LEUPHANA
UNIVERSITÄT LÜNEBURG

Die Maschine auf der Couch. Oder: Was ist schon 'künstlich' an Künstlicher Intelligenz?

Apprich, Clemens

Published in:
Zeitschrift für Medienwissenschaft

DOI:
10.25969/mediarep/12617

Publication date:
2019

Document Version
Verlags-PDF (auch: Version of Record)

Link to publication

Citation for pulished version (APA):
Apprich, C. (2019). Die Maschine auf der Couch. Oder: Was ist schon 'künstlich' an Künstlicher Intelligenz? *Zeitschrift für Medienwissenschaft*, *11*(2), 20-28. https://doi.org/10.25969/mediarep/12617

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 14. Juli. 2025

media/rep/

Repositorium für die Medienwissenschaft

Clemens Apprich

Die Maschine auf der Couch. Oder: Was ist schon ‹künstlich› an Künstlicher Intelligenz?

2019

https://doi.org/10.25969/mediarep/12617

Veröffentlichungsversion / published version
Zeitschriftenartikel / journal article

Empfohlene Zitierung / Suggested Citation:
Apprich, Clemens: Die Maschine auf der Couch. Oder: Was ist schon ‹künstlich› an Künstlicher Intelligenz?. In: *Zeitschrift für Medienwissenschaft*. Heft 21: Künstliche Intelligenzen, Jg. 11 (2019), Nr. 2, S. 20–28. DOI: https://doi.org/10.25969/mediarep/12617.

Nutzungsbedingungen:
Dieser Text wird unter einer Creative Commons BY-NC-ND 4.0/ Lizenz zur Verfügung gestellt. Nähere Auskünfte zu dieser Lizenz finden Sie hier:
https://creativecommons.org/licenses/by-nc-nd/4.0/

Terms of use:
This document is made available under a creative commons BY-NC-ND 4.0/ License. For more information see:
https://creativecommons.org/licenses/by-nc-nd/4.0/

DFG Deutsche Forschungsgemeinschaft

Philipps Universität Marburg

CLEMENS APPRICH

DIE MASCHINE AUF DER COUCH

Oder: Was ist schon ‹künstlich› an Künstlicher Intelligenz?

Es gibt keine künstliche Intelligenz. Freilich, es gibt ein künstliches Hüftgelenk, weil Säugetiere über ein natürliches Pendant verfügen und es somit ersetzt werden kann.[1] Aber was sollte analog dazu künstliche Intelligenz ersetzen? Eine natürliche Intelligenz? Stephen Jay Gould hat in seinem Buch *Der falsch vermessene* Mensch die wissenschaftlichen Schwächen, aber auch politischen Gefahren in der Suche nach einer biologisch fundierten Intelligenz aufgezeigt.[2] Seine historische Studie zeichnet nach, wie Schädel- und Körpervermessung, aber auch der nach wie vor omnipräsente IQ-Test das Phantasma eines auf Biologismus beruhenden Positivismus bedienen. Somit ist auch die Ausrichtung Künstlicher Intelligenz an der Geschichte der Gehirnforschung, wie sie nicht zuletzt in aktuellen medienwissenschaftlichen Debatten erfolgt, zutiefst problematisch.[3] Der in der Neurowissenschaft verankerte Konnektionismus, also die Annahme, dass sich Intelligenz auf ihre hirnphysiologische Verdrahtung reduzieren ließe, erfreut sich mit der Wiederentdeckung künstlicher neuronaler Netzwerke großer Beliebtheit. In ihrem Kern geht diese Vorstellung auf die Frühkybernetik, insbesondere die Arbeit von Warren McCulloch und Walter Pitts, zurück.[4] Zentral für deren Modell ist das Alles-oder-nichts-Prinzip, wonach Neuronen entweder feuern oder nicht feuern, und so im Sinne einer diskreten Aussagelogik (wahr/falsch) repräsentiert und mithilfe elektronischer Halbleiter (an/aus) operationalisiert werden können. Kombiniert zu einem *nervous net* sind künstliche Neuronen dem Modell nach in der Lage, jede Funktion einer Turingmaschine, das heißt die schrittweise Berechnung von Zeichenketten nach bestimmten Regeln, durchzuführen.[5] Im Folgenden werde ich den Spuren einer solchen kybernetischen Bestimmung von Intelligenz nachspüren und zwei Fallstudien aus dem Bereich der computerbasierten Psychotherapie (ELIZA und PARRY) gegenüberstellen. Im Zentrum steht dabei die Frage,

1 Danke an Wolfgang Hagen für den Hinweis und die wunderbare Analogie.

2 Stephen Jay Gould: *Der falsch vermessene Mensch*, Frankfurt/M. 1988.

3 Zur geschichtlichen Einordnung siehe auch Andreas Sudmann: Einleitung, in: Christoph Engemann, ders. (Hg.): *Machine Learning. Medien, Infrastrukturen und Technologien der Künstlichen Intelligenz*, Bielefeld 2018, 9–23, hier 19 f.

4 In ihrem 1943 erschienenen Aufsatz «A Logical Calculus of Ideas Immanent in Nervous Activity» befassen sich der ausgebildete Psychiater (McCulloch) und der Schüler von Rudolf Carnap (Pitts) mit der Möglichkeit, neurophysiologische Erkenntnisse auf logische Operationen von Rechenmaschinen zu übertragen. Warren S. McCulloch, Walter Pitts: A Logical Calculus of Ideas Immanent in Nervous Activity, in: *Bulletin of Mathematical Biophysics*, Vol. 5, 1943, 115–133.

5 McCulloch und Pitts liefern in ihrem Aufsatz keine Beweise dafür, was ihre Netze berechnen können, nur Hinweise, dass es etwas – im Turing'schen Sinn – zu beweisen gibt; siehe Gualtiero Piccinini: The First Computational Theory of Mind and Brain: A Close Look at McCulloch and Pitts's ‹A Logical Calculus of Ideas Immanent in Nervous Activity›, in: *Synthese*, Vol. 141, 2004, 175–215, hier 197 f.

inwieweit das mechanistische Weltbild des Konnektionismus zu einer biologistisch verkürzten Vorstellung von Intelligenz führt und wie diese Vorstellung durch eine psychoanalytische Kritik aufgebrochen werden kann.

I. Das Problem mit der Intelligenz

Wie Orit Halpern in *Beautiful Data* zeigt, handelt es sich bei dem Neuronenmodell von McCulloch und Pitts um den Versuch, eine Rationalität zu entwerfen, die nicht mehr auf den alten Prinzipien von Vernunft und Verstand basiert.[6] McCulloch selbst machte diesen Punkt auf einem kognitionswissenschaftlichen Symposium des California Institute of Technology deutlich, als er die modernen Rechner in Anschluss an Turing als ‹psychotisch› bezeichnete.[7] Für den entschiedenen Gegner psychoanalytischer Ansätze war dies insofern von Relevanz, als damit eine mechanistische Vorstellung von Intelligenz transportiert wurde, die letztlich die Hoffnung enthielt, direkt in die Schaltkreise des Hirns eingreifen zu können.[8] Abseits irgendwelcher *talking cures* sollten Neuronen als hirnphysiologische Basiseinheit in Form von elektronischen Logikgattern nachmodelliert werden, um so ‹Denken› als materiell lösbares Problem der Elektrotechnik zu definieren.[9] «Treating the brain as a Turing machine»[10] umgeht somit die alten psychologischen Fragen nach Bewusstsein und Verstand, indem nicht mehr danach gefragt wird, was Intelligenz *ist*, sondern wie sie *funktioniert*.[11] Aus dieser kybernetischen Sicht ergeben sich allerdings zwei wesentliche Herausforderungen für das McCulloch-Pitts-Modell:

> The first concerns facilitation and extinction, in which antecedent activity temporarily alters responsiveness to subsequent stimulation of one and the same part of the net. The second concerns learning, in which activities concurrent at some previous time have altered the net permanently, so that a stimulus which would previously have been inadequate is now adequate.[12]

Das Problem vorangegangener, nicht erinnerter Aktivität, die das Netz in seinem gegenwärtigen Tun heimsucht, verweist nicht zuletzt auf das Problem des Unbewussten; insofern es auch nicht weiter verwundern mag, wenn ein dezidiert antipsychoanalytisches Programm, wie es neuronale Netze nach McCulloch und Pitts darstellen, mit unbewussten symbolischen Prozessen seine lieben Mühen hat.[13] Das Problem des Lernens, das hier im Zentrum stehen soll, betrifft das Problem der maschinellen Intelligenz bzw. die Frage, inwieweit Maschinen durch schrittweises Lernen so etwas wie menschliche Intelligenz nachahmen können.

Bereits 1950 stellte Alan Turing die Frage, ob Maschinen mithilfe geeigneter Erziehungsmaßnahmen (auch Programmierung genannt) zum Denken gebracht werden können.[14] Zu diesem Zwecke ersetzte er die für ihn unpräzise Frage: «Can machines think?» durch ein Gedankenexperiment:

6 Orit Halpern: *Beautiful Data. A History of Vision and Reason since 1945*, Durham 2014, 145–154.

7 Warren S. McCulloch: Discussion of The General and Logical Theory of Automata by John von Neumann, in: Llyod A. Jeffress (Hg.): *Cerebral Mechanisms in Behavior. The Hixon Symposium*, New York, London 1951, 32–41, hier 32 f.

8 McCulloch war übrigens nicht der Einzige in der Frühgeschichte der Kybernetik: Auch Grey Walter und Ross Ashby gingen von einer mechanistischen Struktur des menschlichen Verstandes aus, der dann eben auch mithilfe von Elektroschocks oder Lobotomie ‹korrigiert› werden könnte. Siehe hierzu Andrew Pickering: *The Cybernetic Brain*, Chicago, London 2011, 171 f.

9 Ein ähnlicher Ansatz, wenn auch aus umgekehrter, das heißt psychoanalytischer Sicht, kommt bekanntermaßen von Friedrich Kittler, der psychische Prozesse ebenfalls als mechanische Effekte elektronischer Medientechnologien beschrieb, vgl. u. a. Friedrich Kittler: *Aufschreibesysteme 1800/1900*, München 1985.

10 McCulloch: Discussion, 32.

11 Halpern: *Beautiful Data*, 148–151.

12 McCulloch u. a.: A Logical Calculus, 117.

13 Siehe hierzu Clemens Apprich: Secret Agents. A Psychoanalytic Critique of Artificial Intelligence and Machine Learning, in: *Digital Culture & Society*, Vol. 4, Nr. 1, 2018, 29–44.

14 Alan Turing: Computing Machinery and Intelligence, in: *Mind*, Vol. 49, 1950, 433–460. Wohlgemerkt spricht Turing nicht von irgendeiner Art Maschine, sondern von digitalen Computern im modernen Sinn (ebd., 435–439).

> The new form of the problem can be described in terms of a game which we call the ‹imitation game›. It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman.[15]

Zwei Punkte sind hier von Interesse: zum einen der Umstand, dass es in diesem, auch Turing-Test genannten Spiel eben nicht darum geht, den_die Fragesteller_in hinters Licht zu führen; so ist sich C über den Aufbau des Tests während des gesamten Spiels bewusst, auch wenn und gerade weil A und B in ihren schriftlichen Antworten lügen dürfen; zum anderen das oft übersehene, aber zentrale Detail, dass Turing das Problem der Intelligenzbestimmung mit dem der Geschlechteridentität einleitet. Er legt nahe, dass A über sein Geschlecht lügen soll, um so den_die Fragesteller_in dazu zu bringen, ihn für B zu halten, während B dem_der Fragesteller_in helfen soll, indem sie möglichst wahrheitsgetreue Antworten gibt.[16] Von der allzu rigiden Geschlechterbinarität einmal abgesehen, klingt in der *queeren* Versuchsanordnung die eingangs gestellte Frage nach der vermeintlichen ‹Natürlichkeit› menschlicher Intelligenz wider, zumal in Turings Beispiel der Computer die Rolle von A übernimmt:

> We now ask the question, ‹What will happen when a machine takes the part of A in this game?› Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman? These questions replace our original, ‹Can machines think?›.[17]

Wie Ulrike Bergermann anführt, geht es Turing nicht darum, die Frage nach dem ontologischen Status Künstlicher Intelligenz zu klären, sondern letztlich darum, die Frage selbst zu denaturalisieren.[18] Wenn Biologie bei der Bestimmung von Intelligenz nicht mehr als Marker dienen soll, bleibt eben nur noch die mittels Fernschreiber ermöglichte Konversation zwischen Mensch und Maschine, um eine Aussage darüber treffen zu können, ob die Maschine nun intelligentes Verhalten zeigt oder nicht.[19]

II. Die sprechende Maschine

Turings Hoffnung, dass Maschinen in absehbarer Zeit eine Konversation auf Englisch halten können, erfüllte sich früher als von ihm erwartet.[20] Das erste überzeugende Gesprächsprogramm war das von Joseph Weizenbaum am MIT zwischen 1964 und 1966 entwickelte Sprachanalysesystem ELIZA.[21] Die Inspiration für das Parsing-Programm kam von George Bernard Shaws Komödie *Pygmalion*, in welcher der Sprachwissenschafter Henry Higgins die arme und durchwegs naiv dargestellte Blumenverkäuferin Eliza Doolittle zu einer Herzogin macht, indem er ihr Englisch im Akzent der Londoner Oberschicht beibringt.[22] Weizenbaum sponn den altgriechischen (und teils misogynen) Stoff weiter: Die Kunstfigur Eliza wird bei ihm zum Sprachprogramm ELIZA, das

15 Turing: Computing Machinery, 433.

16 Judith Genova weist darauf hin, dass diese Konstellation die Frau implizit zu einer «inferior thinker» macht, da von ihr kein strategisches Verhalten erwartet wird; Judith Genova: Turing's Sexual Guessing Game, in: *Social Epistemology*, Vol. 8, Nr. 4, 1994, 313–326, hier 319.

17 Turing: Computing Machinery, 434.

18 Ulrike Bergermann: Biodrag. Turing-Test, KI-Kino und Testosteron, in: Christoph Engemann, Andreas Sudmann (Hg.): *Machine Learning. Medien, Infrastrukturen und Technologien der Künstlichen Intelligenz*, Bielefeld 2018, 339–364, hier 346–349.

19 Der 1990 initiierte Loebner-Prize bedient sich Turings Idee, maschinelle Intelligenz daran zu messen, ob eine menschliche Testperson sie mit einem anderen Menschen verwechselt.

20 Turing ging vom Jahr 2000 aus.

21 Joseph Weizenbaum: ELIZA – A Computer Program for the Study of Natural Language Communication Between Man and Machine, in: *Communications of the ACM*, Vol. 9, Nr. 1, 1966, 36–45.

22 Der Stoff war so brisant, dass die Uraufführung am 16. Oktober 1913 nicht in London, sondern – in deutscher Übersetzung – in Wien stattfand. Das Wiener Publikum war ebenfalls empört; allerdings nicht so sehr wegen der teils derben Sprache, sondern weil sich der Professor und die Schülerin am Ende nicht kriegen.

selbst keine ‹Intelligenz› aufweist, sondern lediglich nach *keywords* im Eingabetext sucht. Um diese zu interpretieren und eine Konversation zu führen, bedarf es eines Skripts, das im Fall von ELIZA einem Rogerianischen Psychotherapeuten nachempfunden war. DOCTOR, wie das Skript hieß, bot sich als Gesprächspartner allein schon deshalb an, weil Carl Rogers Ansatz einer humanistischen Psychologie dem damaligen Klischee einer bloßen *talking cure* entsprach: «The Rogerian psychotherapist is relatively easy to imitate because much of his technique consists of drawing his patient out by reflecting the patient's statements back to him».[23] In der personenzentrierten Gesprächstherapie wird kein Wissen über den_die Patient_in bzw. die Welt, in der er_sie lebt, vorausgesetzt, zumal davon ausgegangen wird, dass alles zur Heilung Notwendige bereits in ihm_ihr vorhanden ist. Bei der technischen Umsetzung des Parsing-Systems hat dies den Vorteil, dass das Programm keine explizite Information speichern muss. Um so etwas wie Verständnis zu simulieren, muss es lediglich die Aussage des_der Patient_in als Frage reformulieren. Dieser als ‹Eliza-Trick› bekannte Mechanismus, der in der Folge bei einer Vielzahl von Chat-Bots zum Einsatz kam, stieß immer wieder auf Kritik – nicht zuletzt bei Weizenbaum selbst, der in einer kritischen Reflexion auf seine Schöpfung betonte, dass es sich bei ELIZA (bzw. dem Skript DOCTOR) eben nicht um ein intelligentes System handelte, sondern um eine Versuchsanordnung, welche die Grenzen des Computers beim Verstehen natürlicher Sprache aufzeigen sollte.[24]

Der aufklärerische Ansatz des Experiments schien aber nicht zu fruchten, wie Weizenbaum anhand einer oft wiederholten Geschichte verdeutlichte. So soll ihn seine Sekretärin, der er bei der Arbeit in jovialer Manier über die Schulter blickte, gebeten haben, den Raum zu verlassen, damit sie sich in Ruhe mit DOCTOR unterhalten könne. In seinem Buch *Computer Power and Human Reason* kommentiert der Computerwissenschaftler die Szene mit deutlicher Verwunderung, war sich die Sekretärin – ganz im Sinne des Turing-Tests – doch der Tatsache bewusst, dass es sich lediglich um ein Programm und nicht um einen echten Therapeuten handelte. Obwohl er aus Erfahrung wusste, dass Menschen eine emotionale Beziehung zu Maschinen entwickeln können und diesen nicht selten menschliche Eigenschaften zuschreiben, war Weizenbaum sichtlich erstaunt über die Tatsache, dass jemand den Computer als eine Person wahrnahm, mit der man sich unterhalten könnte: «What I had not realized is that extremely short exposure to a relatively simple computer program could induce powerful delusional thinking in quite normal people».[25] Dieser als ‹Schock› bezeichnete Moment wurde für ihn nur noch darin übertroffen, dass selbst ernstzunehmende Psychiater_innen ELIZA für psychotherapeutische Zwecke einsetzen wollten.[26] Weizenbaums Kritik richtete sich explizit an Kenneth M. Colby, einen Psychiater, der Kliniken und Zentren mit computerbasierten Therapien versorgen wollte, um so hunderte von Patient_innen in kürzester Zeit behandeln zu können.[27] Die automatisierte Therapie wird laut Colby möglich, weil der menschliche Therapeut – gleich dem Computer – als «information-processor and decision maker»[28] betrachtet werden

23 Joseph Weizenbaum: *Computer Power and Human Reason. From Judgment to Calculation*, New York 1976, 3. Die männliche Konnotation des psychotherapeutischen Skripts, das auf dem an sich unintelligenten weiblichen Programm läuft, wird nicht nur in diesem Zitat mehr als ersichtlich.

24 Vgl. ebd., 8–11. In seinem ursprünglichen ELIZA-Artikel von 1966 war Weizenbaum nicht ganz so skeptisch. So räumte er seinem Programm – ausreichend Speicher und genügend Gesprächssituationen vorausgesetzt – durchaus die Möglichkeit ein, über Zeit lernen und damit ‹intelligenter› werden zu können (vgl. Weizenbaum: ELIZA, 43).

25 Weizenbaum: *Computer Power*, 7.

26 Vgl. ebd., 5.

27 Vgl. Kenneth M. Colby, James B. Watt, John P. Gilbert: A Computer Method of Psychotherapy: Preliminary Communication, in: *Journal of Nervous and Mental Disease*, Vol. 142, Nr. 2, 1966, 148–152, hier 152.

28 Ebd., 151. Auch hier ist der Therapeut durchwegs männlich konnotiert.

kann, der im Verlauf des Gesprächs ein kognitives Modell des_der Patienten_in lernt und anhand des jeweiligen Inputs seinen Output in natürlicher Sprache anpasst. Aufgrund der – zumindest hypothetisch vorhandenen – Lernfähigkeit des Systems wird der enge Rahmen von ELIZA als einem bloßen Sprachanalyseprogramm verlassen und durch die Simulation eines echten, weil kontextabhängigen Gesprächs erweitert.[29]

III. Das schizoide Programm

Glaubt man den Ausführungen von Sherry Turkle, war Kenneth Colby an der ursprünglichen Entwicklung von Weizenbaums ELIZA mitbeteiligt,[30] was die spätere Rivalität der beiden Männer erklären könnte. Während Weizenbaum die Grenzen der Computerwissenschaft in Bezug auf das Verständnis natürlicher Sprache aufzeigen wollte, war Colby vom therapeutischen Nutzen eines objektiven, kostengünstigen und immer verfügbaren Computerprogramms überzeugt. Für ihn konnte eine computerbasierte Psychotherapie dabei helfen, aus der Psychiatrie eine präzise Wissenschaft zu machen.[31] Einige Jahre nach ELIZA schrieb Colby gemeinsam mit einer Gruppe von Informatiker_innen am Stanford Artificial Intelligence Laboratory das Computerprogramm PARRY, das den Dialog zwischen einem Therapeuten und einem paranoiden Schizophrenen (namens Parry) simulierte.[32] Im Gegensatz zu ELIZA bzw. DOCTOR übernahm das Computerprogramm nicht die Rolle des Therapeuten, sondern die des Patienten, was eine andere Gesprächsstrategie erforderlich machte. ELIZA konnte den Input als reformulierte Frage zurückspielen, während PARRY ein Verständnis der Gesprächssituation aufbringen musste: «[T]o understand what information is contained in linguistic expressions, knowledge of syntax and semantics must be combined with beliefs from a conceptual structure capable of making inferences».[33] Anstatt eines rein syntaktischen Parsers, der Probleme damit hat, die Konversation weiterzuführen, sobald ein bestimmtes Wort in der Datenbank fehlt, ging es Colby und seinem Team um die Implementierung von symbolverarbeitenden Regeln, die selbstständig darüber entscheiden, was im Gespräch wichtig und was unwichtig ist. Hierzu wird ein Muster im Sinne eines ‹Glaubenssystems› programmiert, das im Fall von PARRY dem bei Paranoia auftretenden – und laut Colby klinisch nachgewiesenen – Charakterzug eines Verfolgungswahns entspricht. Der simulierte Patient, ein 28 Jahre alter männlicher Protestant mit Spielproblemen, glaubt, dass ihm die Wettmafia auf den Fersen ist. Im Gespräch sucht er deshalb nach expliziten und impliziten Hinweisen für eine mögliche böswillige Absicht seines Gegenübers. Im Gegensatz zu ‹normalem› Kommunikationsverhalten ist die Input-Output-Strategie eines solch «paranoid information-processing system» vor allem von Misstrauen, Angst und Feindseligkeit geprägt, was zu einer konsequenten, weil wahnhaften (Fehl-) Interpretation des Gesprächs führt.[34]

29 In seiner Replik lehnt Weizenbaum diesen Vorschlag ab – «not on the grounds that such a project might be technically infeasible, but on the grounds that it is immoral» (Weizenbaum: *Computer Power*, 269). Sein eigenes Computerprogramm ELIZA kam aber sehr wohl an der Harvard Medical School und am Massachusetts Health Center zum Einsatz, siehe Michael T. McGuire, Stephen Lorch, Gardner C. Quarton: Man-Machine Natural Language Exchanges Based on Selected Features of Unrestricted Input – II. The Use of the Time-Shared Computer as a Research Tool in Studying Dyadic Communication, in: *Journal of Psychiatric Research*, Vol. 5, Nr. 2, 1967, 179–191.

30 Vgl. Sherry Turkle: *Life on the Screen. Identity in the Age of the Internet*, New York 1995, 106.

31 Vgl. Kenneth M. Colby: *Artificial Paranoia: A Computer Simulation of Paranoid Processes*, New York 1975, 23 f.

32 Vgl. Kenneth M. Colby, Sylvia Weber, Franklin D. Hilf: Artificial Paranoia, in: *Artificial Intelligence*, Vol. 2, 1971, 1–25.

33 Colby: *Artificial Paranoia*, 37 f.

34 Colby u. a.: Artificial Paranoia, 5.

PARRY, der u. a. dazu entwickelt wurde, angehende Psychiater_innen in ihrem Umgang mit paranoid-schizophrenen Patient_innen auszubilden, löste eine lebhafte Debatte über die Möglichkeit maschineller Intelligenz aus.[35] Trotz einiger Berichte, wonach PARRY den Turing-Test bestanden hätte,[36] blieben erhebliche Zweifel an der Generalisierbarkeit des schizoiden Programms. So schrieb Colby selbst von einem «one-track mind» seines Patienten, das den Vorteil habe, immer wieder zu seinem Wahnkomplex zurückkehren und so den Verlauf des Gesprächs steuern zu können.[37] Colbys Absicht in der Programmierung von PARRY war daher weniger, ein möglichst überzeugendes (das heißt generalisierbares) Exemplar Künstlicher Intelligenz zu entwickeln, als vielmehr ein empirisch überprüfbares Modell der Paranoia vorzustellen: «I shall propose a postulated structure of symbol-manipulating processes, strategies, functions, or procedures that is capable of producing the observable regularities of the paranoid mode».[38] Hierzu ging er von einer spezifischen Form der Symbolverarbeitung aus, in der das paranoide Verhalten durch eine «underlying organized structure of rules» erzeugt wird, und nicht etwa durch physiologische Ausfälle im Gehirn.[39] Dies ist ein entscheidender Punkt, da die der Paranoia zugrundeliegende Struktur somit einem Algorithmus gleicht, der – ganz im Sinne der Computerwissenschaft – die Art und Weise, wie Symbole verarbeitet werden, organisiert und somit kontrolliert.[40] Gerade weil Menschen und Maschinen bzw. die aufgrund ihres rigiden Wahns einem Computerprogramm ähnelnden Paranoiker_innen symbolverarbeitende Fähigkeiten haben, handelt es sich bei psychisch deviantem Verhalten nicht allein um ein materielles Problem. Colby, der die Psychoanalyse mit einem empirischen Modell der Paranoia untermauern, aber eben nicht grundsätzlich verabschieden wollte, nimmt in dem für die Psychiatrie virulenten Körper-Geist-Problem die gegenteilige Position zu McCulloch und Pitts ein: «We are not comparing the structure of neurons with the structure of transistors; we are comparing the organization of symbol-processing procedures in an algorithm with symbol-processing procedures of the mind-brain».[41] Nicht die Hardware, sondern die Software ist für das Verständnis mentaler Rechenprozesse entscheidend. Zwar beschreibt Colby physiologische Ursachen, welche psychische Störungen hervorrufen können,[42] aber die aus seiner Sicht entscheidende, nämlich therapeutische Frage betrifft nicht die primäre Störung, sondern die nachgelagerte gestörte Verarbeitung von Informationen.[43]

Die Analogie vom Menschen als symbolverarbeitender Maschine impliziert ein funktionales Verständnis psychischer Vorgänge, das informationswissenschaftliche mit psychoanalytischen Ansätzen zu verbinden versucht: «From this standpoint I would postulate a duality at the symbolic level between reasons and causes. That is, a consciously unrecognized reason can operate like a mechanical cause in that it is inaccessible to voluntary modification by symbolic reprogramming».[44] Wird der zugrundeliegende Code dem Bewusstsein zugänglich gemacht, können Modifikationen in der symbolverarbeitenden Struktur vorgenommen werden, ähnlich dem psychoanalytischen Ansatz, unbewusste

35 Siehe die als Anhang publizierten Kommentare zu Colbys Artikel «Modeling a Paranoid Mind», in: *The Behavioral and Brain Sciences*, Vol. 4, 1981, 515–560.

36 Siehe hierzu Margaret A. Boden: *Mind as Machine: A History of Cognitive Science*, Oxford 2006, 370.

37 Colby u. a.: Artificial Paranoia, 9.

38 Colby: *Artificial Paranoia*, 11.

39 Ebd., 100.

40 Siehe hierzu die mittlerweile klassische Definition von Kowalski: Algorithm = logic + control; vgl. Robert Kowalski: Algorithm = logic + control, in: *Communication of the ACM*, Vol. 22, Nr. 7, 1979, 424–436.

41 Colby: *Artificial Paranoia*, 25.

42 Hier unterscheidet sich für Colby ein defekter Transistor nicht so sehr von einem Hirnschaden oder eine durch Drogen induzierte Paranoia.

43 Paranoia ist damit auch nur ein spezifischer Fall psychopathologischen Verhaltens, das durch symbolverarbeitende Prozesse erklärt werden kann, vgl. Colby: *Artificial Paranoia*, 99 f.

44 Ebd., 18.

Vorgänge mithilfe symbolisch-konzeptioneller Techniken an die Oberfläche zu bringen.[45] Damit geht Colby über den rein konnektionistischen Ansatz von McCulloch und Pitts hinaus: Die Intelligenzleistung kann von Hardwareproblemen beeinträchtigt werden, aber die Intelligenzfähigkeit – Denken und Problemlösen – übersteigt die bloße (neuronale) Rechenkapazität, zumal die beste Hardware ohne geeignete Software nicht viel leisten kann. Nun könnte man im Sinne von McCulloch und Pitts einwenden, dass die Möglichkeiten der Software wesentlich von der Hardware vorgegeben werden, Intelligenz also auf die neuronalen Prozesse des Hirns rückführbar ist.[46] Allerdings führt die einseitige Auflösung des *mind-body problem* zu einer biologistischen Verkürzung, wie der Titel von McCullochs Vortrag auf dem bereits angesprochenen Hixon-Symposium nahelegt.[47] Darin beschreibt der Neurophysiologe das menschliche Hirn als informationsverarbeitende Maschine, wobei es ihm weniger darum geht, neuronale Prozesse an sich zu verstehen, als eine Bestimmung des Geistes durch neuronale Mechanismen vorzunehmen. Intelligenz wird auf die Anatomie des Nervensystems reduziert, was die Frage symbolischer Prozesse und ihrer Reprogrammierung (also von Therapie und Lernen) hinfällig macht.[48]

Conclusio

In der mechanistischen Konzeption des Geistes können psychische Konflikte letztlich nur als Störungen eines ansonsten reibungslosen ‹Normalzustandes› aufgefasst werden. Insofern bleibt auch Colbys symbolisch-konzeptioneller Ansatz dem naturalistischen – um nicht zu sagen: behavioristischen – Weltbild seines Gegenübers verhaftet. Der Intellekt steckt bei ihm zwar nicht in der Hardware, dafür aber in der Software, die selbst wiederum auf bestimmten Vorgaben beruht. Die symbolische Verarbeitung ist so lange kein Problem, wie die Zuordnung von Bedeutung klar geregelt ist, das Programm also genau weiß, wie es mit dem jeweiligen Input umzugehen hat. PARRY arbeitet deshalb mit einer ganzen Reihe von Vorannahmen, die Colby in der *naturally-occurring paranoia* als biologisch gegeben ansieht. Der simulierte Patient folgt damit einer spezifischen Vorstellung von Paranoia, die der US-amerikanischen Psychiatrie Anfang der 1970er Jahre entspricht, an der Erklärung derselben als psychoanalytischem Phänomen aber erstaunlich wenig Interesse zeigt.[49] Wie Annette Bitsch betont, war es nicht zuletzt die Lacan'sche Theorie, welche die paranoische Grundposition des modernen Ichs aufzeigen und so dem Glauben an einen natürlichen Intellekt – ob nun in der Hardware oder der Software – den Garaus machen wollte. Das Ich als imaginäre Verkennungsfunktion führt zu psychischen Konflikten, die auch nicht ohne weiteres aufgelöst werden können:

> Wenn Trauma, Entfremdung, Täuschung, Aufspaltung und Verdopplung, unentschuldbare Gefühle von Liebe, Hass, Eifersucht und Angst vor Selbstauflösung am Anfang des Ichs stehen, dann wundert es nicht, dass das Ich stets bedroht ist von Kastrations- und Zerstückelungsängsten, von potentiellen Feinden und Doppelgängern.[50]

45 Der Vergleich mit unbewussten psychischen Prozessen ergibt sich für Colby aus der für das Funktionieren von Computerprogrammen wesentlichen Unterscheidung zwischen kompilierten und interpretierten Modi: «The rough analogy with ever-changing human symbolic behavior lies in suggesting that modifications require change at the source-language level. Otherwise compiled algorithms are inaccessible to second-order monitoring and modification» (ebd., 25).

46 Vgl. McCulloch u. a.: A Logical Calculus, 113 f.

47 Vgl. Warren S. McCulloch: Why the Mind Is in the Head, in: Llyod A. Jeffress (Hg.): *Cerebral Mechanisms in Behavior. The Hixon Symposium*, New York 1951, 42–57.

48 Auch wenn sich seither viel in der informatischen Operationalisierung und kulturwissenschaftlichen Beschreibung intelligenter Systeme getan hat, bleibt der grundsätzliche Gegensatz zwischen symbolischer und konnektionistischer Ausrichtung bei der Konzeption denkender Maschinen auch heute noch bestehen; siehe u. a. Pedro Domingos: *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*, New York 2015.

49 Dies mag auch der Grund dafür sein, dass in Colbys Herleitung einer «Symbol-Processing Theory of the Paranoid Mode» kontinentaleuropäische (oder andere) Ansätze der Psychoanalyse komplett fehlen. Vgl. das gleichnamige Kapitel in: Colby: *Artificial Paranoia*.

50 Annette Bitsch: *Diskrete Gespenster. Die Genealogie des Unbewussten aus der Medientheorie und Philosophie der Zeit*, Bielefeld 2009, 89.

Insofern ist auch PARRYs Angst vor der Wettmafia nicht einfach Ausdruck eines *paranoid mode*, der sich vom *ordinary mode* durch sein *distorted thinking* unterscheidet, sondern entspricht dem zentralen Paradox der Moderne: Das vermeintlich autonome Ich bzw. die Simulation davon steht immer schon in einem Abhängigkeitsverhältnis zum Anderen. Und diese Abhängigkeit führt zu Konflikten, wie dann auch ein viel beachtetes Gespräch zwischen PARRY und ELIZA (bzw. DOCTOR) auf der International Conference on Computer Communication vom 18. September 1972 zeigt:

ELIZA -- Good Evening. Tell me your problems. Please terminate input with a period or a question mark.
PARRY -- People get on my nerves sometimes.
ELIZA -- I am not sure I understand you fully.
PARRY -- You should pay more attention.
ELIZA -- Suppose you should pay more attention.
PARRY -- You're entitled to your own opinion.
ELIZA -- What makes you think I am entitled to my own opinion?

[…]

PARRY -- In my opinion, you have a one-track mind.
ELIZA -- What makes you think I have a one-track mind?
PARRY -- I used to go to the track every week.
ELIZA -- What does that suggest to you?
PARRY -- I avoid racetracks now because the races are fixed.
ELIZA -- Is that the real reason?
PARRY -- I'd rather not discuss it.
ELIZA -- What else would you like to discuss?

[…]

PARRY -- I went to the races.
ELIZA -- I am not sure I understand you fully.
PARRY -- What's wrong with you that you keep on bugging me about this?
ELIZA -- Who are you really talking about?
PARRY -- The way you repeat yourself is getting ridiculous.[51]

Der Zusammenbruch von Kommunikation führt zu Unsinn, da die Symbole beginnen – jenseits exakter Zuschreibungen – für sich zu sprechen.[52] Und gerade dieser von Missverständnissen geprägte Zustand macht den *ordinary mode* einer auf Rechenprozessen laufenden Intelligenz letztlich aus. Hier unterscheiden sich die kommerzialisierten Nachkommen Alexa, Siri oder Duplex auch nicht so sehr von ihren Eltern, zumal es ELIZA und PARRY ja nicht darum ging, ihr Gegenüber zu verwirren, sondern – im Sinne Turings – eine Konversation zwischen Mensch und Computer (oder Computer und Computer) in natürlicher Sprache zu ermöglichen.[53] Was sich im intelligenten und oftmals missverständlichen Verhalten heutiger Sprachassistenten[54] zeigt, ist das Unbewusste als *general intellect*, der von den Nutzer_innen mit Daten aller Art gefüttert und trainiert wird. Marx zufolge handelt es sich bei diesem

51 Vincent Cerf: PARRY Encounters the DOCTOR, in: *Network Working Group*, dort datiert 21.1.1973, https://tools.ietf.org/html/rfc439, gesehen am 30.4.2019.

52 Ein Umstand, den Lacan mit dem informationstheoretischen Begriff *jam* einzufangen versuchte: «Das ist das erste Mal, daß als Grundbegriff die Konfusion als solche auftaucht, jene in der Kommunikation liegende Tendenz aufzuhören, Kommunikation zu sein, das heißt, überhaupt nichts mehr zu kommunizieren» (Jacques Lacan: *Das Ich in der Theorie Freuds und in der Technik der Psychoanalyse. Das Seminar, Buch II*, Wien 2015, 110).

53 Vgl. Meric Dagli: Trusting the Minds Behind the Voices. Vocal Conversational Agents, in: *Medium*, dort datiert 18.6.2017, https://medium.com/@mericda/trusting-the-minds-behind-the-voices-c1c4c4f0584d9, gesehen am 12.4.2019.

54 Genauer wäre wohl der Ausdruck Sprachassistentinnen, da diese – mit der Ausnahme von Google Duplex, wo es beide Geschlechter zur Auswahl gibt – eindeutig als Frauen codiert sind und damit die zweifelhafte Genealogie fortführen; siehe hierzu Luke Munn: *Ferocious Logics: Unmaking the Algorithm*, Lüneburg 2018, 97 f.

‹allgemeinen Wissen› um eine Produktivkraft, deren Kontrolle den realen Lebensprozess bestimmt – «nicht nur in der Form des Wissens, sondern als unmittelbare Organe der gesellschaftlichen Praxis».[55] Insofern ist Künstliche Intelligenz auch weniger künstlich als vielmehr kollektiv. Sie ist, wie das Beispiel der Sprachassistenten zeigt, abhängig von einer Vielzahl an Techniken, Infrastrukturen und Arbeitsschritten.[56] Intelligenz ist nicht etwas, das im Kopf sitzt. Sie ist kein einheitliches Ding, das im Sinne des Konnektionismus nachmodelliert werden könnte. Vielmehr ist Intelligenz äußerlich, ein soziales Ensemble aus menschlichen und nicht-menschlichen Elementen. Die Technologie als vergegenständlichte Wissenskraft ist somit elementarer Bestandteil jeglicher Intelligenz, da sie die Möglichkeit schafft, Wissen zu speichern, zu verarbeiten und weiterzugeben.[57] Dieser transindividuelle Charakter von Intelligenz spiegelt sich in unserem Abhängigkeitsverhältnis zum Anderen wider, ein Verhältnis, das über Technologie vermittelt wird. Und gerade weil es technologisch vermittelt ist, die Welt des Symbolischen also die Welt der Maschine ist,[58] ist auch das in der Künstlichen Intelligenz enthaltene Wissen nur ein Abbild unseres kollektiven Selbst.

55 Karl Marx: *Grundrisse der Kritik der politischen Ökonomie*, Berlin 1953, 602.

56 Dies betrifft nicht nur die Datenproduktion der Nutzer_innen, sondern auch die Arbeit von Minenarbeitern, Reinigungskräften und Programmierer_innen, vgl. Kate Crawford, Vladan Joler: Anatomy of an AI System. The Amazon Echo as an anatomical map of human labor, data and planetary resources, dort datiert 2018, https://anatomyof.ai, gesehen am 30.4.2019.

57 Vgl. Bernard Stiegler: Relational Ecology and the Digital Pharmakon, in: *Culture Machine*, Vol. 13, 2012.

58 Vgl. Lacan: *Das Ich*, 64. Siehe auch Friedrich Kittler: Die Welt des Symbolischen – eine Welt der Maschine, in: ders.: *Draculas Vermächtnis. Technische Schriften*, Leipzig 1993, 58–80.