Ein Bericht von der Konferenz The Science of Consciousness in Taormina – 22. bis 27. Mai 2023

Eine bekannte Geschichte des muslimischen Weisen Nasreddin Hodscha berichtet, dass Nasreddin nachts unter einer Lampe steht und verzweifelt nach etwas sucht. Ein Passant fragt ihn: „Nasreddin, was suchst du?“ Nasreddin antwortet: „Ich suche meine Hausschlüssel.“ „Hast du sie hier verloren?“ „Bestimmt nicht, aber ich suche hier, weil hier das Licht ist.“

Das ist ein häufiges Phänomen: Wir suchen etwas nicht dort, wo es mit großer Wahrscheinlichkeit zu finden ist, sondern dort, wo es am bequemsten zu suchen ist. Etwas Ähnliches, so scheint mir, geschah und geschieht auf den Konferenzen über das Bewusstsein, „The Science of Consciousness“ (TSC). Die jüngste dieser Konferenzen fand vom 22. bis 27. Mai 2023 in Taormina, Sizilien, statt (das vollständige Programm und das Buch mit den Kurzfassungen finden Sie unter https://tsc2023-taormina.it/), und ich hatte das Privileg, im Namen des Scientific and Medical Networks an dieser Konferenz in einer wunderschönen Umgebung teilzunehmen.



Abbildung 1 – Der Ätna von der Porta Catania in Taormina aus gesehen

Ich spreche in der Vergangenheitsform, weil ich die TSC-Konferenzen in den letzten 20 Jahren immer wieder selbst besucht habe und Doktoranden und Postdocs hingeschickt habe. Die letzten europäischen Konferenzen (Helsinki 2017, Interlaken 2019) habe ich selbst besucht. Im Vergleich zu den vorangegangenen europäischen Konferenzen war diese in Sizilien in ihrem Kernplenarprogramm konventioneller. Das heißt: Die meisten, wenn nicht sogar alle Vorträge in dem dicht gedrängten Plenarprogramm gingen von der impliziten materialistischen Annahme aus, dass das Gehirn irgendwie das Bewusstsein produziert: *„Ich bin sicher, wir sind uns alle einig, dass das Bewusstsein vom Gehirn produziert wird„*, war einer der Standardsätze, die einige Redner verwendeten. Wie genau dieser Produktionsprozess ablaufen soll, ist Gegenstand einer kleinen, zivilisierten Debatte, die nicht sehr weit vorangeschritten ist, seit Stuart Hameroff 1997 diese Konferenzreihe in Tucson ins Leben rief. „Neuromythologie“ hat man sie genannt [1].

Ich erinnere mich an einige gewagtere Versuche in Tucson in den frühen 2000er Jahren, als

Henry Stapp noch ein Plenarredner war. Und einige Konferenzteilnehmer, mit denen ich sprach, äußerten ihre Enttäuschung darüber, dass einige fortschrittlichere Versuche entweder gar nicht auf der Konferenz vertreten waren oder auf die acht parallel stattfindenden Sitzungen verwiesen wurden, in denen vielleicht 20 bis 40 Teilnehmer den Vorträgen zuhörten und diskutierten. Sie fanden täglich am Nachmittag statt und waren nach Themen geordnet.

Davon abgesehen waren die Plenarvorträge größtenteils von hohem Niveau, und wenn man sich über den aktuellen Stand auf diesem Gebiet informieren möchte, ist diese Konferenz eine gute Wahl, da sie die Mächtigen und Tüchtigen auf diesem Gebiet versammelt, die dann einen kompetenten Überblick über ihr eigenes Fachgebiet vermitteln. Die meisten von ihnen haben eines oder mehrere wissenschaftliche Bücher veröffentlicht, auf die sie zurückgreifen.

Die Konferenzreihe wurde 1997 ins Leben gerufen und drehte sich stets um die These von Stuart Hameroff, dass Quantenprozesse im Gehirn das Bewusstsein unterstützen, ja sogar erzeugen. Im Wesentlichen besagt sein Modell, dass nicht die Neuronen selbst die algorithmischen Einheiten des Gehirns sind, sondern das Netzwerk der Mikrotubuli, d.h. das Zytoskelett der Neuronen. Hameroff und Kollegen gehen davon aus, dass Mikrotubuli die Basiseinheiten sind, die die Rechenleistung des Gehirns ausmachen, und in ihnen die Tau-Proteine, die als Schalter fungieren. Mikrotubuli haben die einzigartige Eigenschaft, hohle Resonatoren von einer Größe zu bilden, die kohärente Resonanzphänomene ermöglicht. Und Tubulin, das Protein, aus dem unter anderem Mikrotubuli bestehen, dient als Mittel zur Unterstützung von Quantentunnelprozessen, sagt Hameroff. Wenn man die Mikrotubuli als Grundlage der Rechenleistung des Gehirns betrachtet, könnte unser Gehirn 10^{27} Operationen pro Sekunde ausführen, was um den Faktor 10^{11} höher ist als die 10^{16} Operationen pro Sekunde, die möglich sind, wenn man die Neuronen als Grundlage der Rechenleistung unseres Gehirns betrachtet.

In seinem Plenarvortrag am Ende der ersten Sitzung, die diesem Thema gewidmet war, forderte *Stuart Hameroff* eine *Revolution in der Hirn- und Bewusstseinsforschung*. Er vertrat die These, dass wir beim Verständnis des Bewusstseins keine Fortschritte gemacht haben, weil wir von falschen Voraussetzungen ausgegangen sind, nämlich dass Neuronen die grundlegenden Operationseinheiten des Gehirns sind und dass die Vernetzung der Neuronen das Bewusstsein hervorbringt. Dem sei nicht so, sagt Stuart Hameroff, der von Haus aus Anästhesist ist. Alle Narkosemittel und bewusstseinsverändernden Substanzen wie Psychedelika wirken auf das Bewusstsein ein, indem sie die Funktionsfähigkeit der Mikrotubuli beeinträchtigen. Sie stören entweder den Prozess, durch den Tubulin zu Mikrotubuli aggregiert wird, oder indem sie die Schwingungs- und elektrischen Eigenschaften der Moleküle behindern. Mikrotubuli, so Hameroff, sammeln Licht, d.h. sie

funktionieren nicht, indem sie Elektrizität auf herkömmliche Weise leiten – wie ein Draht, wie die gängige Meinung ist, entlang der Axone von Neuronen – sondern indem sie stehende Wellen kohärenter Quantenprozesse von Photonen erzeugen, wie in einem Laser, und andere Mikrotubuli durch kohärente Resonanzphänomene beeinflussen, wie bereits in den 70er Jahren von Fröhlich und seiner Gruppe berechnet und vorhergesagt wurde [2, 3]. Hameroff präsentierte verschiedene Argumente, warum die Standardansicht, dass das Gehirn das Bewusstsein erzeugt, nicht richtig sein kann: Einzellige Organismen wie Pantoffeltierchen können ohne Neuronen und ohne Gehirn fühlen, sich bewegen, paaren und Nahrung aufnehmen. Sie nutzen das Netzwerk der Mikrotubuli, in dem die Tau-Proteine als Schalter fungieren, so Hameroff.

In seiner Zusammenarbeit mit Roger Penrose schlug Stuart Hameroff vor, dass Mikrotubuli das Substrat der orchestrierten Reduktion sind, die Penrose als spontanen Gravitationskollaps eines bestimmten Segments der Raumzeitverzweigung vorhersagte, der das physikalische Äquivalent eines bewussten Moments in Penroses Modell ist [4, 5].

Die anderen Vorträge an diesem Morgen waren praktischerweise um diese allgemeine Idee herum gruppiert: Der erste Vortrag von *Travis Craddock* über *Psychedelika* stellte einige neue Daten vor, die zeigten, dass die meisten Psychedelika tatsächlich das Mikrotubuli-Netzwerk beeinflussen, entweder durch Hemmung von Oszillationen oder durch Beeinträchtigung der Bildung von Mikrotubuli. Meskalin zum Beispiel bindet an Tubulin, ebenso wie Colchicin, ein Medikament zur Behandlung von Gicht, das als Nebenwirkung manchmal Halluzinationen hervorruft. Colchicin stammt übrigens von der Herbstzeitlose (Colchicum autumnale), die in der Homöopathie zur Behandlung von Altersbeschwerden eingesetzt wird.

Jim Al Khalili vom Zentrum für Quantenbiologie an der Universität Surrey erläuterte in hervorragender Weise verschiedene Aspekte der *Nutzung von Quantenprozessen durch das Leben*. Einige dieser Prozesse sind gut belegt, wie z.B. bei der Enzymaktivität und der Photosynthese, der Magnetrezeption bei Vögeln und DNA-Mutationen, aber vielleicht auch bei Krebs und der Entstehung des Lebens. Er wies nachdrücklich darauf hin, dass verschiedene DNA-Mutationen in erster Linie auf Quanten-Tunnelprozesse zurückzuführen sein könnten, die die Erzeugung tautomerer Stellen ermöglichen, sodass Punktmutationen auftreten. Das bedeutet, dass die Basen zwischen den Doppelsträngen der DNA wechseln. Sobald das Leben diese neuartigen DNA-Stränge stabilisiert hat, scheint das neuartige System jedoch zu verhindern, dass quantenmechanische Prozesse die Integrität der DNA weiter stören. Wenn sie versagen, könnte Krebs entstehen. Nebenbei bemerkt, könnte dies neue Denkansätze für Krebs und seine mögliche Behandlung eröffnen.

Jim Al-Khalili ist ein leidenschaftlicher Redner, der offensichtlich davon überzeugt ist, dass

diese quantenbiologische Sichtweise die Rätsel des Lebens und des Bewusstseins löst und damit ein hochkarätiges Beispiel für das naturalistische Unternehmen ist, die Welt ein für alle Mal zu erklären.

Dasselbe gilt für *Anil Seth*, der für den verhinderten Christof Koch einsprang und den Hauptvortrag am Vormittag hielt. Er versuchte seine Zuhörer von seiner These zu überzeugen, dass das Gehirn eine Bayes'sche Vorhersagemaschine ist. Es aktualisiert erlernte und angeborene Vorstellungen über die Welt anhand neuer Informationen und trifft neue Vorhersagen über sie. Diese Vorhersagen und die sie umgebende Gehirnmaschinerie sind das Substrat des Bewusstseins. Sie helfen, das Bewusstsein zu erklären, vorherzusagen und, ja, zu kontrollieren – das war das Wort, das er verwendete. Wir sehen das mechanistische Paradigma in voller Fahrt. Seth hat diese Ansicht in seinem neuen Buch „*Being You – A New Theory of Consciousness*" dargelegt [6]. Diese Redner sind die Lieblinge der Naturalistenszene. Auf seiner Website finden sich Gespräche mit Sam Harris, sein TED-Talk, in dem er erklärt, wie das Gehirn die Realität halluziniert und Bewusstsein erzeugt, und natürlich der Perception Census, eine umfangreiche Online-Studie darüber, wie Menschen verschiedene Reize wahrnehmen.

Ich hatte den Eindruck eines kraftvollen Surfers, der auf der mächtigsten Welle der Welt surft. Aber Wellen haben die unangenehme Angewohnheit, irgendwann zu brechen…

Die Plenarsitzung am Dienstagnachmittag war der *Phänomenologie* gewidmet. *Nicholas Humphrey* sprach über Empfindungsfähigkeit. Humphrey wurde bekannt, weil er einen Affen untersuchte, dessen primärer Sehbereich im Gehirn durch eine Operation zerstört worden war. Doch nach einer Weile konnte sich dieser Affe offenbar in einem Raum voller Hindernisse bewegen, ohne sich zu stoßen, und Erdnüsse finden, die dort ausgelegt waren, was zeigte, dass das Tier immer noch „sehen" konnte, ein Phänomen, das als „Blindsight" bekannt wurde. Er wies darauf hin, dass es eine uralte Sehbahn gibt, die von der Netzhaut bis zum optischen Tectum im Zwischenhirn reicht, dieselbe Bahn, die auch von Fröschen zum „Sehen" benutzt wird. Dabei handelt es sich jedoch um eine unbewusste, aber sehr effektive Art des Sehens. Er benutzte dieses Beispiel, um zu zeigen, dass die Empfindung eine bewusste Apperzeption ist, die vom reinen Wahrnehmungsprozess unterschieden werden muss. (Leibniz pflegte „kleine Wahrnehmungen" ohne Bewusstsein von „Apperzeptionen" zu unterscheiden, die das Bewusstsein hervorbringen, unterstützen und benötigen.) Wenn solche Wahrnehmungen Kopien ihrer selbst, d.h. Repräsentationen höherer Ordnung, hervorbringen, dann führen Wahrnehmungen und Empfindungen zu bewussten kognitiven Operationen und daraus erwächst ein persönliches Selbstgefühl. „*Ich fühle, also bin ich. Du fühlst, also bist du*„, könnte man in diesem Sinne das Diktum von Descartes umformulieren. Damit wäre natürlich auch die Empfindung und damit das Bewusstsein ein kontinuierliches Phänomen, das vielen Tieren ein Bewusstsein verleihen

müsste. Die Voraussetzung dafür ist ein Gehirn, das Rückkopplungsschleifen aufrechterhalten kann, und eine Lebensweise, in der Empfindungen einen Überlebenswert haben können. Es wäre also eine recht junge Erfindung im evolutionären Prozess. Würmer, Quallen und ähnliche Lebewesen wären unbewusst. Bienen, Tintenfische und ähnliche Tiere hätten ein kognitives Bewusstsein, aber keine Empfindung, während die meisten Säugetiere mit einem höheren Gehirn sowohl ein Bewusstsein als auch eine Empfindung hätten. Nach dieser Auffassung wäre ein großer Gedächtnisspeicher, der die Repräsentation von Empfindungen ermöglicht, die Voraussetzung für phänomenales Bewusstsein. Seine Ideen sind in seinem kürzlich erschienenen Buch „Sentience: The Invention of Consciousness" zusammengefasst [7]

Shaun Gallagher, ein Philosoph, sprach über das *minimale Selbst*. Dabei handelt es sich um eine Art präreflexives Selbstbewusstsein, das einen Sinn seiner selbst beinhaltet [8]. Er verwendete Avicennas (Ibn Sinas) Gedankenexperiment des „fliegenden Mannes". Diese Idee wurde von Avicenna im 11. Jahrhundert in seinem Lehrbuch der Psychologie vorgestellt: Die Vorstellung eines neu erschaffenen Menschen, der keine körperlichen Empfindungen hat und in einer Art glückseligem Zustand ohne sensorischen Input im Raum schwebt, sondern nur mit einem Gefühl seiner Selbst ausgestattet ist. Dies ist die phänomenale Erfahrung als solche, wie sie auch von Galen Strawson und dem Phänomenologen Dan Zahavi [9] herausgearbeitet wurde, und ist also nicht sozial konstruiert. Es wäre immer noch ein Selbst und damit anders als der Körper als solcher. Es ist eine Art neu-antiker Drehung des alten idealistischen Arguments, dass sich das Bewusstsein von der Materie unterscheidet. Aber nicht alle werden das übernehmen, denn einige Redner auf der Konferenz waren der Meinung, dass ein solcher fliegender Mann biologisch unmöglich sei. Und weiter geht die Debatte…

Covid veränderte Konferenzen und diese Veränderung war immer noch sichtbar, indem einige Redner Zoom-Vorträge hielten (und nicht wenige trugen Masken), wie der erste am Mittwoch von *Jay Sanguinetti* aus den USA über Hirnstimulation. Er ging davon aus, dass das Gehirn das Bewusstsein erzeugt und dass man durch Stimulation und Knock-out-Studien untersuchen kann, welche Teile des Gehirns für das Bewusstsein verantwortlich sind. Er benutzte eine neuartige Technik der transkraniellen Ultraschallstimulation, die sehr präzise auf bestimmte Bereiche bis hinunter in den Thalamus fokussiert werden kann. Sanguinetti setzte diese Technik in einigen Fällen erfolgreich ein, um Komapatienten zu wecken und den Affekt zu modulieren. Er konnte auch zeigen, dass die Stimulierung des posterioren cingulären Kortex, der Teil des Default-Mode-Netzwerks ist, das selbstbezogene Inhalte wie Bilder und Gedanken verarbeitet, die Ich-Wahrnehmung moduliert. Dieser Bereich wird auch durch Psilocybin und durch transkraniellen Ultraschall gestört. Die Gelassenheit wird durch den Nucleus caudatus, einen Teil der Basalganglien, moduliert, dessen Aktivität bei Langzeitmeditierenden verändert ist. Dies kann durch Ultraschallstimulation nachgeahmt

werden, deren Wirkung bis zu 30 Minuten nach der Stimulation zu beobachten ist. Ich vermute, dass die Erzeugung von Gleichmut wie bei Langzeitmeditierenden eine kontinuierliche Stimulation erfordern würde. Werden wir bald eine Bevölkerung von gestressten Trägern von Ultraschallhelmen sehen? So interessant ich diese Forschung auch fand, so erstaunt war ich doch über den offensichtlichen konzeptionellen Fehler, der anscheinend unbemerkt blieb: Durch die einfache Modulation von Bewusstseinsinhalten oder die Unterbrechung von Prozessen zeigen wir nicht, dass die Gehirnaktivität für den Bewusstseinsprozess kausal ist. Wenn wir die Stromzufuhr zu einem Computerbildschirm unterbrechen, unterbrechen wir den Strom, aber wir würden nicht sagen, dass der Inhalt des Bildschirms durch den Strom erzeugt wird. Elektrizität ist notwendig, aber nicht ausreichend für das Bild auf dem Bildschirm.

Orli Dahan aus Israel hielt einen leidenschaftlichen und inspirierenden Vortrag über das *Bewusstsein während der Geburt*, ihr Promotionsprojekt. Im Wesentlichen zeigte sie anhand von Fragebogen- und Interviewdaten aus einer großen Kohorte von Gebärenden, dass das Gehirn Frauen auf die Geburtserfahrung vorbereitet. Man weiß, dass das Gehirn während der Schwangerschaft Veränderungen durchläuft, indem die graue Substanz zunächst abgebaut wird und dann wieder zunimmt. In der Mitte der Schwangerschaft steigen die Angst- und Furchtwerte, während sie gegen Ende der Schwangerschaft in Verbindung mit einer Hypofrontalität abnehmen. Diese Hypofrontalität, d.h. die Herabregulierung präfrontaler Prozesse, ermöglicht eine Reduzierung der kognitiven Aktivität, sodass das Geburtserlebnis nicht ängstlich erwartet wird und mit weniger Anspannung und Angst angegangen werden kann. Dies ermöglicht es Frauen, sich auf die Entspannung zu konzentrieren, was wiederum die Schmerzen mildert. Aus Dahans Interviews ging hervor, dass Frauen die mit der Geburt verbundenen Schmerzen sogar als positive Erfahrung erleben können. Die derzeitigen Geburtsumgebungen in Krankenhäusern unterstützen diesen natürlichen Prozess jedoch nicht, sondern sind eher störend, mit Lärm, hellem Licht und der Anwesenheit von zu vielen Menschen, was alles Stress und Angst auslöst und zu traumatischen Erfahrungen beiträgt.

Die Plenarsitzungen am Donnerstagmorgen waren den Themen Künstliche Intelligenz und Bewusstsein gewidmet. Den Auftakt bildete ein gemeinsamer Vortrag von *Lenore Blum* und *Manuel Blum* über die *Conscious Turing Machine (CTM)*, d.h. einen riesigen Computer, von dem man annimmt, hofft oder erwartet, dass er irgendwann ein Bewusstsein hat. Die Grundidee ist, dass viele verschiedene spezialisierte Prozessoren verschiedene Eingaben erhalten, Informationen verarbeiten, sie im Kurzzeitgedächtnis ablegen, den Inhalt anderer Speicher verwenden und wieder nach oben weiterleiten. Dies setzt einen Wettbewerbsprozess in Gang, der durch ein Zufallsneuron zwischen den Schichten befeuert wird, das die binären Chunks verzerrt. Das Ergebnis ist die Wahrscheinlichkeit, dass ein Stück verarbeiteter Information, das den Wettbewerb gewinnt, in den Speicher übertragen

wird und schließlich in die Ausgabe gelangt. Das Interessante an dieser Architektur ist, dass es keine zentrale Entscheidungseinheit gibt, sondern nur algorithmische Regeln und Wahrscheinlichkeiten, die die Einheiten von ihren lokalen Positionen abkoppeln. Dadurch entsteht im Langzeitgedächtnis des Systems eine Art Expertenalgorithmus, der auch in der Lage ist, ein Weltmodell zu erzeugen, einschließlich der CTMs selbst. Dieses, so behaupten die Blums, stellt das Bewusstsein dar. Das CTM sendet dieses Weltmodell und diese Repräsentation an alle anderen Prozesse und unterstützt so das Bewusstsein.

Dies ist gleichzeitig die Blumsche Definition von Bewusstsein: Der Empfang einer globalen Übermittlung von Inhalt aus dem Kurzzeitgedächtnis, d.h. sobald ein Prozess global verfügbar ist, soll er bewusst sein. Die Frage, ob das richtig ist, wird pragmatisch beantwortet: Wenn das Modell die Intuition widerspiegelt und wenn es nicht zu kompliziert ist und nicht abstürzt, dann ist es wahrscheinlich gut genug. Interessanterweise stürzten alle deterministischen Modelle ab und nur die probabilistischen überlebten.

Owen Holland von der Universität Edinburgh hielt einen Vortrag, in dem er sich kritisch mit der aktuellen Bewegung der *Künstlichen Intelligenz (KI)* auseinandersetzte. KI-Forscher wüssten zu wenig über Biologie und Bewusstseinsforschung. Künstliche Systeme würden biologische Systeme ergänzen, nicht ersetzen. Aber sie müssen ein Schlüsselelement respektieren: die Emergenz. Hier gibt es zwei Hauptrichtungen: Wie und wann findet Emergenz statt, und wenn sie stattfindet, wie lässt sie sich erklären? Was wichtig sein könnte, sind lokale Aktionen, die zu einer globalen Aufgabe konvergieren. Er veranschaulichte dies mit einem interessanten Video, in dem Roboter, sehr einfache Roboter der ersten Generation aus den 70er Jahren, eine einfache Aufgabe hatten: Sie sollten nur Chips in Paketen von drei Stück sammeln. Die Chips waren chaotisch auf einer Ebene ausgelegt, die durch eine weiche Stoffwand begrenzt war, die es den Robotern ermöglichte, dagegen zu stoßen, ohne die Begrenzung zu zerstören. Auf diese Weise konnten sie die Chips lokal in 3er-Päckchen sammeln, was schließlich zu einem einzigen Haufen von Chips führte, der ordentlich in der Mitte angeordnet war. Dies ist eines von vielen Beispielen dafür, wie ein sehr einfacher Algorithmus zu komplexen oder scheinbar geordneten Strukturen führen kann. Aber ist es ein Beispiel für das Entstehen von Bewusstsein? Ich glaube nicht. Holland machte deutlich, dass Bewusstsein irgendwann entstehen könnte, aber keine der derzeitigen KI-Anwendungen wird dies schaffen.

David Chalmers hielt den zweiten Hauptvortrag über massive Sprachmodule, also sprachfähige KI, und ihr potenzielles Bewusstsein. Chalmers ist eine Art Ikone, sowohl für das Fachgebiet als auch für die Konferenz, da er einer der Mitorganisatoren der ersten Konferenzen in Tucson war. Er präsentierte die faszinierende Komplexität moderner Sprachmodule in der KI, wie z.B. ChatGPT. Sie wurden ursprünglich als Textverstehensmaschinen entwickelt, die auf Millionen von Texten mit einer Billion von

Parametern trainiert wurden. Sie wiesen interessante neue Phänomene auf, wie die Fähigkeit, zu sprechen und zu schreiben, zu programmieren und zu rechnen, praktisch zu argumentieren und zu erklären. Erweiterte Sprachmodule (LLM+) sind sogar in der Lage, wahrzunehmen und zu handeln. Sie enthalten enorme Datenmengen und führen Simulationen durch. Schließlich können Agentenmodelle solche LLM+ für die Planung nutzen.

Die Fragen, die sich dabei stellen, sind: Sind sie sicher, gerecht und verantwortungsvoll? Können sie einen moralischen Status haben? Können sie denken und verstehen? Sind sie Agenten, und haben sie somit ein Bewusstsein?

Nun, im Juni 2022 feuerte Google den Ingenieur Blake Lemoin, weil er öffentlich die Meinung vertrat, dass Googles KI *Lamdba* ein Bewusstsein habe. War oder ist Lambda also bewusst? Nach der Definition von David Chalmers würde dies Empfindungsvermögen und subjektive Erfahrung voraussetzen, d.h. sensorische und/oder affektive, kognitive und agierende Erfahrung. Selbstbewusstsein würde Bewusstsein von sich selbst bedeuten. Es ist wichtig, dies von Intelligenz zu unterscheiden. Intelligenz, selbst übermenschliche Intelligenz, ist nicht identisch mit Bewusstsein. Es könnte Intelligenz ohne Bewusstsein und Bewusstsein ohne Intelligenz geben.

Interessanterweise sagt *Lambda* selbst, dass es kein Bewusstsein habe, wenn man es fragt. (Aber vielleicht lügt es ja?) Dasselbe gilt für andere Systeme, wie z. B. ChatGPT, das verschiedene und manchmal widersprüchliche Antworten auf diese Frage gibt: Manchmal sagt es, es sei bewusst, und manchmal sagt es, es sei nicht bewusst. Nach Ansicht von Chalmers bestehen diese Systeme den Turing-Test nicht und liefern keine schlüssigen Beweise für Bewusstsein. Das liegt seiner Meinung nach an der fehlenden Verkörperung und Biologie. Außerdem haben sie keine Welt- und Selbstmodelle, keine rekurrente Verarbeitung und kein einheitliches Handlungsempfinden. Wenn Hameroff recht hat und das Bewusstsein die Biologie der Mikrotubuli voraussetzt, dann kann keine KI jemals ein Bewusstsein haben.

Da sie keine Sinne und keinen Körper haben, können sie auch kein agierendes Bewusstsein haben. Weltmodelle, wie bei Tieren und Menschen, scheinen eine notwendige Bedingung für Bewusstsein zu sein. LLMs sind jedoch nur Maschinen, die Vorhersagefehler minimieren und kein echtes Verständnis haben. Bender, Gebru und Kollegen, die von Chalmers zitiert werden, haben diese Systeme *„stochastische Papageien“* genannt.

Chalmers lieferte auch die Gegenargumente: LLMs weisen sehr interessante neue Merkmale auf, sodass sie tatsächlich über Weltmodelle verfügen könnten. Neuere Systeme könnten tatsächlich einen globalen Arbeitsraum haben. Aber es fehlt ihnen immer noch an rekurrentem Feedback und sie scheinen keine einheitlichen Handelnden zu sein. Da all

diese Voraussetzungen wie Verkörperung, Weltmodelle, globaler Arbeitsraum, rekurrente Eigenschaften und einheitliches Handeln zwar nicht verwirklicht sind, dies aber womöglich nur vorläufig ist, könnten diese Bedingungen in Zukunft erfüllt sein. Wenn jedoch die Biologie und damit die Mikrotubuli eine Bedingung sind, wird es nie eine bewusste KI geben. Wenn alle anderen Bedingungen erfüllt werden können und die Biologie keine Voraussetzung ist, könnte es sie früher oder später geben.

Das wäre also der Plan für die Zukunft: Eine gute Definition von Bewusstsein. Gegenwärtig gibt Chalmers der Frage ob LLMs Bewusstsein haben eine Chance von 10 %. Wenn alle Bedingungen erfüllt sind, könnte die Wahrscheinlichkeit eines KI-Bewusstseins in 10 Jahren auf 50 % steigen. Seine eigene Schätzung lautet: eine Wahrscheinlichkeit von mehr als 25 % bis 2032. Vor der Beantwortung dieser Frage wäre es jedoch notwendig, das Bewusstsein wirklich zu verstehen. Selbst wenn LLM+ kein menschliches Bewusstsein hätten, könnten sie ein tierisches Bewusstsein haben, und dann würde sich sofort die Frage nach ihrem moralischen Status stellen.

Die Plenarsitzungen am Freitag waren dem tierischen Bewusstsein gewidmet.

Frans de Waal hielt einen sehr bewegenden Vortrag mit vielen Videoclips von seinen Studien an wilden Affen, Bonobos und Schimpansen. Sie zeigten, dass die meisten sozialen Phänomene beim Menschen, die mit Bewusstsein in Verbindung gebracht werden, auch bei Tieren zu finden sind. Er berief sich dabei auf die Wellen-Regel: Was bei Tieren entdeckt wird, finden wir irgendwann auch beim Menschen, was die enge Verbindung zeigt. Der Gestaltpsychologe Köhler wies nach, dass Menschenaffen Probleme nicht nur durch Versuch und Irrtum, sondern durch Nachdenken lösen können. Affen zeigen Perspektivenübernahme, sie zeigen emotionale Ansteckung und Trostverhalten. Traumatisierte Bonobos haben ähnliche Schwierigkeiten mit Empathie wie traumatisierte Menschen. Präriewühlmäuse trösten nur Artgenossen und Geschwister. Die Selbsterkennung im Spiegel, die normalerweise als Zeichen des Selbstbewusstseins gewertet wird, ist nicht nur bei Menschenaffen, sondern auch bei Elefanten und Putzerfischen vorhanden, nicht aber bei Neuweltaffen. Diese können Spiegel zwar nutzen, um beispielsweise verstecktes Futter zu finden, aber nicht, um sich selbst zu erkennen. Der Unterschied zwischen Mensch und Tier ist also nach de Waals Ansicht nicht existent, es gibt nur einen Unterschied im Grad, aber nicht im Prinzip. Wie Menschen hören auch die Menschenaffen auf zu essen, wenn sie spüren, dass ihr Ende naht. Dieses flammende Plädoyer für Bewusstsein bei Tieren inspirierte zur Lektüre seines Buches „*Are we smart enough to know how smart animals are*„. [10]

Im gleichen Sinne zeigte Giorgio Vallortigara faszinierende Experimente mit Hühnern, die ein Verhalten an den Tag legen, das auf Zahlenverständnis schließen lässt. Küken können

mit jeder Art von Objekt aufgezogen werden und akzeptieren dieses Objekt dann als eine Art Mutterersatz. Sie können mit zwei, drei oder mehr Objekten aufgezogen werden. Wenn sie zum Beispiel mit zwei identischen kleinen Puppen aufgezogen werden und nacheinander eine, zwei oder mehr von ihnen hinter einem Sichtschirm versteckt werden, nachdem sie den Küken nacheinander präsentiert wurden, werden sie mit eindeutiger Präferenz den Schirm aufsuchen, hinter dem sich die beiden Objekte verbergen, und wenn sie mit drei von ihnen aufgezogen werden, werden sie sich zu dem Schirm begeben, hinter dem sich drei von ihnen verbergen. Interessant und fast etwas gruselig ist jedoch, dass diese Küken anscheinend zählen können. Wenn sie dazu gebracht werden, drei Objekte zu bevorzugen, und wenn sie sehen, wie ein Objekt hinter einem Schirm versteckt wird, dann vier hinter einem anderen Schirm, und dann zwei der vier wieder weggenommen und hinter den Schirm mit dem einen Objekt gestellt werden, werden sie zu dem Schirm mit den drei Objekten laufen, anscheinend indem sie die entsprechende Arithmetik gemacht haben, während sie beobachten, wie die Objekte hinter Schirme gestellt und wieder weggenommen werden. Verschiedene andere Experimente, die zeigen, dass Tiere ähnlich wie kleine Kinder einfache Logik und Berechnungen durchführen können, waren recht überzeugend. Sein Buch „*Born knowing*“ enthält diese Informationen.[11]

Schließlich zeigte *David Edelman* in seinem Vortrag *„Early origin of consciousness*“ Filme von seinen Experimenten mit Tintenfischen. Während der kambrischen Explosion vor 550 Millionen Jahren entstanden eine Vielzahl verschiedener Tiere und mit ihnen Augen, die in ihrer Verschaltung den unseren verblüffend ähnlich sind. Der Krake hat eine sehr ähnliche Augenverdrahtung: verschiedene Rezeptoren, die in einem Nerv vereinigt sind und in Ganglien oder Neuronen zusammenlaufen. Das Augenlicht entwickelte sich vor etwa 10 Millionen Jahren, um Mobilität und Tiefenerkennung zu ermöglichen und war somit die Voraussetzung für die Beziehung zwischen Räuber und Beute. Edelman geht davon aus, dass die Verbindung solcher Sinneseindrücke und das Festhalten einer solchen Wahrnehmung im Gedächtnis die Grundlage für das primäre Bewusstsein ist. Dies trifft auf viele Tiere zu. Bei höheren Wirbeltieren kommen thalamo-kortikale Schleifen hinzu, die wahrscheinlich das Substrat des Bewusstseins sind. Erforderlich sind schnelle Sinneskanäle, ihre Integration in Wahrnehmungseinheiten, das Vorhalten solcher Wahrnehmungen im Arbeitsgedächtnis und eine Schaltung, die die Wahrnehmung mit dem Gedächtnis verbindet. Das trifft sicherlich auf den Kraken zu, der über ein sehr komplexes visuelles System verfügt, das dem unseren ähnlich ist. Kraken können beobachten und durch Beobachtung lernen, ohne durch Versuch und Irrtum auszuprobieren. Sie können lernen, sich in Labyrinthen anhand von Orientierungspunkten zurechtzufinden, nicht durch Versuch und Irrtum. Sie nutzen also ihr Gedächtnis. Ihre Fernsicht ermöglicht es ihnen, die Zeit einzuschätzen, und sie können Vorhersagen treffen und beispielsweise die Bewegungen ihrer Beute beobachten, was bedeutet, dass sie über Hemmschaltungen verfügen. Denn

wenn ein Raubtier wie der Krake die Zeit richtig einschätzen kann, kann es auf den besten Moment warten, und dieses Warten setzt Hemmschaltungen voraus.

Es gab zwei Postersitzungen. Einige von ihnen waren sehr interessant, und der Vorteil war sicherlich, dass für sie feste Zeit eingeräumt war, sodass viele Besucher vorbeischauten. In der zweiten Sitzung wurde unser Galileo-Poster ausgestellt, mit dem ich für meinen Galileo-Bericht warb, und Stuart Hameroff, ein Mitglied der Galileo-Kommission, zeigte seine Unterstützung durch einen Besuch und zog eine neugierige Menge mit sich.

Ich hatte einen Vortrag in einer der parallelen Sitzungen über Philosophie: *„Lob des Todes – Eine philosophische Kritik des Transhumanismus„*. Darüber werde ich gesondert schreiben. Ich habe das Argument dargelegt, dass das transhumanistische Programm zur Abschaffung des Todes philosophisch gesehen selbstzerstörerisch und widersprüchlich ist. Es ist ein Programm, das von Natur aus neokolonialistisch ist, weil sich nur reiche Menschen die notwendigen Behandlungen leisten können. Es ist ein Programm, das die jetzige Generation gegenüber künftigen Generationen bevorzugt und damit Innovation und Wachstum langsam erstickt. Es wird entweder unweigerlich zu einer Überbevölkerung führen, da die älteren Generationen weiterleben und die jüngeren geboren werden, oder es wird zu einer strengen Bevölkerungskontrolle führen und ist somit ein faschistisches Programm. Und moralisch vernachlässigt es die Tatsache, dass erst die Endlichkeit unseres Lebens unsere Urteile und Entscheidungen wertvoll und moralisch wichtig macht.

Ich habe in diesem Sinne den letzten Tag ausgelassen und stattdessen den Ätna bestiegen (Abb. 2). Da es am Sonntag zuvor einen kleinen Ausbruch gegeben hatte, der den Flugverkehr in Catania zum Erliegen brachte, war der oberste Gipfel tabu, und die Wanderführer achteten darauf, dass niemand die 2.900 m Grenze überschritt. Aber der schwefelhaltige Geruch hätte wohl höhere Aufstiege verhindert. Trotzdem war es beeindruckend, diesen höchsten der aktiven Vulkane Europas besucht zu haben und die warme Erde zu berühren, die den Schnee geschmolzen hatte.



Abbildung 2 – Der Ätna, gesehen von einem Standort auf etwa 2.900 m Höhe vom Südaufstieg aus

Insgesamt könnte man meinen Eindruck von der Taormina-Konferenz folgendermaßen zusammenfassen: Wir kennen jetzt viele Details über das Bewusstsein, aber das Bewusstsein selbst war nicht Teil davon. Vielleicht ist das Problem tatsächlich, wie Stuart Hameroff vorschlug, dass die Neurowissenschaft ein neues Paradigma braucht, vom Neuron zu den Mikrotubuli. Aber vielleicht sollten wir sogar noch weiter gehen: Wir müssen anerkennen, dass das Bewusstsein als eine völlig neue und ontologisch andere Kategorie nicht aus etwas anderem als sich selbst entstehen kann. Und vielleicht brauchen wir ein Modell, in dem das Bewusstsein im Sinne von Bewusstheit und Selbstbewusstsein tatsächlich eine Eigenschaft eines neuronalen Systems sein könnte, die mit ihm vergeht. Aber anders als dieses könnte es noch eine andere Art von Bewusstsein geben, Bewusstsein großgeschrieben sozusagen, Bewusstsein 2. Es repräsentiert, was in älteren Ontologien als Geist oder Seele bezeichnet wurde. Ich habe das bei früheren Gelegenheiten [12, 13] näher erläutert. Davon war auf dieser Konferenz kaum etwas zu hören.

Quellen und Literatur

1. Hasler F. Neuromythologie: eine Streitschrift gegen die Deutungsmacht der Hirnforschung. 5., unveränd. Aufl ed. Bielefeld: transcript; 2015.
2. Fröhlich H, Kremer F. Coherent Excitation in Biological Systems. Berlin: Springer; 1983.
3. Fröhlich H. Long range coherence and the action of enzymes. Nature. 1970;228:1093.
4. Hameroff S, Penrose R. Consciousness i n the universe: A review of the 'Orch OR' theory. Physics in Life Review. 2014;11:39-78.
5. Hameroff S, Penrose R. Conscious events as orchestrated space-time selections. NeuroQuantology. 2003;1:10-35.
6. Seth AK. Being You: A New Theory of Consciousness. New York: Dutton; 2021.
7. Humphrey N. Sentience: The Invention of Consciousness. Oxford: Oxford University Press; 2022.
8. Gallagher S. How the body shapes the mind. Oxford: Clarendon; 2005.
9. Zahavi D. Husserls Phänomenologie. Tübingen: Mohr Siebeck; 2009.
10. Waal FBMd. Are we smart enough to know how smart animals are? New York, London: W.W. Norton; 2016.
11. Vallortigara G. Born Knowing: Imprinting and the Origins of Knowledge. Cambridge, MA: MIT Press; 2021.
12. Walach H. The higher self – spark of the soul, summit of the mind. History of an important concept of transpersonal psychology in the West. International Journal of Transpersonal Studies. 2005;24:16-28. Link
13. Walach H. Mind – Body – Spirituality. Mind and Matter. 2007;5:215-40. Link