Zur Analogie zwischen einer elektronischen Rechenmaschine und dem Gehirn

Von

ERWIN ENGELER

University of Minnesota und IBM-Forschungslaboratorium Zürich

und

AMBROS P. SPEISER

Eidg. Technische Hochschule und IBM-Forschungslaboratorium Zürich

Die Arbeit von K. WEISS «Gibt es eine Analogie zwischen einer elektronischen Rechenmaschine und dem Gehirn?» [1] wurde angeregt durch die Diskussion im Anschluss an einen Vortrag im Physikalischen Kolloquium der ETH. Da einer von uns (A. P. SPEISER) der Referent dieses Vortrages war, sei es uns gestattet, zur Publikation Stellung zu nehmen.

Wir können uns mit wesentlichen Teilen der Ausführungen von K. WEISS nicht einverstanden erklären. Der Autor versucht, einen Beweis zu liefern für die Behauptung, dass der menschliche Geist prinzipiell mehr zu leisten vermag als je eine mathematische Maschine. Dazu werden metamathematische und naturwissenschaftliche Argumente vorgebracht. Zuerst soll auf die metamathematischen Argumente soweit eingegangen werden, als es im Rahmen eines kurzen, nichttechnischen Diskussionsbeitrages möglich ist.

Man betrachte ein formales System der Arithmetik mit vorgelegten (PEANOschen) Axiomen und formalen Schlussregeln. Es ist bekannt (seit GOEDEL, 1931 [2] und A. CHURCH, 1936):

1. Es gibt kein Entscheidungsverfahren für die Beweisbarkeit von arithmetischen Sätzen.
2. Das Resultat (1) lässt sich, mittels Arithmetisierung der Sprache, in der PEANO-Arithmetik als rein zahlentheoretischer Satz formulieren und beweisen.

Diese GOEDELschen Untersuchungen können von Gehirnen verstanden werden, das heisst dem menschlichen Geist ist es möglich, sich die Axiome und Schlussweisen der PEANO-Arithmetik anzueignen. Den formalen Beweisprozessen entsprechen irgendwelche (nicht näher bekannte) «Gehirnprozesse». Andererseits, so wird argumentiert, sei es prinzipiell nicht möglich, in modernen Rechenmaschinen* denselben formalen Beweisprozessen entsprechende Maschinenoperationen gegenüberzustellen.

Diese Behauptung beruht auf den folgenden Missverständnissen:

1. Die gesamten Maschinenabläufe können in der «BOOLEschen Logik» dargestellt werden; das heisst es gibt für jede Maschine einen BOOLEschen Ausdruck, aus dem sich durch Einsetzen der Eingabedaten die von der Maschine erzeugten Ausgabedaten errechnen lassen. [Von dieser Voraussetzung ausgehend, argumentiert der Autor wie folgt: Die BOOLEsche Logik ist entscheidbar. Wäre nun die PEANO-Arithmetik in der Maschine inkorporiert, so wäre auch diese entscheidbar (entgegen (1))]. Diese Voraussetzung ist aber schon bei relativ einfachen Maschinen nicht erfüllt; insbesondere kann man leicht Maschinen konstruieren, für welche es beweisbarerweise kein allgemeines Verfahren gibt, womit zu jedem Eingabewort entschieden werden kann, ob die Maschine nach Eingabe dieses Wortes je wieder stillsteht (und überhaupt ein Resultat produziert) [3], [6].

2. Die GOEDEL-Nummern der Ausdrücke der PEANO-ARITHMETIK lassen sich nur Ausdruck für Ausdruck im Gedächtnis der Maschine darstellen, und die Maschine hat keine systematische Übersicht über diesen Katalog. [Daraus würde folgen, dass in der Maschine, im Gegensatz zum Hirn, das Resultat (2) nicht herleitbar ist.] Dem ist entgegenzuhalten, dass die Numerierung der Formeln für den GOEDELschen Beweis gerade so gewählt werden *muss*, dass die Zuordnung der Nummern zu den Formeln (und ihre Umkehrung) auf einer TURING-Maschine berechenbar ist.

3. In bezug auf die Beweisbarkeit arithmetischer Sätze durch Rechenmaschinen müssen folgende zwei Fragen auseinandergehalten werden: Erstens, gibt es ein Maschinenprogramm, welches, wenn es beliebig lang laufen gelassen wird, in irgendeiner Reihenfolge schliesslich alle Theoreme der PEANO-Arithmetik produziert? [Ein solches Programm kann gefunden werden.] Zweitens, gibt es ein Maschinenprogramm, das für jeden vorgelegten arithmetischen Satz die Beweisbarkeit oder Unbeweisbarkeit entscheidet? [Ein solches Programm kann es beweisbarerweise nicht geben.] K. WEISS scheint diese beiden Fragen nicht genügend scharf auseinanderzuhalten und sieht in der Fähigkeit des Menschen, im Prinzip einmal alle beweisbaren arithmetischen Sätze herzuleiten, einen Gegensatz zur (zugegebenen) Unfähigkeit der Maschine, einen vorgelegten Satz zu *entscheiden*.

Die spezifischen metamathematischen Argumente, welche K. WEISS vorgebracht hat, sind also unhaltbar. Andererseits ist natürlich die Diskussion über den Vergleich Mensch-Maschine durchaus noch nicht abgeschlossen, und die Bemühungen der mathematischen Logiker um eine Präzisierung der Begriffe der Entscheidbarkeit und der Berechenbarkeit werden fortgesetzt. Es ist denkbar, dass Untersuchungen über die Funktionsweise des menschlichen Verstandes einen wertvollen Beitrag zu liefern vermögen. Hingegen ist zu bemerken, dass auch die scharfsinnigsten Vorschläge bis jetzt noch zu keiner Erweiterung des durch die TURING-Maschine abgegrenzten Bereiches geführt haben. Im Gegenteil haben sich alle diese Vorschläge als gleichwertige Definitionen desselben Bereiches erwiesen.

* Für diese prinzipielle Diskussion verstehen wir unter «Rechenmaschine» ihre präzise Idealisierung in Form einer TURING-Maschine [3], [4]. Die TURING-Maschine kann als solche gebaut werden [5] oder lässt sich auf einer normalen digitalen Rechenanlage simulieren.

Nach diesen fundamentalen Einwänden möchten wir zu einigen der Einzelheiten Stellung nehmen, mit denen wir nicht einverstanden sind.

1. Speicherkapazität des Gehirns: Das Gehirn hat etwa 10^{10} Neuronen mit je etwa 100 Synapsen, total also 10^{12} Synapsen. Die plausibelste Erklärung des Gedächtnisses ist, dass sein Sitz in den Synapsen liegt oder eng mit denselben verknüpft ist. Die Frage stellt sich also, wie viele Bits eine Synapse speichern kann. Wenn man annimmt, dass eine Synapse 100 verschiedene Stufen zu unterscheiden in der Lage ist, so gibt das eine Speicherkapazität von $\log_2 100 \approx 7$ Bits pro Synapse. Die Speicherkapazität des Hirns liegt also sicher nicht über $7 \cdot 10^{12}$ Bits. Für eine Zahl von 10^{20} fehlen nicht nur phänomenologische Anhaltspunkte, sondern auch jegliche Erklärung, wo eine so ungeheure Datenmenge (die dem Millionenfachen des Inhaltes der Library of Congress entspricht) untergebracht sein könnte und wie der Mechanismus des Zugriffes beschaffen wäre. Ausserdem ist nicht einzusehen, wozu ein solcher Speicher nützlich wäre. Um ihn zu füllen, müssten nämlich ein ganzes Leben lang Tag und Nacht ununterbrochen $5 \cdot 10^{10}$ Bits pro Sekunde von aussen eingegeben oder innerlich erzeugt werden, was ganz undenkbar ist. Man könnte sich höchstens noch vorstellen, dass die gleiche Information an vielen verschiedenen Stellen zugleich gespeichert ist. Die so entstandene Speicherkapazität wäre aber für den Informationsfluss nicht verwertbar und kann daher auch nicht für die Erklärung einer erhöhten Leistungsfähigkeit herangezogen werden.

2. Anzahl Operationen pro Sekunde: Es ist nicht möglich, dass gleichzeitig 10^{12} Synapsen im Betrieb sind (darunter wäre zu verstehen, dass in allen Synapsen dauernd Impulse weitergeleitet werden). Der Energiehaushalt des Gehirns könnte niemals genug Leistung für 10^{14}—10^{16} Operationen pro Sekunde aufbringen. Die Auffassung, dass das Wesen des Gedächtnisses in einem ununterbrochenen Zirkulieren aller Engramme besteht, ist schon lange verlassen worden. Anderseits führt eine grosse Maschine mit 10^5 Transistoren und 10^7 Magnetkernen pro Sekunde nicht 10^6, sondern vielleicht 10^9 Operationen aus (darunter sind elementare, gleichzeitig ablaufende Operationen in Transistoren oder Kernen zu verstehen).

3. Beschreibung der Neuronen mittels der BOOLEschen Algebra: Ein Neuron hat einen Ausgang, der entweder feuert oder nicht feuert, und (ungefähr) 100 Eingänge, bei denen entweder ein Impuls ankommt oder nicht ankommt. Der Ausgang ist also, in einem einzelnen Zeitmoment betrachtet, eine binäre Funktion von 100 binären Variablen. *Jede* solche Funktion, wie kompliziert sie auch immer sei, kann mittels eines BOOLEschen Ausdrucks definiert werden. Der Fall, dass das Feuern von der geometrischen Anordnung der angeregten Synapsen abhängt, lässt sich leicht einbeziehen. Es ist anzunehmen, dass die Form dieses Ausdruckes nicht konstant ist, sondern unter dem Einfluss der Gehirnfunktion, und besonders in Abhängigkeit der Vorgeschichte der Erregung der Synapsen, zeitlich variiert. Das ist aber eine Eigenschaft, die auch der TURING-Maschine (die als solche mit den heutigen technischen Mitteln leicht gebaut werden kann) zukommt.

4. «Verdrahtung» des Gehirns: Der Verfasser sagt, die «Verdrahtung» des Gehirns repräsentiere eine Logik, die viel komplizierter sei als die BOOLEsche Logik. Dazu ist folgendes zu bemerken: In einer sequentiellen Maschine (also einer Maschine, in welcher Elementarvorgänge mit zeitlicher Staffelung ablaufen) lässt sich, wie wir

oben gesagt haben, im allgemeinen das Resultat nicht durch einen BOOLEschen Ausdruck in den Eingangsvariablen darstellen; das gilt für das Gehirn in gleicher Weise wie für eine Rechenanlage. Wenn man hingegen einen einzelnen Zeitmoment betrachtet, so führt jede denkbare Verbindung von Elementen, die sich durch die BOOLEsche Algebra beschreiben lassen, auf ein neues Element, das sich auch wieder durch die BOOLEsche Algebra beschreiben lässt.

5. Notwendigkeit der Programmierung: Es stimmt keineswegs, dass eine Maschine nichts tun kann, ohne vorher programmiert worden zu sein. An Stelle eines Programmes kann eine fertige Verdrahtung vorgesehen werden, die in der Fabrik hergestellt wird. Weitere Fähigkeiten kann sich die Maschine durch ihre Lernfähigkeit, also in Wechselwirkung mit der Aussenwelt während der Erfüllung ihrer Aufgabe, aneignen, was hinlänglich bestätigt worden ist.

* * *

Abschliessend möchten wir festhalten: Trotz grössten Bemühungen an vielen Orten ist es bis jetzt nicht gelungen, zu beweisen, dass das Gehirn zu Denkabläufen fähig ist, die prinzipiell von keiner Maschine geleistet werden können, und auch der Verfasser hat diesen Beweis nicht erbracht. — Hingegen wollen wir gerne zugeben, dass es sehr fraglich ist, ob es jemals gelingen wird, eine dem menschlichen Hirn äquivalente Maschine zu bauen; denn die Anzahl der dazu nötigen Schaltelemente und Verbindungen ist ausserordentlich hoch, und es ist sehr schwierig, über den Aufbau des Gehirns einen hinreichend detaillierten Aufschluss zu gewinnen.

Literatur

1. K. WEISS: Gibt es eine Analogie zwischen einer elektronischen Rechenmaschine und dem Gehirn? Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich, *108* (1963), S. 359—371.
2. K. GÖDEL: Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme. Monatshefte für Mathematik und Physik, *38* (1931), S. 173—198.
3. A. M. TURING: On Computable Numbers. Proc. London Math. Soc., Series 2, *24* (1936): S. 230 bis 265.
4. S. C. KLEENE: Introduction to Metamathematics. Princeton, N. J., 1952.
5. E. F. MOORE: A Simplified Universal Turing Machine. Proc. ACM, Sept. 8, 1952 (1953).
6. H. HERMES: Aufzählbarkeit, Entscheidbarkeit, Berechenbarkeit. Springer, Berlin (1961).