

Zum Wechselverhältnis von Mensch und 'intelligent' erscheinenden, computerbasierten technischen Systemen

von Jürgen Stahl, Leipzig

Version vom 13. März 2016

In der hier von mir verfolgten Problemstellung geht es mir um die Klärung der Frage: Kommt den heutigen und in künftig absehbarer Entwicklung computerbasierten technischen Systemen die eigenständige Fähigkeit zu, Problemlösungen zu generieren und somit 'Intelligenz'?

Ausgangspunkt dieser Problemstellung war die von Klaus Mainzer im Vortrag am 29.1.2016 in Leipzig gegebene Definition, dass KI entsprechend dem Turing Modell bedeutet, dass Computer/Maschinen ›Probleme‹ lösen. Meine Frage lautet: kann ein Maschinensystem tatsächlich Probleme lösen? Problemlösung bedeutet nach meiner Kenntnis sowohl in der Wissenschaftstheorie, der formalen Logik als auch der Kognitionspsychologie, einen Algorithmus finden eben für ein Problem, für das noch kein solcher existiert. Gibt es einen Algorithmus, so ist es eine Aufgabe. Maschinen führen danach folglich Aufgaben aus, deren Algorithmus durch menschliche Subjekte formuliert wurde. Ich sehe nicht, dass Maschinen Algorithmen finden. Maschinen können diese im Rahmen gegebener Lösungen - entsprechend implementierten Programm - modifizieren, möglicherweise weiter entfalten. Aber sie können nach meiner Kenntnis nicht gegenüber einer neuen, unbekanntem Aufgabenstellung Lösungen anbieten und - das ist m.E. nicht minder entscheidend - diese gar in ihren theoretischen, sozialen, technischen, moralisch-ethischen, juristischen, ästhetischen Konsequenzen bewerten.

Damit ist ein Aspekt des qualitativen Unterschieds menschlicher Subjektivität respektive Intelligenz gegenüber dem den Maschinen zugeschriebenen Phänomen bezeichnet. Es geht um das Phänomen, welches die Transzendentalphilosophie mit dem Begriff der intellektuellen Anschauung bezeichnete, als dem qualitativen Sprung, um von der empirisch basierten Anschauung zu einem theoretisch-begrifflichen Modell eines Erkenntnisobjektes zu kommen. Die interne, begriffliche Repräsentation des Erkenntnisobjektes vollzieht sich mittels eines Aktes subjektiven »Konstruierens«, bis hin zum Aufbau von Allgemeinbegriffen. Dieser qualitativ neue Stand des Wissens gegenüber der sinnlich-anschaulichen (empirischen) Ebene erlaubt den individuell wie kooperativ agierenden Subjekten die Reflexion über diese Lösung, inklusive der Abschätzung der Handlungsalternativen mit ihren Folgen sowie der Formulierung

von Anweisungen (ob im mathematischen, pragmatischen oder moralischen Sinne sei dahingestellt).

Als Unterschied zwischen instinktiver und kognitiver Intelligenz wird in der Literatur die systematische Werkzeugherstellung und die Sprache geltend gemacht. Das bedeutet, der Mensch repräsentiert nicht nur einzelne Aspekte der Umwelt, sondern macht darüber hinaus die Umwelt und sein Verhalten darin zum Gegenstand seiner Reflexion. Und das sowohl im Horizont auf seine Geschichte wie seine Zukunft. Der Mensch verfügt über ein reflexives Bewusstsein, wodurch es ihm möglich ist, bewusst, planend und (kybernetisch gesprochen) regelnd, also mit Rückkopplung auf die Natur, auf sein Verhalten und seine Lebensbedingungen einzuwirken.

Natürlich vollzieht sich die Entwicklung computerbasierter technischer Systeme und ihr Einsatz in der Wechselwirkung mit menschlichem Verhalten in einer gelebten Umwelt, in der den Geräten inzwischen nicht mehr nur ein passiv-beeinflussender Status zukommt. Ihr Einsatz erfolgt aus unterschiedlichen Interessen mit dem Ziel, also *gewollt*, steuernd auf unser Handeln einzuwirken, so dass ihnen zuweilen gar ein bestimmender Einfluss zukommt.

Es kam in der Diskussion am 1.3.2016 der Einwand, das mit der heutigen Art und Weise der Datenverarbeitung als Symbolverarbeitung und nicht lediglich als 1/0 Verarbeitung, ein gänzlich anderes Niveau gegenüber den 80er Jahren und davor erreicht sei. Das kann ich nur zur Kenntnis nehmen. Doch ergibt sich damit eine neue Perspektive zu der in Frage stehenden Problematik über die Rolle der Menschen als das zwecksetzende Subjekt in diesem Prozess? Selbstverständlich sind die Menschen dadurch in einer anderen technisch-kulturellen Umgebung befindlich, bewegen sie sich in komplexeren, qualitativ neuen Interaktionszusammenhängen als aktive Momente zu - in bestimmter Weise - nicht minder aktiven technischen Komponenten. Dennoch sind diese technischen Komponenten durch Menschen konzipiert, in ihren Möglichkeiten durch die hardwareseitige Ausstattung und softwareseitige Implementierung bestimmt. Und ihr Einsatz erfolgt gemäß den von Menschen verfolgten ökonomischen, sozialen, kommunikativen, kulturellen und auch herrschaftlichen Interessen. Das Phänomen einer sich ausbildenden 'Schwarmintelligenz' in Gestalt kommunizierender Gerätesysteme wurde durch H.-G. Gräbe in der Nähe des Organismischen verortet. Das halte ich für bedenkenswert, insofern auf der phänomenalen Seite eine dem Organismischen analoge Struktur und Handlungsweise konstatierbar wäre. Aber - und das ist m.E. zu bedenken - eine

solche Gerätestruktur ist initiiert, implementiert, (mehr oder minder) kontrolliert genutzt durch Menschen. Aber auch dieses und sicher künftig darüber hinausgehende Phänomene erledigen m.E. keineswegs die eingangs gestellt Frage. Und das nicht nur aus der Perspektive, durch wen die Systeme initiiert und als Gesamtsystem kontrolliert werden. Darüber hinausgehend wäre zu klären, ob die Möglichkeit besteht, dass solche technischen Systeme bei genügender Entwicklung der erforderlichen Komponenten und Kapazitäten ein Niveau erreichen könnten, in der das System aus sich heraus eigene, über die gegebene Programmierung hinausgehende Lösung zu kreieren vermag, die das technische System sich dann selbst implementiert und das System damit gleichsam zu einer Selbstentwicklung im technischen wie im 'intellektuellen' Sinne fähig wäre.

Eine solche Entwicklungsmöglichkeit würde ich aus grundsätzlichen Erwägungen heraus verneinen, weil damit eine Entwicklung jenseits der Wechselwirkung Mensch - technisches System prognostiziert würde. Für das technischen System würde eine Selbstentwicklung gleichsam ohne Mensch extrapoliert. Dafür gibt es nach meiner Kenntnis bisher keinen Anhaltspunkt. Anders gesagt: Technische Systeme sind nicht evolutionär im biologischen Sinne entstanden und haben dazu auch kein Potenzial.

Natürlich übertragen wir Menschen technischen Systemen im Zuge der fortlaufenden wissenschaftlich-technischen Entwicklung zunehmend mehr und in technischer Hinsicht auf qualitativ neue Weise Fähigkeiten, die bisher in dieser Qualität einzig Menschen zukamen. Zugleich erweitern diese technischen Systeme z.B. unsere sensorischen Möglichkeiten, die Welt wahrzunehmen und diese durch die Syntheseleistung menschlichen Denkens in Gestalt anschaulicher und begrifflicher Modelle zu erfassen. Aber das bedeutet m.E. nicht, dass solche Gerätesysteme ein 'Eigenleben' auszubilden vermögen, welches über ihre gegebene technische und programmseitige Ausstattung hinaus führen würde. (Was wohl auch nie Sinn sein kann: Man stelle sich vor, ein Kontrollsystem eines Kraftwerks oder der Bahn würde eigenständig Entscheidungsvarianten erarbeiten und vollziehen, die außerhalb der vorgesehenen Aufgaben lägen. Den 'Blackout'-Phantasien wären keine Grenzen gesetzt.) Dabei sind es doch die Menschen, die diese Systeme/Geräte in neue Interaktionszusammenhänge einführen und dadurch neue Handlungsperspektiven eröffnen.

Auch wenn solche Systeme möglicherweise Entscheidungen gegen Menschen treffen - siehe den Spiegelartikel "Lotterie des Sterbens" (4/2016) -, so sind diese Entscheidungen Ergebnis des programmierten Möglichkeitsfeldes für

auszulösende Handlungen. (Das scheint freilich im Falle einer Situation, die insofern inadäquat ist, als sie nicht bedacht wurde und das technische System keine sinnvolle Handlungsanweisung verfolgen kann; aber auch das liegt damit innerhalb des immer begrenzten technisch-softwareseitigen Möglichkeitsfeldes - eben weil das technische System das vorgegebene, vorbedachte Feld nicht selbst erweitern kann). Die Prämisse vom 'vorbedachten' Möglichkeitsfeld schließt auch die Problematik 'lernfähiger' Systeme und der dadurch gegebenen Erweiterung zunächst bestehender Grenzen ein. Das gilt auch für das Data-Mining mittels künstlicher neuronaler Netze. Festzuhalten bleibt auch hier, dass Data-Mining für sich noch kein Wissen bereit stellt, sondern Beziehungen zwischen Daten expliziert. Aus diesen kann durch geeignete Interpretation dann Wissen abgeleitet werden. Indem die Verfahren hierzu nichtlineare Charakteristiken aufweisen, bedeutet dies nicht, dass ihre Formalisierbarkeit nicht Voraussetzung und Bedingung für ihre Entwicklung und Anwendung wäre. Ebenso sollte man zumindest im Hinterkopf haben, dass auch das 'Lernen' neuronaler Netze auf mathematischen Algorithmen beruht.

Ich halte diese prinzipielle Fragestellung der Wechselwirkung Mensch - computerbasiertes

technisches System nicht gebunden an begrenzt überschaubare technisch-technologische Möglichkeiten und auch nicht an die jeweilige gesellschaftliche Form. (Die kommt m.E. zum Tragen, wenn ökonomische, soziale, moralische und rechtliche Dimensionen dieser Entwicklung bedacht werden.) Denn eine Spezifik der Subjektseite in dieser Wechselwirkung liegt darin, dass Menschen formalisierbare wie auch *nicht*formalisierbare kreative Akte in der Interaktion mit anderen Subjekten und ihrer Umwelt vollziehen. Ein für die hier vollzogene Betrachtung wesentlicher Bereich nichtformalisierbarer Akte liegt dabei im *Übergang* von der anschauenden zur begrifflichen Ebene der Erkenntnis. Technische Systeme können einen solchen Übergang nicht vollziehen, eben weil es dafür keinen Algorithmus gibt. Die sich damit verbindenden Problemlösungen sind Leistungen derer, die die Geräte konzipieren, technisch realisieren, softwareseitig erforderliche Programme entwickeln und die Systeme im interessen gebundenen Einsatz zur Nutzung bringen. Das schließt selbstverständlich ein, dass in dieser Wechselwirkung das Subjekt als ein sich in seinen Handlungs- und Denkmöglichkeiten änderndes Subjekt zu begreifen ist. Stellt man diese These in Frage, dann wäre entweder zu zeigen, dass es keine nichtformalisierbaren Klassen von Objekten gibt oder man müsste zeigen, dass

die Formalisierbarkeit keine zwingende Bedingung jeglicher Programmierung ist bzw. sein wird.

Jegliche Programmierung ist begriffliches Abbilden eines Erkenntnisobjektes, stellt mittels der Zeichenfunktionen der Sprache eine Konstruktion auf einer gedanklichen Metaebene dar. Nicht die Sprache an sich trägt die Fähigkeit der Metasprachenbildung in sich. Es sind die kognitiven Prozesse in Verbindung mit den Regeln der Sprachbenutzung, die diese Möglichkeit eröffnen (Klix). Die technischen Systeme verfügen nur insofern über die ihnen übertragenen kognitiven Fähigkeiten, als diese ihnen softwareseitig 'eingegeben' wurden, weil sie formalisierbar waren. Die begrifflich-klassifikatorische Repräsentation eines Erkenntnisobjektes in Gestalt eines Softwareprogramms ist Leistung des erkennend-handelnden, sozial interagierenden Individuums. An dieser Leistung setzen Operationen wie Transformation, Verknüpfen, Verdichten etc. an. Insofern eine Vielzahl kognitiver Handlungen formalisierbar sind, können sie auch durch technische Systeme durchgeführt werden. Nur ist damit die Spezifik einer Problemlösung, das Finden eines neuen Algorithmus, nicht auf das Gerät übergegangen. Wenn ein formalisierbarer Algorithmus gefunden wurde, war das noch immer kreative Leistung von Menschen, mit dem Ziel, einen Erkenntnisgegenstand adäquat zu beschreiben. Vertritt man eine andere Meinung, so wäre zu zeigen, dass die Entwicklung von Software als Form sprachlicher Repräsentation in Verbindung mit der zugrundeliegenden, genutzten technischen Ausstattung in der Lage ist oder sein wird, nicht nur die Technik des sensorischen Abstrahierens als kognitiver Leistung nach zu vollziehen. Es müsste gezeigt werden, dass solche Systeme über die anschaulich-bildliche Repräsentation der Realität in dem Sinne hinaus gehen, dass freizügige multiple Klassifizierungen, d.h. eigenständige kognitive Strukturbildungen erfolgen, die in der Lage sind, sich von den entscheidungsrelevanten Begriffsbildungen zu lösen und zu neuen, die Realität treffenden oder diese verfremdenden Resultaten zu führen.

Im Zentrum der These einer künstlichen Intelligenz steht Turings Postulierung einer Maschine 'intelligenten' Verhaltens, die sich nicht verraten würde, indem sie immer wieder dieselbe Art von Fehlern machte und unfähig wäre, sich selbst zu korrigieren oder durch externe Argumente; aber auch in der Lage wäre, aus Erfahrungen zu lernen. Die Erziehung dieser Maschine solle einem hochbegabten Lehrer anvertraut werden, der wiederum kein detailliertes Wissen von den inneren Abläufen der Maschine haben darf - der Wissensingenieur. Turing sieht,

dass Fehler zu machen, eine unvermeidliche Begleiterscheinung des Mathematikers (Menschen) sei, um zuweilen auf eine völlig neue Lösung zu stoßen. Die durch Maschinen vollzogenen Fehler sieht Turing analog dazu und prognostiziert faktisch den Maschinen eine ihnen eigene Kreativität. Nur: Ist das tatsächlich gegeben? Oder beruhen dann nicht die Algorithmen eben auf durch Menschen erarbeiteten Lösungswegen (zufällige Prozessfehler eingeschlossen)? Und wenn es solche den Maschinen eigene Phänomene gibt, wer trifft die Auswahl über die zu nutzenden Lösungen?

Das Programm "Logic Theorist" gilt als historischer Ausgangspunkt des KI-Programms. Es vermochte logisch-mathematische Ableitungsarbeit zu leisten. Es bewies, dass logisch-mathematische Ableitungsarbeit in Richtung eines vorgegebenen Beweisziels - also ein ganzes Feld geistiger Arbeiten - automatisierbar sind, indem es Ketten logischer Symbole solange umwandelte, bis es auf die gesuchten Theoreme stieß. Dennoch bleibt dabei immer zu verdeutlichen, dass die technischen und intellektuellen Bedingungen dafür durch Menschen in Gestalt von Problemlösung verfasst, programmiert und in Gang gesetzt wurden; die Durchführung der dann formalisierten Abläufe ist ein Abarbeiten von Aufgaben. Die im intellektuellen Sinne kreative Leistung lag davor, nämlich in der Entwicklung eines entsprechenden Programmes.

Das Kriterium zur Unterscheidung zwischen Mensch und computerbasiertem technischen System liegt nicht in der Schnelligkeit, sondern in der Andersartigkeit der Abarbeitung und Entscheidungsfindung und der Weiterentwicklung der Programme durch den Menschen als Initiator dieser Entwicklung. Sicher kann man hierzu formulieren, dass automatisches Beweisen allmählich ein hochtechnologisches Produktionsmittel vorstellt, indem es zur Programmverifikation und zur Überprüfung von Hardwarekonfigurationen in der industriellen Computerherstellung eingesetzt wird. Auch die Position, Computer nicht nur als 'Zahlenfresser', sondern als symbolverarbeitende Systeme und die Computerwissenschaft als eine empirische Wissenschaft aufzufassen, die nicht nur Programme, sondern auch Menschen beim Problemlösen beobachtet, um deren Taten maschinell zu simulieren (Newell/Simon), hebt m.E. den aufgezeigten Unterschied zwischen Mensch und technischem System nicht auf. Es führt gelegentlich eher zu der eigenartigen Verkehrung: durch Beobachten von KI-Programmen zu begreifen, wie menschliches Denken funktioniert, um schließlich Menschen als unzulängliche und deshalb ersetzbare KI-Maschinen zu begreifen. Eine solche auf der phänomenalen Ebene sich bewegende Wertung verwischt den

qualitativen Unterschieden zwischen künstlicher Intelligenz in Gestalt von computerbasierten technischen Systemen und menschlichem Denken. Nur Letzterem kommt Kreativität, Reflexionsvermögen und Selbstbewusstsein über den Erkenntnisvorgang in der tätig-erkennenden Aneignung der Außenwelt als einem entwerfenden, das Erkenntnisobjekt im Zuge der Interiorisation des Wissens konstituierenden Vorgang zu.

Es geht mir in einer solchen Diskussion nicht um Abwertung oder Aufwertung menschlicher Kreativität vs. künstlicher Intelligenz und umgekehrt, sondern darum, die jeweilige Leistungsfähigkeit, also Besonderheit in ihrem gegenseitigen Bezug aufeinander aufzuzeigen. Denn ist es nicht ein Gebrauchswertversprechen, das da zuweilen blendet: Wir bieten ein besseres Hirn, wir bieten fehlerfreie, nicht krank werdende, nicht streikende, nicht in Frührente gehende Anlagen, die alle anstehenden Probleme letztendlich technisch zu lösen vermögen - bis hin zur Versorgung der Kranken, Schwachen, Alten... Im Bewusstsein der Vorteile der neuen Datenverarbeitungs- und Kommunikationstechnologien wende ich mich jedoch gegen eine ersetzungslogische Blickverengung, die den Anspruch erhebt, möglichst alle humanspezifische Fähigkeiten operationalisieren zu können und damit gleichsam die Wirklichkeit zu entsubjektivieren.