wahrnehmung, eine Tatsache, die sehr wahrscheinlich zur Unverständlichkeit des in der Physik nachgewiesenen Welle-Teilchen-Dualismus führt. Wäre unser Tastsinn ebenso sensibel wie der Gesichtssinn, so könnten wir den Lichtdruck spüren und die Teilchen-Natur des Lichtes wäre direkt wahrnehmbar.

Sprache wird in der KI als informativ und nicht performativ begriffen. Sprechen ändert die Welt, ist Handlung und ruft Handlungen hervor, man denke an die Kontaktaufnahme, an Absprachen und Verhandlungen, u.s.w.. Auch wenn in der Situationssemantik versucht wird, solche Aspekte mitzubetrachten, so bleibt sie doch sperrig für subtilere und personen-, situations- und kulturabhängige performative Äußerungen. Linguistische, logische und kognitive Aspekte von Sprache können nicht voneinander getrennt werden, und diese wieder nicht von pragmatischen, überlebensorientierten, sinnstiftenden Inhalten. In seiner Sprechakttheorie stellt AUSTIN das Eingehen von Verpflichtungen in den Vordergrund. Für ihn dient Sprache dem Herstellen von Verständnis und Verbindlichkeiten.

Aus hermeneutischer Sicht (Gadamer, Heidegger) bleibt in der KI unberücksichtigt, daß der individuelle Verstehenshintergrund erst die Bedeutung von Sprache konstituiert, daß Geschriebenes und Gesprochenes also unterschiedlich bei verschiedenen Menschen, Kulturen, in verschiedenen Situationen und zu verschiedenen Zeiten interpretiert wird. Zum Beispiel kann die Äußerung: "im Kühlschrank ist Wasser" abhängig vom Kontext auf verschiedenere Weise interpretiert werden: In der Situation des Essens und bei Hitze wird die Äußerung möglicherweise als das Vorhandensein einer Flasche Mineralwasser interpretiert werden. In anderen Zusammenhängen könnte man das Wasser als Kondenswasser im Kühlschrank, das seine Reparaturbedürftigkeit feststellt, interpretieren. Nimmt man die Äußerung wörtlich im chemischen Sinn, so ist natürlich in jeder Flasche Wein oder Saft 99 % Wasser enthalten, und dennoch wird kaum jemand die Äußerung so verstehen.

Auch aus humanethologischer, ontogenetischer Sicht ist die Vorstellung der KI über Wissen und Sprache kritisierbar. So stellt LENNEBERG durch empirische Versuche fest, daß Kinder Wortbedeutungen verstehen lernen im Zuge eines Differenzierungsprozesses, bei dem zunächst größeren Kategorien Bedeutungen zuordnet werden, die mit wachsendem Verständnis zunehmend differenziert werden.

## **Symbolmanipulation**

Mit der rationalen Rekonstruktion geht das Symbolmanipulationsparadigma einher, also die Vorstellung, daß Wissen in symbolischer Form irgendwie im Kopf gespeichert ist, daß die Mechanismen der Intelligenz im Operieren mit Symbolen bestehen. Zwar erschafft Intelligenz Zeichen und gebraucht sie, wie in der Schrift oder der Arithmetik. Dies wird aber in dem Symbolmanipulationsparadigma als uniformes Medium von Intelligenz mißverstanden. So wird einseitig explizites Denken, also logisches Schließen, Operieren mit Zahlen usw. hervorgehoben, hingegen treten nichtbewußtes oder implizit verfügbares Wissen, gestalthafte Wahrnehmung, integrative Leistungen, ganzheitliches Erfassen von Situationen und Tatbeständen in einem Augenblick in den Hintergrund und sind nur schwer behandelbar. Operieren mit Zahlen, logisches Schließen usw. sind aber Produkt einer langen Entwicklung seit den Ägyptern, Griechen, den arabischen Kulturen - und eigentlich erst seit der Aufklärung so richtig handhabbar durch die Dezimaldarstellung: Die römischen Zahlen erlaubten nicht, beliebig große Zahlen darzustellen, immer neue Zeichen mußten kreiert werden. Damit war keine uniforme Darstellung von Operationen möglich, da diese für jedes neue Zeichen neu zu definieren waren. Algebra, Arithmetik und die heutigen Formen der Logik sind erst seit kurzer Zeit einer intelligenten Menschheitsgeschichte vorstellbar. Intelligenz gab es aber vorher schon und gibt es unabhängig davon.

Geht man nun davon aus, daß das symbolische Manipulationsparadigma eine adäquate Beschreibung von Intelligenz - egal ob menschlicher oder künstlicher - erlaubt, so stehen einige grundsätzliche Klärungen aus: Die Symbole eines Repräsentationssystems von z.B. Künstlicher Intelligenz werden vom menschlichen Betrachter als semantisch interpretierte Zeichen aufgefaßt. Innerhalb des Repräsentationssystems selbst jedoch spielen sie eine rein syntaktisch-formale Rolle. Die Operationen, die Intelligenz konstituieren, operieren auf diesen Repräsentationen und Symbolen ebenfalls in formal manipulierender Weise. Wenden wir dieses Intelligenzverständnis auf das menschliche Gehirn an, so müssen die menschlichen Grundelemente (wobei die Frage, sind es Neuronen oder Neuronenkombinationen, einmal außen vor gelassen sein soll) aber von sich aus eine Semantik tragen, wenn wir nicht einen unendlichen Regreß auf immer neue Interpretieren vornehmen wollen. Aber wer interpretiert und wie wird interpretiert? Ist es das physiologische Material? Und warum nicht elektronisches? Diese Frage wird durch die kognitionswissenschaftliche These der Äquivalenz [Äquivalenz?] von elektronisch oder biologisch sich konstituierender Intelligenz

nicht gelöst.

## **Rationale Methode**

Das Symbolmanipulationsparadigma begünstigt die Vorstellung von unzerteilbaren kleinsten Einheiten von Symbolen und Wissensseinheiten, welche für die Anwendbarkeit der rationalen Methode notwendig ist. Diese Methode hat für Synergie-Effekte allerdings keinen Platz.

Das Vorgehen dieser von BACON und DESCARTES geschaffenen Methode, nämlich das Zergliedern in Einzelteile und die Rekombination mittels uniformer Operationen, beruht auf der Existenz atomarer Einheiten sowie der Annahme, daß Einzelteile kontextunabhängig durch formale Operationen zu größeren Teilen kombinierbar seien. Dadurch endet die Rekonstruktion zu untersuchender Objekte durch sukzessives Zerlegen in kleinere Teile schließlich bei atomaren, und diese können wieder induktiv zusammengefügt werden. Die Geschichte der Verschriftung z.B. jedoch zeigt, daß die uns elementar und unzerlegbar erscheinenden Buchstaben Produkt einer langen und einigermaßen willkürlichen Entwicklung der Definition von Grundeinheiten, nämlich Phonemen und Buchstaben, darstellt, welche in der chinesischen und der japanischen Bedeutungsschrift gar nicht benötigt werden, und in der koreanischen oder der arabischen Schrift zu ganz anderen Grundbausteinen geführt hat. Was wir also für elementare Phoneme halten, ist phonetisch nicht notwendig atomar: oft wären größere Einheiten plausibler, aber es gibt auch keinen Grund, unsere Phoneme nicht weiter zu zerlegen. Auch die Vorstellung, daß sich die Welt in der Kombination von Einzelteilen erschließt, wird z.B. bei der Konstitution von Bedeutung sprachlicher Sätze widerlegt. In natürlichen Sprachen hängt die Bedeutung von Satzteilen von der Bedeutung des gesamten Satzes ab, und die Bedeutung der Kombinationsfunktion ebenso von den Bedeutungen der Satzteile wie von der Bedeutung des Kontextes des Satzes.

Untersuchen wir die genannten Kritikpunkte speziell für wissensbasierte Systeme: Die KI geht hier im Besonderen von der grundsätzlichen Explizierbarkeit von Wissen und Expertentum aus.

Aber erstens ist Wissen immer abhängig von subjektiven Erfahrungen, der historischen, sozialen und kulturellen Herkunft der wissenden Person, und Wissen ist handlungsorientiert. Wissen also ist eine soziale Größe, die zusammenhängt mit Übernahme von Verantwortung. Sie ist stark abhängig von Hintergrundwissen und vom Kontext, in dem die Äußerung geschieht.

Zweitens kann der Experte sein Wissen meist nicht explizieren, das heißt, sehr häufig kann er nur im Problemkontext sein Wissen reproduzieren, er kann es nicht in Einzelteile zerlegen und abstrakt darstellen, er wird auf diesbezügliche Fragen meist mit der Angabe typischer Beispiele reagieren. Überdies erweist sich das Expertentum nicht in regelhaftem Agieren auf bestimmte häufige Situationen, sondern der Experte zeichnet sich gerade dadurch aus, daß er Ausnahmesituationen oder Situationen, in denen sofort gehandelt werden muß,

rasch erkennt. Rasches ganzheitliches Erfassen einer Situation und ihrer evt. Besonderheit kennzeichnet den Experten.

Drittens findet die rationale Rekonstruktion auch den Wissensingenieur, also jenen Informatiker, der das Wissen "aus dem Experten extrahiert" und spezifiziert, um es einer Modellierung zugänglich zu machen, in einer unerfüllbaren Rolle. Theoretisch müßte er ohne Kenntnis des Expertengebietes seiner Aufgabe nachkommen können, praktisch ist er so gar nicht in der Lage, Fragen zu stellen. Man nennt dies "die Blindheit des Wissensingenieurs". Er kann Wissen des Experten nur soweit erfassen, als sein eigener Verstehenshintergrund ihm dieses erlaubt. Der Wissensingenieur also gibt immer konstitutive Anteile in die Rekonstruktion, Anteile seiner Vorstellung und seiner Vorurteile und seiner Fähigkeiten über den Wissensbereich.

Bei der rationalen Rekonstruktion handelt es sich also in Wahrheit um die formale Rekonstruktion der Überzeugung über ein Weltmodell oder ein Expertenmodell durch den Programmierer, besser, um die Formalisierung der reflektierenden Erschließung des Gegenstandsbereichs durch den Programmierer, eingeschränkt durch dessen historisch vermitteltes kontingentes Vorverständnis.

## **II. Die Abgeschlossenheit formaler Theorien**

Bei wissensbasierten Systemen, z.B. Expertensystemen, führt die Abgeschlossenheit zu

- der Unkenntnis des Systems über seine eigenen Grenzen, z.B. würde ein Expertensystem zur Pilzbestimmung eine Schraube als Pilz behandeln und sie entweder kategorisieren oder nicht einordnen können, würde aber nicht bemerken, daß sie kein Pilz ist.

- Zum anderen führt die Abgeschlossenheit zur sog. Nichtrobustheit, d.h. der Eigenschaft, daß bei nur leicht veränderten Eingaben das System mit chaotischen Abweichungen reagieren kann oder mit unerwarteten Fehlern. Unkenntnis des Systems über seine eigenen Grenzen und Nichtrobustheit zusammen führen dazu, daß wissensbasierte Systeme, die Heuristiken benutzen oder bestimmte Annahmen über die Abgeschlossenheit treffen, wie z.B. durch die closed-world-assumption, grundsätzlich nicht reliabel sind. Natürlich kann dem begegnet werden, indem man Eingabebereiche festmacht, innerhalb derer das System korrekt funktioniert, z.B. bei Heuristiken. Dies erfordert jedoch wieder komplizierte mathematische Überlegungen und schränkt den Anwendungsbereich so ein, daß es dann vorteilhafter ist, eine geschlossene Lösung für den Teilbereich zu suchen. In der Praxis versucht man diesem Problem durch eingebaute Erklärungskomponenten zu begegnen, durch die der Benutzer die Vorgehensweise des Systems verfolgen kann. Dies aber erfordert eine gewisse Kompetenz und die Bereitwilligkeit des Benutzers.

Eine weitere Folge der Abgeschlossenheit ist die Unfähigkeit automatischer Beweisverfahren, wirklich interessante oder wichtige mathematische Sätze zu finden.

Der Mensch gewinnt Einsicht aufgrund von Hintergrundtheorien.

Beweiser aber arbeiten ohne Metawelt ihrer eigenen Welt. Daher wäre ein solcher z.B. nicht in der Lage, den Gödel'schen Unvollständigkeitssatz zu finden. Der Gödel'sche Satz besagt, daß in jedem Formalen System, in dem wenigstens die natürlichen Zahlen mit der einfachen Arithmetik ausdrückbar sind, wahre Sätze existieren, die nicht beweisbar sind. Der Beispielsatz G, mit dem Gödel diesen Nachweis führt, behauptet seine eigene Unbeweisbarkeit: sei  $P[k]$  eine Liste von Ausdrücken über einer Variablen  $w$ ,  $B(x)$  eine Liste aller Beweise von Sätzen und wir bilden

G:  $P[k](w) :=$  es gibt kein  $x$  ( $B(x)$  beweist  $P[w](w)$ ) für  $w:=k$ ,

wobei der Ausdruck ( $B(x)$  beweist  $P[w](w)$ ) arithmetisch kodiert ist. Dieser Ausdruck enthält keine Variable mehr und muß daher, will das System nicht widerspruchsvoll sein, nach entsprechender Interpretation der Zeichen entweder wahr oder falsch sein. Ist er wahr, so ist er nicht beweisbar, denn er sagt eben dies von sich aus; falsch kann er nicht sein, sonst würde er etwas Falsches, nämlich sich selbst, beweisen.

Da G also korrekt ist, hat Gödel so einen wahren, aber unbeweisbaren Satz gefunden.

Auf diese Weise sind Metaaussagen über das Formale System in den Satz integriert über arithmetische Aussagen. Die Metaaussagen jedoch melden sich nicht, wenn ein Beweiser diesen Satz erzeugt. Das Finden dieses Satzes erfordert Hintergrundwissen, das von außerhalb des Formalen Systems, in dem der Satz formulierbar und beweisbar ist, kommt. Er würde zwar in einem genügend reichhaltigen formalen System mit allen anderen Sätzen, die in diesem formalen System gelten, erzeugt, jedoch nicht als interessanter Satz, sondern als komplizierter, aber unwesentlicher zahlentheoretischer Satz. Die Interpretation, warum er wichtig ist, würde nicht mitgeliefert. (Man denke an die schönen Bilder aus der Literatur, wie in HESSE's Glasperlenspiel oder an Jonathan SWIFT's Satzmaschine in Gulliver's Reisen.)

Eine solche Metaaussage über das formale System ist dadurch in den oben erwähnten Satz integrierbar, daß diese als arithmetische Aussage in einem zahlentheoretischen Satz kodifiziert ist. Jede Maschine, die den Gödel'schen Unvollständigkeitssatz erzeugte und bewiese, würde diese Interpretation, nämlich welche Aussage hier zahlentheoretisch kodifiziert ist, nicht mitliefern. Auch die Dekodierung wäre nicht auffindbar, man müßte alle möglichen Kodierungen und alle möglichen Interpretationen durchprobieren, wenn man mit der Intention an die von der Maschine gelieferten Sätze herginge, daß diese kodifizierte Metaaussagen enthalten. Die Metainterpretation also wiederum meldet sich nicht. Wie eben beschrieben, ist sogar ein Teil der Interpretation des Satzes im arithmetischen Satz kodiert, er operiert somit im formalen System und

seiner Metawelt gleichzeitig, ein Ebenenwechsel, den das System und der Beweiser in ihm nicht bewältigen können.

Darüberhinaus wären die Gödel'schen Sätze vor 100 Jahren uninteressant gewesen, weil sie in dem historischen Kontext, der damalige Stand der Mathematik und der Logik war, keine Erkenntnis geliefert hätten, mit der man ohne Kenntnis von Hilbert's Programm der Formalisierung der Mathematik etwas hätte anfangen können.

Der Mensch jedoch findet solche wichtigen Sätze durch Intuition aus anderen Theorien, er arbeitet mit Analogien und Kombinationen von Ideen aus den verschiedensten Wissensbereichen. Insofern also ist ein formales Beweissystem blind. Die Bedeutung, nicht die blinde Berechnung, macht die Substanz aus.

Ein anderer Zusammenhang zum Gödelschen Satz ist die seit den 60er Jahren anhaltende philosophische Diskussion um die Frage, ob dieser die starke KI-These widerlege. Die Frage also "Ist der Mensch eine Gödelnummer?" wäre durch den Unvollständigkeitssatz widerlegbar. (Eine sinnvolle Frage nur, wenn ich annehme, daß der Mensch ausschließlich in formallogischen Kategorien denkt.)

### **3. Komplexitätsprobleme**

Es ist eine der Hauptleistungen der Theoretischen Informatik, Formalisierungen für Ressourcenfragen zur Verfügung gestellt zu haben. Sie betreffen vor allem die Komplexitäten des Programmlaufs auf Daten.

Als eine wichtige Methode der KI war Suchen in großen Suchräumen genannt worden. Suchen in Bäumen oder gar in Graphen jedoch führt zur sogenannten kombinatorischen Explosion, also zu einem exponentiellen Suchzeitverhalten, je tiefer man in den Suchraum oder Suchbaum hineingeht. Zum Beispiel beim Suchen nach Beweisen oder beim Versuch, eine vollständige Lösung des Schachproblems zu finden, scheitert ein solches Verfahren an dem notwendigen Aufwand an Zeit und Speicherplatz, welcher im Falle des Schach alle Ressourcen am Material, das das Weltall zur Verfügung stellt, und umso mehr die in Menschenaltern verfügbaren Zeitressourcen sprengt.

Die Komplexität der Manipulation deklarativen Wissens ist oft nicht handhabbar, denn die operationale Bearbeitung explizit repräsentierten Wissens erfordert auch hier Expliztheit und Suchen. Schon elementare Eigenschaften logischer Formeln, wie die Erfüllbarkeit von Formeln in Prädikatenlogik erster Stufe sind bereits unentscheidbar. Schränkt man die Formeln ein auf aussagenlogische, so sind sie immerhin noch nicht handhabbar, denn die Erfüllbarkeit von Formeln der Aussagenlogik ist ein im oben definierten Sinne nichtausführbares Problem. Andere Beispiele sind die Konsistenzüberprüfung von Regelmengen (Regeln sind stark eingeschränkte Formelarten), die Refutation unerfüllbarer Formeln in konjunktiver Normalform beim Resolutionsverfahren, die Subsumtion einfacher Formaler Systeme. Sie alle sind nicht in (deterministischer) Polynomzeit ausführ-

bar. So ist man meist gezwungen, sich mit einer Einschränkung der Sprache zu helfen. Versucht man dies insofern zu automatisieren, daß man Transformationen von der größeren Sprache zu der eingeschränkten durchführt, so ist das aber auch wieder nicht effizient möglich. So ist die Umformung von allgemeinen aussagenlogischen Formeln in disjunktive Normalform wieder ein nicht ausführbares Problem.

Die negativen Ergebnisse der uniformen algorithmischen Behandlung von alltäglichen informatischen Problemen führen gerade dazu, Heuristiken zu Hilfe nehmen zu müssen, mit den erwähnten Folgen für die Robustheit.

### **Konnektionismus**

Hauptunterschied natürlicher Intelligenz und natürlichen Denkens zu Modellbildungen mit symbolischer Repräsentation und rationaler Rekonstruktion ist die Tatsache, daß wir Menschen uns intelligent oder klug in der Welt bewegen, ohne eine Theorie derselben zu besitzen.

Der Konnektionismus kommt ohne rationale Rekonstruktion aus. Er betrachtet kognitive Leistungen als Emergenzphänomene, welche aus der Wechselwirkung hochgeradig untereinander vernetzter einfachster Einheiten entstehen. Konnektionistische Systeme (sogennante "Künstliche Neuronale Netze") versuchen die vernetzte neuronale Struktur von Lebewesen zu kopieren.

So kann Wissen ohne ausdrückliche symbolische Repräsentation erworben und angewendet werden. Konnektionistische Engramme haben keine Konstituentenstruktur und daher keine Kompositionalität. Sie umgehen alle Probleme, die durch die Reflexion aufgeworfen werden. So wird eine Modellierung menschlichen Denkens und Verhaltens versucht, für die das Postulat der vollständigen Explizierbarkeit nicht mehr erhoben werden muß. Sie umgehen damit die oben genannten Probleme. Komplexitätsprobleme der rationalen Rekonstruktion werden eher auf die Ebene der Speichergröße, also auf die Zahl der Recheneinheiten und die des Vernetzungsgrades verlagert.

Die Vorteile der impliziten holistischen Repräsentation und der besseren Lernfähigkeiten konnektionistischer Systeme werden jedoch durch Nachteile erkauft, die vor allem die Kontrolle solcher Systeme betreffen. Symbolische Repräsentationen können exakt im Sinne der Spezifikation sein, das Verhalten konnektionistischer System jedoch kann abgesehen von einfachen Fällen erst im nachhinein retrospektiv überprüft werden. Ausprogrammierte explizite Repräsentationen sind im Prinzip, falls isolierbar und nicht in komplexen Systemen verkoppelt, verifizierbar, das erwünschte Verhalten konnek-

tionistischer Systeme jedoch nicht. Ihre Eigenschaften können nur beobachtet und statistisch erfaßt werden. Es versteht sich von selbst, daß ihr Einsatz in sensiblen Bereichen noch fragwürdiger ist als explizit Ausprogrammiertes.

Was konnektionistischen Systemen vorläufig noch weitgehend fehlt, ist die sensomotorische Einbettung in die Umgebung. Sie ist jedoch wegen der Verteiltheit der Recheneinheiten und der Möglichkeit direkten Zugriffs auf jene nicht absolut ausgeschlossen. Damit ist es im Prinzip möglich, durch Sensoren, Effektoren usw. in unmittelbaren physischen Kontakt mit der Umwelt zu treten. Im Gegensatz zu ausprogrammierten Intelligenzleistungen auf klassischen Rechnern sind sie im Prinzip nicht abgeschlossen. Auch das Zeitkomplexitätsverhalten könnte eine unterschiedliche Kommunikation mit der Umgebung erlauben, da das Netz in jedem Zeittakt auf Eingaben von außen und auf die Propagation des eigenen Systemzustandes reagiert. So sind Realzeitantworten des Systems möglich, wie zum Balancieren eines Stabes oder zur Steuerung eines Gefährtes erforderlich. Für programmierte Fähigkeiten hingegen könnte die Reaktion auf einen Außenreiz ein Problem hoher Komplexität darstellen, welches angemessene Reaktionszeiten ausschließt. Dies mag einen prinzipiellen Unterschied zwischen konnektionistischen Systemen und explizit programmierten Lösungen auf klassischen Rechnern darstellen. Da heute solche Netze noch meist auf symbolischen von-Neumann-Rechnern simuliert werden und somit wieder programmiert sind, gilt noch das Argument, daß die Simulation eine prinzipielle Überschreitung programmierter Fähigkeiten nicht zuläßt. Sobald aber Hardwarelösungen existieren, ist eine andere Einbettung in die und Kommunikation mit der Umgebung denkbar. Die Realzeitantworten sind prinzipiell auch durch Beschleunigung der Rechenzeit von klassischen Rechnern nicht einholbar. Auch die direktere verteilte mögliche Einbettung solcher Netze in die Umgebung macht Reaktionen auf physikalische Zufallseingaben möglich und für diese ist unklar, ob sie Berechenbarkeitseinschränkungen der expliziten Algorithmik nicht sprengen.

Natürliche neuronale Netze jedenfalls sind erheblich komplizierter als alles, was an Konnektionistischen Systemen vorstellbar ist, soweit wir sie überhaupt wissenschaftlich untersucht haben. Die Diversifikation von Neuronen und Neuronengruppen ist sehr groß, Neuronen und Neuronengruppen sind spezialisiert auf spezifische Aufgaben. (Am besten untersucht sind dabei sensorische Schichten von optischen Fähigkeiten.) Die Zahl der Neuronen und ihr Verflechtungsgrad sind physikalisch (noch?) nicht simulierbar. Nur die elektrischen Leitvorgänge sind grob simulierbar, nicht jedoch die chemischen Vorgänge, die eine weitere Diversifikation der Leitungs- und Inhibitionsvorgänge bewirken und ebenfalls lokal unterschiedlich und kontingent sind.

Die heutige Frage ist nicht "wie menschlich sind Maschinen?" sondern "wie maschinell wird der Mensch?" Dies auch insofern heute die menschliche Kommunikation mit ihrer maschinellen Vermittlung gleichsam immer mehr eine unauflösliche Verbindung eingeht, d.h. man kann das Menschliche kaum noch ohne das Maschinelle denken, wie im folgenden dargelegt wird.

## **Datenautobahnen, moderne Netzkommunikation**

### **Internet und Datenautobahnen**

Aus dem programmierbaren Berechnungs- und Transformationsmedium Computer ist heute ein Informationstransport- und Kommunikationsmedium *Netz* geworden.

So ist mittels telnet, ftp, www, gopher, lynx, bulletin boards, verschiedenen Recherchesystemen und Datenbanken (z.B. Bibliothekskataloge), ein gezielter Zugriff auf Programme, Daten und Informationen möglich, die irgendwo auf der Welt gespeichert sind. Die on-line Distanzüberbrückung vergrößert die individuelle Reichweite, und die Recherchemöglichkeiten zur Informationsbeschaffung vervielfachen gleichsam janusköpfig die eigene Verfolgungskraft, die früher auf Bibliotheken, Telefon, persönliche Informationen angewiesen war. Die Verortung von Wissen und Informationen im Netz macht diese beliebig reproduzierbar, erleichtert Täuschung, Fälschung und Diebstahl.

Zusammen mit den vielfältigen Interaktionsmöglichkeiten wird beispielsweise wissenschaftliches Wissen schwerer schützbar und entindividualisiert, wird gleichzeitig weiter verbreitet und unverlässlicher: dies in mehrfacher Hinsicht: gewollte und ungewollte Veränderungen sind leicht möglich; der Gültigkeitsbereich kontingenten Wissens kann unzulässig verallgemeinert werden; veraltetes Wissen überdauert unkontrolliert, wenn entsprechende Dateien nicht mehr gepflegt werden..

Die Selbstdarstellungsmöglichkeiten in www, etc. zeigen potentielle und erzeugen neue entfernte Kommunikationspartner in beliebiger Anzahl. Rasch und mühelos kann mit emailing, voice mail, news- oder chatting groups über grosse Strecken kommuniziert werden, ohne daß die gleichzeitige Anwesenheit der Kommunikationspartner erforderlich ist (ähnlich wie bei der Post), indem die Speichermöglichkeiten des Adressaten genutzt werden. D.h. die zeitliche Synchronisation der Kommunikation wird von räumlichen Medien, den mailboxes und anderen Speichermedien geleistet. Auch durch diese Gruppenbildung im Netz wird eine Vervielfachung von Kontakten erreicht. Zeitlich direkte, aber räumlich getrennte Kommunikation wird mit telnet, talking, conferencing, oder shared activities, wie computer supported cooperative work unterstützt.

Mehr noch werden nämlich mit Hilfe sicherer Kryptierverfahren elektronische Geschäfte, elektronische Überweisungen, elektronische Unterschriften,

elektronisches Geld, elektronische Zertifizierung, - Verhandlungen, - Rechtssicherungen, wie - Urkunden, - Haftung, Lizenzierung, elektronische Wahlen u.v.a.m. im Internet zur Verfügung gestellt. Die Plazierung von Geschäften im Netz erfolgt umstandslos und nahezu kostenlos im Vergleich zur Etablierung eines physischen Geschäfts. Je nach Verkaufsobjekt muß ein Versand abgeschlossen werden oder nicht: virtualisierbare Waren können direkt übers Netz verteilt werden: Bücher, Seiten, Zeitungen, Informationen, know how, Dienstleistungen, Programme, Geld, u.s.w..

Die Verortung nicht nur von Information, Wissen und Wissenschaft, sondern auch von Wirtschaft, Recht, politischen Vorgängen in elektronischen Speicher- und Kommunikationsmedien virtualisiert diese Systeme. Durch die Formalisierung und durch den Wegfall von menschlicher kontingenter Bearbeitung werden diese dadurch schwerer adaptierbar, fester und härter für die betroffenen Menschen. Gleichzeitig macht die Virtualisierung sie auch weicher und angreifbarer. Sie könnten so leicht ihre Bindungskraft verlieren, wenn Menschen Virtuelles weniger ernst nehmen, oder ihnen der Umgang mit der inadäquaten Unabänderlichkeit unmöglich gemacht wird.

## **Virtualität**

### **Der virtuelle Raum:**

Wenn der soziale Raum als Netzwerk von Kommunikationen angesehen wird, so unterscheiden sich reale und virtuelle Räume nur durch den Einschub des Mediums Computer(netzwerk) in die Kommunikationen. Allerdings ändert sich durch die technische Vermittlung die menschliche körperliche Erfahrung.

Virtualisierung bezeichnet einen Transformationsprozeß von vormals physischen Gebilden und Vorgängen, wie Geld oder Kontrakten, Unterschriften in elektronischer Form, notariellen Beglaubigungen, Wahlen etc..

Der durch die Informationstechnik verfügbar gewordene virtuelle (Speicher-) Raum wird gefüllt mit Texten, künstlichen oder manipulierten Bildern, (manipulierten) Filmen, dreidimensionalen bewegten Körpern und virtuellen Innen- und Aussenräumen, in die interaktiv fiktive Reisen unternommen werden können.

Bei den Abenteuerspielen mit Raumschiffen oder in virtuellen Körpern wird das Virtuelle nicht nur beobachtet, sondern im Rollenspiel auch bereits selbst verkörpert. Orte eher denn Vorgänge stellen die Grundelemente des Spieltextes dar. Und sie sind nicht kausal verknüpft, sondern metonymisch. Zwar vergeht beim Spielen Zeit, doch die Spielstruktur ist eher eine räumliche. Das interaktive Spiel gelingt über eine Schnittstelle, derer sich die Spielerin bedienen muß oder mit der sie interagieren muß.

Eine weitere Dimension dieses Ineinandergreifens von Realität und Fiktion entsteht im Cyberspace. Mit Datenhandschuhen und -anzügen, sowie Bild-

schirmkopfhelmen, die mit Detektoren, Sensoren und dem ganzen technischen Arsenal der Telekommunikations ausgestattet sind, werden fiktive Ortswechsel, Treffen, Berührungen, die Wirkung eigener Handlungen sinnlich erlebbar, ohne daß der Ort der realen Handlung verlassen werden muß. Die Simulationstechniken, welche in "Realzeit" funktionieren, verwandeln Nähe, Bewegung, Widerstände, Distanzen in virtuelle Realitäten. Sein und Schein sind verbunden. Der Mensch kommuniziert sehend, tastend und fühlend mit der Maschine, so als ob er sich in dem simulierten Raum bewegte, in ihm aneckte oder handelte, so wie Alfred Jarry es sich mit seiner Liebesmaschine vorgestellt hat.

Und von der unmittelbaren körperlich-sinnlichen fiktiven, aber doch Erfahrung abgesetzt entsteht in einem weiteren Abstraktionsschritt die Vorstellung des schwebenden Flugs durch das Internet, des Surfens im Informationsnetzwerk. Die Cybernautin bewegt sich gleichsam zwischen räumlichen Formationen aus hosts, Software-Systemen und Leitungen zu entfernten Datenräumen, wird dabei behindert durch firewalls, Software- und Kommunikationsschutzvorrichtungen, die selbst wieder Software sind, tarnt sich als Cyberpunk, um gefälschte Daten einzuschleusen, bricht Kryptierungen, um an fremde Daten, Kommunikation oder Software zu gelangen, sie unbemerkt und unbemerkbar zu stehlen. Der sinnliche Mensch tritt hier zurück gegenüber dem virtuellen Raum, er äußert sich nur noch im Tippen auf der Tastatur, in den Mausbewegungen. Dafür feiert sein Geist die Erfolge seiner abstrakten Vorstellungen und Fähigkeiten.

In unserem Zusammenhang von Interesse ist auch die Agentenprogrammierung. Das sind Programme, die selbständig im Netz Aktionen durchführen, in Abhängigkeit von Vorfindlichem Informationen suchen und kombinieren u.s.w.. Agenten werden auch in Zusammenhang mit Arbeitsflüssen menschenähnlich modelliert. Sie sind dann innerhalb des Netzes von Menschen nicht unterscheidbar.

Weithin bekannt geworden sind bereits die negativen Auswirkungen der beliebigen Vervielfachung durch Agenten im Netz: Die kombinatorische Explosion von Agenten kann das Zumüllen der Leitungen und der privaten mailboxes mit unbrauchbaren Informationen bewirken. Solche Effekte, aber auch die Überallgegenwärtigkeit des Gespeicherten im Netz haben information filtering auf den Plan gerückt. Dies kann etwa wieder mittels lernender Agenten geschehen, oder aber mit menschlichen Instanzen: ein neues Berufsfeld eröffnet sich hier. Die Macht der Filterinstanzen jedoch ist groß, ihr eigener bias, ihre Interessen, ihre Blindheit haben Einfluß auf die Verfügung von Wissen.

Agenten sind Programme, die von netizens in Gang gesetzt oder getriggert selbständig Aufgaben ausführen. Sie können selbst neue Agenten anstoßen oder gar erzeugen und so explosionsartige Vermehrungen ihrer Aktionen herbeiführen. Agenten können wesentliche features von im Netz operierenden Menschen einprogrammiert haben, so daß sie im Netz von menschlichen

Nutzern nicht unterscheidbar sind.

## **Die virtuelle Identität**

Letztendlich ist unklar, ob es sich bei den netizens, den Bürgern der Netze, um die an ihrem Terminal tippenden Menschen handelt, oder um die als Agenten, homepages oder virtuelle Identitäten von jenen in Gang gesetzten Spiegelungen. Ähnlich können menschliche Nutzer sich tarnen oder geborgte Identitäten im Netz simulieren. Bekannt wurde beispielsweise Julia, die von einem männlichen Nutzer in Gang gesetzte virtuelle Frau, die mit vielen anderen netizens in Kommunikation trat und alsbald ihr Eigenleben zu führen begann, sodaß ihr Kreator sie an einer Krankheit sterben lassen wollte. Dies war wegen der übermächtigen Hilfsbereitschaft der Netzpartner nicht möglich, sodaß sie ein weiteres Jahr im Netz ihr Leben fristete, bis ihr erschöpfter Kreator die simulierte Existenz aufdeckte.

Das Netz und die netizens also erzeugen im freien Spiel einen virtuellen Raum von simulierten und manipulierten Vorstellungen, bewirken eine Vervielfältigung von Identitäten, die einer ontischen Existenz nicht mehr bedürfen. Deutlicher noch werden diese Virtualisierungen an stillen und bewegten Bildern, die den Unterschied zwischen Erfindung, Manipulation und Abbild nicht mehr zeigen und so ihre eigene Welt erzeugen, deren Rückführung ins oder Konfrontation mit dem wirklichen Leben Brüche schafft.

Die Virtualisierung von vormals physischen Systemen oder von Teilen entkoppelt oder befreit von deren Bindungen. Dies führt zur Entbindung der Subjekte von Verantwortung, zur Entethisierung. Die Wirkung ist die Ablösung von und Indifferenz gegenüber Orten und dem realen Leben. Doch was ist dann das reale Leben, findet es nicht auch im Cyberspace statt?

Die virtuellen Intelligenzen scheinen Element der technischen Evolution oder der angeblichen Netzdemokratie zu sein; die Benutzer bemerken nicht, daß diese von (meist wirtschaftlichen) Entscheidungen von Politik und Industrie abhängen und getriggert sind. Zu fragen wäre: Ist die reale technische Entwicklung demokratisch legitimiert?

Die Virtualität von Identitäten und Agenten im Netz hat diese Fragen bereits an die Cyberexistenz abgegeben. Es interessiert nicht mehr, wie der Mensch denkt, sondern welche Erscheinungsformen, Aktionen und Identitäten er im Netz darstellen und ausführen kann.

Die Ununterscheidbarkeit virtueller und wirklicher Agenten ist ethisch und sozial relevant. Es besteht die Gefahr, daß Menschen keinen Unterschied mehr sehen, sich geborgter Identitäten bedienen, vielfältige Identitäten benutzen und daß sie sich weiter mit Computern und vernetzter Intelligenz identifizieren, und

somit wirklich ununterscheidbar werden.