

Mensch und Automat

(«Das Goetheanum» Nr. 24 / 12.06.1983)

Können Maschinen denken? Kann der Mensch denken? Diese zwei Fragen können von einer naiven, weil mit der eigenen Leistung nie rechnenden Wissenschaft mit ja oder nein beantwortet werden, je nachdem, was man unter «Denken» versteht. Und es ist vielleicht der eigentliche Sinn dieser Fragen, das Augenmerk auf das Denken zu lenken, das ausserhalb der Geisteswissenschaft in seiner Phänomenologie und in seinem Ursprung nicht untersucht werden konnte, weil dazu die Bedingungen gefehlt haben: die Vertiefung oder Erhöhung des Denkens selbst, ohne das es nicht möglich ist, das *erscheinende Denken* zu betrachten - von derselben Ebene «sieht» man es nicht, allein von oben her.

Jedenfalls können die obigen Fragen nur denkende Wesen stellen. Maschinen haben kein «Interesse» daran. Um sich mit solchen Fragen befassen zu können, gehört die Struktur der Bewusstseinsseele dazu, die heute für eine grosse Anzahl von Menschen gegeben ist. Sie ist dadurch charakterisierbar, dass der Mensch auf sein Bewusstsein und auf seine Inhalte schauen und sie als Vergangenheiten zu erkennen kann: als Gedachtes, Wahrgenommenes, Vorgestelltes. *Geschaut* auf diese Vergangenheit wird von einer Instanz her, die selbstverständlich nicht zu ihr gehört - von der Vergangenheit her ist sie selbst nicht entdeckbar -, sondern die auf der Ebene der Gegenwärtigkeit liegt. Das Bewusstsein erlebt sich da nicht im zeitlosen, nicht zählbaren Vorgang des Verstehens, des gedanklich Intuitiven, aber die Ebene ist für das Bewusstsein punktuell berührbar. Sie enthält die zeitlosen Vorgänge des Denkens, Wahrnehmens, Vorstellens, deren Produkte bewusst werden: die Aussenseite oder Erscheinung der Vorgänge, deren Innenseite sie selbst sind. Im Bewusstsein hat sich das Vergangenheitsmoment klar von der Gegenwärtigkeit getrennt. Der Mensch erlebt bewusst das erstere; von den Vorgängen kann er eine verschleierte Erfahrung am Rand seines Bewusstseins haben. Die *Fähigkeiten* des Denkens, Wahrnehmens, Vorstellens bleiben ihm *überbewusst*, wie auch die des Sprechens: er weiss es nicht, wie er es tut, wie er Laute, Worte, Sätze bildet, die Physiologie des Sprechens, die Grammatik, die Syntax sind nicht bewusst beim Ausüben des Sprechens.

Wenn auch das Innenerlebnis im Bewusstseinsakt bloss ist, weiss davon jeder Mensch, und es ist auch jedem Denkenden klar, dass keine Maschine solches Erlebnis haben kann. Die einfache Uhr *weiss* nicht, wieviel Uhr es ist und die komplizierteste Computer-Uhr, die «sprechen», auf Fragen eingehen kann, *weiss* von allen ihren Leistungen auch nichts, obwohl sie dem Menschen vieles angeben kann, wie die Uhr die Zeit. Auch was eine Zeichenserie in einem Enzephalogramm bedeutet, kann nur durch das Innenerlebnis der Versuchsperson festgestellt werden. Niemand denkt ernsthaft, dass ein Computer ein Innenleben hat; ein jeder weiss, dass er selbst ein Innenleben hat; es ist aber problematisch zu wissen, ob ein anderer Mensch solche Erlebnisse hat. Anders ausgedrückt: Wie kann man unterscheiden, ob man mit einem Menschen oder einem perfekten Computer ein «Gespräch» - eventuell durch Schreibmaschine - führt?

Der Computer ahmt die Aussenseite des menschlichen Denkens nach: «so pflegt der Mensch zu denken, im Denken vorzugehen». Der Nachdruck liegt bei «pflegt», das heisst, es handelt sich um Denkverfahren, die *bekannt* sind. *Weil* sie bekannt sind, können sie von ihrem *formalen* Wesen, von aussen her abgebildet werden. Die Abbildbarkeit hat demnach Bedingungen. Diese lauten:

1. Der Denkvorgang musste einmal von einem Menschen hervorgebracht werden.
2. Der Mensch hat den Denkvorgang für richtig und wertvoll beurteilt, und zwar ohne auf formale und äusserliche Merkmale zu achten.
3. Im Denkprozess können nur im voraus bestimmte Begriffe, Begrifflichkeiten und Denkooperationen vorkommen.

4. Die «Sprache» des Computers und seines eventuell menschlichen Gesprächspartners darf in den Operationen nur eindeutige Ausdrücke, d.h. Zeichen und Symbole verwenden. «Vielleicht», «obwohl», «nichtsdestoweniger», «allerdings», «aber» usw. haben als Begrifflichkeiten für den Computer keine Bedeutung.

Damit ist für die Unterscheidungsmöglichkeit «Mensch oder Computer» an *äußeren Merkmalen* gesorgt.

1. Der Computer wird keine neuen Denkopoperationen «erfinden», obwohl er die vorgegebenen schneller, konsequenter und ohne Irrtum ausführen kann; eben weil er nicht denkt und auch komplizierte Aufgaben zu lösen imstande ist, denen der Mensch gar nicht gewachsen ist.
2. Ob eine vom Computer ausgeführte Operation richtig, wertvoll, zweckmässig usw. ist, muss *letztlich* ein Mensch beurteilen.
3. Wenn ein Computer programmiert ist (und das muss er sein, wenn er arbeiten soll), kann er selbstverständlich keine anderen, etwa neuen Begriffe verwenden, als die im Programm enthaltenen. Falls er das tun würde, müsste übrigens der Mensch ihm schnell im Verstehen nachkommen.
4. Jede Antwort auf eine zwei- oder mehrdeutige Rede würde den Computer «verraten», er würde den Text stets im - dem Programm nach - nächstliegenden Sinne «auffassen» und beantworten. Weder ein Witz, noch ein Theaterstück, noch ein Gedicht - alles mehrdeutige Texte - sind für ihn «greifbar». Er würde den Witz immer «ernst» nehmen. Wie es manche Menschen tun.

Man kann vieles gegen die obigen Gedanken einwenden. Es gibt Märchen, Musik, Witz und Gedichte erzeugende Computer, es gibt scheinbar nicht eindeutige Computer-«Sprachen»; wenn man aber genau verfolgt, was in diesen Fällen passiert, wird man die Richtigkeit der erwähnten Kriterien bestätigen können. Es ist auch klar, dass die verwendeten Ausdrücke nicht gut definiert sind, also keineswegs *eindeutig* sind, selbst die Eindeutigkeit ist kaum zu bestimmen. Eindeutig sind die naturwissenschaftlich-technischen Begriffe, wie «spezifisches Gewicht» oder «elektrischer Widerstand»; aber wer würde ein Wort von Rilke, wie «Jeder Engel ist schrecklich», nicht auch eindeutig empfinden, wenn auch auf einer ganz anderen Ebene als der der technischen Begriffe, auch wenn die Bedeutung des Satzes zu einer Vielfalt zerfällt, falls man sie auf der *Informationsebene* des Technisch-Wissenschaftlichen zu erklären versuchen würde. Dasselbe wäre von einer musikalischen Phrase zu sagen. Das ist aber ein allgemeines Problem der undefinierbarkeit und stammt aus dem intuitiven oder überbewussten Charakter und Ursprung der grundlegenden Begriffe, wie «Sein», «und», «ich», «jetzt» usw., die eo ipso undefinierbar sind - womit sollte man sie definieren? Trotzdem sind die Kriterien verwendbar und können das Denken orientieren.

Es ist aber offensichtlich, dass auch sehr viele Menschen nicht imstande sind, den vier Kriterien menschlich zu entsprechen: sie können keine für sie neuen Begriffe oder Gedankengänge hervorbringen; sie können nicht entscheiden, ob das Denken wirklich Denken ist; sie können keine für sie neuen Begriffe verstehen und keine Mehrdeutigkeit akzeptieren, zum Beispiel einen Witz «verstehen». Diese Erfahrung ist sehr bedenklich und erklärt, warum für die menschlichen Bewusstseinstätigkeiten mechanische bzw. Automaten-Modelle entstehen konnten. Für das *Entstehen* der spezifisch menschlichen Fähigkeiten - Sprechen, Denken, ichbewusstes Handeln - beim Kind versagt jedes Modell und muss auch versagen. Denn diese Fähigkeiten beruhen auf dem Verstehen, auf der Intuition. Dieses ist unabbildbar und unsignalisierbar, wie auch das Miss- und das Nicht-Verstehen. Es ist nicht möglich, *ein formales Merkmal* dafür anzugeben, *wann* der Mensch versteht. Denn das Verstehen *ist ein Innenerlebnis*, wie das Denken selbst. Daher sind sie nicht abbildbar oder formalisierbar. Denken, Verstehen kann nicht erklärt, nur vorexerziert und eventuell entfacht werden. Von welchen

menschlichen Tätigkeiten weiss der Mensch, wie er sie ausführt? Es geht um das Sprechen, Denken, darum, zielbewusst eine Bewegung auszuführen; alle geschehen überbewusst. Wenn man sie durch eine Maschine nachahmen lässt, so geschieht es keineswegs so, wie es im Menschen vor sich geht: Mit welchem Recht kann behauptet werden, dass es im Menschen wie in der Maschine gemacht wird, auch wenn man von den vier Kriterien absieht?

Die Nachahmung ist dadurch möglich, dass diese Tätigkeiten *worthaft* sind und daher, wie alles, das *worthaft* ist, eine erscheinende Aussenseite - Zeichen und Regeln - oder eine formale Art neben der Innenseite, die die Quelle und das Primäre ist, haben. Die Innenseite hat keine formalen Merkmale, deshalb kann sie nicht abgebildet werden. Die Nachahmung der Aussenseite entspricht etwa dem, dass ein Mensch ganz unbedacht, aus Routine etwas «sagt», fragt oder tut. Das Verstehen ist nicht nachahmbar, seine Produkte aber sind abbildbar und daher von aussen nachzuzahlen.

Solange der Mensch ein klares Bewusstsein über das Wesen des Computers und über sein eigenes hat, ist der äussere Automat nicht bedenklich. Viel schwerwiegender ist, wenn der Mensch selber viele automatische Elemente in seinem Verhalten, Denken, Sprechen, Tun zeigt. Der innere Automat entsteht dadurch, dass die Innenseite der Bewusstseinstätigkeiten nicht bemerkt oder geleugnet, danach auch weggelassen wird, wie im Automaten. Von diesem Nicht-Erfahren her stammen die Theorien des Denkens, die selbstaufhebend sind, durch den Widerspruch das Denken als einen determinierten oder Zufallsprozess auffassen und *zugleich* als Theorien, vom Denken formuliert, Gültigkeit, Wahrheit beanspruchen. Denn ist das Denken wirklich so, wie die Theorien es meinen, so bleibt kein Forum, das über die Richtigkeit der Theorien entscheiden könnte.

Was im Automaten geschieht, hat nichts zu tun mit den menschlichen Tätigkeiten, die nachgemacht werden. So ist zum Beispiel für den Automaten ein Zeichen, Wort oder Text kein Zeichen, Wort, Text im menschlichen Sinne, sondern ersetzt einen Knopf oder ein Schaltelement. Jedes Programm ist durch Schaltelemente ersetzbar. Auch Knöpfe sind belehrbar: Beispiel dafür ist der alternativ ein- und ausschaltende Druckknopf, je nach der vorgefundenen Lage. Aber auch auf einen Menschen können Zeichen, Worte, Texte wie ein Knopfdruck wirken. Maschinen können selbststeuernd sein, Fehler im eigenen Mechanismus entdecken und eliminieren. Das als *Selbsterfahren* - *wer* würde erfahren und *was!* - zu deuten ist ein Denkfehler, ähnlich dem, nach welchem die Uhr wüsste, wieviel Uhr es ist - sie zeigt es ja dem Menschen an.

Der Mensch nimmt begrifflich stark durchsetzt wahr, daher ist für ihn ein Gegenstand aus beliebiger Sicht, von unten, oben usw. erkennbar und auch vor einem beliebigen Hintergrund. Durch diesen Zug der Begrifflichkeit wird die grösste Schwierigkeit im «Wahrnehmen» von Robotmaschinen verursacht, weil in diesen die denkerische Komponente fehlt. Das bedeutet, dass, um einen Gegenstand «erkennen» zu lernen, dieser von allen möglichen Perspektiven und vor allen Hintergründen der Maschine «gezeigt» werden müsste beim «Lernen» - eine hoffnungslose Aufgabe -, denn die Kamera selbst kann den Gegenstand von der Umgebung und vom Hintergrund nicht unterscheidend abheben, wie es der Mensch tut durch die Begrifflichkeit. Er weiss unter allen Umständen, wo der Gegenstand endet. Die phantastischen Wahrnehmungsleistungen von Robotern in den Science-Fiction-Romanen existieren nur auf dem Papier. Das Sprechen ist weit eindeutiger und leichter nachahmbar als das Wahrnehmen, oder, anders ausgedrückt: das Denken ist zum Sprechen im formalen Sinne weit weniger notwendig als zum Wahrnehmen.

In einer Umgebung von Maschinen muss sich der Mensch maschinenartig benehmen, das heisst, ohne zu denken richtig handeln können, wie zum Beispiel beim Autofahren. Das Automatische oder Gewohnheitsmässige aber dehnt sich viel weiter und tiefer in ihm aus: alles, das assoziativ, reagierend in ihm ist, alles, das im Verhalten nicht durch das Ich, sondern durch das Unterbewusste bestimmt wird, gehört zum inneren Automaten. Durch einen eindeutigen Text vorbestimmten Themas kann nicht festgestellt werden, ob man mit

einem Menschen spricht, so automatenhaft, so routinemässig würde das menschliche Gespräch. Allerdings müsste der Mensch den Unterschied merken, wenn er dem Gesprächspartner gegenübersteht, auch wenn dieser eine technisch perfekte Puppe ist, eine gesteigerte Olympia aus *Hoffmanns Erzählungen*, falls er einen entwickelten Du-Sinn hat.

Die Entwicklung zum Automaten hin ist das Gegenteil einer positiven menschlichen Entwicklung. Denn der Mensch verwirklicht sich um so mehr, je mehr er improvisatorisch zu leben vermag. Weil er selbst ein worthaftes Wesen ist, hat er eine Erscheinungsseite und diese kann nachgeahmt werden, bzw. es liegt in ihr die Möglichkeit des Automatischen. «Wir würden in einer gewissen Beziehung Denkmaschinen sein, Wissenschafts-Maschinen, wenn zu unserem physischen und Ätherleib nicht das Ich und der Astralleib käme. Aber so wie unser physischer Leib und Ätherleib sind, sind sie auch in gewisser Beziehung automatisch. Sie sind im Grunde genommen als solche nur dadurch auf der Erde noch entwicklungsfähig, dass sie vom Ich bewohnt werden.» (GA 176, Vortrag vom 25.9.1917). Weil sie vom Ich bewohnt werden, bilden sich vom Ätherleib und Astralleib freie Kräfte, wodurch sich das Ich im physischen Leib artikulieren kann. Selbst die *geformten* Teile des Astralischen und der Ichhaftigkeit - das Ego oder Alltags-Ich - bedrohen den Menschen, dass er durch sie, wenn das Ich sie nicht beherrscht, in das automatische, unmenschliche, wortfeindliche Wesen hereinbezogen wird.

Dass Computer möglich geworden sind, hängt damit zusammen, dass der Mensch durch die Struktur der Bewusstseinsseele weitgehend zum «trägerlosen» oder wortlosen Denken fähig ist (siehe Nr. 7/1982). Diese Fähigkeit bewusst auszubilden, wäre sein Weg zu seiner wahren Selbstverwirklichung. Eine Anwendung dieser Fähigkeit besteht darin, dass er Denkopoperationen untersuchen, analytisch erfassen und mechanisch nachahmen kann. Die Gefahr der nicht bewusst ergriffenen Fähigkeit liegt darin, dass der Mensch sich selbst als einen Mechanismus auffasst und dadurch auch ein Mechanismus wird.