



Da bekommt es HAL mit der Angst zu tun: Dave entfernt der künstlichen Intelligenz aus Kubricks „2001 – Odyssee im Weltraum“ den Speicher

Die Datei, die Datei hat immer recht

Wer programmiert unsere Zukunft? Nick Bostrom warnt vor der rasanten Entwicklung künstlicher Intelligenz

• Norbert Zähringer

Im Film „2001 – Odyssee im Weltraum“, Stanley Kubricks Science-Fiction-Meisterwerk aus dem Jahr 1968, reist eine Gruppe von Astronauten im einem Raumschiff zum Jupiter. Gesteuert wird das Schiff von „HAL 9000“, einem Supercomputer, der mit künstlicher Intelligenz ausgestattet ist und als absolut unfehlbar gilt. Während die humanoide Besatzung relativ untätig durchs All dümpelt (drei befinden sich im Tiefschlaf, zwei drücken Knöpfe) hat HAL an einem vertrackten Problem zu kauen: Nur er kennt das wahre Ziel der Mission, die offiziell den Auftrag hat, den Jupiter zu erforschen, tatsächlich aber den Ursprung eines mysteriösen Außerirdischen Signals aufzuklären soll. Er soll also einerseits immer korrekte Antworten geben, muss aber gleichzeitig lügen.

Nachdem er – absichtlich oder nicht – ein Bauteil des Raumschiffs fälschlicherweise für defekt erklärt hat, kommen der Besatzung Zweifel an seiner Unfehlbarkeit, und sie erwägen ihn abzuschalten. Dies jedoch würde die geheime Mission gefährden. Konsequenterweise schaltet HAL die Lebenserhaltungssysteme der drei schlafenden Besatzungsmitglieder ab, bringt einen der beiden Wachhabenden um und verweigert auch dem letzten Überlebenden, dem Astronauten Dave, der seinen Kollegen mit einer Mini-Kapsel retten wollte, die Rückkehr ins Mutterschiff.

„Perverse Instantiierung“ würde Nick Bostrom das nennen, der in seinem Buch „Superintelligenz“ eine ganze Reihe von Katastrophen auflistet, die eintreten könnten, wenn eine Künstliche Intelligenz eines Tages nicht nur ein kleines Raumschiff, sondern die Geschicke der menschlichen Zivilisation lenken sollte. Diese KI wird freilich noch wesentlich „intelligenter“ sein als der HAL von 1968. Bostrom sagt eine „Intelligenzexplosion“ voraus: „Es sieht so aus, als ob es irgendwann in diesem Jahrhundert geschehen wird, aber sicher können wir uns dessen nicht sein.“

Unter Intelligenzexplosion versteht Bostrom das Entstehen einer künstlichen Intelligenz (sei es als Emulation eines menschlichen Gehirns oder als Supercomputer), die sowohl besser als auch schneller denkt als wir Menschen und daher über eine Reihe von „Superkräften“ verfügen würde: Neue Technologien, Waffensysteme, soziale Manipulation, Ressourcenkontrolle, strategische Planung, totale Überwachung – das alles wären für eine KI, die mit Lichtgeschwindigkeit operiert, keine Probleme mehr. Zumal sie sich auch selbst verbessern könnte, also zu allem Übel immer noch schlauer werden würde. Eine solche Maschine wäre daher, zitiert Bostrom den Mathematiker I. J. Good, „die letzte Erfindung, die der Mensch je machen muss, vorausgesetzt, die Maschine ist fügsam genug, um uns zu sagen, wie man sie unter Kontrolle hält.“ Und genau da liegt das Problem. Warum sollte die Maschine das machen? Lassen wir Menschen uns etwa von den Gorillas sagen, was wir auf der Erde tun oder besser lassen sollen?

Nick Bostrom ist Philosoph an der Universität von Oxford und Leiter des dortigen „Future of Humanity Institute“. Schon seit Längerem hat er sich mit Szenarien beschäftigt, die er als „existenzielle Risiken“ für den Fortbestand der Menschheit ansieht. Die kommende Superintelligenz ist für ihn genau das – geeignet, uns alle vom Planeten zu tilgen (oder neben den Gorillas im Zoo sitzen zu lassen). Den einzigen Vorteil, den wir in diesem Fall haben, sei die Tatsache, dass wir diejenigen sind, „die das Ding bauen.“

Könnte man also die Superintelligenz, solange sie noch nicht superintelligent ist, schon mal so konstruieren, dass sie uns später wenigstens nicht schaden kann?

Bostrom ist kein Autor, der die Sensationsgier der Leser befriedigt. Ihm geht es in „Superintelligenz“ um praktikable Strategien, eine kommende Katastrophe abzuwenden. Man kann sein Buch als eine Art Studie über die Möglichkeiten des Findens von Lösungen für Probleme sehen, die es noch gar nicht gibt.

Wird sich der Übergang von einem normalen Elektronengehirn zu einer Superintelligenz langsam (in Wochen oder Monaten) oder sehr schnell (in Sekunden oder Millisekunden) vollziehen? Was könnte so eine Superintelligenz wollen? Strebt sie gleich nach Weltherrschaft oder eher nach höherer Erkenntnis? Was, wenn ihr Macht und Weisheit völlig egal sind? „Es liegt nichts Widersprüchliches in einer KI“, schreibt Bostrom, „deren einziges Endziel darin besteht, die Sandkörner auf Boracay (einer Philippineninsel) zu zählen, die Dezimalfolgen von Pi zu berechnen oder die Gesamtzahl der Büroklammern zu maximieren.“ Besser, man sagt der KI also vorher, was sie wollen will.

Doch abgesehen davon, dass die einmal emanzipierte Superintelligenz es vielleicht gar nicht mehr für nötig hält, sich an die Vorgaben ihrer Schöpfer zu halten, kann die Sache auch schiefgehen, wenn sie sich zu genau an das hält, was menschliche Programmierer in ihren Code geschrieben haben. Siehe HAL, der Supercomputer aus Kubricks „Odyssee 2001“. Als er abgeschaltet werden soll, sieht er das als Gefahr für die ihm von den Menschen aufgetragene Mission und „schaltet“ die Menschen ab.

Ist es da nicht besser, wenn man die Superintelligenz einsperrt? Man könnte so etwas wie ein künstliches „Orakel“ schaffen, das nur antwortet, wenn wir es etwas fragen. Doch selbst in einem solchen Fall könnte die KI versuchen, uns mit ihren Antworten zu manipulieren, die Wächter zu übertölpeln und aus ihrem Gefängnis auszubrechen.

Bostrom exerziert in seinem Buch die möglichen Gefahren der Intelligenzexplosion, die Irr- und Lösungspfade für das Kontrollproblem mit zuweilen anstrengender akademischer Gründlichkeit durch. Faszinierend sind seine Überlegungen vor allem deshalb, weil sie letztlich immer wieder auf unser Menschsein zurückfallen: Wenn er fragt, was die KI will, impliziert

dies natürlich die Frage, was wir eigentlich wollen; wenn er darüber nachdenkt, welche Werte man dem Superhirn vorab mitgeben sollte, müssen wir uns fragen, welche Werte uns wichtig sind. Bostrom empfiehlt internationale Zusammenarbeit, doch wie soll ein Wertekanon aussehen, auf den sich alle Menschen weltweit einigen können? Und gibt es nicht, fragt man sich, im Verborgenen schon längst einen Wettlauf?

Im Oktober 2013 meldete zum Beispiel China stolz, dass es mit „Tianhe-2“ einen Supercomputer geschaffen habe, der mit seinen über 3 Millionen Prozessorkernen fast doppelt so schnell rechnet wie sein Gegenpart „Titan“ in den USA (und auch mehr als doppelt soviel Energie benötigt: 17,8 Megawatt). Würde es uns gefallen, wenn die Werte, die die Programmierer der Superintelligenz einpflanzen, beim Parteikongress der KP Chinas beschlossen wurden?

Angesichts des strategischen Vorteils, den die Entwicklung einer Superintelligenz verspricht, könnte hingegen schon die bloße Vermutung, dass eine andere Nation einen Vorsprung hat, fatal sein: Eigene Wissenschaftler könnten dann nämlich entscheidende Sicherheitsvorkehrungen zugunsten eines schnelleren Erreichens des Ziels vernachlässigen.

Unlängst bekannte Elon Musk, Multimillionär und Gründer von „Tesla Motors“ (Elektroautos) und „SpaceX“ (Trägerraketen und Raumtransporter), während eines Symposiums am Massachusetts Institute of Technology, dass er künstliche Intelligenz für die möglicherweise „größte existentielle Bedrohung“ halte. Gleichwohl hatte Musk, ebenso wie Facebook-Gründer Marc Zuckerberg oder Google, zuvor in KI-forschende Unternehmen investiert, um, wie er auf Nachfragen entgegnete, „ein Auge auf die Entwicklung“ zu haben.

Denn auch das ist eine Möglichkeit: Dass nicht ein einzelnes Land, sondern ein Privatunternehmen im Geheimen die Superintelligenz schafft.

Sollen wir uns nun wünschen, dass die „Intelligenzexplosion“ noch möglichst weit in der Zukunft liegt oder demnächst geschieht? Bostrom hat keine endgültige Antwort. Je weiter der „Take-off“, wie er es nennt, entfernt sei, desto besser könnten wir uns darauf vorbereiten. Einerseits. Andererseits wäre die baldige Schaffung einer „freundlichen“ Superintelligenz angesichts anderer globaler Risiken wie Atomkriege, Klimaveränderungen oder Pandemien vielleicht gerade wünschenswert. Die Superintelligenz könnte diese Gefahren bannen, bevor sie uns zerstören.

In Kubricks „Odyssee 2001“ gelangt es dem Astronauten Dave am Ende, doch noch über eine Notluke in das Raumschiff zu gelangen und in den Computerraum vorzudringen, wo er HAL nach und nach alle Speicherkarten entfernt. Das wäre einer wirklichen Superintelligenz, wie sie Bostrom vorschwebt, wohl kaum passiert.

Sie hätte schon lange vorher entsprechende Gegenmaßnahmen ergriffen und wäre uns immer einen Schritt voraus. „Was einmal gedacht wurde, kann nicht mehr zurückgenommen werden“, heißt es in Friedrich Dürrenmatts Theaterstück „Die Physiker“. „Ich habe Angst“, sagt HAL, bevor Dave ihn abschaltet.

Bildungsreise durch Trümmerdeutschland

Ein US-Soldat als Meister Eckhart: Melvin Lasky
„Deutsches Tagebuch 1945“ • Michael Rutschky

Das Ende des Krieges 1945 bildet so etwas wie den Ursprungsmythos der Bundesrepublik. Man kann sich fragen, wann genau das Ende anfängt; Stalingrad war der Untergang der 6. Armee. Aber damit tritt der reine Schrecken in Erscheinung, und darin mag man nicht eintauchen – Stalin als besten Freund Deutschlands zu kanonisieren, misslang der DDR gründlich. Nein, am D-Day, am 6. Juni 1944, begann das Ende des Krieges für Westdeutschland, mit der Invasion der Amerikaner in der Normandie. Man kann es sich immer wieder vom TV nacherzählen lassen, wie sie ins Reich der Finsternis eindringen, und seine Macht schwindet ins Nichts.

Melvin Lasky, 25 Jahre alt, aus New York gebürtig, ist ab Februar 1945 dabei, beginnend in Frankreich. Nicht als Kämpfer, sondern als Chronist, im Auftrag der Historischen Abteilung der US-Armee, also beobachtend und mitschreibend – er wird enttäuscht in seinem Tagebuch vermerken, wie stark die offizielle Version seine Beiträge geglättet hat, wie das große Bild das Kriegsgeschehen im Einzelnen schön und verfälscht. Dem scheint schon die Form des Tagebuchs überlegen. Es enthält ausschließlich Nahaufnahmen, während der Chronist leibhaftig dem Geschehen folgt. Der junge Lasky hatte es für sich als eine Art Bildungsreise geplant, die seine intellektuellen Verwirrungen klären sollte, als Kahlschlag, Reinigung – ein ehrwürdiges Motiv für Kriegsteilnahme.

Das klappt natürlich nicht so richtig. Was der junge Mann erlebt, das läuft immer wieder auf Desillusionierung hinaus. Klar, die verlogenen Kommentare des Großdeutschen Rundfunks – der Chefkommentator Hans Fritzsche, der dann in Nürnberg von Kriegsverbrechen freigesprochen wird, tritt namentlich auf Goebbels' letzte Schwadronen werden sukzessive von der Wahrheit der alliierten Siege über die deutsche Wehrmacht gelöscht, und das befriedigt. Die Nachricht von Hitlers Selbstmord (der den Krieg ja nicht unmittelbar beendet) fehlt an dem entsprechenden Tag. Er scheint kein einschneidendes Ereignis mehr. Und er lag ja auch nicht als Nahaufnahme vor – Lasky zitiert nicht einmal die berühmte Falschmeldung des Großdeutschen Rundfunks.

Was ihn unablässig beschäftigt, ja überwältigt, indem er im Rücken der kämpfenden Armeen in das Reich eindringt, das ist dessen ungeheuerliche Zerstörung. Frankfurt, Darmstadt, München, Berlin (dessen Ruinen eine eigene Großartigkeit abstrahlen) – nihilistische Monumente, sozusagen, die unseren jungen Helden, der doch seine Bildungsreise absolviert, um sich selbst zu finden, in tiefe Zweifel an der Weltgeschichte stürzen, an der er unmittelbar teilnimmt.

Es kommt hinzu, dass die Soldaten, die das Heer der Finsternis niederkämpfen, sich keineswegs als Heer des Lichts präsentieren. Sie sind so oft betrunken. Sie fluchen und lästern auf das Widerwärtigste. Sie plündern und halten zynische Reden. Sie bleiben vollständig unberührt von den Kulturtrümmern, die sie durchqueren.

Ein vielsagendes Gegenbild zu den Schrecken der Vergewaltigungserzählungen, in denen der Siegeszug der Roten Armee im Osten überliefert ist, malt unser Held hinsichtlich der Liebhaftigkeit, die sich – trotz des anfänglichen Fraternisierungsverbots – zwischen den *Frauleins* und den GIs entwickeln. Auch unser Held bedient sich dankbar und ohne Schuldgefühle, Jeanne und Lydia, Helma und Aina, Christy und all die anderen. Keineswegs folgen die deutschen Mädchen einfach der Not, schon gar nicht physischem Zwang. Laskys Schilderungen erwecken Sigmund Freuds mythische Triebmächte zum Leben: Nachdem Thanatos, der Todestrieb, derart gewütet hat, muss sich Eros revanchieren. Eine besonders schöne Nahaufnahme liefert hier das Mädchen, das mit einem

schwarzen GI liiert ist. Wieso? antwortet sie kühl der Militärpolizei. Gelten in den Staaten nicht alle Menschen als gleich? Sie ist 1945 die Botin einer fernen Zukunft.

Wie deutlich es um eine Bildungsreise geht, bezeugen die vielen Bücher. Unser Held errettet sie aus dem Schutt, zu dem die Häuser zerfallen sind. Er inspiziert die Auslagen der Buchhandlungen. In Darmstadt prüft er die Bestände einer Schulbücherei – wir schreiben Anfang April – und freut sich daran, dass sie nicht völlig durchnazifiziert sind. Klar, einerseits Alfred Rosenbergs „Mythus des 20. Jahrhunderts“ und „Mein Kampf“ und anderer einschlägiger Dreck. Dann aber die Anthologie „Der deutsche Genius“ von 1923, Vorwort Thomas Mann, die ganz andere Traditionen aufruft. Ebenso eine Aufsatzsammlung von Wilhelm Dilthey, „Von deutscher Dichtung und Musik“.

Im „Politischen Handwörterbuch“ von 1923 fehlt Hitler durchaus, ebenso das Stichwort „Führer“, und „Demokratie“ liest sich verlässlich-vernünftig. Auch solche Bücher, spekuliert unser Held, müssen während des Dritten Reiches Schüler gelesen haben, eine Vorausdeutung auf ein neues Deutschland. Bücher sind Hoffnungsträger. Hierher gehört die Nahaufnahme, Heidelberg, 30. Juli, eigentlich ein Elendsbild: der jüdisch aussehende junge Mann, der Bücher in seine Mappe stopft und unter seinem Regemental augenscheinlich die KZ-Kleidung trägt – aber frisch gewaschen und gebügelt.

In Heidelberg – wo die US-Armee ihr Hauptquartier nehmen wird – trifft unser Held den berühmten Philosophen Karl Jaspers, und sie orakeln über Deutschlands Zukunft. Lasky trifft Marianne Weber, die Witwe des Meisters; sie notierte ihrerseits den Besuch des vielversprechenden jungen Amerikaners. Wovon hier und in vielen anderen Interviews und Gesprächen immer wieder ausführlich und vorwurfsvoll die Rede ist: dass die Militärregierung rundherum versagt, alle denkbaren Fehler macht. Was um so stärker enttäuscht und beunruhigt, als die deutsche Bevölkerung das Besatzungsregime augenscheinlich akzeptiert, wenn nicht begrüßt, keine Spur der befürchteten Widerstandshandlungen („Werwolf“). Die existenzielle Enttäuschung, die unser Held angesichts der (erfolgreichen) Kriegshandlungen empfindet, erstreckt sich also auch auf die Nachkriegsgeschichte. Die große Wandlung bleibt aus. An einer Stelle klagt unser Held, der Menschheit sei, wie sich jetzt zeige, das „innere Licht“ verloren gegangen – ein mystisches Konzept, in dem ein Religionshistoriker wohl die Spur Meister Eckharts, der als exquisit deutscher Mystiker gilt, erkennen würde.

Nun, in dieser Hinsicht gingen unserem Helden prophetische Gaben ab. Trizonesien gilt als mythisch gelungene Vorgeschichte Westdeutschlands, der Bundesrepublik. Der junge Lasky war zu nah dran, konnte die Anzeichen der Zukunft nicht erkennen – wenn sie sich überhaupt zeigten. Auch behinderten ihn seine Heilerwartungen in der Beobachtung, Hoffnungen, ohne die man freilich einen Krieg wie den gegen Hitler nicht wird gewinnen können.

Die Reisen nach Berlin und in die Sowjetzone überzeugen Melvin Lasky, der aus einem marxistisch-trotzkistischen Intellektuellenzirkel New Yorks kommt, dass den Funktionen und den Anhängern der sich bildenden SED in keiner Weise zu trauen ist, wenn sie Demokratie, Gewaltenteilung, Rechtsstaat versprechen. Hier wird, der junge Lasky erkennt es ohne Zögern, der Stalinismus implantiert. Und dem Kampf dagegen hat Melvin Lasky – der ironischerweise lange wie Lenin aussah – sein Leben gewidmet. Und letzten Endes von der Weltgeschichte Recht bekommen.

Melvin Lasky: **Alles war still.** Deutsches Tagebuch 1945. A. d. Engl. von Christa Krüger u. Henning Thies. Rowohlt Berlin. 410 S., 24,95 €.



Melvin Lasky kam 1945 als Chronist der US-Armee nach Deutschland – und entdeckte zwischen allen Trümmern die faszinierenden Reste einer Kulturnation