

# Künstliche Intelligenz - Eine Frage der Zeit?

M. Burkhart

Kaum jemand scheint heutzutage ernsthaft daran zu zweifeln, dass es Wissenschaftlern eines Tages gelingen wird, intelligente Roboter zu entwickeln. Nachdem die Maschinen den Menschen in körperlicher Hinsicht schon lange ersetzt haben, dringen sie immer mehr auch in Bereiche vor, die bisher dem Menschen auf Grund seiner einzigartigen Intelligenz vorbehalten waren. In der Science-Fiction gehört künstliche Intelligenz zum ganz normalen Alltag. Steven Spielbergs Film „A. I.“ diskutiert bereits die gesellschaftlichen Probleme, die auftreten könnten, wenn sich intelligente, fühlende „Mechas“ unter die Menschen mischen. „The Matrix“ beschreibt eine Zukunft, in welcher sich die Maschinen emanzipieren und von der Bevormundung durch die Menschen befreien. Stanislaw Lem unterstreicht diese Vision mit den Worten von GOLEM, dem hyperintelligenten Computer, welcher Vorlesungen für die dummen Menschen hält: *„Sollte diese Tendenz sich noch wenigstens hundert Jahre lang fortsetzen, so werdet ihr selbst am Ende die dümmsten Punkte auf dem mit technischer Raffinesse ausgestatteten Erdboden sein.“* ([1] Seite 151). Ausdrücke wie „Der Computer ist schuld“, „Jetzt lernt der PC sprechen“ oder „Mein Computer *spinnt*“ zeigen, dass dem Computer umgangssprachlich bereits menschliche Eigenschaften zugeschrieben werden. Daraus resultiert eine grosse gesellschaftliche Unsicherheit darüber, was Computer eigentlich können und wie sie funktionieren.

## Algorithmen

Computer arbeiten Algorithmen ab. Doch was sind Algorithmen? Was können sie? Alan M. Turing hat in den 30er Jahren ein mathematisches Modell für den intuitiven Begriff des Algorithmus geschaffen [2]. Kern dieses Modells ist die sogenannte Turingmaschine. Diese Turingmaschine besitzt interne Zustände, in welchen sie sich befindet, und hat ein (unendliches) Speicherband zur Verfügung, welches sie beschreiben und lesen kann. Aufgrund von Regeln wechselt die Turingmaschine in andere Zustände und führt dabei Aktionen auf dem Speicherband aus. Ein Algorithmus ist nun alles, was eine solche Turingmaschine berechnen kann.

Seit Turings Definition der Turingmaschine sind einige Jahrzehnte vergangen und die ersten Computer wurden tatsächlich gebaut. Sie haben unsere Gesellschaft erobert und unsere Arbeit sowie unseren Alltag verändert. Ihre Leistung verdoppelt sich gemäss Moores Gesetz seit Dekaden ungefähr alle 18 Monate. Auch die Speicherkapazität wächst exponentiell. Und trotzdem ist jeder heutige Computer mitsamt seinen Programmen nichts weiter als eine komplizierte Turingmaschine. Im Sinne der Berechenbarkeitstheorie sind heutige Computer und Turingmaschinen äquivalent. Das bedeutet, dass ein Computer nicht mehr und nicht weniger berechnen kann, als Turings Modell von Algorithmen. Turing war nicht der einzige, der sich mit der Definition des Begriffs Algorithmus auseinandersetzte. Andere Systeme, wie Lambda-calculus, rekursive Funktionen oder Markov-Algorithmen wurden vorgeschlagen. Es stellte sich jedoch heraus, dass alle diese Modelle mathematisch äquivalent sind. Das heisst, sie beschreiben denselben Begriff des Algorithmus wie die Turingmaschine.

## Die Optimisten

Die Frage drängt sich auf, ob unser Gehirn nicht auch bloss ein Computer ist, der einen komplizierten Algorithmus ausführt? Bereits im Jahre 1680 hegte Leibniz die Phantasie vom „Gott als Uhrmacher“ und dem Gehirn als „Uhrwerk“. Er stellte sich vor, wie man

schrumpfen und in das Gehirn eines Menschen einsteigen könnte. Dort würde man all die Pumpen, Kolben, Zahnräder und Hebel arbeiten sehen. Nur durch deren Beschreibung könnte die gesamte Funktion des Menschen erklärt werden, ohne lästige Begriffe wie Geist, Wille oder Bewusstsein. Die heutigen Anhänger der starken KI sprechen natürlich nicht von mechanischen Bauteilen, doch unterscheiden sich ihre Ansichten nicht wesentlich von der von Leibniz. Sie sind überzeugt, dass unser Gehirn nichts weiter als ein riesiger neuronaler Computer ist. Sämtliche Hirnfunktionen liessen sich im Prinzip algorithmisch simulieren. Früher oder später werde man die der Intelligenz zu Grunde liegenden Algorithmen gewiss entdecken oder das Bewusstsein werde sich von selbst einstellen: „*Programme und Maschinen werden auf gleiche Weise Empfindungen erwecken: als Nebenprodukt ihrer Struktur, der Art wie sie organisiert sind - nicht durch direkte Einprogrammierung.*“ ([3], Seite 721). Wenn man den Computern etwas mehr Zeit gibt komplexer und grösser zu werden, werden sie dann eines Tages auch Intelligenz und ein Bewusstsein entwickeln? Warum auch nicht? Die Evolution soll den Menschen und sein Gehirn schliesslich nach demselben Prinzip hervorgebracht haben.

### **Ein Loch ist im Eimer**

1931 hat Kurt Gödel die mathematische Welt mit seinem verblüffenden Unvollständigkeitssatz erschüttert. Dieser Satz sagt aus, dass jedes widerspruchsfreie mathematische System, welches mächtig genug ist einfache arithmetische Aussagen zu machen, wahre Aussagen enthält, die sich weder beweisen noch widerlegen lassen. Solche Aussagen gelten als „unentscheidbar“. Übertragen auf die theoretische Informatik bedeutet dies, dass Probleme existieren, die zwar lösbar sind, deren Lösung aber nicht algorithmisch herbeigeführt werden kann. Beispiele für solche Probleme sind das Halteproblem, das Lösen diophantischer Gleichungen, die Überdeckung der Ebene mit Vielecken oder sämtliche Probleme der nicht-rekursiven Mathematik. Was bedeutet dieses „Loch im Eimer“ der Algorithmik nun für die Suche nach der künstlichen Intelligenz? In seinem Buch „*Computerdenken - Des Kaisers neue Kleider*“ [4] zeichnet der Mathematikprofessor Roger Penrose ein interessantes Bild von der Erkenntnis mathematischer Wahrheit. Er sagt, dass sich viele Wahrheiten nur aufgrund einer „Einsicht“ als wahr ansehen lassen, genau so wie eine Gödelsche Aussage von allen Mathematikern als wahr anerkannt wird, obwohl sie eben nicht beweisbar ist. Gemäss dem *Reflexionsprinzip* der Logik kann sich der Mensch der „Bedeutung“ von Axiomen und Ableitungsregeln bewusst werden und daraus Erkenntnisse gewinnen, aus welchen sich wiederum mathematische Aussagen konstruieren lassen, die auf formalem Wege (über Axiome und Regeln) nicht erreicht werden können. Für Penrose existiert eine platonische Ideenwelt der Mathematik, welche nur unserem Geist zugänglich ist. Viele Lösungen von Problemen lassen sich nur auf dieser „höheren“, semantischen, geistigen Stufe finden: „*Der Begriff der mathematischen Wahrheit geht über das ganze Konzept des Formalismus hinaus. An ihr ist etwas Absolutes und 'Gottgegebenes'.*“ (Seite 109). Wenn sich schon die klar definierte und logische Welt der Mathematik den Algorithmen verschliesst, wie viel mehr dann die Prozesse, die beim menschlichen Denken ablaufen! Gerade dem Bewusstsein und der Intelligenz, die laut Penrose nötig sind um mathematische Wahrheit zu erkennen, müssen starke nicht-algorithmische Prozesse zu Grunde liegen. Man weiss bis heute immer noch sehr wenig über solche Prozesse. Sicher ist einzig, dass Computer in ihrer heutigen Form keinerlei Zugang zu solchen Vorgängen haben, weil sie in ihrer algorithmischen Beschaffenheit gefangen sind.

### **Neuronale Netze**

Wenn nun herkömmliche Computer streng algorithmisch funktionieren, wir Menschen aber auch nicht-algorithmische Leistungen vollbringen können, muss das Geheimnis der Intelligenz wohl irgendwo im menschlichen Gehirn verborgen sein!? Die Neuroinformatik analysiert die Funktionsweise des Gehirns und versucht, diese mittels künstlicher Modelle nachzuahmen. Die Grundbausteine unseres Gehirns sind Neuronen. Diese Neuronen sind

spezielle Nervenzellen, die durch Synapsen miteinander verbunden sind. Ein ganzes Netz solcher Neuronen kann nun Berechnungen durchführen. Zu den am besten verstandenen Bereichen im Gehirn gehört der Weg der visuellen Informationen. Die Kette der Verarbeitung beginnt, wenn Licht auf die Retina fällt. Schritt für Schritt werden danach die Signale durch verschiedene Regionen im Hirn geschleust, die Informationen wie Farbe, Kontrast, Kanten, räumliche Eigenschaften, Orientierung oder Bewegung extrahieren. Nur sind alle diese Komponenten der visuellen Informationsverarbeitung rein algorithmische Funktionen. Zusätzlich sind sie sehr lokal und jeweils auf eine kleine Teilaufgabe spezialisiert. Es können nirgends im Gehirn Regionen gefunden werden, in denen alle diese Teil-Informationen zusammenfließen und als Ganzes wahrgenommen werden. Trotzdem erfahren Menschen die sie umgebende Umwelt als eine nahtlose Einheit. Dieses bis anhin ungelöste Problem ist unter dem Namen „Binding Problem“ bekannt. Alle Simulationen von neuronalen Strukturen beschränken sich auf die Realisierung solcher algorithmischer Komponenten. Die gesuchte Intelligenz müsste allerdings genau an der Stelle in Erscheinung treten, wo die Resultate dieser Komponenten zusammengefügt werden. Ein weiteres Problem der neuronalen Netze besteht darin, dass sie sich prinzipiell beliebig genau mit herkömmlichen Computern simulieren lassen. Das bedeutet, dass neuronale Netze keine nicht-algorithmischen Funktionen ausführen können. Wenn sie das tatsächlich könnten, könnte jede Turingmaschine das auch, indem sie einfach ein neuronales Netz simulieren würde. Die Performance-Einbusse der Software-Simulation ist diesbezüglich vollkommen irrelevant. Ein einfacher intelligenter Gedanke, welcher zehn Jahre benötigen würde, um von einem Computer gedacht zu werden, wäre dennoch ein intelligenter Gedanke.

### **Quantencomputer**

Im Gegensatz zu den Bits der klassischen Computer, die immer exakt 0 oder 1 darstellen, können die Qubits (quantum bits) der Quantencomputer eine Superposition sämtlicher möglicher Zustände annehmen. Eine Berechnung mit diesen Qubits würde dann alle diese Zustände zugleich miteinbeziehen. Das Resultat einer solchen Berechnung ist jedoch probabilistisch und verlangt nach speziell auf Quantencomputer angepassten Algorithmen. Die enorme Parallelität eines Quantencomputers macht es potentiell möglich, NP-schwierige Probleme in polynomieller Zeit zu lösen. Auf der Suche nach künstlicher Intelligenz führt diese enorme Steigerung an Rechenkraft jedoch nicht weiter, denn nach David Deutsch kann ein Quantencomputer keine nicht-algorithmischen Probleme lösen, womit er wiederum äquivalent zu einer (wenn auch nicht-deterministischen) Turingmaschine wäre.

### **Physik der Zukunft**

Vielleicht weiss die Wissenschaft einfach noch zuwenig über die detaillierte Funktionsweise des Gehirns. Möglicherweise muss man noch tiefer vordringen, noch kleinere Strukturen analysieren, um auf die gesuchten Intelligenzmuster zu stossen. Penrose sucht auch in der Quantenmechanik nach Erklärungen für mögliche „geistige“ Gehirnfunktionen. Doch selbst der Quantenzustand eines Systems entwickelt sich gemäss der Schrödingergleichung vollkommen deterministisch und lässt daher keinen „freien Willen“ als Steuerzentrale zu. Der einzige Indeterminismus der Quantentheorie tritt auf, wenn die Wellenfunktion eines Zustandes aufgrund einer „Beobachtung“ in eine probabilistische Alternative kollabiert. Doch auch dabei geschieht nichts Magisches und Penrose kommt zum Schluss, dass die Zukunft neue Theorien liefern müsse, damit man dem Phänomen des Geistes auf die Spur kommen könne: *„Wir wissen, dass auf der submikroskopischen Ebene die Quantengesetze herrschen; aber auf der Ebene von Billardkugeln gilt die klassische Physik. Irgendwo dazwischen, würde ich behaupten, müssen wir das neue Gesetz begreifen, damit wir sehen, wie die Quantenwelt in die klassische übergeht. Auch denke ich, dass wir dieses neue Gesetz brauchen, wenn wir überhaupt jemals geistige Phänomene verstehen wollen!“* ([4] Seite 291)

### **Ich und mein Gehirn**

Die starke KI ist eine sehr materialistische These. Nach ihr beruhen alle geistigen Prozesse einzig auf der Vernetzung von Neuronen. Es wird mit dem Märchen der „Seele“ aufgeräumt, man braucht keine metaphysische Welt mehr, es lässt sich alles in „Hardware“ realisieren. Die allgemeine Ansicht, dass Menschen ihre Handlungen durch einen „freien Willen“ beeinflussen können, wird als trügerisch angesehen. Alles was sie tun ist vorbestimmt. Entgegen dem Mainstream der materialistischen Theorien hat der Hirnforscher und Nobelpreisträger John C. Eccles seine Theorie des *dualistischen Interaktionismus* [5] entwickelt. Dualistisch darum, weil Eccles zusätzlich zur physischen Gesamtheit des Gehirns den Geist als eigenständige, real existierende, immaterielle Instanz betrachtet. Interaktionismus bedeutet dabei, dass diese beiden Entitäten - materielles Gehirn und immaterieller Geist - miteinander kommunizieren. Der Geist beinhaltet die Identität und den freien Willen eines Menschen. Als Bindeglied zwischen Gehirn und Geist zieht Eccles die Quantenmechanik heran: *„Die Hypothese der Wechselwirkung von Geist und Gehirn lautet, dass mentale Ereignisse über ein quantenmechanisches Wahrscheinlichkeitsfeld die Wahrscheinlichkeit der Emission von Vesikeln aus präsynaptischen Vesikelgittern ändern.“* (Seite 114). Weil bei dieser Interaktion nur Information und keine Energie fließt, wird der Energieerhaltungssatz der Physik nicht verletzt. Der Geist benutzt den Körper mitsamt seinem Gehirn nur als Werkzeug, um mit der sichtbaren Welt zu interagieren. Unsere Neuronen bereiten die von den Sinnen gewonnenen Informationen durch Berechnungen auf und machen sie dann dem Geist zugänglich. Unser freier Wille entwickelt daraufhin vielleicht eine mentale Absicht, welche wiederum im Gehirn Wahrscheinlichkeitsfelder verändert und so Einfluss auf das Feuern bestimmter Neuronen nimmt. Wie ein Puppenspieler alle Fäden seiner Marionette in der Hand hält, so laufen beim Geist die Resultate aller Hirnregionen zusammen: *„Man muss erkennen, dass die vollständigen visuellen Bilder im Geist erfahren werden, der sie aus der Analyse in der Sehrinde zusammensetzen scheint.“* (Seite 258). Die Lösung des „Binding Problem“ wäre somit auf einer geistigen Ebene angesiedelt.

Die Intelligenz und das Bewusstsein eines Menschen scheinen in einer transzendenten, unserer Physik bis anhin nicht zugänglichen Dimension, zu Hause zu sein. Die Funktionsweise eines metaphysischen Geistes ist nicht an die bekannte Physik oder gar die Algorithmik gebunden. Dabei kann das Gehirn sehr wohl ein vollkommen deterministisch und algorithmisch funktionierender neuronaler Computer sein, denn es ist schliesslich nur das Gefäß eines autonomen, nicht-algorithmischen Geistes. Die Strukturen der Intelligenz sind damit aber für die materialistische Wissenschaft bis auf Weiteres unauffindbar geworden.

## **literatur**

[1] Stanislaw Lem: Also sprach GOLEM, suhrkamp taschenbuch 1266, 1986

[2] Alan M. Turing: On computable numbers, with an application to the Entscheidungsproblem, Proc. London Math Soc.

[3] Douglas R. Hofstadter: Gödel Escher Bach - Ein Endloses Geflochtenes Band, dtv, München, 1999

[4] Roger Penrose: Computerdenken, Des Kaisers neue Kleider, Spektrum der Wissenschaft, Heidelberg, 1991

[5] John C. Eccles: Wie das Selbst sein Gehirn steuert, Piper Verlag GmbH, München, 1994