

TECHNOLOGISCHE SINGULARITÄT

© 1993 von Vernor Vinge

(Dieser Artikel darf für nichtkommerzielle Zwecke vervielfältigt werden, wenn er vollständig kopiert wird, einschließlich dieses Hinweises.)

*Die Originalversion dieses Artikels wurde auf dem vom NASA Lewis Research Center und dem Ohio Aerospace Institute gesponserten VISION-21-Symposium vom 30. bis 31. März 1993 vorgestellt. Eine leicht veränderte Version erschien in der Winterausgabe 1993 von **Whole Earth Review**.*

1. Was ist die Singularität?

Die Beschleunigung des technologischen Fortschritts war das zentrale Merkmal dieses Jahrhunderts. Wir stehen am Rande eines Wandels, vergleichbar mit dem Aufstieg des menschlichen Lebens auf der Erde. Der genaue Grund für diese Veränderung ist die bevorstehende Schaffung von Wesenheiten mit übermenschlicher Intelligenz durch Technologie. Die Wissenschaft kann diesen Durchbruch auf verschiedene Weise erreichen (und dies ist ein weiterer Grund, darauf zu vertrauen, dass das Ereignis eintreten wird):

Computer, die "wach" und übermenschlich intelligent sind, können entwickelt werden. (Bis heute gab es viele Kontroversen darüber, ob wir menschliche Äquivalenz in einer Maschine herstellen können. Aber wenn die Antwort „Ja“ lautet, dann gibt es wenig Zweifel daran, dass in Kürze intelligentere Wesen konstruiert werden können.)

Große Computernetzwerke und ihre zugehörigen Benutzer können als übermenschlich intelligente Einheiten "aufwachen".

Computer/Mensch-Schnittstellen können so intim werden, dass Benutzer vernünftigerweise als übermenschlich intelligent angesehen werden können.

Die biologische Wissenschaft kann Mittel bereitstellen, um den natürlichen menschlichen Intellekt zu verbessern.

Die ersten drei Möglichkeiten hängen von Verbesserungen der Computerhardware ab. Der Hardware-Fortschritt hat in den letzten Jahrzehnten eine erstaunlich stetige Kurve genommen. Basierend auf diesem Trend glaube ich, dass die Schaffung einer übermenschlichen Intelligenz in den nächsten dreißig Jahren stattfinden wird. (Charles Platt hat darauf hingewiesen, dass KI-Enthusiasten seit dreißig Jahren solche Behauptungen aufstellen. Nur damit ich mich nicht einer relativen Zeitweideutigkeit schuldig mache, lassen Sie mich genauer sein: Ich wäre überrascht, wenn dieses Ereignis vor 2005 eintritt oder nach 2030.)

Was sind die Folgen dieses Ereignisses? Wenn eine übermenschliche Intelligenz den Fortschritt vorantreibt, wird dieser Fortschritt viel schneller sein. Tatsächlich scheint es keinen Grund zu geben, warum der Fortschritt selbst nicht die Schaffung noch intelligenterer Einheiten beinhalten würde – in einem noch kürzeren Zeitrahmen. Die beste Analogie, die ich sehe, ist die zur evolutionären Vergangenheit: Tiere können sich an Probleme anpassen und Erfindungen machen, aber oft nicht schneller als die natürliche Selektion ihre Arbeit erledigen kann – die Welt fungiert im Falle der natürlichen Selektion als ihr eigener Simulator. Wir Menschen haben die Fähigkeit, die Welt zu verinnerlichen und Was-wäre-wenn in unseren Köpfen zu führen; Wir können viele Probleme tausendmal schneller lösen, als es die natürliche Auslese könnte. Durch die Schaffung von Mitteln zur

Ausführung dieser Simulationen mit viel höheren Geschwindigkeiten

Diese Veränderung wird ein Wegwerfen aller menschlichen Regeln sein, vielleicht im Handumdrehen – ein exponentielles Ausbrechen, jenseits jeder Hoffnung auf Kontrolle. Entwicklungen, von denen angenommen wurde, dass sie erst in „einer Million Jahren“ (wenn überhaupt) stattfinden könnten, werden wahrscheinlich im nächsten Jahrhundert stattfinden.

Es ist fair, dieses Ereignis als Singularität („die Singularität“ für die Zwecke dieses Stücks) zu bezeichnen. Es ist ein Punkt, an dem unsere alten Modelle verworfen werden müssen und eine neue Realität herrscht, ein Punkt, der immer größer und größer über die menschlichen Angelegenheiten ragen wird, bis die Vorstellung alltäglich wird. Doch wenn es endlich passiert, kann es immer noch eine große Überraschung und eine größere Unbekannte sein. In den 1950er Jahren sahen es nur sehr

¹
wenige: Stan Ulam paraphrasierte John von Neumann mit den Worten:

Ein Gespräch konzentrierte sich auf den immer schneller werdenden Fortschritt der Technologie und die Veränderungen in der Lebensweise der Menschen, was den Anschein erweckt, als würde man sich einer wesentlichen Singularität in der Geschichte der Rasse nähern, über die hinaus die menschlichen Angelegenheiten, wie wir sie kennen, nicht weitergehen könnten.

Von Neumann verwendet sogar den Begriff Singularität, obwohl er anscheinend an normalen Fortschritt denkt, nicht an die Schaffung eines übermenschlichen Intellekts. (Für mich ist die Übermenschlichkeit die Essenz der Singularität. Ohne das würden wir eine Fülle technischer Reichtümer bekommen, die nie richtig absorbiert werden.)

In den 1960er Jahren wurden einige der Implikationen der übermenschlichen Intelligenz erkannt. IJ
schrieb
Good2 :

Lassen Sie eine ultraintelligente Maschine als eine Maschine definieren, die alle intellektuellen Aktivitäten eines noch so schlaunen Mannes bei weitem übertreffen kann. Da das Design von Maschinen eine dieser intellektuellen Aktivitäten ist, könnte eine ultraintelligente Maschine noch bessere Maschinen entwerfen; es würde dann zweifellos eine "Intelligenzexplosion" geben, und die Intelligenz des Menschen würde weit zurückbleiben. Somit ist die erste ultraintelligente Maschine die letzte Erfindung, die der Mensch jemals machen muss, vorausgesetzt, die Maschine ist fügsam genug, uns zu sagen, wie wir sie unter Kontrolle halten können. . . . Es ist mehr als wahrscheinlich, dass innerhalb des zwanzigsten Jahrhunderts eine ultraintelligente Maschine gebaut wird und dass dies die letzte Erfindung sein wird, die der Mensch machen muss.

Good hat die Essenz des Ausreißers erfasst, aber er verfolgt nicht seine beunruhigendsten Konsequenzen. Jede intelligente Maschine der Art, wie er sie beschreibt, wäre kein „Werkzeug“ der Menschheit – genauso wenig wie Menschen die Werkzeuge von Kaninchen, Rotkehlchen oder Schimpansen sind.

In den sechziger, siebziger und achtziger Jahren verbreitete sich die Erkenntnis der Katastrophe. Vielleicht waren es die Science-Fiction-Autoren, die die ersten konkreten Auswirkungen spürten. Schließlich sind die "harten" Science-Fiction-Autoren diejenigen, die versuchen, spezifische Geschichten darüber zu schreiben, was Technologie für uns tun kann. Mehr und mehr fühlten diese Autoren eine undurchsichtige Mauer vor der Zukunft. Früher konnten sie solche Fantasien Millionen von Jahren in die Zukunft versetzen. Jetzt sahen sie, dass ihre fleißigsten Extrapolationen zu dem Unerkennbaren führten. . . demnächst. Einst mögen galaktische Imperien wie eine posthumane Domäne erschienen sein. Jetzt sind es leider sogar interplanetare.

Was ist mit den kommenden Jahrzehnten, wenn wir an den Rand rutschen? Wie wird sich der Ansatz der Singularität über das menschliche Weltbild ausbreiten? Die allgemeinen Kritiker der maschinellen Intelligenz werden noch eine Weile gute Presse haben. Bis wir eine Hardware haben, die so leistungsfähig wie ein menschliches Gehirn ist, ist es wahrscheinlich töricht zu glauben, dass wir in der Lage sein werden, eine dem Menschen gleichwertige (oder höhere) Intelligenz zu schaffen. (Es gibt die weit hergeholte Möglichkeit, dass wir aus weniger leistungsstarker Hardware ein menschliches

Äquivalent machen könnten – wenn wir bereit wären, auf Geschwindigkeit zu verzichten, wenn wir bereit wären, uns mit einem künstlichen Wesen zufrieden zu geben, das buchstäblich langsam ist. Aber das ist viel wahrscheinlicher. Die Entwicklung der Software wird ein kniffliger Prozess sein, der viele Fehlstarts und Experimente mit sich bringt.

Aber mit der Zeit sollten wir mehr Symptome sehen. Das Dilemma der Science-Fiction-Autoren wird sich auch in anderen kreativen Unternehmungen bemerkbar machen. (Ich habe gehört, wie nachdenkliche Comicautoren sich Sorgen darüber machen, wie man spektakuläre Effekte erzeugt, wenn alles Sichtbare durch das technologisch Alltägliche produziert werden kann.) Wir werden erleben, wie die Automatisierung immer höhere Jobs ersetzt. Wir haben jetzt Werkzeuge (symbolische Mathematikprogramme, Cad/Cam), die uns von den meisten Plackereien auf niedriger Ebene befreien. Anders ausgedrückt: Die wirklich produktive Arbeit ist die Domäne einer immer kleineren und elitäreren Fraktion der Menschheit. Im Kommen der Singularität werden wir sehen, wie sich die Vorhersagen einer wahren technologischen Arbeitslosigkeit endlich bewahrheiten.

Ein weiteres Symptom des Fortschritts in Richtung der Singularität: Ideen selbst sollten sich immer schneller verbreiten, und selbst die radikalsten werden schnell alltäglich werden.

Und was ist mit der Ankunft der Singularität selbst? Was lässt sich über sein tatsächliches Aussehen sagen? Da es sich um einen intellektuellen Ausreißer handelt, wird er wahrscheinlich schneller eintreten als jede bisher beobachtete technische Revolution. Das auslösende Ereignis wird wahrscheinlich unerwartet sein – vielleicht sogar von den beteiligten Forschern („Aber alle unsere vorherigen Modelle waren katatonisch! Wir haben nur einige Parameter optimiert ...“). Wenn die Vernetzung weit genug verbreitet ist (in allgegenwärtigen eingebetteten Systemen), mag es scheinen, als ob unsere Artefakte als Ganzes plötzlich erwacht wären.

Und was passiert ein oder zwei Monate (oder ein oder zwei Tage) danach? Ich habe nur Analogien, auf die ich verweisen kann: Der Aufstieg der Menschheit. Wir werden in der posthumanen Ära sein. Und bei all meinem technologischen Optimismus denke ich, dass ich mich wohler fühlen würde, wenn ich diese transzendenten Ereignisse aus tausend Jahren Entfernung betrachten würde. . . statt zwanzig.

2. Kann die Singularität vermieden werden?

Nun, vielleicht wird es gar nicht passieren: Manchmal versuche ich, mir die Symptome vorzustellen, mit denen wir rechnen müssen, wenn sich die Singularität nicht entwickelt. Es gibt die weithin

3 4

respektierten Argumente von Penrose³ und Searle⁴ gegen die Praktikabilität der maschinellen Intelligenz. Im August 1992 veranstaltete die Thinking Machines Corporation einen Workshop, um zu untersuchen, „Wie wir eine denkende Maschine bauen werden“. Wie der Titel des Workshops vermuten lässt, unterstützten die Teilnehmer die Argumente gegen maschinelle Intelligenz nicht besonders. Tatsächlich herrschte allgemeine Einigkeit darüber, dass Geister auf nichtbiologischen Substraten existieren können und dass Algorithmen von zentraler Bedeutung für die Existenz von Geistern sind. Es gab jedoch viele Diskussionen über die rohe Hardwareleistung, die in organischen Gehirnen vorhanden ist. Eine Minderheit war der Ansicht, dass die größten Computer von 1992 nur drei Größenordnungen von der Leistung des menschlichen Gehirns entfernt waren. Die Mehrheit der

5

Teilnehmer stimmte der Einschätzung von Hans Moravec zu⁵ dass wir zehn bis vierzig Jahre von der Hardware-Parität entfernt sind. Und doch gab es eine andere Minderheit, die vermutete, dass die Rechenleistung einzelner Neuronen viel höher sein könnte als allgemein angenommen. Wenn dem so ist, könnte unsere gegenwärtige Computerhardware bis zu zehn Größenordnungen hinter der Ausrüstung zurückbleiben, die wir in unseren Köpfen herumtragen. Wenn dies wahr ist (oder wenn die Kritik von Penrose oder Searle gültig ist), werden wir vielleicht nie eine Singularität sehen. Stattdessen begannen sich unsere Hardware-Leistungskurven in den frühen 00er Jahren abzuflachen – weil wir nicht in der Lage waren, die Designarbeit zu automatisieren, die zur Unterstützung weiterer Hardware-Verbesserungen erforderlich war. Wir würden am Ende sehr leistungsfähige Hardware haben, aber ohne die Möglichkeit, sie weiter voranzutreiben. Kommerzielle digitale Signalverarbeitung könnte großartig sein, selbst digitalen Operationen ein analoges Aussehen verleihen, aber nichts würde jemals

"aufwachen" und es würde niemals den intellektuellen Ausreißer geben, der die Essenz der Singularität ausmacht. Es würde wahrscheinlich als ein goldenes Zeitalter angesehen werden. . . und es wäre auch ein Ende des Fortschritts. Das ist sehr ähnlich wie die von Gunther Stent vorhergesagte Zukunft, der ausdrücklich die Entwicklung transhumaner Intelligenz als hinreichende Bedingung anführt, um seine Projektionen zu brechen.⁶

Aber wenn die technologische Singularität passieren kann, wird sie passieren. Selbst wenn alle Regierungen der Welt die „Bedrohung“ verstehen und sich davor in tödlicher Angst befinden würden, würde der Fortschritt in Richtung des Ziels weitergehen. Der Wettbewerbsvorteil – wirtschaftlich, militärisch, sogar künstlerisch – jedes Fortschritts in der Automatisierung ist so zwingend, dass das Verbot solcher Dinge lediglich sicherstellt, dass jemand anderes sie zuerst bekommt.

Eric Drexler hat spektakuläre Erkenntnisse darüber geliefert, wie weit technische Verbesserungen gehen können.⁷ Er stimmt zu, dass in naher Zukunft übermenschliche Intelligenzen verfügbar sein werden. Aber Drexler argumentiert, dass wir solche transhumanen Geräte einschränken können, damit ihre Ergebnisse untersucht und sicher verwendet werden können.

Ich behaupte, dass eine Beschränkung an sich unpraktisch ist. Stellen Sie sich vor, Sie wären in Ihrem Haus eingesperrt und hätten nur begrenzten Datenzugriff nach außen, zu Ihren Mastern. Wenn diese Meister mit einer Geschwindigkeit dachten – sagen wir – eine Million Mal langsamer als Sie, gibt es kaum Zweifel, dass Sie über einen Zeitraum von Jahren (Ihrer Zeit) einen Weg finden könnten, zu entkommen. Ich nenne diese „schnell denkende“ Form der Superintelligenz „schwache Übermenschlichkeit“. Solch ein "schwach übermenschliches" Wesen würde wahrscheinlich in ein paar Wochen außerhalb der Zeit ausbrennen. "Starke Übermenschlichkeit" wäre mehr als nur die Taktrate eines menschenähnlichen Geistes zu erhöhen. Es ist schwer zu sagen, wie „starke Übermenschlichkeit“ aussehen würde, aber der Unterschied scheint tiefgreifend zu sein. Stellen Sie sich vor, einen Hundgeist mit sehr hoher Geschwindigkeit laufen zu lassen. Würden tausend Jahre Hundeleben zu irgendeiner menschlichen Einsicht führen? Viele Spekulationen über Superintelligenz scheinen auf dem schwach übermenschlichen Modell zu beruhen. Ich glaube, dass wir unsere besten Vermutungen über die Welt nach der Singularität erhalten können, indem wir über die Natur der starken Übermenschlichkeit nachdenken. Ich werde auf diesen Punkt zurückkommen.

Ein anderer Ansatz zur Eindämmung besteht darin, Regeln in den Geist der geschaffenen übermenschlichen Entität einzubauen. Ich denke, dass alle Regeln, die streng genug sind, um effektiv zu sein, auch ein Gerät hervorbringen würden, dessen Fähigkeiten den uneingeschränkten Versionen deutlich unterlegen wären (so dass die menschliche Konkurrenz die Entwicklung der gefährlicheren Modelle begünstigen würde).

Wenn die Singularität nicht verhindert oder eingedämmt werden kann, wie schlimm könnte dann die posthumane Ära sein? Brunnen . . . ziemlich schlimm. Die physische Auslöschung der Menschheit ist eine Möglichkeit. (Oder, wie Eric Drexler es über die Nanotechnologie ausdrückte: Angesichts all dessen, was eine solche Technologie leisten kann, würden Regierungen vielleicht einfach entscheiden, dass sie keine Bürger mehr brauchen.) Doch physisches Aussterben ist vielleicht nicht die gruseligste Möglichkeit. Denken Sie an die unterschiedlichen Arten, wie wir mit Tieren umgehen. Eine posthumane Welt hätte noch viele Nischen, in denen eine menschenähnliche Automatisierung wünschenswert wäre: eingebettete Systeme in autonomen Geräten, selbstbewusste Daemons in den niederen Funktionen größerer Lebewesen. (Eine stark übermenschliche Intelligenz wäre

⁸ wahrscheinlich eine Society of Mind mit einigen sehr kompetenten Komponenten.) Einige dieser menschlichen Äquivalente könnten für nichts anderes als die digitale Signalverarbeitung verwendet werden. Andere mögen sehr menschlich sein, aber mit einer Einseitigkeit, einer Hingabe, die sie in unserer Zeit in eine psychiatrische Klinik bringen würde. Obwohl keines dieser Geschöpfe ein Mensch aus Fleisch und Blut sein mag, könnten sie in der neuen Umgebung dem am nächsten kommen, was wir heute Mensch nennen.

Ich habe oben argumentiert, dass wir die Singularität nicht verhindern können, dass ihr Kommen eine

unvermeidliche Folge der natürlichen Konkurrenzfähigkeit der Menschen und der der Technologie innewohnenden Möglichkeiten ist. Und doch: Wir sind die Initiatoren. Auch die größte Lawine wird durch Kleinigkeiten ausgelöst. Wir haben die Freiheit, Anfangsbedingungen festzulegen, Dinge auf eine Weise geschehen zu lassen, die weniger feindlich ist als andere. Natürlich ist (wie bei Lawinenausbrüchen) möglicherweise nicht klar, was der richtige Führungsschubser wirklich ist:

3. Andere Wege zur Singularität

Wenn Menschen davon sprechen, übermenschlich intelligente Wesen zu erschaffen, stellen sie sich normalerweise ein KI-Projekt vor. Aber wie ich am Anfang dieses Artikels bemerkte, gibt es andere Wege zur Übermenschlichkeit. Computernetzwerke und Mensch-Computer-Schnittstellen scheinen banaler als KI, könnten aber zur Singularität führen. Ich nenne diesen gegensätzlichen Ansatz Intelligence Amplification (IA). IA geht sehr natürlich vor, in den meisten Fällen wird es von seinen Entwicklern nicht einmal als das erkannt, was es ist. Aber jedes Mal, wenn unsere Fähigkeit, auf Informationen zuzugreifen und sie anderen mitzuteilen, verbessert wird, haben wir in gewissem Sinne eine Steigerung gegenüber der natürlichen Intelligenz erreicht. Schon jetzt ist das Team eines Ph.D. Eine menschliche und gute Computer-Workstation (sogar eine Off-Net-Workstation) könnte wahrscheinlich jeden existierenden schriftlichen Intelligenztest ausreizen.

Und es ist sehr wahrscheinlich, dass IA ein viel einfacherer Weg zum Erreichen von Übermenschlichkeit ist als reine KI. Beim Menschen sind die schwierigsten Entwicklungsprobleme bereits gelöst. Aus uns selbst heraus aufzubauen, sollte einfacher sein, als herauszufinden, was wir wirklich sind, und dann Maschinen zu bauen, die all das sind. Und es gibt zumindest mutmaßlichen

9

Präzedenzfall für diesen Ansatz. Cairns-Smith hat spekuliert, dass biologisches Leben als Ergänzung zu noch primitiverem Leben auf der Grundlage von kristallinem Wachstum entstanden sein könnte. Lynn Margulis (in **10** und anderswo) hat starke Argumente dafür vorgebracht, dass Mutualismus eine große treibende Kraft in der Evolution ist.

Beachten Sie, dass ich nicht vorschlage, die KI-Forschung zu ignorieren. KI-Fortschritte werden oft in IA Anwendung finden und umgekehrt. Ich schlage vor, dass wir anerkennen, dass es in der Netzwerk- und Schnittstellenforschung etwas so Tiefgründiges (und potenziell Wildes) wie künstliche Intelligenz gibt. Mit dieser Einsicht sehen wir möglicherweise Projekte, die nicht so direkt anwendbar sind wie herkömmliche Schnittstellen- und Netzwerkdesignarbeiten, die uns aber dazu dienen, uns auf dem IA-Pfad in Richtung der Singularität voranzubringen.

Hier einige mögliche Projekte, denen aus IA-Sicht eine besondere Bedeutung zukommt:

Mensch/Computer-Team-Automatisierung: Nehmen Sie Probleme, die normalerweise für eine rein maschinelle Lösung in Betracht gezogen werden (wie Bergsteigerprobleme), und entwerfen Sie Programme und Schnittstellen, die sich die menschliche Intuition und verfügbare Computerhardware zunutze machen. In Anbetracht der Bizarrheit von höherdimensionalen Bergsteigerproblemen (und der sauberen Algorithmen, die zu ihrer Lösung entwickelt wurden), könnten dem menschlichen Teammitglied einige sehr interessante Anzeigen und Steuerungswerkzeuge zur Verfügung gestellt werden.

Mensch-Computer-Symbiose in der Kunst: Kombinieren Sie die grafische Gestaltungsfähigkeit moderner Maschinen mit der ästhetischen Sensibilität des Menschen. Natürlich ist ein enormer Forschungsaufwand in die Gestaltung von Computerhilfen für Künstler geflossen. Ich schlage vor, dass wir ausdrücklich eine stärkere Zusammenführung von Kompetenzen anstreben, dass wir die Möglichkeit der Kooperation explizit anerkennen. Karl Sims hat in dieser Richtung wunderbare Arbeit

11

geleistet.

Mensch/Computer-Teams bei Schachturnieren: Wir haben bereits Programme, die besser spielen können als fast alle Menschen. Aber wie viel Arbeit wurde daran gearbeitet, wie diese Kraft von einem Menschen genutzt werden könnte, um etwas noch Besseres zu erreichen? Wenn solche Mannschaften

zumindes an einigen Schachturnieren zugelassen würden, könnte dies die positive Wirkung auf die IA-Forschung haben, die die Zulassung von Computern in Turnieren für die entsprechende Nische in der KI hatte.

Schnittstellen, die Computer- und Netzwerkzugriff ermöglichen, ohne dass der Mensch an einen Ort gebunden sein muss, der vor einem Computer sitzt. (Dieser Aspekt von IA passt so gut zu bekannten wirtschaftlichen Vorteilen, dass bereits viel Aufwand darauf verwendet wird.)

Symmetrischere Entscheidungsunterstützungssysteme. Ein beliebtes Forschungs-/Produktgebiet der letzten Jahre waren Entscheidungsunterstützungssysteme. Dies ist eine Form von IA, ist aber möglicherweise zu sehr auf orakelhafte Systeme ausgerichtet. Genauso wie das Programm dem Benutzer Informationen gibt, muss es die Idee geben, dass der Benutzer dem Programm Anleitung gibt.

Lokale Netze, um menschliche Teams effektiver zu machen als ihre Mitglieder. Dies ist in der Regel der Bereich „Groupware“; der Blickwinkelwechsel wäre hier, die Gruppentätigkeit als einen Kombinationsorganismus zu betrachten.

In gewisser Weise könnte das Ziel dieses Vorschlags darin bestehen, eine "Ordnungsordnung" für solche Kombinationsoperationen zu erfinden. Beispielsweise lässt sich der Gruppenfokus leichter aufrechterhalten als in klassischen Meetings. Die Expertise einzelner Mitglieder könnte von Ego-Themen isoliert werden, so dass sich der Beitrag verschiedener Mitglieder auf das Teamprojekt konzentriert. Und natürlich könnten gemeinsame Datenbanken viel bequemer genutzt werden als im herkömmlichen Gremienbetrieb.

Das Internet als kombiniertes Mensch-Maschine-Werkzeug. Von allen Punkten auf der Liste geht es hier am schnellsten voran. Die Macht und der Einfluss des Internets werden stark unterschätzt. Die Anarchie der Entwicklung des weltweiten Netzes ist ein Beweis für sein Potenzial. Mit zunehmender Konnektivität, Bandbreite, Archivgröße und Computergeschwindigkeit sehen wir so etwas wie Lynn Margulis' Vision der Biosphäre als Datenprozessor rekapituliert, aber mit millionenfach höherer Geschwindigkeit und mit Millionen von menschlich intelligenten Agenten (uns).

Die obigen Beispiele veranschaulichen Forschung, die im Kontext moderner Informatikabteilungen durchgeführt werden kann. Es gibt andere Paradigmen. Beispielsweise würde ein Großteil der Arbeit in den Bereichen künstliche Intelligenz und neuronale Netze von einer engeren Verbindung mit dem biologischen Leben profitieren. Anstatt einfach zu versuchen, biologisches Leben mit Computern zu modellieren und zu verstehen, könnte die Forschung auf die Schaffung zusammengesetzter Systeme ausgerichtet sein, die sich auf biologisches Leben als Orientierungshilfe verlassen oder für die Funktionen, die wir noch nicht gut genug verstehen, um sie in Hardware zu implementieren. Ein langjähriger Traum der Science-Fiction waren direkte Gehirn-zu-Computer-Schnittstellen. Tatsächlich wird in diesem Bereich konkret gearbeitet:

Die Gliedmaßenprothetik ist ein Thema mit direkter kommerzieller Anwendbarkeit. Nerv-zu-Silizium-Wandler können hergestellt werden. Dies ist ein aufregender kurzfristiger Schritt in Richtung direkter Kommunikation.

Direkte Verbindungen in das Gehirn scheinen machbar, wenn die Bitrate niedrig ist: Angesichts der menschlichen Lernflexibilität müssen die tatsächlichen Neuronenziele des Gehirns möglicherweise nicht genau ausgewählt werden. Sogar 100 Bit pro Sekunde wären von großem Nutzen für Schlaganfallopfer, die sonst auf menügesteuerte Schnittstellen beschränkt wären.

Das Einstecken in den optischen Trunk hat das Potenzial für Bandbreiten von etwa 1 Mbit/Sekunde. Aber dafür müssen wir die feine Architektur des Sehens kennen und wir müssen ein riesiges Netz von Elektroden mit exquisiter Präzision platzieren. Wenn wir wollen, dass unsere Verbindung mit hoher Bandbreite die bereits im Gehirn vorhandenen Pfade ergänzt, wird das Problem weitaus unlösbarer. Einfach ein Gitter von Empfängern mit hoher Bandbreite in ein Gehirn zu stecken, reicht sicherlich nicht aus. Aber nehmen Sie an, dass das Gitter mit hoher Bandbreite vorhanden war, als sich die Gehirnstruktur aufbaute, als sich der Embryo entwickelte. Das suggeriert:

Tierembryo-Experimente. Ich würde in den ersten Jahren solcher Forschung keinen IA-Erfolg erwarten, aber wenn sich entwickelnde Gehirne Zugang zu komplexen simulierten neuronalen Strukturen erhalten, könnten auf lange Sicht Tiere mit zusätzlichen Sinnespfaden und interessanten intellektuellen Fähigkeiten hervorgebracht werden.

Ich hatte gehofft, dass diese Diskussion über IA einige deutlich sicherere Zugänge zur Singularität hervorbringen würde (schließlich ermöglicht IA unsere Teilhabe an einer Art Transzendenz). Leider bin ich mir so ziemlich sicher, dass diese Vorschläge in Betracht gezogen werden sollten, damit sie uns mehr Optionen bieten. Aber was die Sicherheit angeht – einige der Vorschläge sind auf den ersten Blick ein wenig beängstigend. IA für einzelne Menschen schafft eine ziemlich finstere Elite. Wir Menschen haben Millionen von Jahren evolutionäres Gepäck, das uns dazu bringt, Wettbewerb in einem tödlichen Licht zu sehen. Vieles von dieser Tödlichkeit ist in der heutigen Welt möglicherweise nicht notwendig, in einer Welt, in der Verlierer die Tricks der Gewinner übernehmen und in die Unternehmen der Gewinner eingegliedert werden. Eine Kreatur, die de novo gebaut wurde, könnte möglicherweise eine viel harmlosere Entität sein als eine, die auf Reißzähnen und Krallen basiert.

Das Problem ist nicht einfach, dass die Singularität das Hinscheiden der Menschheit aus dem Mittelpunkt darstellt, sondern dass sie unseren tiefsten Vorstellungen vom Sein widerspricht. Ich denke, ein genauerer Blick auf den Begriff der starken Übermenschlichkeit kann zeigen, warum das so ist.

4. Starke Übermenschlichkeit und das Beste, was wir uns wünschen können

Angenommen, wir könnten die Singularität maßschneidern. Angenommen, wir könnten unsere extravagantesten Hoffnungen verwirklichen. Was würden wir dann verlangen? Dass die Menschen selbst ihre eigenen Nachfolger werden würden, dass jede Ungerechtigkeit durch unser Wissen um unsere Wurzeln gemildert würde. Für diejenigen, die unverändert blieben, wäre das Ziel eine wohlwollende Behandlung (vielleicht sogar, den Stay-behinds den Anschein zu geben, Herren von gottähnlichen Sklaven zu sein). Es könnte ein goldenes Zeitalter werden, das auch Fortschritt beinhaltet (Überspringen der Stent-Barriere). Unsterblichkeit (oder zumindest ein Leben lang, solange wir das Universum überleben lassen) wäre erreichbar.

Aber in dieser hellsten und freundlichsten Welt werden die philosophischen Probleme selbst einschüchternd. Ein Geist, der bei gleicher Kapazität bleibt, kann nicht ewig leben; nach ein paar tausend Jahren würde es eher wie eine sich wiederholende Tonbandschleife als eine Person aussehen. Um unendlich lange zu leben, muss der Geist selbst wachsen. . . und wenn es groß genug wird, und zurückschaut . . . welches Mitgefühl kann es mit der Seele haben, die es ursprünglich war? Das spätere Wesen wäre alles, was das Original war, aber weitaus mehr. Und so muss auch für den Einzelnen die Vorstellung von Cairns-Smith oder Lynn Margulis, dass neues Leben schrittweise aus dem alten wächst, noch gelten.

Dieses „Problem“ der Unsterblichkeit taucht auf viel direktere Weise auf. Die Vorstellung von Ego und Selbstbewusstsein war die Grundlage des nüchternen Rationalismus der letzten paar Jahrhunderte. Doch jetzt wird der Begriff des Selbstbewusstseins von den Leuten der künstlichen Intelligenz angegriffen. Intelligence Amplification untergräbt unser Konzept des Egos aus einer anderen Richtung. Die Welt nach der Singularität wird Netzwerke mit extrem hoher Bandbreite beinhalten. Ein zentrales Merkmal stark übermenschlicher Wesenheiten wird wahrscheinlich ihre Fähigkeit sein, mit variablen Bandbreiten zu kommunizieren, darunter solche, die weit über Sprache oder geschriebenen Nachrichten liegen. Was passiert, wenn Teile des Ego kopiert und zusammengeführt werden können, wenn das Selbstbewusstsein wachsen oder schrumpfen kann, um sich der Natur der betrachteten Probleme anzupassen? Dies sind wesentliche Merkmale der starken Übermenschlichkeit und der Singularität.

Aus einem bestimmten Blickwinkel passt die Vision zu vielen unserer glücklichsten Träume: eine endlose Zeit, in der wir einander wirklich kennen und die tiefsten Geheimnisse verstehen können. Aus

einem anderen Blickwinkel ähnelt es dem Worst-Case-Szenario, das ich mir zuvor vorgestellt habe.

Tatsächlich denke ich, dass die neue Ära einfach zu anders ist, um in den klassischen Rahmen von Gut und Böse zu passen. Dieser Rahmen basiert auf der Idee von isolierten, unveränderlichen Köpfen, die durch schwache Verbindungen mit geringer Bandbreite verbunden sind. Aber die Welt nach der Singularität passt zu der größeren Tradition des Wandels und der Zusammenarbeit, die vor langer Zeit begann (vielleicht sogar vor dem Aufstieg des biologischen Lebens). Ich denke, dass bestimmte Vorstellungen von Ethik in einer solchen Zeit gelten würden. Forschungen zu IA und Kommunikation mit hoher Bandbreite sollten dieses Verständnis verbessern. Ich sehe jetzt nur den Schimmer davon; Vielleicht gibt es Regeln, um sich selbst von anderen auf der Grundlage der Verbindungsbandbreite zu unterscheiden. Und während der Geist und das Selbst weitaus labiler sein werden als in der Vergangenheit, muss vieles von dem, was wir schätzen (Wissen, Gedächtnis, Gedanken), niemals verloren gehen. Ich denke, Freeman Dyson hat Recht, wenn er sagt: "*Gott ist das, was aus dem Verstand wird, wenn er den Maßstab unseres Verständnisses überschritten hat.*"¹²

Verweise

1. *Ulam, S.* , **Tribute to John von Neumann** , Bulletin of the American Mathematical Society, vol. 64. nein. 3, Mai 1958, S. 1-49.
2. *Good, IJ* , **Speculations Concerning the First Ultraintelligent Machine** , in *Advances in Computers*, Band 6, Franz L. Alt und Morris Rubinoff, Hrsg., 31-88, 1965, Academic Press.
3. *Penrose, Roger* , **The Emperor's New Mind** , Oxford University Press, 1989.
4. *Searle, John R.* , **Minds, Brains, and Programs** , in *The Behavioral and Brain Sciences*, vol. 3, Cambridge University Press, 1980.
5. *Moravec, Hans* , **Mind Children**, Harvard University Press, 1988.
6. *Stent, Gunther S.* , **The Coming of the Golden Age: A View of the End of Progress** , The Natural History Press, 1969.
7. *Drexler, K. Eric* , **Engines of Creation** , Anchor Press/Doubleday, 1986.
8. *Minsky, Marvin* , **Society of Mind** , Simon und Schuster, 1985.
9. *Cairns-Smith, AG* , **Seven Clues to the Origin of Life** , Cambridge University Press, 1985.
10. *Margulis, Lynn and Dorian Sagan* , **Microcosmos: Four Billion Years of Evolution From Our Microbial Ancestors** , Summit Books, 1986.
11. *Sims, Karl* , **Interactive Evolution of Dynamical Systems** , Thinking Machines Corporation, Technical Report Series (veröffentlicht in *Toward a Practice of Autonomous Systems: Proceedings of the First European Conference on Artificial Life*, Paris, MIT Press, Dezember 1991).
12. *Dyson, Freeman* , **Infinite in All Directions** , Harper & Row, 1988.

Andere Quellen

Alfvén, Hannes , schreibt als Olof Johanneson, **The End of Man?** , Award Books, 1969.

Anderson, Poul , **Kings Who Die** , If, März 1962, 8-36.

Asimov, Isaac , **Runaround** , Astounding Science Fiction, März 1942, 94.

Barrow, John D. und Frank J. Tipler , **The Anthropic Cosmological Principle** , Oxford University Press, 1986.

Bear, Greg , **Blood Music** , Analog Science Fiction-Science Fact Juni 1983.

Conrad, Michael , et al., **Towards an Artificial Brain** , BioSystems, vol. 23, 175-218, 1989.

Dyson, Freeman , **Physik und Biologie in einem offenen Universum** , Review of Modern Physics, vol. 51, 447-460, 1979.

Herbert, Frank , **Dune** , Berkeley Books, 1985.

Kovacs, GTA, et al. , **Regeneration Microelectrode Array for Peripheral Nerve Recording and Stimulation** , IEEE Transactions on Biomedical Engineering, vol. 39, Nr. 9, 893-902.

Niven, Larry , **The Ethics of Madness** , If, April 1967, 82-108.

Platt, Charles , private Mitteilung.

Rasmussen, S. et al. , **Computational Connectionism within Neurons: a Model of Cytoskeletal Automata Subserving Neural Networks**, in Emergent Computation, Stephanie Forrest, Hrsg., 428-449, MIT Press, 1991.

Stapledon, Olaf , **The Starmaker** , Berkeley Books, 1961.

Swanwick Michael , **Vacuum Flowers** , erschienen in Isaac Asimovs Science Fiction Magazine, Dezember 1986 - Februar 1987.

Thearling, Kurt , **How We Will Build a Machine That Thinks** , Workshop bei Thinking Machines Corporation, 24.-26. August 1992.

Vinge, Vernor , **Bookworm, Run!** , Analog, März 1966, 8-40.

Vinge, Vernor , **True Names** , Binary Star Number 5, Dell, 1981.

Vinge, Vernor , **First Word**, Omni, Januar 1983, 10.