

Peter Asaro

Computer als Modelle des Geistes

Über Simulation und das Gehirn als Modell des Designs von Computern¹

After the war, together with a small group of selected engineers and mathematicians, Johnny built, at the Institute for Advanced Study, an experimental electronic calculator, popularly known as Joniac, which eventually became the pilot model for similar machines all over the country. Some of the basic principles developed in the Joniac are used even today in the fastest and most modern calculators. To design the machine, Johnny and his co-workers tried to imitate some of the known operations of the live brain. This is the aspect which led him to study neurology, to seek out men in the fields of neurology and psychiatry, to attend meetings on these subjects, and, eventually, to give lectures to groups on the possibilities of copying an extremely simplified model of the living brain for man-made machines.²

Klara von Neumann, Ehefrau von John von Neumann

Turing knew perfectly well what the job he had to do, which was to manufacture or design a machine that would do the complicated sort of mathematics that had to be done in the Mathematical Division of NPL. But he had all sorts of interesting things that he liked to do: for example, he was really quite obsessed with knowing how the human brain worked and the possible correspondence with what he was doing on computers. [...] Turing thought that the machine should be made quite simple, and at the same time should make everything possible that could be done. His particular purpose was to permit the writing of programs that modify programs, not in the simple way now common but rather in the way that people think.³

Ted Newman, Kollege von Alan Turing

1. Einleitung

Das Ziel des Artikels ist es, einige der Weisen, in denen die Beziehung zwischen Gehirn und Computer als ein Modellieren verstanden werden kann, sowie die Bedeutung von ‚Simulation‘ in der Beziehung zwischen Modellen, Computern und Gehirnen zu untersuchen. Die ziemlich umfangreiche Literatur über Modelle und Simulationen in den Naturwissenschaften fokussiert bevorzugt die physikalischen Wissenschaften, weniger Geist und Gehirn. Während die Kognitionswissenschaften häufig die Begriffe des Modellierens und der Simulation verwenden, sind sie in der Benutzung dieser Begriffe und in der Prüfung ihrer impliziten wissenschaftlichen Bedeutungen enttäuschend inkonsistent.

Mein Zugang besteht in der Erörterung der frühen Überlagerung von Gehirnmodellen und Rechenmodellen in der Kybernetik; und dies in der Absicht, deren Bedeutung für laufende Diskussionen in den Kognitionswissenschaften zu klären. Ich bin der Ansicht, dass die Erforschung der historischen Bedeutungszusammenhänge, in denen das Gehirn und der Computer wechselweise als Modelle für einander dienen, ein guter Anfang ist für die Erörterung der Frage, welche Entwicklung die Kognitionswissenschaft künftig einschlagen soll.

‚Modell‘ ist in diesem Zusammenhang ein spannender Begriff, weil man davon auch das Verbum ‚modellieren‘ ableiten kann, und weil ‚Modell‘ und ‚Modellieren‘ in verschiedenen Satzkontexten, die durch Präpositionen beeinflusst werden, verwendbar sind. X modelliert Y, X ist ein Modell von (of) Y, X ist ein Modell für (for) Y, Y ist auf der Basis von X modelliert, Y ist nach X modelliert und schließlich gibt Y ein Modell für Z ab (*models for*).⁴ So können wir sagen, dass das Gehirn ein *Modell für die Struktur* des Computers war, oder auch: Der Computer wurde *auf der Basis* des Gehirns modelliert. Dies in dem Sinn, dass die Entwickler und Designer der ersten Computer, wie John von Neumann, das biologische Gehirn in ähnlicher Weise auffassten wie der bildende Künstler sein Aktmodell, dass sie also den Computer nach dem Bild, das sie sich vom Gehirn machten, konstruierten. Ist das Werk vollendet, könnten wir geneigt sein, umgekehrt die Skulptur des Künstlers als eine Art Modell jenes Menschen anzusehen, der vor ihm posiert hat. Und analog könnten wir den Computer für ein Modell des menschlichen Gehirns halten. Bei bestimmten elektronischen Werkzeugen scheint es, dass sie Modelle der Leistung und des Verhaltens des Gehirns darstellen, wie zum Beispiel W. Ross Ashbys Homöostaten, die als Modelle der adaptiven Eigenschaften des Gehirns gedacht waren, oder W. Grey Walters Schildkröten, die Modelle der dynamischen Antriebe des lebenden Gehirns sein sollten. In noch einem anderen Sinn werden die ersten Computer von einigen beteiligten Forschern, wie von Alan Turing, als *Modelle (von) der fundamentalen Struktur* des Gehirns angesehen, in dem Sinn, dass der digitale Computer ein

technisch erzeugtes Werkzeug sei, das nach den gleichen Prinzipien wie das Gehirn arbeite und das benützt werden könne, um Theorien über die höheren (höherstufigen) Funktionen des Gehirns zu testen.

In jedem dieser Fälle waren der tatsächliche Prozess des Modellierens und die Konstruktion des Werkzeugs einzigartig und komplex und bezogen sowohl konkrete als auch abstrakte Modelle des Gehirns und des Verhaltens mit ein. Ich habe die Eigenart dieser Arbeitsmodelle, wie sie u. a. durch Homöostaten und Schildkröten repräsentiert sind, an anderem Ort dargestellt.⁵ Dieser Artikel wird den theoretischen Modellen und den Arbeitsmodellen eine weitere Klasse hinzufügen: die ‚Simulation‘, ein Begriff, der in ähnlich vielfältiger Weise benutzt wird wie ‚Modell‘ und ‚modellieren‘. Für viele Autoren ist ‚Simulation‘ einfach ein anderes Wort für Modell, für andere ist es das, was ich das Arbeitsmodell nenne. In einem präziseren Sprachgebrauch bezieht sich ‚Simulation‘ allgemein auf eine besondere Klasse von Modellen, die durch deren Gebrauch definiert ist: spezifische rechnerische Modelle, die dazu benutzt werden, um sich dem Verhalten eines Systems von Gleichungen anzunähern, das mit analytischen Techniken nicht zu lösen ist. Wegen der Rechenleistung des Computers werden Simulationen zu einer der bedeutendsten Arten von Modellen in der naturwissenschaftlichen Praxis. Dennoch nehmen sie eine seltsame Position vis-à-vis von Arbeitsmodellen wie Homöostaten und Schildkröten ein. Tatsächlich könnte man von rechnerischen Simulationen behaupten, sie würden die Unterscheidung zwischen theoretischen Modellen und Arbeitsmodellen verwischen. Einerseits erscheinen Simulationen wie automatisierte Theorien, andererseits scheinen Simulationen eine besonders ergiebige Form von Modellen, nämlich solche, die etwas *tun*: daher der Begriff ‚Arbeitsmodelle‘.

Ich beginne mit einem Rückblick auf einige besonders wichtige Aspekte der Arbeiten von John von Neumann und Alan Turing, die beide mit der theoretischen Grundlegung der mathematischen Berechnung und dem Design der ersten *General Purpose Computer* mit gespeicherten Programmen befasst waren. Sie legten auch die theoretischen Grundlagen für die Wege, auf denen Computer Gehirn und Geist simulieren könnten. Der Rückblick zeigt, dass sich das Design der ersten *General Purpose Computer* stark auf das Neuronenmodell von McCulloch und Pitts⁶ und auf andere Aspekte der Neurophysiologie stützte. Das heißt, die ersten Computer wurden auf der Basis des Gehirns modelliert, oder präziser, auf der Basis eines theoretischen Neuronenmodells. In einem bestimmten Sinn wurde die mathematische Theorie der Berechenbarkeit auf der Basis der menschlichen Praxis, mathematische Rechnungen auszuführen, modelliert. Diese grundlegenden Formen des Modellierens werden nicht oft diskutiert. Dennoch glaube ich, dass sie erheblichen Einfluss auf unsere theoretischen Intuitionen bezüglich der komplexeren Formen des

Modellierens des Computers nach dem Gehirn haben. Dies trifft besonders auf die Frühzeit zu, als man über Computer noch gern als ‚riesige Gehirne‘ sprach.

Unter den Werkzeugen des Modellierens erfordert der Computer als ein einzigartiges Artefakt aus mehreren Gründen besondere Aufmerksamkeit. Natürlich ist der Begriff der symbolischen Simulation eng verbunden mit Vorstellungen über die Natur der Berechnung und die technologische Leistung eines funktionierenden Computers. Überdies sind diese Vorstellungen wie auch die Konzeptualisierung und das Design der ersten Computer in vielfältiger und komplexer Weise mit Modellen des Gehirns und des Geistes verknüpft. Der Zweck dieses Artikels besteht auch darin, diese vielfältigen und komplexen Verbindungen als Beispiele verschiedener Typen des wissenschaftlichen Modellierens und wissenschaftlicher Modelle sichtbar zu machen. So können wir bei der Behandlung von Computer und Gehirn Beispiele für jeden bisher erwähnten Typ von Modellen finden, genau so wie für die Simulation.

Bevor wir unsere historische Reise zu den elektronischen Gehirnen der 1940er Jahre beginnen, lohnt es sich, über die Sichtweise auf den Geist, welche die kognitive Psychologie seit damals bestimmte, nachzudenken. Die Auffassung, der Geist *sei* ein Computer, wurde abgekürzt im Begriff Computationalismus. Diese Sichtweise wurde vor einigen Jahren folgendermaßen zusammengefasst:

„Are minds computers? [...] Computationalism – the view that mental states are computational states – is based on the conviction that there are program descriptions of mental processes and that, at least in principle, it is possible for computers, that is, machines of a particular kind, to possess mentality.“⁷

Trotz unterschiedlicher Formulierungen wurzelt Computationalismus in der grundlegenden Analogie zwischen mentalen Zuständen des Gehirns und Rechner-Zuständen. Die unterschiedlichen Formulierungen sind das Ergebnis unterschiedlicher Definitionen und Auffassungen dieser beiden Typen von Zuständen. Allgemein beruht die Definition von Rechner-Zuständen auf den digitalen seriellen Rechnern mit gespeichertem Programm. Zwei Mathematiker, die den größten Einfluss auf das Design der ersten Rechner dieses Typs hatten, nämlich Alan Turing und John von Neumann, waren auch mit dem Aufbau eines synthetischen Geistes befasst. Wir werden sehen, dass sie ein gemeinsames Verständnis von Berechnung hatten, sich jedoch in ihrer Ansicht, wie ein Computer das Gehirn simulieren könnte, unterschieden. Davon noch stärker unterschieden waren Ross Ashbys Versuche, mentale Phänomene zu modellieren. Ashby hielt ‚Informationsverarbeitung‘ für zentral für den adaptiven Geist, aber er sah darin nicht zwangsläufig eine Form der Berechnung. Daher baute er analoge Simulationen (Arbeitsmodelle) des Gehirns, während

Turing und von Neumann unterschiedliche Arten symbolischer Simulation (theoretische Modelle) erfinden wollten.

Zuerst sollte die Unterscheidung zwischen Simulation als Arbeitsmodell und Simulation als automatisiertes theoretisches Modell geklärt werden. Die Nützlichkeit der Unterscheidung verschiedener Typen von Modellen und des Modellierens wird sichtbar, wenn wir die frühe Geschichte des Computers und die ersten Versuche, das Gehirn mit diesen Maschinen zu simulieren, untersuchen. Dabei werde ich mich mit der Automatentheorie beschäftigen, die von John von Neumann entwickelt wurde, als er die EDVAC baute, begründetermaßen der erste Computer mit gespeicherten Programmen. Wir werden uns auch mit von Neumanns Gebrauch des Neuronenmodells von McCulloch und Pitts in diesem Design beschäftigen, und mit seinen späteren Ideen, das Gehirn auf einem Computer zu simulieren. Danach stelle ich Alan Turings Bemühungen um die *Universelle Turing Maschine* als ein Modell von Geist und Gehirn dar. Das wird uns zu seinem Vorschlag Turings gegenüber Ashby führen, den Homöostaten auf seinem ACE Computer zu simulieren und zu seinem direkten Vergleich dieser beiden Typen von Simulation. Ich schließe dann mit den Konsequenzen einer solchen Sichtweise auf Modell und Simulation für den Computationalismus.

2. Analoge und symbolische Simulationen

Es ist vorteilhaft, sorgfältig zwischen ‚theoretischen Modellen‘ und ‚Arbeitsmodellen‘ zu unterscheiden. Während der Begriff des theoretischen Modells seine Wurzeln in der normativen Philosophie der Wissenschaft hat, sind sowohl ‚theoretisches Modell‘ als auch ‚Arbeitsmodell‘ für eine deskriptive Studie wissenschaftlicher Praxis nützlich. Der Unterschied besteht darin, dass Arbeitsmodelle etwas in der Welt tun, sie haben materielle Wirkungsmacht, und diese Wirkungsmacht ist unabhängig von menschlichem Handeln. Theoretische Modelle können auch Wirkungsmacht besitzen, doch handelt es sich jedenfalls um normativ-disziplinäre Wirkungsmacht,⁸ die menschliches Handeln erfordert, das normativen Regeln und Einschränkungen entsprechen will. Überlegen wir nun, wie diese Unterscheidung zwischen theoretischen Modellen und Arbeitsmodellen mit gegenwärtigen Diskussionen der ‚Simulation‘ zusammenhängt.

In einem Artikel über rechnerische Simulationen diskutiert Winsberg⁹ drei traditionelle Versuche, sie zu behandeln: erstens als Metapher, zweitens als Experiment und drittens als einen mittleren Modus. Simulation als Metapher anzusehen bedeutet, dass Simulationen lediglich *brute-force-number-crunching*-Prozeduren seien, die immer dann benutzt werden, wenn analytische Techniken unmöglich erscheinen.

Anders ausgedrückt handelt es sich um eine degenerierte Form des Theoretisierens. Die Sichtweise, Simulationen seien ein Experiment und der Computer sei die experimentelle Vorgabe, unterstellt, dass es eine mimetische Beziehung zwischen der Simulation und dem Simulierten gibt in der Weise, dass die Simulation in der Lage ist, das Reale zu imitieren und als Platzhalter zu agieren. Der dritte Modus nimmt an, dass Simulationen eine völlig andere Angelegenheit seien: ein Verfahren, irgendwo zwischen Experiment und Theoriebildung.

Winsberg unterscheidet sorgsam zwischen jenen Simulationen, in denen analytische Lösungen Ausdrücke in geschlossener Form produzieren, und solchen Simulationen, die numerische Methoden dazu verwenden, große Mengen von Zahlen zu produzieren, welche die gewöhnlichen Werkzeuge der experimentellen Praxis für die Analyse benötigen: Visualisierung, Statistik, *Data Mining* etc.¹⁰ Der Unterschied besteht weniger in den mimetischen Eigenschaften der Mathematik und der Rechenverfahren, als in den Praktiken, die Wissenschaftler benutzen, um sich mit den Modellen zu beschäftigen und mit ihnen zu arbeiten. Die numerischen Methoden ähneln eher Experimenten als Theorien, weil die gleichen Praktiken wie bei experimentellen Untersuchungen benutzt werden.

In dem Sinne, wie ich Arbeitsmodelle beschrieben habe, wurden diese numerischen Simulationen benutzt, um Phänomene zu erzeugen und zu synthetisieren, die untersucht und erforscht werden können und mit denen man experimentieren kann etc., – sie sind selbst Gegenstand empirischer Untersuchungen. Darin besteht der Unterschied zu *theoretischen* Modellen oder Simulationen, die auf von einer Theorie abgeleiteten Gleichungssystemen basieren und der direkten mathematischen Analyse zugänglich sind. In diesem analytischen Sinn bedarf es nicht der Datenanalysetechniken der experimentellen Praxis, um Strukturen und Muster in den Phänomenen auszumachen. Diese Modelle produzieren Daten nur als eine Anwendung der Theorie für einen gegebenen Fall, und die gewünschten Ergebnisse können leicht aus den Formeln erzielt werden. Diese Unterscheidung entspricht auch sehr schön der Unterscheidung zwischen ‚analytisch‘ und ‚synthetisch‘. In diesem Fall bemühen sich analytische Simulationen darum, lokale Modelle aus einer allgemeinen Theorie auf formalem Weg zu gewinnen, während synthetische Simulationen versuchen, die Lücken fehlender Theorie und fehlender Daten zu schließen, indem sie etwas Neues erzeugen, das manipuliert werden kann, mit dem man experimentieren kann und das benutzt werden kann, um Daten zu erzeugen in der Absicht, Theorien zu entwerfen und zu testen.

Ein großer Teil der philosophischen Literatur über Simulation scheint Winsberg blockiert zu sein durch Visualisierung als Schlüsselaspekt, der Simulationen interessant mache. Dies wird im Begriff der mimetischen Qualitäten in den Repräsentationen gefasst. Wie Isomorphismus sind mimetische Beziehungen als eine objek-

tive Weise gedacht, die genaueren Beziehungen zwischen realem und simuliertem System auszudrücken. Sie unterliegen aber der zusätzlichen Anforderung, auch die graphischen Aspekte des Originals zu erhalten:

„The extensive use of realistic images in simulation is a stepping stone that simulationists use in order to make inferences from their data. It is also a tool they use in order to draw comparisons between simulation results and real systems; a move that is part of the process of sanctioning of their results. It is the drawing of inferences and sanctioning of results that give rise to the interesting philosophical connections between simulation and experimental practice.“¹¹

Sicherlich können die graphischen Aspekte solcher Modelle visuelle Untersuchungspraktiken ermöglichen. Wie Winsberg festhält, können Visualisierungen von vielen mathematischen Modellen geleistet werden. Für die Eigenschaft, die sie zu guten Modellen macht, ist das nebensächlich. Vertraut man zu sehr auf die Kraft mimetischer Eigenschaften, wird man tendenziell dahin geführt, Simulation als rein und wahrlich experimentell anzusehen und den Begriff des „numerischen Experiments“ wortwörtlich zu nehmen und die Computersimulation als Stellvertreter und Platzhalter für reale Phänomene zu interpretieren. Aus einer erst zu beweisenden Voraussetzung wird die Schlussfolgerung gezogen, dass die Simulation den realen Phänomenen sehr nahe komme.¹² Dies aber würde bedeuten, es komme nicht darauf an, dass die Simulation in mancher Hinsicht als Stellvertreter fungiert, sondern in welcher Hinsicht, in welchem Ausmaß und mit welcher Genauigkeit sie das tut.

Besonders wenn wir uns für den epistemischen Status von Modellen und Simulationen interessieren, sind diese Fragen sehr wichtig, und wir dürfen nicht für gesichert nehmen, dass die Tatsache, dass wir an Simulationen Experimente ausführen können, garantiere, dass diese den gleichen epistemischen Status und einen viel weniger metaphysischen Status hätten als reale Experimente. Demnach bedeutet dies nicht, dass ihnen der experimentelle Status vollkommen fehlt oder dass sie metaphysisch schwach sind.

Mit der Überlegung, dass Simulationen ein neuer Modus wissenschaftlicher Praxis zwischen Theoriebildung und Experiment sein könnten, behauptet Winsberg, dass dies eher eine Gelegenheit sei, um über Simulation nachzudenken zu beginnen, jedoch nicht sie zu erklären:

„What is of interest philosophically is to understand (a) how it is that what is at root a theoretical enterprise, takes on characteristics of experimentation, (b) what those characteristics are – at the abstract, reconstructed level, (c) what consequences there are of such a hybrid for our understanding of the

nature of modelling, theorizing, and experimenting, and (d) how simulation produces knowledge and what kind of knowledge that is.¹³

Auf der Suche nach einem Kompromiss für Simulation betont Winsberg ausdrücklich, dass Techniken und Praktiken des Modellierens ihre eigenen epistemischen Verdienste haben können, unabhängig von der Theorie, die ihnen zugrunde liegt:

„[T]he credibility of [a] model comes not only from the credentials supplied to it by the governing theory, but also from the antecedently established credentials of the model building techniques developed over an extended tradition of employment. That is what I mean when I say that these techniques have their own life; they carry with them their own history of prior successes and accomplishments, and, when properly used, they can bring to the table independent warrant for belief in the models that are used to build.“¹⁴

Für Winsberg ist also plausibel, dass die epistemische Grundlage für Simulationen sowohl aus der Theorie als auch aus der Praxis des Modellierens kommt. Es ist jedenfalls wichtig festzuhalten, dass er auch glaubt, dass Simulationen nur in dem Maß autonom (oder semi-autonom, wie er meint) von Theorie sind, in dem sie unabhängige epistemische Grundlagen besitzen, die in der Tradition der Praxis wurzeln. Ich stimme dem zu und möchte lediglich hinzufügen, dass Arbeitsmodelle Modelle sind, die im Rahmen einer Tradition der Praxis des Modellierens erzeugt werden und diese unterstützen.

Bevor ich die ersten Versuche bespreche, den Computer für die Simulation von Geist und Gehirn zu benutzen, möchte ich mich mit einer weiteren Frage zur Simulation beschäftigen, nämlich: Wie vermag ein Modell jenes System, das es modelliert, zu repräsentieren? Während der Konnektionismus-Debatte in den 1990er Jahren gab es ein wiederkehrendes Thema des Streits, das auf der Unterscheidung zwischen ‚analogen‘ und ‚digitalen‘ Rechnern beruhte. Möglicherweise wurde diese Unterscheidung überstrapaziert oder besser: die wirkliche Natur der Unterscheidung wurde nicht immer vollends erkannt. Eine Ausnahme bildet hier Trenholme.¹⁵ Indem er einen gedanklichen Bogen spannt, der die Ideen von Kenneth Craik, Norbert Wiener, Philip Johnson-Laird und Rodney Brooks einbezieht, präsentiert Trenholme eine Sichtweise von Simulation, von der ich glaube, dass sie mit dem Arbeitsmodell kompatibel ist. Ich hoffe, dass die Anwendung seiner Sichtweise auf die Arbeit von John von Neumann und Alan Turing zu den frühen Computern eine neue Perspektive darauf ergibt, wie der Computer in verschiedener Weise als eine Simulation von Geist und Gehirn aufgefasst wurde.

Kurz gefasst besteht die Idee darin, dass obgleich alle synthetischen Simulationen Arbeitsmodelle sind – im allgemeinen Sinn des Automatisch-Seins, genau so wie

alle Modelle ‚repräsentational‘ im allgemeinen Sinn sind – sich die Art ihres Modellseins auf signifikante Weise unterscheiden kann. Eine *analoge* Simulation modelliert direkt ein natürliches System, während eine *symbolische* Simulation ein intermediäres symbolisches System benutzt und daher ein indirektes Modell darstellt. Dieses Moment des Indirekten ist hoch signifikant in dem Maß, als es den relevanten Aspekt einer Simulation aus dem Gebiet der Arbeitsmodelle, epistemischen Artefakte und materiellen Wirkungsmacht in das Gebiet der Modelle und normativ-disziplinären Wirkungsmacht verschiebt. Was also Trenholme eine naturalistische analoge Simulation nennt, ist das, was ich Arbeitsmodell nannte. Und was er eine symbolische Simulation nennt, ist eine spezielle Art eines theoretischen Modells, eine Art, deren disziplinäre Handlungsträgerschaft von menschlichen Händen und menschlichem Geist auf automatische Berechnungen übertragen wird.

Trenholme beginnt mit einer Klarstellung. Während in der damaligen Debatte viel in den Begriffen ‚analog‘ versus ‚digital‘ ausgedrückt wurde, kehrt er sorgfältig heraus, dass das eigentliche Thema im Unterschied von *analogen* und *symbolischen* Repräsentationen besteht. Wie wir in der Diskussion von John von Neumanns Automatentheorie sehen werden, ist ‚analog‘ versus ‚digital‘ eine Angelegenheit, wie *Zahlen* in einem Computer repräsentiert werden. Nichts wird darüber ausgesagt, ob und wie diese Zahlen in der Lage sind, irgendetwas anderes über sich selbst hinaus zu repräsentieren. Das eigentliche Problem besteht darin, ob Zahlen überhaupt ein essenzieller Teil einer Simulation sind. Daher sollte die Dichotomie ‚analog‘ versus ‚digital‘ die Idee erfassen, dass analoge Simulationen keine symbolischen Repräsentationen erforderlich machen oder auf ihnen beruhen. Eher dienen sie als Simulationen vorrangig oder vollständig kraft ihrer kausalen Struktur. Trenholme nennt sie „naturalistische analoge Simulationen“, um sie von den sorglosen Definitionen analoger Berechnungen zu unterscheiden.

Trenholme argumentiert weiters, dass naturalistische analoge Simulationen¹⁶ nicht in derselben Weise Repräsentationseigenschaften haben wie symbolische Simulationen. Dies beruht auf einem zusätzlichen *Mapping* in dem Sinn, wie es an anderer Stelle diskutiert wurde.¹⁷ Kurz gefasst beziehen *analoge* Simulationen die kausale Struktur eines natürlichen Phänomens auf die kausale Struktur einer Simulation durch einen Isomorphismus – tatsächlich wird die lockere „Ähnlichkeits“-Relation aufgerufen (zusammen mit probabilistischer Kausaltheorie). Eine *symbolische* Simulation dagegen wird erzielt oder ‚gemapped‘ aus einer formalen Theorie natürlicher Phänomene, und dann wird sie ‚gemapped‘ oder abgebildet in eine rechnerische Simulation der formalen Theorie. Der symbolischen Simulation fehlt also der Isomorphismus zwischen Kausalstrukturen. Stattdessen haben symbolische Simulationen Repräsentationseigenschaften im Sinne von semantischen Relationen wie Denotation. Analogen Simulationen fehlt diese Schicht reiner Repräsentation:

„Symbolic simulation is thus a two-stage affair: first the mapping of inference structure of the theory onto hardware states which defines symbolic computation; second, the mapping of inference structure of the theory onto hardware states which (under appropriate conditions) qualifies the processing as a symbolic simulation. Analog simulation, in contrast, is defined by a single mapping from causal relations among elements of the simulation to causal relations among elements of the simulated phenomenon.“¹⁸

Während eine symbolische Simulation auf *irgendeiner* Ebene auf Kausalstrukturen beruht, hängt sie von der *symbolischen* Ebene ab, um die Simulation auszuführen. Es gibt einen allgemeinen Sinn von Repräsentation, den beide Typen von Simulation einbeziehen. Das heißt, dass diese Simulation in gewissem Sinne eine Repräsentation des simulierten Systems ist, aber dass die Art der Repräsentation nicht notwendigerweise symbolische Repräsentation ist.

Beispielsweise „repräsentiert“ das Quecksilberbarometer Luftdruck *während seiner Arbeit* ohne Einbeziehung symbolischer Systeme – das Symbolische tritt erst hinzu, wenn das Ergebnis abgelesen wird. Analoge Simulationen beruhen auf Kausalstruktur, weniger auf symbolischen Relationen, um eine gute Simulation zu erzielen. Diese Kausalstrukturen sind die materiellen Handlungsträger in Arbeitsmodellen. Einerseits müssen diese Kausalstrukturen in bestimmter Weise eingegrenzt werden, um eine valide Simulation zu konstituieren (sie sind disziplinäre materielle Wirkungsmacht), während es andererseits immer Wege gibt, auf denen sie unbeschränkt und offen sind, in dem Sinn, dass sie sich auf die potenziell unendlichen Kausalbeziehungen in der Welt beziehen und immer neue emergente Eigenschaften zeigen können. Daher ist hier dieser doppelte Aspekt materieller Wirkungsmacht in analogen Simulationen zu berücksichtigen, der in symbolischen Simulationen weitgehend fehlt, welche versuchen, das Verhalten des physikalischen Systems zu beschränken, um lediglich das präzise symbolische System, das intendiert ist, zu produzieren. Wenn sich eine symbolische Simulation irregulär verhält, so ist ihr Output tatsächlich bedeutungslos.

Symbolische Repräsentation ist eine sehr spezifische Form der Repräsentation, in der Symbole durch Denotation und Referenz repräsentieren sollen. Symbole alleine haben keine Kausalstruktur, und um eine automatische Simulation auf ihrer Grundlage durchzuführen, bedarf es eines laut Newell und Simon „physikalischen Symbolsystems“, also eines Computers:

„*The Physical Symbol System Hypothesis*. A physical symbol system has the necessary and sufficient means for general intelligent action. By ‚necessary‘ we mean that any system that exhibits general intelligence will prove upon analysis to be a physical symbol system. By ‚sufficient‘ we mean that any

physical symbol system of sufficient size can be organized further to exhibit general intelligence.“¹⁹

Simulationen auf der Grundlage symbolischer Repräsentationen sind daher so eng mit der Entwicklung des modernen Computers verknüpft, dass sie häufig einfach ‚Computersimulation‘ genannt werden. Es gibt auch andere Ausdrücke, die auf die Begriffe des Modellierens und der Simulation anspielen, aber sie verschleiern das Problem eher:

„Certain ways of speaking – for example, saying that a theory is ‚modeled on a computer‘ – risk conflating the two mappings, thus blurring the distinction between analog and symbolic simulation, as does the notion of representation when applied to analog simulations. [...] Thus we may say that a user (or an observer) recognizes an analog simulation as such when properties in the analog device are held to *represent* the corresponding properties (those that play the corresponding causal role under the causal-structural isomorphism). In the case of symbolic simulation, once the user identifies the phenomenon to be simulated, the term of the theory coded into the computer may be held to *represent* the relevant features of the phenomenon under its standard (intended) interpretation; here the notion of *representation* stands for language-world semantical relations such as *reference* and *denotation*. An obvious distinction can be made between these semantical relations and the notion of isomorphism of causal structure used in characterizing naturalistic analog simulation.“²⁰

Sobald die Unterscheidung zwischen analoger und symbolischer Repräsentation klar ist, wird erkennbar, wie dieses Argument mit den ausgedehnten Fragen betreffend Computationalismus verknüpft ist.²¹

An diesem Punkt ist darauf hinzuweisen, dass es mindestens zwei signifikant unterschiedliche Wege gibt, den Isomorphismus²² zwischen Simulation und dem System oder dem Prozess, das oder den sie simuliert, zu entwickeln. Der eine Weg ist die einfache Korrespondenz von Input und Output, oder was zuerst von Ashby Black-box-Simulation genannt wurde. Eine Black-box-Simulation liegt vor, wenn es ähnliche Outputs nach ähnlichen Inputs gibt. Im Fall *analoger* Simulationen sind Inputs und Outputs kausal verknüpft, in *symbolischen* Simulationen sind sie symbolisch verknüpft. Die Sache ist im zweiten Typ von Isomorphismus komplizierter: Hier wird angenommen, dass der *interne Prozess* der Simulation und das System oder der Prozess, welche simuliert werden, korrespondieren. In dieser Form der Simulation zählen nicht nur die Input-Output-Relationen, sondern auch die internen Zustände und Prozesse der Simulation. Der Unterschied zwischen analogen und symbolischen Simulationen wird daran klarer.²³ Während symbolische Simulationen bestimmte symbolische Prozesse realisieren müssen, besteht eine Unabhängigkeit zwischen symbolischen Prozessen und ihrer physikalischen Rea-

lisierung. Das nennt man gewöhnlich multiple Realisierbarkeit, was bedeutet, dass verschiedene Kausalsysteme den gleichen symbolischen Prozess realisieren können. Bei analogen Simulationen, denen eine unabhängige symbolische Ebene fehlt, müssen die Kausalprozesse, die die Simulation realisieren, isomorph zum System sein, das simuliert wird, um als Simulation zu gelten.

Gehen wir nun von dieser eher abstrakten Diskussion der Simulation über zu konkreten historischen Beispielen. Wir können in den Ideen von John von Neumann und Alan Turing zwei verschiedene Arten von symbolischen Simulationen erkennen. Für von Neumann soll der Computer benutzt werden, um ein *physikalisches Modell* des Gehirns zu simulieren, während für Turing der Computer den gleichen grundlegenden Algorithmus oder ein Programm simulieren sollte, der oder das ein menschliches Gehirn simuliert.

3. Analogien, Ziffern und Zahlen

Die Unterscheidung von ‚analog‘ und ‚digital‘ ist John von Neumann geschuldet. Seine Arbeiten über die ersten speicherprogrammierten Computer und seine Sicht auf die Beziehung von Computer und Gehirn bieten eine komplexe Geschichte, die zahlreiche Schichten von Analogie, Modellieren und Simulation einbezieht. Eine vollständige Darstellung der Entwicklung seines Denkens geht über die Ziele dieses Artikels weit hinaus, aber es lohnt sich zu überlegen, wie die Sicht von Arbeitsmodellen und Simulationen zu von Neumanns Positionen passt.

Es ist klar, dass Neumanns Entwurf des ersten *Computer Memory* (Gedächtnis/Speicher) stark auf seiner Rezeption neurophysiologischer Forschung und explizit auf der Arbeit von McCulloch und Pitts²⁴ und deren Konzeption neuronaler Netzwerke beruhte. Das ergibt sich schon aus seinem Antrag²⁵ zum Design der EDVAC (1945). Die Geschichte wird noch komplexer, wenn wir seine Theorie der Automaten²⁶ (von 1948), seinen Brief an Norbert Wiener²⁷ (von 1946) über die Simulation des Gehirns und seine späteren Reflexionen über das Verhältnis von Gehirn und Computer anlässlich der posthum veröffentlichten *Silliman Lectures*²⁸ einbeziehen. In diesem Abschnitt werde ich kurz auf einige der Neumannschen Gedanken zur Modellierung des Computers nach dem Gehirn und zur Simulation des Gehirns mittels Computer eingehen.

Der Schlüssel zum Verständnis der Ideen von Neumann ist, dass der Computer für ihn ein besonderer Typ eines repräsentationalen Systems ist. Im Fall seiner Automatentheorie ist der Computer grundsätzlich ein automatisches System für das Repräsentieren und Manipulieren von *Zahlen*. Zahlen sind nicht das gleiche wie Quantitäten oder Ziffern, sie sind abstrakte mathematische Einheiten. Quantitäten

und Ziffern sind konkrete Arten, Zahlen für praktische Zwecke wiederzugeben. Um Mathematik zu automatisieren ist es daher notwendig, ein physikalisches System zu entwickeln, das Zahlen repräsentieren kann, und die Wahl zwischen Quantitäten oder Ziffern ist eine offene Frage, bevor die technologischen Möglichkeiten erwogen werden.

Am 20. September 1948 präsentiert John von Neumann auf dem *Hixon Symposium on Cerebral Mechanisms in Behavior* eine Theorie der Automaten. In seinem Text *The General and Logical Theory of Automata* spricht er ausdrücklich über die Notwendigkeit einer strengen Theorie der Berechnung und entwirft eine formale Theorie der Automaten (axiomatisch festgelegte rechnerische Mechanismen). Er beginnt damit, zwei allgemeine Klassen von Automaten danach zu unterscheiden, wie sie Zahlen repräsentieren. Jene Klasse von Automaten, die auf dem Analogprinzip gründet, repräsentiert Zahlen durch Analogie, das heißt durch bestimmte physikalische Quantitäten, ähnlich wie ein Thermometer die aktuelle Temperatur in seiner Umgebung durch die Quecksilbersäule repräsentiert. Wenn wir zum Beispiel zwei Zahlen durch elektrischen Strom in zwei Stromkreisen repräsentieren, können wir diese beiden Zahlen addieren, indem die beiden Stromkreise in geeigneter Weise kombiniert werden, und das Resultat entspricht dem Strom der kombinierten Stromkreise. Es ist möglich, alle einfachen arithmetischen Operationen (+, -, *, /) auf diese Weise auszuführen.

Automaten, die hingegen auf dem Digitalprinzip beruhen, repräsentieren Zahlen nicht als physikalische Quantitäten, sondern als Aggregate numerischer Ziffern in der Weise, wie Menschen sie auf ein Blatt Papier schreiben oder mit ihren Fingern zählen (das entspricht auch der Etymologie von ‚digital‘). Solch ein Automat könnte eine Scheibe (*dial*) mit zehn Positionen haben, die 0–9 repräsentieren, oder aber auch eine Serie solcher Scheiben für Einer-Stellen, Zehner-Stellen, Hunderter-Stellen etc. im Sinn des Dezimalsystems. Die digitale Repräsentation, die von fast allen modernen Computern benutzt wird, ist ein binäres System, in dem Verdrahtungen elektrische Spannungen von zwei hinreichend unterschiedlichen Größen übertragen und eine Menge von kanonischen Schaltkreisen (Grundschaltkreisen) bestehen, um mathematische und logische Operationen an den binären Repräsentationen auszuführen.

Die beiden Typen von Automaten, die John von Neumann unterscheidet und beschreibt, entsprechen Trenholmes Unterscheidung²⁹ von ‚analog‘ versus ‚digital‘, obwohl Trenholme sie für die Unterscheidung von Simulationstypen als falsch kritisiert. Das heißt, in jedem Fall ist man bemüht, Zahlen zu repräsentieren. Und solange niemand bemüht ist, reine Mathematik zu simulieren, sind beide Typen symbolische Simulationen und im *Doppel-Mapping* verfangen. Selbst wenn wir einen analogen Computer wie einen *Differential Analyzer* verwenden würden,

müssten wir immer noch eine Theorie darüber haben, wie die Kausalstruktur des Analogcomputers die erwünschte mathematische Berechnung realisiert, und auch eine Theorie darüber, wie sich diese Berechnungen auf die fraglichen natürlichen Phänomene beziehen. Dies besagt, dass der *Differential Analyzer* keine Simulation ist, zum Beispiel eines hydrodynamischen Systems, sondern ein Automat für die Lösung einer Menge von Gleichungen, die von einer Theorie dieses hydrodynamischen Systems abgebildet werden.

Die hauptsächliche Konsequenz der unterschiedlichen Arten von numerischer Repräsentation in Automaten ist ihre praktische Realisierung, wenn elektronische Computer tatsächlich gebaut werden. Analogcomputer sind durch die konstante Einführung von Störungen in ihre Schaltkreise sehr fehleranfällig, obgleich sie den Vorteil hoher Präzision haben. Wenn dividiert wird, zum Beispiel $5/7$, kann das Ergebnis nicht mit einer finiten Anzahl von Ziffern ausgedrückt werden, und so wird das Ergebnis eines Digitalrechners durch die Zahl der Ziffern, die auf einmal dargestellt werden können, begrenzt sein, während ein Analogcomputer das Ergebnis mit unbegrenzter Präzision darstellen kann.³⁰ Aber dies ist lediglich ein Ideal. Tatsächlich war es recht schwierig, analoge Schaltkreise dazu zu bringen, Berechnungen mit hoher Genauigkeit auszuführen, da kleine Veränderungen in den physikalischen Größen ein Rauschen im System bewirken, verursacht durch Quellen außerhalb der Berechnung. Während also die analoge Division $5/7$ theoretisch absolut präzise sein könnte, begrenzen schon kleine Wahrscheinlichkeiten sehr kleiner Störungen in den Schaltkreisen die Präzision enorm. Von Neumann bestand darauf, dass es die Schwierigkeit des technischen Problems sei, einen verlässlichen Signal-Rausch-Abstand aufrecht zu erhalten, der den Analogcomputer davon abhält, präziser zu rechnen als wenige Dezimalstellen genau. Das Problem tritt bei Digital-Rechnern nicht auf. Diese haben den weiteren Vorteil, dass Genauigkeit unbegrenzt und ökonomisch erhöht werden kann, indem einfach die Anzahl der Ziffern, die durch das Duplizieren von Komponenten repräsentiert werden, erhöht wird.

Während über diese Frage häufig hinweggesehen wird, behauptet John von Neumann selbst, dass es keinen essenziellen mathematischen Unterschied zwischen den beiden Arten von Automaten gebe, obwohl große praktische Differenzen bestünden. Vor allem könne jegliche Art von Berechnung, die – im Prinzip – auf einem analogen Automaten ausgeführt werden kann, auch auf einem digitalen Automaten ausgeführt werden, und umgekehrt. Neumann argumentiert, dass Organismen – natürliche Automaten – tatsächlich *gemischte* Automaten sind, die beide Prinzipien für verschiedene und spezifische Funktionen nutzen, während der Hochgeschwindigkeits-Computer eine strikt digitale Maschine ist. Mit Bedacht zeigt er, dass der Versuch, eine Theorie des Geistes auf die Arten der Berechnungen zu gründen, die in digitalen Computern zur Verfügung stehen, nicht ihre Anwendbar-

keit auf das lebende Gehirn beeinträchtigen sollte, da die neuronale Aktivität sich als digital erwiesen habe, und selbst wenn dies nicht so wäre, die Theorie jedenfalls durch eine passend konstruierte analoge Maschine realisiert werden könnte. Mit anderen Worten, eine rechnerische Simulation des Gehirns wird eine symbolische Simulation sein, solange die Simulation auf den Berechnungen von Modellen der theoretischen Physik beruht, unabhängig davon, ob sie mit analogen oder digitalen Computern durchgeführt wird.

4. Die Modellierung des Computers auf der Basis des Gehirn-Modells

These repeated excursions into biological information processing and the interdisciplinary study of cybernetics have been ignored in previous accounts of von Neumann's computing, yet they clearly shaped his ideas.³¹

Dieses Kapitel schweift von unseren Diskussionen um Simulationen ab. Jedenfalls ist es insofern relevant, als es die Praxis des Modellierens in der Entwicklung des Computers als einer Maschine betrifft. Im Besonderen erzählt es die Geschichte, wie der Computer auf der Basis des Gehirns modelliert wurde, oder wenigstens auf der Basis bestimmter Theorien darüber, wie das Gehirn und seine Neuronen arbeiten. Es war ein wenig mehr nötig als die numerische Repräsentation, um den programmierbaren Computer technisch zu realisieren. Namentlich musste der programmierbare Computer seine eigenen Interaktionen als Zahlen repräsentieren, die er dann speichern, aus dem Speicher abrufen, verändern etc. kann. Mit anderen Worten, bestimmte Zahlen sind *Repräsentationen von Anweisungen für den Computer*, die er interpretieren und ausführen soll. Das war der große Design-Sprung, der Rechenmaschinen zu Universalcomputern veränderte. Zugleich ist es der Kern einer tiefen Konfusion bezüglich des repräsentationalen Charakters der Berechnung. Dieser Kern wurde früh fruchtbar gemacht durch die Verbindungen, die zwischen dem Rechnerspeicher (Gedächtnis) und der neuronalen Struktur des Gehirns etabliert werden. Aus von Neumanns Beschreibung des EDVAC-Designs wird klar, dass er auf das Gehirnmodell und auf das McCulloch-Pitts-Neuron sah, um sich für sein Design des ersten Computer-Gedächtnis/Speichers zu inspirieren. Dies führte zu einer besonderen Art der Analogie zwischen dem Computer und dem Geist/Gehirn, die die Proponenten des Computationalismus bald aufnahmen, die John von Neumann selbst aber letztlich zurückgewiesen hätte.

Während verschiedene elektromechanische Rechner vor, während und nach dem Zweiten Weltkrieg gebaut wurden, bestand das hauptsächliche technische Hindernis, einen *General Purpose Computer* (i. e. eine arbeitende Universelle Turing-

Maschine) zu bauen darin, wie man ihm ein Gedächtnis resp. einen Speicher geben sollte.³² Es war ein Problem der Organisation und zugleich eines der technischen Realisierung. Allgemein wird angenommen, dass es zuerst von John von Neumann am *Institute for Advanced Study* gelöst wurde. Von Neumann löste es, indem er eine Gedächtnis/Speicher-Einheit – er nannte sie *organ* – als einen zentralen Teil der Computer-Architektur vorstellte. Das Gedächtnis/Speicher-Organ war von numerischen Zählern durch das Repräsentieren oder Speichern von logischen Instruktionen oder Codes einerseits und Zahlen andererseits unterschieden. So wurde es dem Computer möglich, Teillösungen zu errechnen, Zwischenergebnisse zu speichern, sich selbst neu zu konfigurieren und eine neue Funktion auszuführen, indem Befehle im Speicher befolgt wurden und die Rechnung mit den Instruktionen, die als Zahlen abgespeichert sind, wieder aufzunehmen.³³ Das hatte klare Vorteile gegenüber der Notwendigkeit, für jeden Schritt einer Berechnung die Maschine „händisch“ zu konfigurieren, oder der Notwendigkeit, eine Rechenmaschine zu bauen, die kompliziert genug wäre, um für alle nötigen Schritte eines anspruchsvollen rechenaufwändigen Problems programmiert zu werden.

Eine oft übersehene Tatsache im Zusammenhang mit frühem Computerdesign ist das Ausmaß, in dem sich John von Neumann auf das McCulloch-Pitts-Modell bezog, und die neurologische Sprache, in der er sein Design beschrieb. Das ist ein klarer Fall von „Modellieren auf der Basis von...“ – also des Transfers einer Theorie in einer Domäne hin zum Design einer Technologie in einer anderen Domäne. Dass John von Neumann sein Design nach neurophysiologischen Theorien und nach dem McCulloch-Pitts-Neuron modellierte, wird durch seine Art, den Computer zu beschreiben, eindeutig belegt:

„The three specific parts CA [Central Arithmetical organ], CC [Central Control organ], and M [Memory organ] correspond to the *associative* neurons in the human nervous system. It remains to discuss the equivalents of the *sensory* or *afferent* and the *motor* or *efferent* neurons. These are the *input* and the *output* organs of the device, and we shall now consider them briefly.“³⁴

Das Naheliegendste, das zu dieser Passage gesagt werden kann, ist, dass sich John von Neumann nicht auf die Teile der Computer-Architektur als ‚Einheiten‘, ‚Komponenten‘, ‚Module‘ oder ähnliche, in ihrer Konnotation neutrale Ingenieurs-Begriffe bezieht. Als Mathematiker wählte er den semantisch aufgeladenen Begriff ‚Organe‘ aus der Beschreibung biologischer Systeme. Darüber hinaus werden die Teile der Computer-Architektur explizit als *mit spezifischen Elementen des menschlichen Gehirns korrespondierend* verstanden. Dieses Thema führt von Neumann bei der Beschreibung der Maschine aus:

„It is worth mentioning, that the neurons of the higher animals are definitely elements in the above sense. They have all-or-none character, that is two states: Quiescent and excited. They fulfil the requirements of [section] 4.1 with an interesting variant: An excited neuron emits the standard stimulus along many lines (axons). Such a line can, however, be connected in two different ways to the next neuron: First: In an *excitatory synapsis*, so that the stimulus causes the excitation of that neuron. Second: In an *inhibitory synapsis*, so that the stimulus absolutely prevents the excitation of the neuron by any other (excitatory) synapsis. The neuron also has a definite reaction time, between the reception of a stimulus and the emission of the stimuli caused by it, the *synaptic delay*. Following W. Pitts and W. S. McCulloch [*sic*] („A logical calculus of the ideas immanent in nervous activity,“ *Bull. Math. Biophysics*, vol. 5 [1943], pp. 115–133.) We ignore the more complicated aspects of neuron functioning: Thresholds, temporal summation, relative inhibition, changes of the threshold by after effects of stimulation beyond synaptic delay, *etc.* It is, however, convenient to consider occasionally neurons with fixed thresholds 2 and 3, that is neurons which can be excited only by (simultaneous) stimuli on 2 or 3 excitatory synapses (and none on an inhibitory synapsis). It is easily seen, that these simplified neuron functions can be imitated by telegraph relays or by vacuum tubes. Although the nervous system is presumably asynchronous (for the synaptic delays), precise synaptic delays can be obtained by using synchronous setups.“³⁵

An dieser Passage wird vollends klar, dass John von Neumann ernsthafte Überlegungen über Struktur und Funktion biologischer Neuronen anstellt, wenngleich er sie in der Art von McCullochs und Pitts Neuronenmodell idealisiert. Wir sehen, wie er die funktionale Identität von neuronaler Aktivität und mathematischer Logik (eine Vereinheitlichung von neuronalen Mechanismen und Turing Maschinen) als ein Modellierungswerkzeug zum Bau des elektronischen Computers verwendet. Seine Nützlichkeit als Modell ist selbstverständlich metaphorisch oder analogisch, aber diese Art Modell ist insofern mächtig, als es Design-Entscheidungen beeinflusst und nicht nur die Konstruktion von Design-Alternativen. In diesem Sinne ist es werthaltig oder normativ.

„The analogs of human neurons, discussed in 4.2-4.3 and again referred to at the end of 5.1, seem to provide elements of just the kind postulated at the end of 6.1. We propose to use them accordingly for the purpose described there: as the constituent elements of the device, for the duration of the preliminary discussion. We must therefore give a precise account of the properties which we postulate for these elements.“³⁶

Aber dieses Modell war nicht das einzige Kriterium, das bei den Design-Entscheidungen wirksam werden sollte:

„At this point the following observation is necessary. In the human nervous system the conduction times along the lines (axons) can be longer than the synaptic delays, hence our above procedure of neglecting them aside of t would be unsound. In the actually intended vacuum tube interpretation, however, this procedure is justified: t is to be about a microsecond, an electro-magnetic impulse travels in this time 300 meters, and as the lines are likely to be short compared to this, the conduction times may indeed be neglected.“³⁷

Es scheint hiermit eindeutig und klar, dass von Neumanns Verständnis von physikalischer Berechnung nahezu gänzlich vom McCulloch-Pitts-Modell umschrieben wurde. Welche Schlussfolgerungen können aus der hiermit belegten Bedeutung des McCulloch-Pitts-Modells für das Design des ersten Computers gezogen werden? Die erste besteht darin, dass der Computer von seiner grundlegenden Idee her eine Art Modell des Gehirns war. Daher ist es ein Fehler zu meinen, dass die *Artificial Intelligence*-Forschung oder die Kognitionswissenschaften eine Analogie zwischen Geist und Computer „entdeckt“ hätten. Es gab sie bereits. Stattdessen rekonfigurierten sie diese Analogie im Bemühen, Computerprogramme als psychologische Theorien zu entwickeln. Eine Möglichkeit, das Problem des Computationalismus aufzugreifen, besteht in der Frage, ob der Computer selbst unabhängig vom Programm oder der Simulation, die er ausführt, ein Modell des Geistes oder ein gutes Modell des Geistes darstellt. Dies lässt die Frage offen, ob der Geist tatsächlich ein Computer ist, oder der Computer Geist. Aber nahezu niemand glaubt, dass jeder Computer ein Geist ist. Turing scheint der Ansicht gewesen zu sein, dass der Geist eine Universelle Maschine ist, und möglicherweise auch umgekehrt, dass Universelle Maschinen wenigstens dazu fähig wären, Geist zu sein. Wir werden diese Ansichten kurz prüfen. Es ist dazu lediglich zu überlegen, ob es für den Computer eine Möglichkeit gibt, das Gehirn so effektiv zu simulieren, dass er selber eine Art Geist wird.

5. Die Simulation des Gehirns auf dem Computer

An dieser Stelle ist es wert zu überlegen, wie John von Neumann sich den Computer als Simulation des Gehirns vorstellte. Obwohl er seine Automatentheorie auf einer strengen Unterscheidung zwischen analoger und digitaler numerischer Repräsentation basierte, und den Gedächtnis/Speicher des Computers mit gespeichertem Programm nach McCullochs und Pitts essenziell digitalem Neuronenmodell modellierte, war er sich in anderen Schriften darüber im Klaren, dass das Gehirn weitaus komplizierter ist als diese theoretischen Idealisierungen unterstellen. Als

Ergebnis würde die Simulation des Gehirns durch einen Computer viel schwerer zu realisieren sein.

Am 29. November 1946, nach der zweiten Macy-Konferenz, schrieb von Neumann an den Mathematiker und Kybernetiker Norbert Wiener einen Brief, in dem er die Situation bezüglich der Theorie der biologischen Informationsverarbeitung und die Replikation der Fähigkeiten des Gehirns in einem Computer bewertete. Dieser Brief ist aus vielen Gründen interessant und verdient eine genauere Untersuchung, als sie für unsere Zwecke hier angestellt wird. Trotz seines frühen Enthusiasmus war von Neumann vielleicht der erste, der die Grenzen des Computers für die Simulation des Gehirns erkannte. Er artikuliert die Schwächen dieser Zugangsweise klar in seinem Brief an Norbert Wiener. Hier argumentiert er, dass ein direktes Modellieren der physikalischen Struktur des Gehirns so kompliziert wäre, dass es fast hoffnungslos sei; eine Idee, die er noch klarer in seiner späteren Arbeit ausdrücken sollte. Er schließt aus dieser Situation, dass es ein viel besserer Ansatz wäre, sich detaillierter zytologischer Arbeit zuzuwenden. Im Besonderen schlägt er vor, einfache Organismen in allen atomaren Details (wortwörtlich) zu verstehen und zu diesem Zweck mit der Untersuchung von bakteriophagen Viren zu beginnen.

Im folgenden Brief nennt von Neumann einen der hauptsächlichen Gründe für die Schwierigkeit, eine Analogie zwischen Gehirnen und Computern herzustellen:

„Besides, the [brain] system is not even digital (*i.e.* neural): It is intimately connected to a very complex analogy; (*i.e.* humoral or hormonal) system, and almost every feedback loop goes through both sectors. If not through the ‚outside‘ world (*i.e.* the epidermis or within the digestive system) as well.“³⁸

Hier gibt es zwei wichtige Punkte: Erstens, von Neumann ist sich der eingebetteten und situierten Natur des menschlichen Gehirns bewusst und auch dessen, dass die Informationsverarbeitung des Gehirns essenziell in Rückkopplungsschleifen mit der Umwelt erfolgt. Unglücklicherweise lenkt von Neumann seine Energien nicht darauf, dieses Problem weiter zu verfolgen. Zweitens, dass die angenommene saubere Unterscheidung zwischen ‚analog‘ und ‚digital‘ im lebenden Gehirn nicht so klar ist. Das ist die Richtung, der John von Neumann weiterhin große Aufmerksamkeit widmet.

Eine bedeutende Konsequenz des McCulloch-Pitts-Modells war, den digitalen Charakter des Neuronen-Verhaltens im Gehirn als ein Grundelement für das Verständnis seiner Organisation festzusetzen. Während diese Schrift große Verdienste für die Formalisierung neuronaler Netzwerke hatte und sie unter Zuhilfenahme mathematischer Logik behandelte, war diese Idealisierung letztlich eine große Vereinfachung. Von Neumann behandelt den Fall in seinem *Silliman Lecture* Manuskript niederschmetternd. An einer Stelle kritisiert er die Vorstellung, dass der

Mechanismus neuronaler Exzitation und Inhibition als eine einfache Summenfunktion zu behandeln wäre:

„It may well be that certain nerve pulse combinations will stimulate a given neuron not simply by virtue of their number, but also by virtue of the spatial relations of the synapses on a single nerve cell, and the combinations of stimulations on these that are effective (that generate a response pulse in the last-mentioned neuron) are characterized not only by their number but also by their coverage of certain special regions on that neuron (on its body or its dendrite system, *cf.* above), by the spatial relations of such regions to each other, and by even more complicated quantitative and geometrical relationships that might be relevant.“³⁹

Während es verführerisch ist, die Dinge so zu behandeln, als seien alle Inputs für ein Neuron gleich, ist tatsächlich die komplexe dreidimensionale Geometrie der synoptischen Verbindungen zu den Dendriten der Neuronen relevant für die elektrochemischen Prozesse, die einen Puls auslösen. Ebenso ist das Ideal des synchronen Timing der neuronalen Aktivität, das essenziell für McCullochs und Pitts' Behauptung ist, nämlich dass vielschichtige Netzwerke als Äquivalent zu logischen Sätzen behandelt werden können, unrichtig:

„On all these matters certain (more or less incomplete) bodies of observation exist, and they all indicate that the individual neuron may be – at least in suitable special situations – a much more complicated mechanism than the dogmatic description in terms of stimulus-response, following the simple patterns of elementary logical operations, can express.“⁴⁰

Während von Neumann also die Vorschläge, Neuronen als logische Einheiten zu behandeln, welche Summierungen ausführen, für das Design des Computer-Gedächtnis/Speichers als nützlich betrachtete, war er tief beunruhigt davon, wie entfernt diese Idealisierung vom realen Gehirn war, wenn es um den Bau einer Simulation ginge. Die Neuronen führten nicht nur die idealistischen Funktionen, die McCulloch und Pitts für erforderlich erklärt hatten, eindeutig nicht aus. Auch von Neumanns eigene Unterscheidungen zwischen analogen und digitalen Automaten passten auf das Gehirn zwar in komplexer, aber keineswegs in direkter Weise:

„The observation I wish to make is this: processes which go through the nervous system may, as I pointed out before, change their character from digital to analog, and back to digital, *etc.*, repeatedly. Nerve pulses, *i.e.* the digital part of the mechanism, may control a particular stage of such a process, *e.g.* the contraction of a specific muscle or the secretion of a specific chemical. This phenomenon is one belonging to the analog class, but it may be the

origin of a train of nerve pulses which are due to its being sensed by suitable inner receptors. When such nerve pulses are being generated, we are back in the digital line of progression again. As mentioned above, such changes from a digital process to an analog one, and back again to a digital one, may alternate several times. Thus the nerve-pulse part of the system, which is digital, and the one involving chemical changes or mechanical distortions due to muscular contractions, which is of the analog type, may, by alternating with each other, give any particular process a mixed character.“⁴¹

Schließlich führten diese Komplexitäten John von Neumann zu der Annahme, die Erforschung des Gehirns würde zu einer neuen Mathematik führen.

Von Neumann beginnt seinen Brief an Norbert Wiener mit einer vernichtenden Kritik seiner eigenen Anstrengungen und ebenso der Bemühungen Wieners, Turings und McCulloch-Pitts, eine substanzielle Theorie der Informationsverarbeitung im Gehirn zu formulieren.

„Our thoughts – I mean yours and Pitts’ and mine – were so mainly focused on the subject of neurology, and more specifically on the human nervous system and there primarily on the central nervous system. Thus, in trying to understand the function of automata and the general principles governing them, we selected for prompt action the most complicated object under the sun—literally. [...] The difficulties are almost too obvious to mention: They reside in the exceptional complexity of the human nervous system, and indeed of any nervous system.“⁴²

An dieser Stelle wird klar, dass von Neumann den Mangel an formalisiertem wissenschaftlichem Verständnis des Verhaltens der Neuronen als Haupthindernis ansah, um synthetische Automaten-Gehirne zu konstruieren. Die Konsequenz seiner mehrjährigen vergeblichen Anstrengungen, diese Ideen klarer zu machen, war Erbitterung:

„What seems worth emphasizing to me is, however, that after the great positive contribution of Turing-cum-Pitts-and-McCulloch is assimilated, the situation is rather worse than better than before. Indeed, these authors have demonstrated in absolute and hopeless generality, that anything and everything Brouwerian can be done by an appropriate mechanism and specifically by a neural mechanism – and that even one, definite mechanism can be universal.“⁴³

Das ist vielleicht die vernichtendste Kritik von Turings Projekt vor dem Aufkommen der *Artificial Intelligence*-Forschung. Es ist außerdem die klarste Aussage über die paradoxe Beziehung zwischen dem Universalcomputer und dem synthetischen

Gehirn, welches wir im nächsten Abschnitt behandeln. Die Kritik lautet wie folgt: Die Gleichsetzung des Universalcomputers mit den Funktionsweisen des Gehirns ist zugleich unendlich mächtig und essenziell schwach. Die „absolute und hoffnungslose Allgemeinheit“ der Theorie besteht darin, dass sie zwar Prozesse des Geistes replizieren kann, wenn diese formal definierbar sind (nach Brouwer), sie aber absolut keine Einsicht in die Natur, Struktur oder Organisation dieser Prozesse gibt. Der Universalcomputer kann alle Maschinen imitieren, auch die Gehirn-Maschine, aber er sagt uns nichts über Gehirne. Was stattdessen benötigt wird, ist eine strenge Theorie des biologischen Gehirns:

„Inverting the argument: Nothing that we may know or learn about the functioning of the organism can give, without ‚microscopic‘, cytological work any clues regarding the further details of the neural mechanism.“⁴⁴

Schließlich lässt uns die abstrakte Untersuchung neuronaler Netzwerke mit dem grundsätzlichen Problem des empirischen Verständnisses des Gehirns zurück. Und auf diese Weise ist die Simulation des Gehirns auf dem Computer durch das theoretische Wissen der Neurowissenschaft begrenzt. Das überrascht nicht sehr, wenn wir von Neumanns Ansatz der Gehirnsimulation als eine Form der symbolischen Simulation und damit als *theoretische* Simulation anerkennen. Das heißt, seine Idee einer Computersimulation des Gehirns ist vollständig abhängig von einem theoretischen Modell des Gehirns, das in eine symbolische Repräsentation auf einem Computer umgelegt wird. Daher ist sie ein klares Beispiel für die Anpassung zwischen Simulation und theoretischem Modell.

Wenden wir uns nun einer anderen Auffassung zu, wie ein Computer das Gehirn simulieren kann, in der es sich nichtsdestoweniger um eine symbolische Simulation handelt, die auf einem theoretischen Modell basiert.

6. Der Computer als universelle modellierende Maschine

*It is possible to invent a single machine which can be used to compute any computable sequence.*⁴⁵

Nachdem wir die komplexe Beziehung zwischen Computer und Gehirn bei John von Neumann untersucht haben, wenden wir uns nun der geradlinigeren, wenngleich abstrakteren Beziehung zwischen Computer und Geist bei Alan Turing zu. Vereinfacht gesagt, war Turing mit der Idee einer Universellen Maschine beschäftigt (nun Universelle Turing Maschine/UTM genannt). Die Idee der UTM, wie sie bei

Turing 1936 beschrieben wird,⁴⁶ ist die eines abstrakten formalen Computers, der in der Lage ist, jede berechenbare Funktion zu berechnen. An anderer Stelle spricht Turing über die Fähigkeit der UTM, jede andere Maschine zu „modellieren“. Diese Art des Sprechens ist durch unsere frühere Unterscheidung zwischen analoger und symbolischer Simulation umschrieben. Die UTM ist eine rein symbolische Vorstellung, und tatsächlich funktionierende Computer sind bloß Annäherungen an diesen mathematischen Formalismus. Turing versuchte dennoch, reale Computer auf der Basis dieses Formalismus zu entwerfen und zu bauen. Am bedeutendsten wurden die *Advanced Computational Engine* (ACE) und später die *Mark I* in Manchester.

In nahezu jedem Vortrag und Artikel über Berechnung zwischen 1944 bis 1950 berief sich Turing darauf, dass die Ähnlichkeit zwischen seiner UTM und einem programmierbaren digitalen Computer von großer Bedeutung sei:

„Some years ago I was researching on what might now be described as an investigation of the theoretical possibilities and limitations of digital computing machines. I considered a type of machine which had a central mechanism, and an infinite memory which was contained on an infinite tape. This type of machine appeared to be sufficiently general. One of my conclusions was that the idea of a ‚rule of thumb‘ process and a ‚machine process‘ were synonymous. The expression ‚machine process‘ of course means one which could be carried out by the type of machine I was considering. It was essential in these theoretical arguments that the memory should be infinite. It can easily be shown that otherwise the machine can only execute periodic operations. Machines such as the ACE may be regarded as practical versions of this same type of machine. There is at least a very close analogy.“⁴⁷

An dieser und anderen Stellen wird klar, dass Turing Universalität als das entscheidende Element der UTMs ansah, d. h. ihre Fähigkeit, andere Maschinen zu modellieren. Das Design von Rechenmaschinen sollte daher nach der mathematischen Theorie der UTMs modelliert werden. Turing sah den einzigen Unterschied zwischen dem mathematischen Formalismus der UTMs und dem physikalischen Computer in der begrenzten Größe des Gedächtnis/Speichers und in anderen physikalischen Begrenztheiten einer Maschine, wie zum Beispiel Zeit. Abgesehen von diesen praktischen Einschränkungen kann Turings Idee eines Computers als die einer ‚Universellen Modellier-Maschine‘ bezeichnet werden. Sein Fokus auf universelles Modellieren bestimmte auch seine Überlegungen zu Lernen und Intelligenz und deren Modellieren auf dem Computer. Turing entwarf das Problem als das eines unorganisierten Systems, das organisiert wird. Aber anstatt eine thermodynamische Interpretation von Organisieren – wie es Kybernetiker tun würden – zu wählen, versuchte er zu zeigen, wie sich unorganisierte Systeme effizient in UTMs organisieren könnten, mittels Verstärkung durch Belohnung und Strafe.

Turing wandte die gleiche Vorstellung von Universellen Modellier-Maschinen, die er bezüglich des Designs von Computern hatte, auf das menschliche Gehirn an:

„All of this suggests that the cortex of the infant is an unorganized machine, which can be organized by suitable interfering training. The organizing might result in the modification of the machine into a universal machine or something like it. This would mean that the adult will obey orders given in appropriate language, even if they were very complicated; he would have no common sense, and would obey the most ridiculous orders unflinchingly. When all his orders had been fulfilled he would sink into a comatose state or perhaps obey some standing order, such as eating. Creatures not unlike this can be found, but most people behave quite differently under many circumstance. However, the resemblance to the universal machine is still very great, and suggests to us that the step from the unorganized infant to a universal machine is one which should be understood. When this has been mastered we shall be in a much better position to consider how the organizing process might have been modified to produce a more normal type of mind.“⁴⁸

Für Turing scheint der Schlüssel für die Geheimnisse des Geistes bei den UTMs gelegen zu haben. Der Geist hatte für ihn die Fähigkeit, andere Maschinen zu modellieren und war dabei ein universelles Modell besonderer Art, wenngleich nicht vollständig identisch mit dem un kreativen Computer.

Dieser Zugang zu Intelligenz und Geist kann gut mit jenem von W. Ross Ashby kontrastiert werden, der Verhalten nicht als ein logisches Befolgen von Regeln ansah, sondern als Trajektorien im Phasenraum. Er trachtete danach, seine Ideen in Maschinen zu verkörpern, mit denen er direkt interagieren konnte, wie mit dem Homöostat und später mit DAMS. Lernen war für Ashby ein nie endender Prozess, in dem das Ziel das Überleben ist und in dem die Umwelt sich ständig ändert. Modellieren war eine Konsequenz daraus, nicht notwendigerweise die Ursache. Turing hingegen sah Lernen als eine Suche nach einem einzigen stabilen Ziel – den organisierten universellen Computer. Für Turing waren es die Symbolmuster im Gedächtnis/Speicher, die symbolischen „instruction tables“ (Befehlsfolgen), die die Geheimnisse des Geistes enthalten. Für Ashby waren es die Muster der Interaktionen und die Feedback-Schleifen zwischen System und Umwelt. In beiden Fällen war es möglich, eine Maschine zu bauen, die die jeweilige Vorstellung verkörperte. Für Turing wurde die Maschine der Inbegriff der symbolischen Simulation des Geistes, für Ashby dagegen war sie der Inbegriff der analogen Simulation von mentalem Verhalten. Dieser Unterschied kann am besten anhand des Briefwechsels zwischen Ashby und Turing verstanden werden.

7. Die Simulation des Homöostaten

Als Ashby im Frühjahr 1946 von seinem Dienst im *Royal Medical Corps* in Indien zurückkehrte, war er für ein Jahrzehnt mit Lernmechanismen befasst. Er begann darüber nachzudenken, eine Maschine zu konstruieren, um die Prinzipien, die er als essenziell für Anpassung hielt, zu demonstrieren. Im Jahr 1946 sind seine Aufzeichnungen voll mit mathematischen Formalismen für verschiedene, Stabilität suchende Systeme. Diese wechseln fallweise zu Diagrammen von einfachen und dann komplexeren elektrischen Schaltungen, um das Verhalten eines Sets von Gleichungen in einem System zu erkennen, in das direkt manuell eingegriffen werden kann.

Im Herbst 1946 schrieb Ashby an Turing über die adaptive Maschine, die er entwickelte.⁴⁹ Wie mit allen anderen Maschinen und Prozessen glaubte Turing, dass ein programmierbarer Computer wie ACE Ashbys Homöostaten modellieren könnte, wie er in seiner Antwort an Ashby am 20. November 1946 erklärte.

„The ACE will be used, as you suggest, in the first instance in an entirely disciplined manner, similar to the action of the lower centres, although the reflexes will be extremely complicated. The disciplined action carries with it the disagreeable feature, which you mentioned, that it will be entirely uncritical when anything goes wrong. It will also be necessarily devoid of anything that could be called originality. There is, however, no reason why the machine should always be used in such a manner: there is nothing in its construction which obliges us to do so. It would be quite possible for the machine to try out variations of behaviour and accept or reject them in the manner you describe and I have been hoping to make the machine do this. This is possible because, without altering the design of the machine itself, it can, in theory at any rate, be used as a model of any other machine, by making it remember a suitable set of instructions. The ACE is in fact, analogous to the ‚universal machine‘ described in my paper on computable numbers. This theoretical possibility is attainable in practice, in all reasonable cases, at worst at the expense of operating slightly slower than a machine specially designed for the purpose in question. Thus, although the brain may in fact operate by changing its neuron circuits by the growth of axons and dendrites, we could nevertheless make a model, within the ACE, in which this possibility was allowed for, but in which the actual construction of the ACE did not alter, but only the remembered data, describing the mode of behaviour applicable at any time. I feel that you would be well advised to take advantage of this principle, and do your experiments on the ACE, instead of building a special machine.“⁵⁰

Dieser Brief Turings hielt Ashby jedoch nicht ab, weiterhin die Entwicklung einer eigenen Maschine zu betreiben. Und tatsächlich trat er schon kurz nachdem er den Brief erhalten hatte mit einem der grundlegenden Elemente des Homöostat-Schalt-

kreises hervor. Der Unterschied in Perspektive und Zugang zwischen Ashby und Turing kann nun deutlich gemacht werden. Von besonderem Interesse ist Turings Insistieren darauf, dass der universelle Charakter von ACE bedeute, dass es keinen Sinn hätte, eine spezielle Maschine zu bauen, wie Ashby vorschlug, auch wenn er allgemein die Bedeutung des Lernens für intelligentes Verhalten einräumte. Es gibt auch eine Spannung zwischen dem streng determinierten Verhalten von ACE und dem adaptiven Verhalten des Gehirns. Zu überlegen, wie ACE den Homöostaten modellieren könnte, bringt die Unterschiede zwischen analoger und symbolischer Berechnung in den Fokus.

Während Turings Brief beiläufig von der Rede über ACE, die das Gehirn modelliert, übergeht zu ACE, die den Homöostat modelliert, sollten wir innehalten, um den Unterschied zu überlegen. Dieser Unterschied ist der Kern der Unterscheidung von ‚analog‘ versus ‚digital‘, insofern Turing nicht daran glaubte, dass der nichtprogrammierte Computer ein Modell des Gehirns sei. Dies deshalb, weil der Computer zu streng und zu diszipliniert sei, um gehirnnählich zu sein, obgleich er einige Ähnlichkeit mit den „lower centers“ (tieferen Schichten des Gehirns) aufweise. Nötig seien die Befehlstabellen, also das Programm, das die Gehirnmaschine definiert. Während Turing die vage Idee hatte, dieses Programm könnte auf konditioniertem Verhalten basiert werden, hatte er kaum eine Vorstellung, wie er seine Maschine programmieren könnte, damit sie sich so verhalte. Das heißt, ACE könnte eine symbolische Simulation des Gehirns ausführen, aber nur wenn man ihm die geeignete Organisation gibt.

Ashbys Homöostat ist jedenfalls selbst ein Modell des Gehirns. Er ist dies aber aufgrund seiner kausalen, und nicht seiner symbolischen Struktur. Was Turing jedoch im Homöostaten sah, war das Potenzial, das Gehirn indirekt zu modellieren. Weil der Homöostat eine Maschine ist, deren Design verstanden ist, sollte es möglich sein, ihn symbolisch auf dem ACE zu simulieren. Weil der Homöostat ein simuliertes Gehirn ist, würde seine Simulation auf dem ACE eine Simulation des Gehirns sein (gleichsam eine Generation entfernt). Die Ebene des Mittelbaren ist in diesem Fall der Homöostat selbst, als eine Theorie des adaptiven Verhaltens des Gehirns. Jegliche Simulation auf der ACE würde nur so gut sein können wie die ihm zugrunde liegende Theorie. Sie sagt uns etwas über den Homöostaten, aber sie sagt nur indirekt etwas über das Gehirn. Wir können uns jedoch weiter fragen, welche Vorteile die Direktheit der analogen Simulation hat, die uns der Homöostat bietet. Es gibt verschiedene Gründe, warum der analoge Homöostat ein besseres Modell des Gehirns ist als der auf der ACE simulierte, symbolische Homöostat.

Zunächst waren es praktische Fragen, mit denen Ashby und Turing zu tun hatten, dann methodologische Fragen der wissenschaftlichen Praxis. Und schließlich

gab es die Frage der Effektivität der beiden Simulationen als Demonstrationen. Diesen Fragen wende ich mich nun zu.

Zuerst stellt sich die Frage der Praktikabilität, die beiden Simulationen zu bauen. Es zeigte sich, dass der Homöostat weitaus leichter zu bauen war als die ACE. Pilot ACE war bis Mai 1950 nicht voll operationsfähig, zwei Jahre *nachdem* der Homöostat der *EEG Society* in Bristol vorgestellt worden war. (Von Neumanns Maschine war nicht voll operationsfähig bis 1952, obgleich dies mehr von ihren Teilen abhing, die laufend neu entwickelt wurden.) Darüber hinaus war der Homöostat eine weitaus ökonomischere Lösung als die ACE, auch wenn er nur eine Anwendung hatte. In diesem Sinn war Ashby ein intellektueller Entrepreneur, der auf seinem Küchentisch aus Restbeständen aus dem Krieg in einem selbstfinanzierten Projekt eine bedeutende Maschine baute. Während dies damals richtig war, ist heute das Gegenteil richtig, da Computer auf fast allen Schreibtischen stehen und es außerordentlich schwierig wäre, die analogen Schaltelemente, die Ashby verwendete, zu finden. Darin liegt der wahre Vorteil der Simulation durch symbolische Berechnung: Weil der Computer nahezu jedermanns Simulationsbedürfnisse abdecken kann und weil er in großer Zahl billig produziert werden kann, ist er trotz seiner sehr viel komplexeren Struktur billiger und leichter zu bauen als spezialisierte Modelle. Soviel war Turing 1946 klar. Weniger klar war die Qualität der symbolischen Simulation im Vergleich mit der analogen.

In seiner Analyse der Simulation argumentiert Winsberg,⁵¹ es sei häufig der Fall, dass die der Simulation zugrunde liegende mathematische Theorie nicht in Frage stehe, lediglich die lokalen Ergebnisse einer spezifischen Situation. Eine der großen Vorteile der Simulationen für diese Fälle sei die Fähigkeit, ein Modell zu visualisieren, eine Fähigkeit, die Hughes die mimetische Eigenschaft nennt.⁵² Dieser Vorteil beruhe auf der Fähigkeit der Forscherinnen und Forscher, beim Verstehen lokaler Situationen experimentelle Praktiken (besonders Beobachtung) und Labortechniken einzusetzen, statt nur auf mathematische Praktiken und analytische Techniken angewiesen zu sein. Dennoch hängen diese mimetischen Eigenschaften von der Entwicklung von Instrumenten innerhalb der Simulation ab, die diese Eigenschaften sichtbar machen. Die symbolischen Prozesse seien ziemlich opak und direkter Beobachtung nicht zugänglich. Dies bedeutet zusätzliche Ebenen der Mittelbarkeit der Interpretation und Visualisierung der Daten, die aus der Simulation gewonnen werden. Aus einer epistemischen Perspektive gerät man immer weiter weg von den fraglichen Phänomenen. Denken wir an den Homöostaten und seine ACE-Simulation. Mit dem Homöostaten kann man direkt interagieren, man kann ihn direkt beobachten, jegliche Interaktion mit der ACE-Simulation muss dagegen vorprogrammiert werden und die Ergebnisse aus dem symbolischen Output der Maschine interpretiert werden. Während wir heute ein einfaches Java Programm

schreiben können, das uns den Homöostaten farblich visualisiert, war dies 1946 kaum vorstellbar – was genau den Punkt belegt, dass jede solche symbolische Visualisierung eine weitere Ebene der Interpretation einführt.

Es gibt weitere wichtige praktische Fragen der Zeit, welche in indirekter Weise durch die technologische Computerentwicklung verdunkelt werden. Erstens gibt es Fragen der mathematischen Komplexität und der Rechenzeit. Das Verhalten der Schaltkreise des Homöostaten mit Gleichungen aus der Elektrotechnik oder der Physik zu simulieren würde für eine Maschine wie die ACE wahrscheinlich viel zu lange gedauert haben, um sie in Echtzeit auszuführen (gemeint ist hier die ungefähre Dauer, die der Homöostat selbst benötigt). Die Bedeutung von Echtzeit-Simulationen ist mit den oben erwähnten mimetischen Eigenschaften verknüpft. Und zwar derart, dass man mit einem Echtzeit-System auf Weisen interagieren kann, die selbst wieder in Nicht-Echtzeit-Systemen simuliert werden müssen. Nicht immer benötigt man solches in den Wissenschaften, aber es ist außerordentlich wichtig für das Verständnis bestimmter Phänomene. Echtzeit bezieht sich hier auf den zeitlichen Bezugsrahmen des Beobachters und seine Fähigkeit, auf dem System zu agieren, daraus Feedback zu gewinnen und die Beziehung zwischen beiden auszuwerten. Diese Feedback-Schleife muss für den Beobachter in Echtzeit erfolgen, auch wenn der zeitliche Bezugsrahmen des sinnlichen Phänomens stark beschleunigt (wie bei astronomischen Phänomenen, etwa einer Planetenbewegung) oder stark verlangsamt (wie bei Hochgeschwindigkeitsphänomenen wie der Proteinfaltung) ist. Die Bedeutung dieses Aspekts von analogen Modellen wird noch augenscheinlicher, wenn wir die Idee eines Regulators als Modell in Betracht ziehen.

8. Schlussfolgerungen: Simulation und Computationalismus

Von Neumann sah das Potenzial der Mathematik, detaillierte physikalische Modelle zu simulieren und schließlich die Steuerung enorm komplexer Systeme mit vielen dynamisch interdependenten mathematischen Berechnungen zu automatisieren, die für die Simulation dieser Systeme benötigt werden, und um Mathematikern zu ermöglichen, immer komplexere und detailliertere Simulationen zu entwickeln. Auch das Gehirn war ein natürliches Phänomen, das auf diese Weise simuliert werden könnte, wenn es denn einmal wissenschaftlich verstanden wäre. Der Unterschied zwischen Turing und von Neumann bestand darin, dass von Neumann der Ansicht war, dass Turings Konzept der Universellen Maschine nahezu nichts für unser Verständnis des Geistes beitrage, weil es uns nichts über das Gehirn sage. Ohne ein klares Verständnis der Operationen des Gehirns würde ein Computer niemals einen künstlichen Geist verkörpern können.

Von Neumann gelangte schließlich zur Auffassung, dass die nächste Aufgabe darin bestünde, die *detaillierte Mikrostruktur des Gehirns* zu modellieren. Für ihn war der Computer nicht selbst ein Gehirn, obwohl sein eigenes Design für den EDVAC Computer stark auf Analogien mit Gehirnstruktur und neuronalen Funktionen beruhte. Die wichtigen Aspekte des Computer-Designs waren technische Überlegungen der Skalierbarkeit, der Reliabilität und der numerischen Genauigkeit, zumal Computer mit größerem Gedächtnis/Speicher und höherer Prozessorgeschwindigkeit gebaut werden sollten. Dies deshalb, weil John von Neumann rasch erkannte, dass die Skala der Simulationen, die für Gehirne benötigt werden, sehr große Zahlen von Berechnungen erfordern; selbst kleine und seltene Rechenfehler würden zum Scheitern führen. Deshalb war er bezüglich eines raschen Fortschritts der Simulation des Gehirns zusehends skeptischer.

Turings Sichtweise des Computers war jene der Universellen Modelliermaschine. Aber statt sich um das Modellieren des Gehirns im physischen Detail zu sorgen, suchte er das Gehirn symbolisch in seinem Verhalten zu simulieren. Sowohl Turing als auch von Neumann bemühten sich um symbolische Simulationen des Gehirns, wählten aber sehr unterschiedliche Zugangsweisen. Ashby blieb während der 1950er Jahre den analogen Simulationen verschrieben, aber er scheint sie nach dem Scheitern von DAMS und den wachsenden Fähigkeiten der digitalen Computer aufgegeben zu haben.

Eine andere Möglichkeit, den Unterschied zwischen analogen und symbolischen Simulationen zu fassen, besteht darin, in Betracht zu ziehen, dass die Symbole einer Simulation innerhalb eines Computers nicht in derselben Weise mit der Welt verbunden sind, wie dies bei den Kausalstrukturen einer analogen Simulation der Fall ist. In Bezug auf den Homöostaten waren diese Verbindungen für Ashbys theoretische Sichtweisen des Geistes hoch relevant. Turings Anspruch, dass diese Unterschiede irrelevant seien, verraten somit die signifikante theoretische Disparität zwischen den beiden Wissenschaftlern. Die beste und einfachste Möglichkeit, den Unterschied zwischen Ashby und Turing zu charakterisieren, besteht hinsichtlich des Verhältnisses von System und Umwelt, welches für Turing irrelevant, für Ashby aber zentral war. Dies ist ein Unterschied von sowohl theoretischer als auch praktischer Bedeutung. Theoretisch geht Ashby von der reichen Komplexität der Verbindungen zwischen System und Umwelt aus, während Turing die Umwelt nur in den einfachsten und idealisiertesten Begriffen behandelt. Praktisch war Ashby daran interessiert, eine Maschine zu bauen, mit der ein Beobachter in Echtzeit direkt interagieren konnte. Solche direkte Interaktion war für seine pädagogischen wie rhetorischen Absichten und für seine Wünsche als Experimentator essenziell. Für Turing war der logische Beweis die endgültige Form des Zeigens; Experimentieren hatte keine Priorität, und die Details der physikalischen Maschinerie waren für ihn belanglos.

Die Unterschiede traten mit der Zeit sogar noch deutlicher hervor. Als Turing sich dann weiteren Überlegungen zu Lernen und Adaption zuwandte, äußerte er sich freimütiger über seine Gründe, Fragen allgemeiner Intelligenz, menschlichen Verhaltens und menschen-ähnliche Körper für Computer zu meiden.

(Aus dem Amerikanischen von Albert Müller)

Anmerkungen

- 1 Dieser Beitrag ging aus der noch ungedruckten Dissertation des Autors hervor. Vgl. Peter Asaro, *On the Origins of the Synthetic Mind. Working Models, Mechanisms, and Simulations*, Diss. Champaign, Urbana 2006.
Der Übersetzer dieses Textes (Albert Müller) verdankt einiges zur Frage der Übersetzbarkeit von ‚agency‘ dem Übersetzer von Andrew Pickering’s *Kybernetik und neue Ontologien*, Berlin 2007; vgl. Gustav Roßler, Nachwort des Übersetzers, in: ebd., 177–184. ‚Mind‘ wurde grundsätzlich mit ‚Geist‘ übersetzt, ‚memory‘ dort, wo es notwendig schien, mit ‚Gedächtnis/Speicher‘.
- 2 Klara von Neumann, Preface to John von Neumann, *The Computer and the Brain*, New Haven, CN 1958, viii.
- 3 Ted Newman, *Memories of the Pilot ACE*, in: *Computer Resurrection* 9 (Spring 1994), 12.
- 4 In diesem Beispiel habe ich versucht, X für das Modell stehen zu lassen, Y für den modellierten Gegenstand und Z für den „agent“, der die Modellierung ausführt. Es wäre günstiger „für Z“ in alle Fälle von Präpositionen aufzunehmen.
- 5 Peter Asaro, *Working Models and the Synthetic Method: Electronic Brains as Mediators Between Neurons and Behavior*, in: *Science Studies* 19/1 (2006), 12–34.
- 6 Warren S. McCulloch u. Walter Pitts, *A Logical Calculus of the Ideas Immanent in Nervous Activity*, in: *Bulletin of Mathematical Biophysics*, 5 (1943), 115–133. Wieder abgedruckt in Rook McCulloch, Hg., *The Collected Works of Warren S. McCulloch*, Bd. 1., Salinas, CA 1989, 343–361.
- 7 Matthias Scheutz, Hg., *Computationalism: New Directions*, Cambridge, MA 2002, ix.
- 8 Andrew Pickering, *The Mangle of Practice: Time, Agency, and Science*, Chicago 1995.
- 9 Eric Winsberg, *Simulated Experiments: Methodology for a Virtual World*, in: *Philosophy of Science* 70 (2003), 105–125.
- 10 Ebd., 111.
- 11 Ebd., 113.
- 12 Ebd., 115.
- 13 Ebd., 118.
- 14 Ebd., 122.
- 15 Russell Trenholme, *Analog Simulation*, in: *Philosophy of Science* 61 (1994), 115–131.
- 16 Ich nenne sie vereinfacht analoge Simulationen, werde sie aber sorgfältig von analogen Repräsentationen und analogen Berechnungen unterscheiden, wie es auch notwendig ist.
- 17 Asaro, *On the Origins of the Synthetic Mind*.
- 18 Trenholme, *Analog Simulation*, 119.
- 19 Allen Newell u. Herbert Simon, *Computer Science as Empirical Inquiry: Symbols and Search*, in: *Communications of the Association of Computing Machinery* 19 (March 1976), 113–126. Wieder abgedruckt in John Haugeland, Hg., *Mind Design: Philosophy, Psychology, Artificial Intelligence*. Cambridge, MA 1981, 35–66, hier: 41.
- 20 Trenholme, *Analog Simulation*, 119.
- 21 Beispielsweise zielt Searle in seinem berühmten Chinesisches-Zimmer-Gedankenexperiment genau auf diese semantischen Aspekte von Symbolen. Er möchte zeigen, dass Intentionalität ein essenzielles Element symbolischer Repräsentation ist und der Computationalismus oder wenigstens die starke

- Artificial Intelligence bei ihrer Erklärung scheitert; vgl. John Searle, Minds, Brains, and Programs, in: Behavioral and Brain Sciences 3 (1980), 417–424. Wieder abgedruckt in Haugeland, Mind Design, 282–306.
- 22 Ich benütze ‚Isomorphismus‘ nicht im strengen Sinn der Korrespondenz, sondern eher im weiteren Sinn von Ähnlichkeit, wie Teller dies vorschlägt; vgl. Paul Teller, Twilight of the Perfect Model, in: Erkenntnis, 55/3 (2001), 393–415.
- 23 Bei Black-Box-Simulationen interessieren uns interne Prozesse nicht, und so macht es keinen Unterschied, ob die Simulation symbolisch, kausal, oder durch Magie zustande kommt.
- 24 McCulloch u. Pitts, A logical Calculus.
- 25 John von Neumann, First Draft of a Report on the EDVAC (1945), Moore School University of Pennsylvania, in: IEEE Annals of the History of Computing, 15/4 (1993), 27–75.
- 26 John von Neumann, The General and Logical Theory of Automata, in: L. A. Jeffress, Hg., Cerebral Mechanisms in Behavior. The Hixon Symposium, New York 1951; wieder abgedruckt in: William Aspray u. Arthur Burks, Hg., Papers of John von Neumann on Computers and Computing Theory, Cambridge, MA 1987, 391–431.
- 27 John von Neumann, Letter to Norbert Wiener (1946), in: Proceedings of Symposia in Applied Mathematics, 52 (1997), 507.
- 28 John von Neumann, The Computer and the Brain, New Haven 1958, 54–55.
- 29 Trenholme, Analog Simulation.
- 30 Programmierer sprechen von einem Trunkierungsfehler, wenn die Anzahl der für ein Ergebnis benötigten Ziffern die Anzahl der Ziffern, die die Maschine darstellen kann, übersteigt. Bei einer Division führt dies zu einem Abschneiden oder Auf- und Abrunden der Dezimalpositionen, nachdem die letzte verfügbare Stelle benutzt wurde. Nehmen wir eine Maschine, die zwei zehnstellige Zahlen multiplizieren soll, aber nur zehn Ziffern zugleich darstellen kann. Das Ergebnis der Multiplikation wäre mindestens 18-stellig, aber die Maschine kann bloß die Hälfte davon darstellen. Daher wäre entweder überhaupt kein sinnvolles Ergebnis zu erzielen, oder die Maschine konvertiert die Zahlen in wissenschaftliche Notierung, schneidet sie ab und verliert damit Genauigkeit.
- 31 William Aspray, John von Neumann and the Origins of Modern Computing, Cambridge, MA 1990, 189.
- 32 Die frühen Rechenmaschinen waren entweder hoch spezialisiert, um eine einzelne Klasse von Funktionen auszuführen, oder mussten – wie die ENIAC – händisch programmiert werden, indem ganze Wände von Schaltern und Kabelnetze arrangiert wurden, ähnlich wie bei der händischen Telefonvermittlung. Beispielsweise benutzte Howard Aikens Mark I Automatic Controlled Sequence Calculator, der 1944 für IBM gebaut wurde, 72 Drehzähler, um Zahlen mechanisch zu speichern, hatte aber keine Mittel, um die Rechenbefehle selbst zu speichern – die ‚Funktion‘ war in der mechanischen Konfiguration des Geräts verankert.
- 33 Herman H. Goldstine u. John von Neumann, On the Principles of Large Scale Computing Machines, given on May 15, 1946 to Mathematical Computing Advisory Panel of the Navy’s Office of Research and Inventions, in: William Aspray u. Arthur Burks, Hg., Papers of John von Neumann on Computers and Computing Theory. Cambridge, MA 1987.
- 34 von Neumann, First Draft of a Report, 20, Hervorhebung durch den Autor.
- 35 Ebd., 24.
- 36 Ebd., 30.
- 37 Ebd., 30.
- 38 von Neumann, Letter to Norbert Wiener.
- 39 von Neumann, The Computer and the Brain, 54–55.
- 40 Ebd., 56.
- 41 Ebd., 68–69.
- 42 von Neumann, Letter to Norbert Wiener, 506–507.
- 43 Ebd., 507.
- 44 Ebd.
- 45 Alan M. Turing, On computable numbers, with an application to the *Entscheidungsproblem* (1936), in: Martin Davis, Hg., The Undecidable: Basic Papers on Undecidable Propositions, Unsolvable Problems and Computable Functions, New York 1965, 127.

- 46 Turing, On computable numbers.
- 47 Alan M. Turing, Lecture to the London Mathematical Society on 20 February, 1947, in: Darrel C. Ince, Hg., Collected Works of A. M. Turing: Mechanical Intelligence, New York 1992, 87–105, hier 106–7.
- 48 Alan M. Turing, Intelligent Machinery, Report to the Executive Committee of the National Physical Laboratory (1948), in: Darrel C. Ince, Hg., Collected Works of A. M. Turing: Mechanical Intelligence, New York 1992, 107-127, hier: 120.
- 49 Dieser Brief ist verschollen.
- 50 Alan M. Turing, Letter to W. Ross Ashby, Turing Archives, 11–20–46, Cambridge University.
- 51 Winsberg, Simulated Experiments.
- 52 R. I. G. Hughes, Models and Representations, in: Philosophy of Science 64 (1997), 325–336.