

Der moderne Mensch ist im einzelnen
so kenntnisreich wie im ganzen unwissend und blind.
Robert Musil

Das Evangelium des robo sapiens

Eine Sammelrezension von Neuerscheinungen zur künstlichen Intelligenz

Peter Brödner

WZN-Institut Arbeit und Technik, Gelsenkirchen

1

„Vor Ablauf des nächsten Jahrhunderts wird der Mensch seine Stellung als das intelligenteste und leistungsfähigste Wesen auf Erden verloren haben.“ Mit diesen Worten läutet der amerikanische Computerexperte Ray Kurzweil den eher unrühmlichen Abgang des homo sapiens als Krone der Schöpfung ein (18). Eine höhere Intelligenz, vom Menschen selbst in Gestalt mikroskopisch kleiner Roboter aus spezialisierten, zu Berechnungen fähigen Molekülen („Nanobots“) geschaffen, bestimmt künftig, im „postbiologischen Zeitalter“, das Schicksal der Welt: „Computer werden einen freien Willen haben. Sie werden spirituelle Erfahrungen für sich reklamieren. Und die Menschen, deren Denken noch immer von der Arbeit organischer Neuronen abhängt – werden ihnen glauben“ (24). Die Maschine unterwirft ihren Schöpfer, und das schon ab dem Jahr 2020.

Die Gegenstimme ließ nicht lange auf sich warten. In der Internet-Zeitschrift „Wired“ meldete sich Bill Joy, „Chief Scientist“ bei Sun Microsystems und Berater von US Präsident Clinton, zu Wort und verlangte flugs ein Moratorium für diese neben der Gentechnik „gefährlichste Technologie“ des 21. Jahrhunderts (nachgedruckt in der FAZ vom 6. 6. 2000).

Nun wird die Diskussion um Technischen Fortschritt immer schon beherrscht von Heilsverkündern und Apokalyptikern. Beide haben, wie Hans Magnus Enzensberger kürzlich in einem Essay über Computer und Medien trefflich bemerkte, „etwas gemeinsam: Propheten sind gegen Tatsachen immun.“ Auch die frohen Botschaften von der „künstlichen Intelligenz“ lesen sich immer wieder gleich. Nach Verkündigung der jede Vorstellung sprengenden Leistungen der Computersysteme folgt der Appell an den Wunderglauben des verblüfften Publikums. Mangels prüffähiger Tatsachen werden allerlei Hoffnungen und Ängste aufgerührt, entgegenstehende Erkenntnisse werden ignoriert. Ein starker Eindruck von déjà vu stellt sich ein: Das Muster bleibt, es wechselt nur die technische Basis.

Bereits Ende der fünfziger Jahre sagte Herbert Simon voraus, „daß binnen zehn Jahren ein digitaler Rechner Schachweltmeister sein wird, es sei denn, das Reglement schlosse ihn aus; daß binnen zehn Jahren ein digitaler Rechner einen wichtigen neuen mathematischen Satz entdecken und beweisen wird; daß binnen zehn Jahren die meisten psychologischen Theorien

die Form von Computerprogrammen oder von qualitativen Aussagen über die kennzeichnenden Eigenschaften von Computerprogrammen annehmen werden“ (Simon / Newell 1958, 6). Technische Basis war hier der „General Problem Solver“, ein Programm, das mittels formaler Symbolverarbeitung ein umfassendes Zweck-Mittel-System des untersuchten Problems in der Gestalt eines Entscheidungsbaumes systematisch zu analysieren imstande ist. Pech nur, daß sich so nichts davon bewahrheitet hat.

Mit Systemen der formalen Symbolverarbeitung lassen sich, das lehrt die nunmehr über vierzigjährige Geschichte der „künstlichen Intelligenz“, auf ganz bestimmten Problemfeldern erstaunliche, gleichwohl nur begrenzte Erfolge erzielen. Bestimmte Aufgabentypen sind einer formalen Symbolmanipulation durchaus zugänglich, andere sind es nicht. Es sind, wie Hubert Dreyfus (1979) schon damals gegen die Expertenzunft zeigte, vor allem wohlstrukturierte Probleme, seien sie auch noch so kompliziert, deren Lösung auf explizites Wissen zurückgeführt und durch Methoden formaler Symbolmanipulation automatisiert werden kann. Hier, insbesondere etwa bei Konfigurationsproblemen, bei mathematischen Beweisverfahren oder auch beim Schach, feiern die klassischen KI-Systeme ihre Erfolge. Aber das ist es dann auch.

Auffallend schwach bleiben die Leistungen dieser klassischen Methoden in nur schwach strukturierten und kaum formalisierbaren Problemfeldern wie beispielsweise bei der Simulation des gesunden Menschenverstandes, bei stark kontextabhängigem Handeln oder beim Verstehen natürlicher Sprache. In den achtziger Jahren erleben daher, wiederum begleitet von großem publizistischem Getöse, so genannte adaptive Netzwerke oder künstliche neuronale Netze (KNN) eine Renaissance. Sie nehmen sich die Arbeitsweise von Nervensystemen direkt zum Vorbild und vermögen vor allem bei der Musteridentifikation und Klassifikation, bei denen es auf formal schwer faßbare Ähnlichkeiten ankommt, Erstaunliches zu vollbringen. Sie versagen freilich, eben aufgrund ihrer behaviouristischen Natur, bei der logischen Verarbeitung zusammengesetzter symbolischer Ausdrücke.

Und nun also Ray Kurzweil, der seine „Methoden der Intelligenz“ ungeachtet dieser schlechten Erfahrungen auf drei „einfache Paradigmen“ gründet („Wie man mittels dreier einfacher Paradigmen eine intelligente Maschine baut“; 124ff, Anhang): Erstens auf das Prinzip der Rekursion und die rekursive Suche in Entscheidungsbäumen (s.o. formale Symbolmanipulation), zweitens auf die Funktionsweise lernfähiger neuronaler Netze (s.o. KNN) und drittens auf den Einsatz evolutionärer Algorithmen. Zusammengenommen erwartet er von ihnen, daß sie von selbst intelligente Problemlösungen in den unterschiedlichsten Kontexten zustande bringen – vorausgesetzt, sie können über hinreichend vielfältiges Weltwissen verfügen. Kein Wort verliert er freilich darüber, ob und wie diese ganz divergenten Softwarekonzepte überhaupt miteinander verknüpft werden können.

Kurzweil gründet seine Vision auf sein „Gesetz vom steigenden Ertragszuwachs der Evolution“ (35ff, 164ff): Dabei sieht er in der kulturellen Entwicklung die sich ständig beschleunigende Fortsetzung der Evolution mit anderen Mitteln, ohne freilich für die rasch wachsende Ordnung außer vereinzelt Indikatoren auch nur die geringste Plausibilität anzuführen. Die zahlreichen Fehlentwicklungen, Sackgassen, und Regressionen der biologischen und kulturellen Entwicklung bleiben außer Betracht. Es geht nur nach vorn, und das immer schneller. Immerhin gibt es das „Moore'sche Gesetz“, ein empirischer Sachverhalt, demzufolge sich die Zahl der elektronischen Strukturelemente und damit auch die Rechenleistung von integrierten Schaltkreisen etwa alle 24 Monate verdoppelt. Auch wenn bei fortschreitender Miniaturisierung die elektronisch effektiven Strukturen die Größenordnung einzelner Atome erreichen, könnten diese, so Kurzweil, fürderhin auch in der

dritten Dimension räumlich angeordnet werden. Auf dieser Basis sagt er voraus, daß um das Jahr 2020 die „Komplexität“ (er meint wohl die Zahl der Schaltelemente) handelsüblicher Computerprozessoren („die Hardware der Intelligenz“, 163) diejenige des Gehirns mit seinen schätzungsweise 10^{11} Neuronen übertreffen wird.

Insgesamt hinterläßt die Lektüre des rund 500 Seiten starken Wälzers den Eindruck eines permanenten Wechselbades zwischen der Fülle des Stoffs und der Ignoranz des Themas, zwischen dem Plauderton der Darstellung und der Kühnheit der Behauptungen, zwischen der Vielzahl lesenswerter Fakten und ihrer Belanglosigkeit für die Realisierung „künstlicher Intelligenz“. Nach dem Motto: „Eine starke Behauptung ist besser als ein schwacher Beweis“ strotzt das Buch von weitreichenden, aber letztlich unbewiesenen Aussagen, so etwa das „Gesetz vom wachsenden Ertrag der Evolution“, das „Verstehen natürlicher Sprache“ durch Computer und das „simultandolmetschende Telefon“ (19, 21) oder die Gleichsetzung von neuronalen Schaltzuständen des Gehirns mit Bewußtsein, dessen ganzer Bewußtseins- und Gedächtnisinhalt sich folglich „komplett scannen“ und auf die nanotechnisch geschaffene Hardware „herunterladen“ läßt (90ff, 191ff). Prägnanter kann das intellektuelle Elend des Reduktionismus kaum zum Ausdruck gebracht werden.

Intelligenz und Bewußtsein beruhen auf weit mächtigeren Voraussetzungen, auch wenn sie noch niemand genau kennt. Gelegentlich reißt selbst Kurzweil bekannte Probleme der Realisierung „künstlicher Intelligenz“ an, bleibt dann aber ernst zu nehmende Lösungsansätze schuldig. Immer wenn es dabei wirklich interessant zu werden scheint, wirft er Nebelkerzen aus einer Vielzahl belangloser Details. Das typische Werk eines Kopfes voll Wissen, aber ohne jeden Verstand. Mit dem hoch entwickelten Stand des wissenschaftlichen Diskurses um die „künstliche Intelligenz“ und den scharfsinnigen Argumenten ihrer Kritiker setzt er sich nirgends ernsthaft auseinander. Der letztlich alles entscheidende Punkt bleibt ihm völlig verborgen: Wie gelangen wir von der Physik zur Semantik, wie vermag das Gehirn seinen Erregungsmustern oder „Schaltzuständen“ angemessene, kontextsensitive Bedeutung zuzuweisen (der Kern von Wahrnehmung und Verstehen)? Wie lernt es, komplexe Situationen zu bewerten? Wie vermag es Bewußtsein (sogar seiner selbst) zu erzeugen? – Gewiß nicht nur eine Frage von schneller Hardware und ein paar „einfachen Paradigmen“.

2

Unter ganz ähnlichen (und noch anderen) Schwächen leidet auch das Buch des zweiten Evangelisten, Hans Moravec, Professor für Robotik an der Carnegie Mellon University. Wie Kurzweil gründet er seine Hoffnungen auf MIPS und *modus ponens*, auf rasche Leistungssteigerung extrem miniaturisierter Hardware und effektive Suche in logischen Entscheidungsbäumen. Sie verhelfen der Robotik – bei Moravec der Inbegriff „künstlicher Intelligenz“ – nach „dreißig Jahren Stagnation“ zu neuer Blüte: „Rechenleistung ist für den Geist, was die Lokomotive für den Zug ist“ (83). Auch er beruft sich dabei auf die Leistung des Supercomputers „Deep Blue“ im Schachturnier gegen Kasparow als (wenn auch untauglichen) Beleg. Und wie Kurzweil (nur weit simpler in der Argumentation) setzt er bei der Realisierung „künstlicher Intelligenz“ auf lernfähige Software nach Art neuronaler Netze.

Auf dieser Grundlage entwirft Moravec eine Generationenfolge universaler Roboter, die im Abstand von jeweils zehn Jahren zu immer höheren Intelligenzleistungen fähig sein sollen. Sie beginnt im Jahre 2010 auf „Eidechseniveau“ mit der Auswertung komplizierter dreidimensionaler Bilder und dem Verstehen einfacher Sprache. Über Zwischenstufen soll sie dann im Jahre 2040 „Menschenniveau“ erreicht haben mit der Fähigkeit, aufgrund „übermenschlicher Denkprogramme ... gleichzeitig die Welt simulieren und über ihre

Simulation nachdenken“, dabei u.a. neue Programme entwerfen zu können (169). Sie vermöchten dann auch natürliche Sprachen zu verstehen und auf der Basis von „Psychologiemodellen“ Bewußtsein zu entwickeln und Gefühle zu zeigen (178ff).

Soweit bringt er nichts wirklich Neues vor. In den restlichen zwei Fünfteln des Buches verliert sich Moravec dann in einem Dschungel verschlungener Cybermystik, den ein „normaler“ Verstand nicht mehr zu durchdringen vermag. Es beginnt noch eher harmlos mit der Wiederaufnahme des alten Traums, in dem die „intelligenten“ Roboter voll automatisierte Fabriken bauen, worin sie sich selbst ohne Menschenhand reproduzieren, jede Folgegeneration kleiner und klüger, versteht sich. Der Wettbewerb der Roboterfabriken führe dann zum „Ende des Kapitalismus“ und verweise die Menschen auf ihren Platz „reichen Müßiggangs“ unter „Schweizer Verhältnissen“ (!) (206ff). Die Roboter verwandelten sich allmählich in „Exes“ mit dem Drang, den Weltraum zu erobern (226ff); dazu bauten sie „Orbitalbrücken“ zu anderen Himmelskörpern. Mit dem Vordringen der Exes in den Weltraum werde das „Zeitalter des Geistes“ geboren, freilich unter den Wehen „geistiger Kriege“ zwischen Exes (255). Es entstünden „Informationsmuster im Cyberspace“ (257), am Ende losgelöst von materiellem Substrat: „Körper ade“ (265ff).

Das ist Science Fiction, die sich als Science ausgibt und dabei zur miserablen Fiction gerät. Ihr fehlt jede innere Entwicklungslogik und Plausibilität des Geschehens. Es ist das konfuse Werk eines Kopfes, der sich durch allerlei unverdaute Lektüre selbst verwirrt hat. Richtige Science Fiction liefert weit Erhellenderes und Plausibleres über mögliche technische Entwicklungen.

3

Da schlägt das schon etwas ältere Buch von Eric Drexler et al., obgleich ebenfalls als Zukunftsvision angelegt, weit nüchternere und realitätsnähere Töne an. Es bietet zunächst eine gut verständliche Einführung in das weite Feld der Nanotechnologie und die Vielfalt der Methoden und Werkzeuge, mit denen Moleküle gezielt auf- und umgebaut und sogar einzelne Atome manipuliert werden können. Das alles spielt sich in der Größenordnung von Nanometern (10^{-9} m) ab, daher der Name.

Sinn dieser Eingriffe in die Materie auf molekularer und atomarer Ebene ist die Herstellung besserer Produkte (wie Hochleistungscomputer, hoch belastbare, aber federleichte Konstruktionen oder „intelligente“ Werkstoffe) zu sinkenden Kosten und mit geringerem Rohstoffverbrauch. Darauf werden große Hoffnungen gesetzt: Produktion ausreichender Nahrung auf kleinstem Raum, Vermeidung und Abbau von Giftmüll, Wasseraufbereitung und Reinigung der Atmosphäre oder Immunisierung gegen Seuchen. Trotz euphorischem Grundtenor werden dabei auch Sicherheitsrisiken, mögliche Unfälle und Mißbrauchsgefahren ausführlich erörtert.

Von besonderem Interesse im vorliegenden Kontext ist ein weit voraus eilendes Szenario nanotechnischer Produktionsprozesse u.a. zur Herstellung von „Nanocomputern“. Dabei wird ausdrücklich die Replikation solcher zu Berechnungen fähiger molekularer Gebilde ins Auge gefaßt, ein Prozeß, in dem diese bei entsprechender Versorgung mit Rohstoffen und Energie Duplikate von sich selbst herzustellen imstande sein sollen. Leider bleiben Funktionsweise, Herstellungsprozeß und Replikation der Nanocomputer auch in diesem Szenario weitgehend im Dunkeln. Ihre Plausibilität oder gar Machbarkeit lassen sich so nicht beurteilen. Diese Vagheit der Beschreibung ist umso bedauerlicher, als Nanocomputer und Replikation bei

Kurzweil (unter ausdrücklicher Berufung auf Drexler) und dann auch bei Joy eine zentrale Rolle spielen. So steht im Grunde deren ganze Vision auf äußerst wackeligen Füßen.

Aufgeschreckt durch die Bücher von Kurzweil und Moravec nimmt sich dann Bill Joy der Sache an und sieht, ganz in der Rolle des Apokalyptikers, den Himmel über dem homo sapiens einstürzen. In seinem „Wired“-Artikel übernimmt er ohne Wenn und Aber die Visionen der „nanotechnologischen Revolution“ mit ihren sich selbst replizierenden „Nanobots“. Ohne in der Sache selbst etwas hinzuzufügen, beschwört er wortreich die Gefahr der Verdrängung des Menschen als Folge unkontrollierter Vermehrung dieser superintelligenten molekularen Computer und stellt sie in eine Reihe mit den Gefahren der Gentechnik.

Seine Einlassung gipfelt in der Forderung nach einem Moratorium: „Mit der Gentechnik, der Nanotechnologie und der Robotik öffnen wir eine neue Büchse der Pandora ... Diese Technologien sind zu mächtig, als daß wir uns in der zur Verfügung stehenden Zeit vor ihnen schützen könnten. Und selbst wenn wir solche Schutzschilde entwickeln könnten, wären die Nebenwirkungen ihrer Entwicklung mindestens ebenso gefährlich wie die Technologien, vor den sie uns schützen sollen ... Die einzig realistische Perspektive, die ich sehe, lautet Verzicht: Wir müssen auf die Entwicklung allzu gefährlicher Technologien verzichten und unserer Suche nach bestimmten Formen des Wissens Grenzen setzen“ (FAZ vom 6. 6. 2000).

Bei aller gebotenen Vorsicht im Umgang mit Technik, das erscheint nun doch als völlig überzogene Panikreaktion auf Visionen, die noch auf lange Sicht jeder Realisierungsmöglichkeit entbehren – ganz abgesehen davon, daß Moratorien Erkenntnisfortschritte noch niemals haben aufhalten können (da hat Kurzweil ganz recht). Nanotechnische Miniaturisierung und entsprechende Steigerung der Rechenleistung (an der nicht zu zweifeln ist) führen per se nicht im mindesten zu mehr Intelligenz und die derzeitigen Bemühungen um softwaretechnische „Methoden der Intelligenz“ zeitigen vergleichsweise derart klägliche Resultate, daß eher die ganze Vision in Frage steht als daß deren Gefahren beschworen werden müßten. Auch wenn Computer Meisterschach spielen oder gesprochene Sprache algorithmisch in Text umsetzen, den Sinn der Sätze verstehen sie damit noch keineswegs, und Bewußtsein ist ihnen damit nicht gegeben.

4

Soweit in aller Kürze die Auseinandersetzung in der Sache. Zum ausgewachsenen öffentlichen Ärgernis wird die Kurzweil-Joy-Debatte freilich erst dadurch, daß diesen weltfremden Heils- bzw. Unheilsverkündern im Flaggschiff des deutschen Feuilletons breiter Raum gewährt wird. Wohlgemerkt: ohne ihnen in der Sache etwas Adäquates entgegenzusetzen, was, wie hier argumentiert, dringend geboten wäre. Wer seine Zukunft an solchen sich wissenschaftlich gebenden Gaukeleien orientiert, hat sie schon verspielt, bevor sie recht begonnen hat.

Grundsätzlich ist es in höchstem Maße zu begrüßen, daß sich das Feuilleton endlich der eminent wichtigen öffentlichen Auseinandersetzung um derart bedeutsame, die Gesellschaft sicherlich prägende technische Zukunftsentwicklungen wie die Computer-, die Nano- oder die Gentechnik öffnet, insbesondere ihren Chancen und Risiken. Das darf aber nicht auf dem niedrigst möglichen Niveau geschehen, ganz ohne Rücksicht auf den hohen Entwicklungsstand des wissenschaftlichen Diskurses um „künstliche Intelligenz“ und das Gewicht der Argumente ihrer Kritiker. Wie kläglich das Niveau derzeit ist, läßt sich auch an der 6-seitigen Präsentation der letzten Sequenz des menschlichen Genoms in der FAZ

ermessen, eine Strukturabbildung, die dort wahrhaftig als Text („Das Buch des Lebens“) vorgestellt wird. Was ist davon zu halten, wenn Mitherausgeber Frank Schirrmacher als Philologe und Philosoph nicht einmal ein Bild von einem Text zu unterscheiden weiß? Auch hier dreht es sich, wen wundert es, wieder um die zentrale Frage nach dem Zusammenhang von Physik und Semantik, von Struktur und Bedeutung (die hier niemand im einzelnen kennt).

Eine öffentliche Debatte auf der Höhe der Zeit hätte stattdessen unterschiedliche Positionen und Perspektiven zu Kognition und „künstlicher Intelligenz“ zu Wort kommen zu lassen. Hier geht es um mehr als eine Disziplin, und künftige Entwicklungsperspektiven dürfen keineswegs Computerexperten überlassen bleiben (die selbst in ihrer eigenen Zunft umstritten sind). Neben Problemen der technischen Machbarkeit wären vor allem Fragen der sozialen Wünschbarkeit zu erörtern, dabei immer wieder die jeweils erkennbaren Chancen und Risiken sorgsam abzuwägen, um gegebenenfalls Nützliches zu befördern und Unerwünschtem entgegen zu wirken. Der fruchtlose Streit zwischen Heilspropheten und Apokalyptikern hilft da kein bißchen weiter. Welches Niveau, auch im Feuilleton, dabei zu erreichen ist, mag etwa der Abschluß der SPIEGEL-Serie über das 21. Jahrhundert in Heft 1/2001 zur Erforschung des Bewußtseins demonstrieren (im übrigen sei auf die ausgewählte Literatur am Schluß verwiesen).

Wie es aber derzeit um Technik im Feuilleton zumeist bestellt ist, muß man das Schlimmste befürchten. Statt um öffentlichen Diskurs und Aufklärung – zugegeben eine anstrengende Sache – geht es um bloße Effekthascherei, um erhöhte Aufmerksamkeit. Und je weniger man von der Sache weiß, desto lauter das Getöse. Dafür ist offenbar kein Preis zu hoch, nicht einmal der, sich selbst zum Narren zu machen und das Publikum an der Nase herumzuführen. Wird es das aber nicht gelegentlich merken?

Weiterführende Literatur

Brödner, P., 1997: Der überlistete Odysseus. Über das zerrüttete Verhältnis von Menschen und Maschinen, Berlin: edition sigma

Dreyfus, H. L., 1979: What Computers Can't Do. The Limits of Artificial Intelligence, New York: Harper & Row

Edelman, G. M., 1995: Göttliche Luft, vernichtendes Feuer. Wie der Geist im Gehirn entsteht, München: Piper

Winograd, T.; Flores, F., 1986: Understanding Computers and Cognition. A New Foundation for Design, Norwood: Ablex Publ.

Simon, H. A.; Newell, A., 1958: Heuristic Problem Solving. The Next Advance in Operations Research, Operations Research 6

Besprochene Neuerscheinungen

Ray Kurzweil: Homo sapiens. Leben im 21. Jahrhundert – was bleibt vom Menschen? Köln: Kiepenheuer & Witsch 1999, 512 Seiten

Hans Moravec: Computer übernehmen die Macht. Vom Siegeszug der künstlichen Intelligenz, Hamburg: Hoffmann und Campe 1999, 352 Seiten

K. Eric Drexler und Chris Peterson unter Mitarbeit von Gayle Pergamit: Experiment Zukunft: Die Nanotechnologische Revolution, Bonn: Addison-Wesley 1994, 322 Seiten

Bill Joy: Warum die Zukunft uns nicht braucht, leicht gekürzte Übersetzung eines ursprünglich in der Internet-Zeitschrift „Wired“ publizierten Artikels in der FAZ vom 6. 6. 2000