

KREATIVITÄT UND KÜNSTLICHE INTELLIGENZ

Einige Bemerkungen zu einer Kritik algorithmischer Rationalität

Forschungsprogramme zur Künstlichen Intelligenz feiern scheinbar spektakuläre Erfolge, nicht nur auf den Gebieten selbstlernender Maschinen, in Bezug auf diagnostische Expertensysteme oder mathematische Beweise, sondern auch mit Blick auf die kreative Modellierung von visuellen Prozessen, die musikalische Komposition oder das Storytelling, um nur einige Beispiele zu nennen. Jüngst erzielte ein von einem Computerprogramm gemaltes fiktionales Porträt (*Portrait of Edmond De Belamy*) bei Christie's eine Summe von 432.500 US-Dollar.¹ Seit Längerem experimentiert der amerikanische Komponist und Informatiker David Cope mit dem Kompositionsprogramm Emily Howell, das er als virtuelle Komponistin mit Biografie und Lebensdauer konzipiert hat.² Schon mehr als drei Jahrzehnte sind Morphingprogramme in Entwurfsprozessen der Architektur im Einsatz, und spätestens seit den 1970er Jahren wird mit automatisierten Dichtungen aus dem Computer experimentiert, die von menschlichen Leser_innen nicht mehr als solche erkannt werden. Gilt Kreativität und insbesondere «Kunst-Machen» als eine der letzten Domänen der Unterscheidung zwischen Mensch und Computer, so scheint nunmehr auch diese Bastion gefallen zu sein. Augenscheinlich sind wir mit der Hypothese konfrontiert, die Welt, den Menschen, seine Kultur sowie Kommunikation und Kunst unter das einheitliche Schema ihrer Mathematisierung zu stellen.

Dass sich alle Phänomene zuletzt einer Rationalisierung fügen und sich unter das einheitliche Schema der Mathematik pressen lassen, bildet eine der Grundannahmen sowohl der Computerwissenschaft als auch überhaupt der Moderne.³ Der Lackmustest der daraus hervorgehenden Verwandlung des Denkens und der «ästhetischen Erfahrung» beruht dabei auf den Chancen einer angemessenen Simulation von Sprache und Imagination sowie der Praxis dessen, was wir unter dem chronisch unscharfen Begriff der «Kreativität» verstehen. Kunst und künstlerische Produktion fungieren dafür als ausgezeichnetes Paradigma. Nicht die *creatio* an sich steht zur Debatte – «Neues» kann auch

¹ Christie's: Is artificial intelligence set to become art's next medium?, dort datiert 12.12.2018, www.christies.com/features/A-collaboration-between-two-artists-one-human-one-a-machine-9332-1.aspx, gesehen am 22.05.2019.

² Vgl. Website von David Cope: Emily Howell, www.artsites.ucsc.edu/faculty/cope/emily-howell.htm, gesehen am 22.05.2019.

³ Vgl. den Beitrag Paul Browns in Harald Cohen, Frieder Nake, David C. Brown u. a. im Kapitel Evaluation of Creative Aesthetics, in: Jon McCormack, Mark d'Inverno (Hg.): *Computers and Creativity*, Berlin, Heidelberg 2012, 95–111, hier 105: «I implicitly believe that everything is rationally explainable».

durch Zufall entstehen –, sondern die Frage, inwieweit wir Produkte aus dem Computer als ästhetische Gestaltungen akzeptieren oder, wie offensichtlich bei *Edmond Belamy*, als ›Kunst‹ anschauen wollen. Daraus folgt nicht, dass *Edmond Belamy* Kunst ›ist‹, noch dass seine Programmierer Hugo Caselles-Dupré, Pierre Fautrel und Gauthier Vernier oder Teile der Kunstöffentlichkeit es bereits als solche betrachten, vielmehr demonstriert das Bild, zu welchen ›Schöpfungen‹ Computer in der Lage sind und wie schwer es fällt, diese angemessen von der Originalität humaner Akteure zu unterscheiden.

Im Folgenden wird die Frage nach einer möglichen Verunsicherung unserer Distinktionen dahingehend gestellt, dass die mathematischen Grundlagen informatischer Modelle und ihrer Beziehung zu Kunst und Kreativität reflektiert werden, um sie gleichzeitig einer philosophischen Geltungskritik zu unterziehen. Diese Geltungskritik zielt darauf ab, die zugrundeliegende Algorithmik im Hinblick auf ihre Grenzen auszuloten, die zugleich die Grenzen der ›Turing-Berechenbarkeit‹ beschreiben. Sie gelten im Prinzip für alle Programme, gleich ob wir es mit einer starken oder schwachen KI oder mit der von Frieder Nake sogenannten ›heroischen‹ künstlichen Kreativität oder einer ›kollaborativen‹ zu tun haben.⁴ Das schließt im Besonderen die Frage ein, inwieweit kreative oder künstlerische Prozesse und ihre Beurteilung oder Evaluation überhaupt im Register einer Alternative zwischen ›Berechenbarkeit‹ und ›Nichtberechenbarkeit‹ betrachtet werden können.

Philosophische Kritiken algorithmischer Rationalisierungsprogramme – eine kursorische Lektüre

Bis dato liegen eine Reihe philosophischer Geltungsanalysen algorithmischer Rationalisierungsprogramme vor, vor allem in Ansehung von Computermodellen des Geistes. Sie sind hier relevant, weil Kreativität als eines der basalen humanen Vermögen angesehen wird. Bekanntlich hatten Warren McCulloch und Walter Pitts in *Logical Calculus of the Ideas Immanent in Nervous Activity* von 1943 bereits eine Homologie zwischen logischen Strukturen und der synaptischen Aktivität von Nervenzellen postuliert, was John von Neumann, trotz aller Skepsis, in seinem posthum veröffentlichten Text *Die Rechenmaschine und das Gehirn* von 1956 zu der Bemerkung veranlasste: ›Man macht zunächst die Beobachtung, dass das Nervensystem in erster Linie digital arbeitet‹. Diese Beobachtung gelte deshalb, weil sich Nervenzellen entweder in einem Erregungszustand oder keinem Erregungszustand befänden: ›Daher ist die ursprüngliche Behauptung, dass das Nervensystem prima facie digitaler Natur ist, gerechtfertigt‹.⁵ Dass Gehirne Computern gleichen – eine der Grundannahmen vieler Computerwissenschaftler_innen bis heute⁶ –, ergibt sich dann aus der Äquivalenz zweier binärer Ordnungen, der digital funktionierenden Computersysteme und des Gehirns als Vernetzung entweder ›feuernder‹ oder ›nicht feuernder‹ Synapsen – ein Ansatz, der jedoch philosophisch als ›zu leicht‹ empfunden werden

⁴ Vgl. den Beitrag Frieder Nakes in ebd., 106. Ferner Jon McCormack, Mark d'Inverno: Heroic versus Collaborative AI for the Arts, in: *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence*, Palo Alto 2015, 2438–2444, online unter research.gold.ac.uk/id/eprint/12795, gesehen am 22.5.2019.

⁵ John von Neumann: *Die Rechenmaschine und das Gehirn*, München 1991, 44, 48.

⁶ Siehe etwa Jürgen Schmidhuber: A Formal Theorie of Creativity to Model the Creation of Art, in: McCormack u. a. (Hg.): *Computers and Creativity*, 323–337.

muss, weil er die neuronale Aktivität allein auf Signalübertragung beschränkt und weder Gliazellen⁷ noch die Plastizität des Gehirns angemessen berücksichtigt. Hatten Jerry Fodor⁸ und Hilary Putnam⁹ die Position McCullochs und von Neumanns zudem linguistisch weiterentwickelt, wendete sich vor allem Putnam später davon ab: In *Reason, Truth, and History* demonstrierte er anhand des aus der Science-Fiction-Literatur bekannten «Gehirne im Tank»-Modells, dass ein isoliertes Gehirn seine eigenen referenziellen Beziehungen nicht zu beurteilen und damit auch keine konsistenten Kriterien für Wahrheit, Realität und Bedeutung aufzustellen vermag.¹⁰ Weil sie weder mit einer Geschichte noch mit der Außenwelt verknüpft sind, sei es für Retortengehirne unmöglich zu entscheiden, wo sie sich befänden (in einem Tank, einem Körper etc.), sodass sie auch nicht wissen können, ob ihre Ideen über sich selbst wahr oder falsch seien. Das Gleiche gilt offenbar für Computerprogramme, seien sie KI-gestützt oder nicht, weshalb die Frage des Bewusstseins von Maschinen nicht systematisch gelöst werden kann. Ähnliches wäre von Alan Turings <Test> zu sagen, der notorisch Denken und Rechnen gleichsetzt, was Martin Heidegger schon in den frühen 1950er Jahren als das eigentliche Verhängnis der Moderne kritisiert und in eine direkte Linie mit seiner Technik- und Wissenschaftskritik gestellt hat, deren entscheidender Satz in der bewusst zugespitzten Polemik mündet: «Die Wissenschaft denkt nicht».¹¹ Daran anschließend hatte Hubert L. Dreyfus¹² gegen Turing und mit Maurice Merleau-Ponty darauf hingewiesen, dass wir in erster Linie leiblich denken und folglich Bewusstsein eine Funktion des gesamten Körpers darstellt, nicht eines einzelnen Gehirns – eine Kritik, wie sie später ebenfalls von Alva Noë und Markus Gabriel wiederholt werden sollte.¹³ Mit Blick auf Intentionalität und Verstehen hat darüber hinaus John Searle mit seinem Gedankenexperiment vom «Chinesischen Zimmer» einen «Anti-Turing-Test» aufgestellt, der vom Bild einer geschlossenen Black Box ausgeht, worin eine Person, ausgestattet mit einem Regelwerk zur formalen Prozessierung von Schriftmarken, auf einen Input aus chinesischen Schriftzeichen mit einem offenbar richtigen Output antwortet.¹⁴ Aus der Korrektheit der Resultate, so Searle, kann jedoch nicht geschlossen werden, dass die einsitzende Person die erteilten Befehle versteht: Semantik bildet keine Funktion einer Syntax. Ebenso hatte Robert Brandom in seinem Buch *Between Saying and Doing* darauf bestanden, dass bedeutungsvolles Handeln ein Verständnis praktischer Konsequenzen einschließt, welche nicht wiederum aus formal-algorithmischen Inferenzen folgen können. Vielmehr erwiesen sich mögliche Folgen stets als kontextuell offen, was Kriterien der Unterscheidung zwischen relevanten und irrelevanten Implikationen voraussetzt, die ihrerseits in Semantiken fundiert sind, welche nicht ausschließlich intrinsischen Logiken gehorchen.¹⁵

Sämtlichen dieser Kritiken ist zu eigen, dass sie die Geltungsgrenzen von Maschinenintelligenzen aus der prinzipiellen Inkompatibilität zwischen Syntax und Semantik oder genauer: aus der Nichtableitbarkeit bestimmter, für die menschlichen Sprachen charakteristischer Interpretationen und deren

⁷ Vgl. Jonah Lehrer, Andrew Koob: *The Root of Thought. What do Glial Cells do?* Interview mit Andrew Koob in: *Scientific American*, dort datiert 27.10.2009, www.scientificamerican.com/article.cfm?id=the-root-of-thought-what, gesehen 22.5.2019.

⁸ Vgl. Jerry A. Fodor: *The Language of Thought*, Cambridge 1975.

⁹ Vgl. Hilary Putnam: *Geist und Maschine*, in: Walther C. Zimmerli, Stefan Wolf (Hg.): *Künstliche Intelligenz*, Stuttgart 1994, 146–183.

¹⁰ Vgl. Hilary Putnam: *Vernunft, Wahrheit und Geschichte*, Frankfurt / M. 1990, 29–40.

¹¹ Martin Heidegger: *Was heißt denken?*, Tübingen 1984, 4.

¹² Vgl. Hubert L. Dreyfus: *Was Computer nicht können. Die Grenzen künstlicher Intelligenz*, Frankfurt / M. 1989, 183 ff.

¹³ Vgl. Alva Noë: *Du bist nicht dein Gehirn. Eine radikale Philosophie des Bewusstseins*, 3. Aufl., München 2011; Markus Gabriel: *Der Sinn des Denkens*, Berlin 2018, 91 ff.

¹⁴ Vgl. John R. Searle: *Geist, Gehirn, Programm*, in: Zimmerli u. a. (Hg.): *Künstliche Intelligenz*, 232–265.

¹⁵ Vgl. Robert Brandom: *Between Saying and Doing*, Oxford 2008, 77–82, 117–140.

Pragmatiken herleiten. Sie variieren damit die Heidegger'sche Intuition einer grundsätzlichen Differenz von Denken (als sinnbasiert) und Rechnen (als formal regelgeleitet). Dem entsprechen ebenfalls jene philosophischen Kritiken, die im Speziellen die ethischen Konsequenzen von Computerentscheidungen, sei es auf dem Gebiet von KI oder der Robotik, und Ähnliches diskutieren, insbesondere dort, wo es um Verhaltenskontrolle, Gesichtserkennung, medizinische Diagnosen, Gerichtsentscheidungen oder automatisch geführte Kriege geht.¹⁶ So richtig diese Einwendungen sind, so wenig gehen sie allerdings <aufs Ganze> der impliziten Prämissen der angewendeten Programme, die weniger in der Informatik und ihrer funktionalistischen Lösungspragmatiken zu suchen sind, als vielmehr in den Fundamenten der Mathematik. Dieser Hinweis, vor allem an die Adresse der Kultur- und Medienwissenschaften, ist deshalb wichtig, weil sich die Referenzen dort noch immer an Friedrich Kittlers Diktum orientieren, digital-algorithmische Systeme aus dem spezifischen Beitrag der Informatik zu lesen. Kittler hatte sogar polemisch von einem Analphabetismus bei denen gesprochen, die diese aktuellen <Sprachen> nicht beherrschten, und damit die Debatte um digitale Systeme auf der Ebene informatischer Verfügungsphantasmen sistiert. Die am Paradigma von Machtanalysen eingeübte Perspektive berührt jedoch die elementare philosophische Differenz zwischen Genesis und Geltung nicht. Denn Geltungskritik im Bereich der Informatik zielt auf jene Grundlagen, wie sie die Mathematik <formaler Systeme> legt, die ihrerseits eine Frucht der <Grundlagenkrise> zu Anfang des 20. Jahrhunderts darstellt, zu deren tiefsten Resultaten u. a. die beiden Gödel'schen Unvollständigkeitssätze sowie die Begründung der rekursiven Mathematik und das Modell der Turingmaschine gehörten. Alle *computation* sowie Algorithmen sind, auch wenn sie im Endlichen operieren, nichts anderes als komplexe <Mathematikmaschinen>, deren *logische* Grenzen in gewissen Unentscheidbarkeiten liegen. Ihr Ausweis eröffnet in einem kantischen Sinne eine <Kritik algorithmischer Rationalität>, die im Wesentlichen eine (apriorische) <Kritik> ihrer formalen mathematischen Fundamente und nicht ihrer (aposteriorischen) empirischen Programme bildet.

Kritik <algorithmischer Rationalität>

Dabei sei unter <algorithmischer Rationalität> jenes komplexe Ensemble formaler Operationen verstanden, wie es sich von Anfang an in einem mathematischen Universum situiert. Es trifft zu einem gewissen Grade schon für die <Digitalisierung> selbst zu, soweit sie sich in Diskretisierungen, d. h. der Zerlegung der Welt in disjunkte Einheiten, erfüllt.¹⁷ Es handelt sich also um ein Intervalldenken, das aufgrund seiner syntaktischen Strukturierung eine Codierung derjenigen Aufgabengebiete oder Fragestellungen gestattet, die mittels Regeln transformierbar gemacht werden können. Für ihre algorithmische Aufbereitung kommen zwei weitere Elemente hinzu: *erstens* Daten als Werte derjenigen Funktionen, die als <rekursive Funktionen> diese Daten einer <Berechenbarkeit>

¹⁶ Vgl. exemplarisch die Beiträge in: Patrick Lin, Ryan Jenkins, Keith Abney (Hg.): *Robot Ethics 2.0. From Autonomus Cars to Artificial Intelligence*, Oxford 2017.

¹⁷ Vgl. Nelson Goodman: *Sprachen der Kunst*, Frankfurt / M. 1995, 154–157.

zuführen, sowie *zweitens* ihre Programmierung, die deren Programme in <universelle Turingmaschinen> verwandeln, die wiederum als mathematische Formalisierungen der Algorithmik überhaupt fungieren. Doch sei zugleich angefügt, dass Mathematik und Mathematisierung keineswegs im einheitlichen Terrain von <Digitalisierung> und <Algorithmik> aufgehen. Letztere sind Teilbereiche der Ersteren, nicht umgekehrt, vielmehr beschreiben Digitalisierung und Algorithmik mathematische Provinzen, in denen allein die Kategorien der Entscheidbarkeit und Berechenbarkeit dominieren.¹⁸ Wir bekommen es folglich mit einer *Arithmetisierung von Problemen* zu tun, die zuvor schon formalisiert worden sein müssen, was sich u. a. darin ausdrückt, dass nur solche Fragen behandelt werden, die sich *numerisch*, d. h. auf der Basis algebraischer Gleichungen oder durch Approximation und Ähnliches lösen lassen. Hingegen bildet die Mathematik im Ganzen eine *Strukturwissenschaft*, die ebenso die Theorie von Räumen und Relationen wie von «nicht-rekursiven» Problemen umfasst,¹⁹ so dass sich bereits das Mathematische nicht ohne Weiteres einer durchgängigen *computation* fügt. Wo hingegen Informatik und Computerisierung als Hilfswissenschaften mit dem Mathematischen selbst gleichgesetzt werden, sind wir mit einem restringierten Code konfrontiert, wie er im Rahmen der Moderne eine hegemoniale Gestalt angenommen hat. Er impliziert, das Reale ausschließlich im Raster des Digitalismus und seiner algorithmischen Netze zu verhandeln und damit ganz dem Horizont von Berechenbarkeit und Entscheidbarkeit zu subordinieren, deren universaler mathematischer Repräsentant bekanntlich die Turingmaschine bildet.

Es scheint, als sei dieses Modell grenzenlos anwendbar, jedenfalls solange, wie sich eine vollständige <Diskretisierung> oder <Dataierung> der Wirklichkeit sowie eine schrittweise Zerlegung von Phänomenen und Prozessen in operative Einheiten vornehmen lassen, d. h., solange man einen *bestimmten Begriff von Welt* unterstellt, nämlich einen, der mit dem Semiotischen oder Skripturalen zusammenfällt. Dann schafft die unbegrenzte Applikation mathematischer Algorithmen ein kompaktes endliches und abzählbares Zeichenuniversum, ohne auf ein <Außen>, eine prinzipielle Nichtberechenbarkeit im Sinne einer <Nicht-rekursivität> oder eines <nicht mathematischen Anderen> zu stoßen. Woher ein solches beziehen?

Gödel'sche Unentscheidbarkeit und der formale Ort einer Nichtformalisierbarkeit

Eine erste Annäherung bietet die Diskussion um die sogenannte Metamathematik der 1930er und 40er Jahre und ihre Bedeutung für die Anfänge der Computerwissenschaft, die, anders als heute, vor allem einer engen Kollaboration zwischen Mathematikern und Ingenieuren unter Führung Ersterer entsprang. Sie bedingt eine mathematische Analyse des Mathematischen selbst, für die insbesondere die Gödel'schen Unvollständigkeitstheoreme und das Turing'sche

¹⁸ Vgl. Hans Hermes: *Aufzählbarkeit, Entscheidbarkeit, Berechenbarkeit*, 2. Aufl., Heidelberg, New York 1971, 1–32, 33–59, 95–114.

¹⁹ Vgl. Roger Penrose: *Computerdenken. Die Debatte um künstliche Intelligenz, Bewußtsein und die Gesetze der Physik*, Heidelberg 1991, 125–139.

Halteproblem entstehen, die sich im Übrigen als äquivalent erwiesen haben. Besonders die Gödeltheoreme besagen, dass sich in jedem formalen System, das mächtig genug ist, die Arithmetik zu umfassen, wahre Aussagen formulieren lassen, die sich weder beweisen noch widerlegen lassen.²⁰ Im selben Sinne hatte Turing bewiesen, dass es keine definitive Methode und damit auch keine vollkommene Turingmaschine geben kann, die die berechenbaren von nicht berechenbaren Funktionen zu scheiden vermag,²¹ wobei zu berücksichtigen ist, dass zwischen den Begriffen «berechenbare Funktion», «Turingmaschine» und «Algorithmus» sowie zwischen «Entscheidbarkeit» und «Beweisbarkeit» mathematisch ein enger Zusammenhang besteht.²² Schemenhaft zeichnet sich damit eine *erste innere Grenze des Mathematischen* im Sinne der Unmöglichkeit ihrer eigenen Mathematisierbarkeit ab, denn deutlich wird, dass die Mathematik, als ein loses Ensemble formaler Theorien, sich nicht wiederum vollständig ins Schema formaler Systeme bzw. abstrakter, d.h. auf reiner Syntax basierender, Maschinen pressen lässt. Anders gewendet: Das Computierbare und das Mathematische bleiben disparat, woraus John Randolph Lucas in «Minds, Machines and Gödel» sowie später Roger Penrose in seinen beiden Büchern *The Emperor's New Mind* und *Shadows of the Mind*²³ eine prinzipielle Kritik aller möglichen *artificial intelligence*-Programme abgeleitet haben: «A machine cannot be a complete and adequate model of the mind».²⁴

Die philosophischen Konsequenzen der Gödelsätze für das Problem Künstlicher Intelligenzen wie auch der Status der Lucas'schen und Penrose'schen Kritiken bleiben umstritten,²⁵ zumal sie selbst nicht als Beweise, sondern lediglich als Indikatoren dafür angesehen werden können, dass menschliches Denken auf etwas anderem als auf maschinellen Operationen beruht. Dabei mochte Gödel selbst jegliche Folgerung aus seinen Ergebnissen allein auf das beschränkt wissen, was später als «Gödel'sche Disjunktion» bekannt geworden ist,²⁶ denn entweder kann das humane mathematische Denken durch keinen Algorithmus umfasst werden und ist also von jeder möglichen Maschine zu unterscheiden, oder es gibt tatsächlich absolut unentscheidbare Probleme, die durch keine Methode, weder durch menschliche Rechnungen noch durch algorithmische Verfahren, gelöst werden können.²⁷ Denn wenn Denken algorithmisch simulierbar wäre, existiert nach den Unvollständigkeitstheoremen eine nicht zu überspringende Grenze, sodass es für immer unlösbare mathematische Probleme sowohl für humane Mathematiker_innen als auch Computersysteme geben muss, oder aber zwischen dem Menschlich-Mathematischen, d.h. derjenigen Mathematik, die das menschliche Denken einzusehen vermag, und seiner Mediatisierung durch *computations* besteht eine unüberbrückbare Kluft.

Leider ist für eine eingehendere Analyse dieses Problems an dieser Stelle kein Platz, dennoch mündet offenbar die Anwendung der Unvollständigkeitssätze auf *artificial intelligence*-Programme in eine weitere Unentscheidbarkeit, die das gängige Postulat der Computerwissenschaft, formale Maschinen und Gehirne seien gleich, unterhöhlt. Denn dass eine definitive Demarkation für

²⁰ Vgl. Ernest Nagel, James R. Newman: *Der Gödelsche Beweis*, München 2007.

²¹ Vgl. Alan Turing: Über berechenbare Zahlen mit einer Anwendung auf das Entscheidungsproblem, in: ders.: *Intelligence Service*, hg. v. Bernhard Dotzler, Friedrich A. Kittler, Berlin 1987, 17–60, hier 36 ff.

²² Vgl. George Dyson: *Turings Kathedrale. Die Ursprünge des digitalen Zeitalters*, Berlin 2014, 360.

²³ Vgl. Penrose: *Computerdenken*; ders.: *Schatten des Geistes*, Heidelberg, Berlin, Oxford 1995.

²⁴ John R. Lucas: *Minds, Machines, and Gödel*, in: *Philosophy*, Vol. 36, 1961, 112–137, hier 113.

²⁵ Vgl. z. B. Dieter Wandschneider: Die Gödeltheoreme und das Problem Künstlicher Intelligenz, in: *Ethik & Sozialwissenschaft*, Bd. 1, 1990, 107–114. Auch: Jason Megill: The Lucas-Penrose Argument about Gödel's Theorem, in: *Internet Encyclopedia of Philosophy*, www.iep.utm.edu/lp-arguel/, gesehen am 22.5.2019.

²⁶ Vgl. Leon Horsten, Philip Welch: *Gödel's Disjunction. The Scope and Limits of Mathematical Knowledge*, Oxford 2016.

²⁷ Vgl. Kurt Gödel: Some Basic Theorems on the Foundations of Mathematics and their Implications, in: ders.: *Collected Works*, hg. v. Solomon Feferman, John W. Dawson Jr., Stephen C. Kleene u. a., Bd. 3, Oxford 1995, 304–323.

Künstliche Intelligenzen besteht, kann logisch so wenig geschlossen werden wie das Umgekehrte, sodass es sich lediglich um eine Hypothese handelt, die, analog zur Church'schen These, die eine Identität zwischen humanen Rechnungen und ihren algorithmischen Formalisierungen postuliert, unbewiesen bleiben muss. Doch während die Church'sche These wahrscheinlich ist, erscheint die *artificial intelligence*-Hypothese hochgradig unwahrscheinlich, weil konstitutive Begriffe wie «Sinn», «Verkörperung» oder auch die soziale Manifestation von «Wissen» im Rahmen dieser erst gar nicht rekonstruierbar erscheinen.

Weitere Folgerungen und die Unverzichtbarkeit nicht algorithmischer Kreativität

Waren diese und ähnliche Fragen der frühen Phase der Computerisierung noch geläufig, droht ihr Bewusstsein heute unter dem enthusiastischen Eindruck der Erfolge von *deep learning* oder *deep neural networks* zunehmend zu schwinden. Denn dem neuerdings in den Computerwissenschaften angeschlagenen «hohen Ton» einer Theologisierung – Ray Kurzweils *The Singularity Is Near*,²⁸ in direkter Anspielung auf die Johannes-Apokalypse – liegt eine systematische Verkennung dieser und ähnlicher mathematischer Bedingtheiten zugrunde. So lassen sich in unmittelbarer Folge der Unentscheidbarkeitstheoreme noch *weitere Geltungsgrenzen* markieren, denn soweit Turing sein Modell einer Turingmaschine als allgemeine Theorie des Algorithmus konzipierte, ergibt sich ein nicht zu schließender Abstand zwischen Berechenbarkeit und Nichtberechenbarkeit, der *als Differenz* nicht selbst wieder einer Algorithmisierung zugeführt werden kann. «Es ist möglich, eine einzige Maschine zu erfinden, die dazu verwendet werden kann, jede berechenbare Folge zu errechnen», heißt es deshalb in Turgings grundlegendem Aufsatz «On Computable Numbers» von 1936,²⁹ und doch ist die Anzahl möglicher Maschinen aufgrund ihres endlichen Charakters «abzählbar», wohingegen die Menge aller Funktionen «überabzählbar» ist. Es gibt folglich «nicht berechenbare Funktionen», und zwar «wesentlich mehr», woraus resultiert, dass das allgemeine Berechenbarkeitsproblem, die Frage nämlich, was genau berechenbar ist und was nicht, algorithmisch nicht zu lösen ist.

Mit der Begründung der Theorie der Turingmaschinen taucht somit von Anfang an eine *zweite* nicht auflösbare Barriere der *computation* auf. Zudem hatte Turing in seiner Dissertation «Systems of Logic Based on Ordinals» versucht, die Konsequenzen aus der Gödel'schen Unvollständigkeit mittels einer Abfolge von «Sprachen» mit wachsender «Vollständigkeit» abzumildern, indem ein erwiesenermaßen vollständiges System L zu einem ebenfalls vollständigen System L', dieses wiederum zu L'' etc. so erweitert wird, dass eine Art Stufenhierarchie entsteht. Doch zeigt sich hier dieselbe Misslichkeit, dass die Übergänge von L zu L' zu L'' etc. nicht wiederum maschinell vollzogen werden können, vielmehr sprach Turing, so wörtlich, von einem «Orakel».³⁰ Anders formuliert: Für den Übergang zwischen formalen Sprachen gibt es kein

²⁸ Vgl. Ray Kurzweil: *The Singularity Is Near*, New York 2005.

²⁹ Turing: Über berechenbare Zahlen, 31.

³⁰ Alan Turing: Systems of Logic based on Ordinals, in: *Proceedings of the London Mathematical Society*, 2. Folge, Vol. 45, Nr. 3, 1939, 161–228, hier 161 ff., 172 ff.

allgemeines Gesetz, sondern einzig eine Intuition oder kreative Erfindung, die zugleich deutlich macht, dass es nicht *einen einzigen* Übergang gibt, sondern unbestimmt viele.

Wir sind folglich mit einem System deterministischer Maschinen konfrontiert, zwischen denen ein nicht deterministischer <Sprung> besteht. Er kann nicht anders als ästhetisch überbrückt werden, wenn wir unter dem Ästhetischen jene grundlosen Inventionen der *creatio* verstehen, die Kant der <freien> bzw. <reflektierenden> Urteilskraft zuschlug.³¹ Dann erweist sich die Mathematik selbst als eine schöpferische Tätigkeit, die ihren Grund sowohl im Mechanischen als auch in gewissen nicht logischen Spontaneitäten besitzt, denn die <Erfindung> des Mathematischen ist nicht ihrerseits das Resultat einer mathematischen Deduktion, vielmehr fußt sie ebenso wohl auf formalen Konstruktionen wie auf einer Serie von schöpferischen Zugängen, die *als* solche wiederum keinem mathematischen Kalkül gehorchen, sondern buchstäblich übergangslosen Übergängen. Sie wurzeln, anders gesagt, in einer <Poetik der Findungen>, die mithin eine *dritte* Grenze markiert und das Problem der möglichen Differenz zwischen Kreativität und Algorithmik aufwirft. John von Neumann hatte denselben Umstand so ausgedrückt: «[M]an kann ein Instrument bauen, das alles kann, was machbar ist, aber man kann kein Instrument bauen, das einem sagt, ob es machbar ist»,³² denn die Frage, ob etwas in einem Typus machbar ist, gehört einem anderen Typus an.

Naiver Kunst- und Kreativitätsbegriff

Gewöhnlich werden, wo sich die Frage nach der Kreativität stellt, und zwar sowohl in ihrem Verhältnis zur Berechenbarkeit als auch zu Mathematik, Programmierung und Kunst, von den Computerwissenschaften zwei argumentative Strategien verfolgt. Erstens wird das stereotype Narrativ vorgelegt, das davon erzählt, wie Maschinen Menschen übertreffen, wenn etwa Deep Blue oder Alpha Go die besten Schach- und Go-Meister schlagen. Zweitens ist auffallend, dass die Standardbeispiele gleichzeitig eine subkutane Erzählung mittransportieren, die durch ihre Wahl selbst induziert wird. Denn sämtlich adressieren sie Bereiche, die wohldefinierbar sind und sich prinzipiell mathematisieren lassen. Dass z.B. Schachspielen einer elaborierten Kombinatorik bedarf statt des kühnen Regelbruchs, liegt auf der Hand, doch einen einfachen Gedanken zu verstehen, Witze machen oder scheinbar Unzusammenhängendes sinnvoll miteinander zu kombinieren, um zu neuen Einsichten zu gelangen, verlangt andere Semantiken als die, die mit einer Prädikatenlogik 2. Stufe oder mehrwertigen Logiken kompatibel wären. Eine <Kritik algorithmischer Rationalität> setzt hier an. Sie zielt auf die notwendige Einschränkung des Geltungsbereichs des Berechenbaren selbst, um die Programme nicht nur zu trivialisieren, sondern sie dem widersprüchlichen metaphysischen Projekt einer durchgängigen Mathematisierung der Welt zuzuweisen.

³¹ Immanuel Kant: Kritik der Urteilskraft und Schriften zur Naturphilosophie, in: ders.: *Werke* in 12 Bänden. Bd. 9, Wiesbaden 1956, hier A 24.

³² John von Neumann: *Rigorous Theories of Control and Information*, in: ders.: *Theory of Self-Reproducing Automata*, hg. v. Arthur W. Burks, Urbana 1966, 42–56, hier 51, Übers. DM.

Dabei beruht ihr tautologischer Anspruch zuletzt auf einem Missverständnis dessen, was unter <Kunst> oder <Kreativität> zu verstehen ist. Regelmäßig verbinden Forschungsprojekte zur artifiziellen Kreativität und Computerkunst beide mit der Hervorbringung von «Neuem» oder der Erzeugung und Erkennung «überraschender» oder «interessanter» Muster,³³ obgleich diese Ausdrücke in dem Sinne «pseudo-exakt» erscheinen, als sie entweder nicht definiert sind oder mit Bezug auf ihre Differenz zu anderen Programmen und deren Möglichkeiten erklärt werden.³⁴ Zudem bedürfen die Begriffe des Neuen, der Überraschung oder des Interesses überall noch der humanen Urteilsbildung, von der nicht klar ist, ob sie sich überhaupt mit automatischen Evaluationen deckt oder sich «computieren» lässt.³⁵ So rekurren zahlreiche Autoren des Bandes *Computers and Creativity* einschließlich der Herausgeber auf systemtheoretischen oder evolutionsbiologischen Ansätzen wie «Emergenz»,³⁶ «Mutation»³⁷ oder der Fähigkeit «non-humaner Agenten», Daten durch Lernprozesse zu komprimieren und auf andere Bereiche übertragen zu können – bis hin zu der allereinfachsten, den Lexika des 19. Jahrhunderts entnommenen Formeln, Kunst beruhe auf solchen «Imaginationen», die imstande seien, Werke von dauerhafter «Schönheit» und «emotionaler Kraft» zu erschaffen.³⁸ Die erstaunliche Simplität der Definitionen orientiert sich sämtlich an Vorstellungen, die nicht nur bereits vor mehr als 100 Jahren von den künstlerischen Avantgarden des 20. Jahrhunderts hinweggefegt wurden, vielmehr ahnen sie nicht einmal etwas von einer spezifisch epistemologischen Dimension des Ästhetischen.³⁹ Konsequent blenden sie aus, was Kunst allererst zu Kunst macht: Reflexivität als Aufschließung eines *anderen* Wissens. Stattdessen wird im Zeichen einer Präferenz für Rationalismus und *hard sciences* eine direkte Verbindung zwischen «natürlichen» Kreativitäten wie der Entwicklung des Lebens und der «sozialen» bzw. «historischen» Virulenz der Künste gezo-gen, ungeachtet wesentlicher Inkompatibilitäten.⁴⁰

Wenn wir im Gegenzug auf der *Persistenz einer ästhetischen Reflexivität* als Kriterium für Kunst bestehen, so deshalb, weil diese sich nicht in Wahrnehmungsextensionen, subjektiven Ausdrucksexzessen oder der Darstellung von Grenzphänomenen erschöpft, so wenig wie sich die Eigenschaften von Kunst ausschließlich über ihre «Werke» erschließen lassen, vielmehr «er-eignen» diese immer auch eine Verwandlung und Verschiebung von Kunst selbst. Kunst ist stets *Kunst über Kunst*; sie impliziert daher in jedem Akt und Artefakt eine Transformation des Ästhetischen selbst, wohingegen sich die meisten Modelle einer *artificial creativity* an Kontinuitäten orientieren und einem anachronistischen Geniekult des 19. Jahrhunderts verhaftet bleiben, der sich seinerseits einer Vulgarisierung der kantischen Definition des *Genies* verdankt. Nicht nur säkularisiert dessen Vorstellung das *Theologoumenon* einer *creatio ex nihilo*, das im Kern algorithmisch durch Probabilistik und Datenanalyse nachgeahmt wird, sondern der «Schöpfungsakt» wird formal entscheidbar, wenn er vermöge neuronaler Netze und selbstlernender Algorithmen mit Mustererkennung und

³³ Siehe Margaret A. Boden: *Die Flügel des Geistes. Kreativität und künstliche Intelligenz*, München 1992, 44 ff., sowie die Beiträge in dies., Mark d'Inverno, Jon McCormack (Hg.): *Computational Creativity. An Interdisciplinary Approach* (= Dagstuhl Seminar Proceedings, Vol. 09291), Dagstuhl 2009, online unter drops.dagstuhl.de/opus/volltexte/2009/2224/, gesehen am 22.5.2019.

³⁴ Vgl. Schmidhuber: *A Formal Theory of Creativity*, 323.

³⁵ Philip Galanter neigt, ebenso wie Frieder Nake, zu einer generellen Skepsis. Vgl. Philip Galanter: *Computational Aesthetic Evaluation. Past and Future*, in: McCormack u. a. (Hg.): *Computers and Creativity*, 255–293, und Cohen u. a.: *Evaluation of Creative Aesthetics*, 100 ff., 108 f.

³⁶ Jon McCormack, Mark d'Inverno: *Preface*, in: McCormack u. a. (Hg.): *Computers and Creativity*, vii.

³⁷ Ebd., viiif.

³⁸ McCormack, d'Inverno: *Heroic versus Collaborative AI for the Arts*, 2438, Übers. DM.

³⁹ Vgl. Dieter Mersch: *Epistemologien des Ästhetischen*, Zürich, Berlin 2015.

⁴⁰ Vgl. exemplarisch Oliver Brown: *Generative and Adaptive Creativity. A Unified Approach to Creativity in Nature, Humans, and Machine*, in: McCormack u. a. (Hg.): *Computers and Creativity*, 361–382.

Stilanalysen existierender Kunstwerke unterfüttert und statistisch variiert wird. Zuletzt bleiben die automatischen Erfindungen von gewissen Randomisierungen abhängig, sodass jede *artificial creativity* lauter sinnlose <Sprünge> inszeniert, gegen die, mit Turings Orakel und dem Enigma der mathematischen Intuition, die bereits erwähnte <Poetik der Findungen> zu setzen wäre, die *andere Arten von <Sprüngen>* privilegiert. Denn was Sprung heißt, springt nicht ziellos, sondern wie die kreative Überschreitung irgendetwas Neues erzeugt, vielmehr erweist sich diese stets als rückgebunden an ein Ganzes, das die existenziellen wie kulturellen Bedingungen des Humanen und seiner Grenzen inkludiert. Kunst und Kreativität bleiben notwendig *bezogen*: Es gibt keine Lebensform, die sich nicht gleichzeitig in Paradoxa verstrickte, weshalb die Relativität der *creatio* systematisch mit jener *reflectio* verquickt bleibt, welche die Paradoxa sowohl entdeckt wie aufzusprenken sucht. Denn Kreativität bildet kein erratisches Vermögen noch eine generative Maschine, sondern *ein Prinzip der Freiheit*, das den historischen Einschlüssen Aufschließungen und daher auch immer andere <Aufschlüsse> gewährt.
