

Ist Künstliche Intelligenz seriöse Wissenschaft?

Maxi Bayer
{maxib@uni-koblenz.de}

Universität Koblenz-Landau
Institut für Informatik

Zusammenfassung Ziel dieser Arbeit¹ ist es, herauszustellen, ob und inwiefern es sich bei der Erforschung und Entwicklung von sogenannter *Künstlicher Intelligenz (KI)* tatsächlich um eine Wissenschaft handelt, die sich Mitteln bedient, die den Anspruch auf Seriösität rechtfertigen. Auf dem Weg zu einer Antwort soll eingangs der Begriff „Intelligenz“ geklärt, sowie ein Überblick über mögliche Umsetzungen gegeben bzw. verschiedene Kategorien zur Umsetzung von Künstlicher Intelligenz vorgestellt werden. Durch einen kleinen Rückblick auf die Anfänge und die entscheidenden Entwicklungsschritte des Forschungsgebietes der KI bis zum heutigen Stand der Wissenschaft soll eine weitere Grundlage zur Einschätzung bieten. Schlußendlich werden Autoren vorgestellt, die ihre Sichtweise auf die Problematik der KI als Wissenschaft verdeutlichen, indem bestehende Ansätze, Realisierungen sowie Probleme aufgezeigt und kritisch bewertet werden.

1 Einführung

Wenn man von *Künstlicher Intelligenz* spricht, sollte vorher geklärt sein, was eigentlich unter *Intelligenz* zu verstehen ist. Hierzu gibt es vielfältige Definitionen, die jedoch die gleiche Kernaussage besitzen. „Der Begriff Intelligenz umfaßt allgemein die geistigen Fähigkeiten des Menschen. Dazu gehören insbesondere Denkvermögen, Auffassungsgabe, Rationalität, Logik, Urteilsvermögen und Kreativität.“ [8]

Im Speziellen wird *Intelligenz* als die Fähigkeit beschrieben, eine Leistung in einer unbekanntem Situation zu erbringen, indem die Zusammenhänge der Lage spontan verstanden werden.[6] Eine *Künstliche Intelligenz* soll also diese genannten menschlichen Attribute und Fähigkeiten in sich vereinen, ohne selbst „menschlich“ sondern vielmehr vom Menschen selbst geschaffen zu sein.

Wie bzw. in welcher Form KI realisiert werden soll, läßt sich in vier Kategorien einordnen. So sei KI ein System, das (a)...denkt wie ein Mensch oder (b)...handelt, wie ein Mensch oder (c)...rational denkt oder (d)...rational handelt. [7]

In diese Kategorien lassen sich theoretische Ansätze einordnen, die ebenfalls die unterschiedlichen Vorstellungen von der Realisierung und die differenzierten Herangehensweisen verdeutlichen. Im Folgenden wird zu jeder Kategorie ein entsprechender Ansatz vorgestellt.

¹ erstellt im Rahmen des Informatik-Seminars „Logik auf Abwegen: Irrglaube, Lüge, Täuschung“ unter Leitung von Jun.-Prof. Dr. Bernhard Beckert an der Universität Koblenz im Sommersemester 2004

1.1 (a) Kognitives Modellieren

Das interdisziplinäre Gebiet der Kognitiven Wissenschaften (= Wissenschaften der „Erkenntnis“) bringt Computermodelle aus der KI mit den experimentellen Techniken der Psychologie zusammen, um genaue und überprüfbare Theorien über die Arbeitsweise des menschlichen Denkens zu konstruieren. Wenn man erreichen möchte, dass ein Programm *denkt, wie ein Mensch*, benötigen wir eine Methode, um festzustellen, *wie* Menschen denken. Es gibt zwei Möglichkeiten, um einen Einblick in die Arbeitsweise des menschlichen Verstandes zu erhalten. Durch Selbstbeobachtung könnten wir versuchen, unsere Gedanken aufzufangen, während sie vorbeiziehen oder man führt psychologische Experimente durch - das ist der fundiertere Weg. Wenn die Theorie des Denkens bzw. des Verstandes erst einmal ausreichend präzisiert worden ist, scheint es möglich, diese Theorie auch als Computerprogramm auszudrücken. Falls das Ein- und Ausgabe- sowie das zeitliche Verhalten korrespondierend zu dem eines Menschen ist, wäre das ein Beweis dafür, dass ein gewisser Teil des Mechanismus innerhalb des Programms genauso gut in *uns* arbeiten könnte.

Die Unterscheidung der beiden Ansätze der KI und den Kognitiven Wissenschaften erlaubt beiden, sich schneller zu entwickeln. Diese beiden Felder befruchten sich sozusagen weiterhin gegenseitig, heute vor allem auf dem Gebiet des Rechnersehens und der natürlichen Sprache. [7]

1.2 (b) Der Turing Test

Der Turing-Test ist nach seinem Entwickler, dem britischen Mathematiker und Kryptiker Alan Turing (1912 - 1954) benannt. Dieser Test sollte eine befriedigende Definition von *Intelligenz* bereitstellen. Anstatt eine lange und wahrscheinlich höchst kontroverse Liste an benötigten Voraussetzungen für Intelligenz vorzuschlagen, regte er diesen Test an, der auf Unverwechselbarkeit von intelligenten Wesen (= den Menschen) basieren soll. Dabei stellt eine Testperson einige schriftliche Anfragen, die danach dem Computer übergeben werden. Der Test für die Maschine ist bestanden, wenn die Testperson aus den ihr vorgelegten Antworten nicht sagen kann, ob diese von einem Menschen oder einem Computer stammen. Es ist offensichtlich, dass es, um diesen Test zu bestehen, eines größeren Programmieraufwands bedarf und folgende Fähigkeiten von Seiten des Programmes voraussetzt:

- **natürliche Sprachverarbeitung** um erfolgreich in Englisch zu kommunizieren
- **Wissensrepräsentation** um zu speichern, was es weiß oder hört
- **automatische Schlußfolgerungen** um die gespeicherten Informationen nutzen zu können, um Fragen zu beantworten und neue Entscheidungen zu treffen
- **maschinelles Lernen** um neue Zustände/Umwände zu übernehmen und um Muster zu erkennen und zu extrahieren

Dieser Test ist auch fünfzig Jahre nach seiner Entwicklung noch relevant für heutige KI-Forscher. Jedoch stellt es sich als äußerst schwierig heraus, ihn zu bestehen, vor allem, da man glaubt, es sei noch wichtiger, die zugrundeliegenden Prinzipien von Intelligenz zu erforschen anstatt sie einfach zu duplizieren.[7]

1.3 (c) Die Gesetze des Denkens - „Laws of Thought“

Dieser Ansatz geht zurück bis in die Antike und Aristoteles war einer der ersten, der versuchte, das „Denken“ genauer zu definieren. Sein sogenannter Syllogismus² stellte Muster bereit, um aus korrekten Annahmen und Voraussetzungen wiederum korrekte Schlußfolgerungen ziehen zu können. Beispielhaft dafür ist die bekannte Satzfolge „Sokrates ist ein Mensch; alle Menschen sind sterblich; also ist Sokrates sterblich“. Man schlug vor, die Gesetze des Denkens als grundlegende Operation zu betrachten, die zur Steuerung unserer Gedanken ergo des menschlichen Denkens angenommen werden kann. Das Studium dieser Operationen führte zur Entstehung des Fachgebietes der Logik. Bereits im 19. Jahrhundert entwickelten Logiker eine präzise Notation für alle Aussagen, die man über jegliche Dinge in dieser Welt und deren Zusammenhänge untereinander machen konnte. (Im Gegensatz zu den üblichen arithmetischen Notationen, die hauptsächlich Gleichheits- oder Ungleichheitsaussagen über Zahlen liefern.)

Schon seit 1965 existierten Programme, die im Prinzip jegliches entscheidbare Problem - beschrieben in logischer Notation - lösen konnten. Innerhalb der KI hofft man, auf solche logischen Programme bauen zu können, um intelligente Systeme zu kreieren. [7]

1.4 (d) Der rationale Agent

Ein Agent³ ist jemand, der agiert, also etwas tut. Ein rationaler Agent soll so handeln, so dass er das beste Resultat erzielt. Man erwartet von einem rationalen Agenten, dass er Attribute besitzt, die ihn klar von den reinen „Programmen“ abgrenzt. Dazu gehören Berechnungen unter autonomer Kontrolle, Wahrnehmung ihrer Umgebung, das Bestehen über eine verlängerte Zeitperiode hinweg, die Aufnahme von Veränderungen und in der Lage zu sein, Ziele anderer aufzugreifen.

Der im vorigen Absatz vorgestellte logische Ansatz zu den „Laws of Thought“ ist zum Teil in einen Agenten integriert. Und zwar in der Weise, dass *eine* Möglichkeit, sich für ein bestimmtes (rationales) Handeln zu entscheiden, über den Weg führt, logisch zu ergründen, mit welcher Alternative man am ehesten zum Ziel kommt und genau diese dann anzuwenden. Korrekte Schlußfolgerungen sind jedoch nicht alles beim rationalen Handeln. Es gibt Situationen, in denen kann man nichts Korrektes tun, muss aber trotzdem handeln, d.h. rationales Handeln beinhaltet nicht immer Schlußfolgerungen. Zum Beispiel ist das sofortige Zurückschrecken von einer heißen Herdplatte eine Reflexhandlung, die normalerweise von mehr Erfolg gekrönt ist, als eine langsamere Handlung, die nach gründlicher Überlegung getätigt wird.

Wir benötigen die Fähigkeit, Wissen darzustellen und damit zu argumentieren, weil es uns in die Lage versetzt, gute Entscheidungen in einer umfassenden und großen Anzahl von Situationen zu treffen. Wir brauchen ebenso die Fähigkeit, verständliche Sätze in natürlicher Sprache zu produzieren, da uns das Wiedergeben dieser Sätze hilft, in unserer komplexen Gesellschaft zurechtzukommen. Wir benötigen den Prozeß des Lernens,

² Die Syllogismen bilden den Kern der klassischen Logik des Aristoteles. Sie sind ein Katalog von Typen von logischen Schlussfolgerungen. Diese Folgerungen sind immer nach dem gleichen Muster aufgebaut. Jeweils zwei Prämissen (Voraussetzungen), genannt *Obersatz* und *Untersatz* ergeben eine Konklusion (Schlußfolgerung).[8]

³ kommt aus dem Lateinischen „agere“ für *tun*

um besser zu verstehen, wie die Welt um uns herum funktioniert, denn das wiederum versetzt uns in die Lage effektivere Strategien entwickeln zu können, um mit ihr umzugehen. Aus diesen Gründen hat das Studium der KI in Form der Entwicklung eines rationalen Agenten zwei Vorteile:

Erstens ist es allgemeiner als der „Laws of Thought“-Ansatz, weil korrekte Schlußfolgerungen zu ziehen nur einer von vielen Möglichkeiten bzw. Mechanismen ist, um Rationalität zu erreichen. Zweitens ist es zugänglicher für die wissenschaftliche Entwicklung als Ansätze, die auf menschlichem Verhalten oder Denken basieren, da das Maß der Rationalität klar definiert und komplett allgemein gehalten ist.

Einen wichtigen Punkt sollte man sich jedoch immer im Hinterkopf behalten: Man wird sehr bald sehen, dass das Erreichen von perfekter Rationalität - also immer das Richtige zu tun - nicht in komplizierten Umgebungen realisierbar ist. Die Anforderungen an den Computer sind einfach zu hoch. [7]

2 Historie

Die erste Arbeit, die nun allgemein als KI anerkannt wird, stammt von Warren McCulloch und Walter Pitts (von 1943). Sie zogen drei Quellen heran: Wissen über die grundlegende Physiologie und Funktion von Neuronen im Gehirn; eine formale Analyse von Aussagenlogik und Turings Theorie der Berechnung. Sie schlugen ein Modell in Form von künstlichen Neuronen vor, in dem jedes einzelne Neuron als „on“ oder „off“ charakterisiert ist. McCulloch und Pitts zeigten zum Beispiel, dass jegliche berechenbare Funktion durch eine Art Netzwerk von verbundenen Neuronen berechnet werden könne und dass alle logischen Verbundoperatoren (AND, OR, NOT etc.) durch einfache Netzstrukturen implementiert werden können.

Die zwei im Fachbereich Mathematik der Universität Princeton promovierten Studenten Marvin Minsky und Dean Edmonds bauten 1951 den ersten neuronalen Netzwerk-Computer namens *SNARC*. Dieser benutzte 3000 Vakuum-Röhren und einen zusätzlichen automatischen Steuerungsmechanismus aus einem B-24 Bomber, um ein Netzwerk von 40 Neuronen zu simulieren.

Es war Minsky selbst, der später einflußreiche Theoreme bzw. Lehrsätze bewies, die die Grenzen der „Neuronalen Netzwerk“-Forschung aufzeigten.

Ebenfalls in Princeton graduierte John McCarthy, der danach an das Dartmouth College im Bundesstaat New Hampshire zog, welches die offizielle Geburtsstätte des Fachgebietes der KI wurde. Er initiierte im Sommer 1956 einen zweimonatigen Workshop, zu dem U.S.-Forscher aus den Gebieten der Automatentheorie, Neuronalen Netzen und dem Studium der Intelligenz eingeladen wurden. Wichtige Persönlichkeiten konnten sich dabei untereinander kennenlernen; die folgenden 20 Jahre sollte das Fachgebiet durch die dort aufgetretenen Personen sowie durch deren Studenten und Kollegen von MIT⁴, CMU⁵, Stanford und IBM dominiert werden. Das wahrscheinlich am längsten währende Ergebnis aus diesem Workshop ist die Prägung und Einigung auf den Begriff „*Artificial Intelligence*“ (*Künstliche Intelligenz*).

In der frühen Beginnerphase der KI konnte man viele Erfolge verbuchen. Noch wenige

⁴ Massachusetts Institute of Technology

⁵ Carnegie Mellon University

Jahre zuvor hatte man Computer lediglich als arithmetische Maschinen betrachtet - jetzt war es gerade zu erstaunlich, wenn er irgendetwas annähernd Cleveres tat. Allen Newell und Herbert Simon feierten 1961 frühe Erfolge mit ihrem *General Problem Solver* (*GPS*), ein Programm, das menschliche „Problemlösungs-Protokolle“ imitieren sollte, d.h. den Ablauf der menschlichen Problemlösung nachahmte. Innerhalb der beschränkten Varianten, mit denen es umgehen konnte, stellte sich heraus, dass die Reihenfolge, in der das Programm Unterziele und mögliche Handlungen erwägte, identisch war zu der Herangehensweise eines Menschen, der bei den gleichen Problemen ansetzte. *GPS* war wahrscheinlich das erste Programm welches dem „Menschlich Denken“-Ansatz Gestalt gab. Der große Erfolg und die im Anschluß folgenden Programme in Form von Erkenntnismodellen brachten Newell und Simon 1976 dazu, die berühmte „*Physical Symbol System Thesis*“ aufzustellen, welche besagt, dass ein physikalisches Symbol-System die notwendigen und ausreichenden Möglichkeiten für allgemein intelligentes Handeln beinhaltet. Das soll bedeuten, dass jedes System (ob menschlich oder maschinell), das Intelligenz hervorbringt, auf Basis von Datenstrukturen arbeitet, die wiederum aus Symbolen bestehen.

Als John McCarthy 1958 von Dartmouth zum MIT umzog, leistete er drei entscheidende Beiträge in einem Jahr: er definierte die high-level⁶ Sprache LISP, die zur dominanten Programmiersprache in der KI werden sollte. Außerdem publizierte McCarthy ein wissenschaftliches Paper mit dem Titel „Programs with Common Sense“ (dt.: „Programme mit Allgemeinem Verstand“), in dem er den sogenannten *Advice Taker* vorstellte - ein hypothetisches Programm, das als das erste komplette KI-System betrachtet werden kann. Dieses Programm verwendete eine Wissensbasis, um nach Lösungen für bestimmte Probleme zu suchen. Der *Advice Taker* verkörperte generelles Wissen über die Welt. McCarthy zeigte zum Beispiel, wie einige einfache Axiome das Programm in die Lage versetzten, einen Plan zu entwickeln, zum Flughafen zu fahren, um ein Flugzeug zu nehmen. Außerdem war das Programm so konstruiert, dass es neue Axiome innerhalb des normalen Rechengangs akzeptierte, was ihm erlaubte, erweiterte Kompetenzen in neuen Gebieten zu erreichen, ohne dafür umprogrammiert zu werden. Der *Advice Taker* verkörperte somit auch die zentralen Prinzipien der Wissensrepräsentation und Argumentation: es ist sinnvoll und nützlich eine formale und explizite Repräsentation der Welt - und von der Art und Weise, wie Handlungen diese Welt beeinflussen - zu haben und in der Lage zu sein, diese Repräsentationen durch sich ergebende Prozesse zu manipulieren.

KI-Forscher waren von Anfang an nicht besonders zurückhaltend, was Vorhersagen über kommende Erfolge und Ergebnisse auf dem Gebiet der Künstlichen Intelligenz angeht. Das große Selbstbewußtsein begründet sich in der vielversprechenden Performanz früher KI-Systeme an einfachen Beispielen. In fast allen Fällen jedoch stellte sich heraus, dass diese Systeme kläglich versagten, wenn man sie auf eine breitere Auswahl und schwierigere Probleme ansetzte. **Die erste Schwierigkeit** kam auf, weil die meisten frühen Programme kaum oder gar kein Wissen über den Inhalt hatten; sie funktionierten aufgrund von simplen syntaktischen Manipulationen. Eine dafür typische Geschichte

⁶ *high-level*: das Design der Sprache orientiert sich an den Bedürfnissen der Problemlösung, im Gegensatz dazu *low-level*: Sprache orientiert sich relativ nah an der eigentlichen Architektur der Maschine [3]

ereignete sich in den 50er Jahren, als der Nationale Forschungsausschuß der Vereinigten Staaten die Arbeit an maschinellen Übersetzern vorantrieb, um die Übersetzung von russischen wissenschaftlichen Arbeiten kurz nach dem Start der Sputnik⁷ 1957 zu beschleunigen. Man hatte angenommen, dass einfache syntaktische Umformungen basierend auf russischer und englischer Grammatik sowie Wortersetzungen mithilfe eines elektronischen Wörterbuchs ausreichen würden, um den genauen Sinn der Sätze zu entschlüsseln. Fakt ist jedoch, dass eine korrekte Übersetzung generelles Wissen über den Inhalt erfordert, um Mehrdeutigkeiten auszuschließen und den tatsächlichen Inhalt des Satzes aufzubauen. Beispielhaft ist der berühmte Satz „The spirit is willing but the flesh is weak.“ (dt.: Der Geist ist willig aber das Fleisch ist schwach.), bei dem die Übersetzung ins Russische und nochmalige Rückübersetzung ins Englische schlußendlich ergab: „The vodka is good but the meat is rotten.“ (dt.: Der Wodka ist gut aber das Fleisch ist verdorben.). Auch heute ist die maschinelle Übersetzung noch ein unvollkommenes aber weit verbreitetes Werkzeug für technische, kommerzielle, politische und Internetdokumente.

Die zweite Schwierigkeit waren die Eigenheiten vieler Probleme, die die KI versuchte zu lösen. Die meisten der frühen KI-Programme lösten Probleme, indem sie verschiedene Kombinationen von Schritten durchprobierten, bis die Lösung gefunden wurde. Diese Strategie funktionierte nur bei minimaler Anzahl von möglichen Handlungen und kurzen Lösungssequenzen. Bevor die Theorie der berechenbaren Komplexität entwickelt wurde, herrschte die weit verbreitete Annahme, dass das „Hochskalieren“ auf größere Probleme lediglich eine Frage von schnellerer Hardware und größerem Speicher sei. Der Optimismus der diese Entwicklung begleitete, wurde bald erstickt, als Forscher daran scheiterten, Theorien zu beweisen, die mehr als einige Dutzend Fakten beinhalteten. Der Fakt, dass ein Programm im Prinzip eine Lösung finden kann, bedeutet nicht, dass das Programm jenen Mechanismus enthält, den es braucht, um auch in der Praxis eine Lösung zu finden. [7] Die Illusion und der Glaube an die unbegrenzte Leistungsfähigkeit des Computers war nicht ausschließlich auf problemlösende Programme beschränkt. Frühe Experimente in sogenannter *Machine Evolution*⁸ basierten auf der Vorstellung, dass man durch geeignete Folgen von kleinen Mutationen an einem Programmcode wiederum ein Programm generieren könnte, das für jede denkbare einfache Aufgabe eine gute Performanz bietet. Daraus entstand die Idee, zufällige Mutationen auszuprobieren und einen Auswahl-Prozess zu verwenden, um nützlich erscheinende Mutationen zu sichern/speichern. Trotz tausender Stunden an CPU-Zeit war fast kein Fortschritt festzustellen. Moderne Algorithmen dieser Art benutzen mittlerweile bessere Repräsentationen und zeigen mehr Erfolg.

Die dritte Schwierigkeit ergab sich aufgrund der generellen Einschränkungen bezüglich der grundlegenden Strukturen, die benutzt wurden, um intelligentes Verhalten zu generieren. Zum Beispiel bewiesen Minsky und Papert in ihrem Buch „Perceptrons“⁹

⁷ Sputnik 1 (russisch: Gefährte) war der erste künstliche Satellit in der Erdumlaufbahn. [?]

⁸ heute: *genetic algorithms* (dt.: genetische Algorithmen) genannt; nach R.M. Friedberg et al. 1959, *A learning machine: Part II*. IBM Journal of Research and Development 3(3), 282-287 [7]

⁹ eine einfache Form eines neuronalen Netzwerkes [7]

(1969)¹⁰, dass - obwohl Perzeptronen alles mögliche lernen - sie nur sehr wenig wiedergeben konnten. Im Speziellen war man nicht in der Lage ein „Zweier-Eingabe“-Perzeptron darauf zu trainieren, zu erkennen, wenn die beiden Eingaben unterschiedlich voneinander waren.

Das Bild von der Problemlösung, welches sich während des ersten Jahrzehnts der KI-Forschung aufgebaut hatte, bestand weitestgehend aus den generellen Bestrebungen von Such-Mechanismen, elementar begründete Schritte zusammenzufassen, um komplette Lösungen zu finden. Solche Ansätze hat man als „weak methods“ (dt.: schwache Methoden) bezeichnet, weil sie, obwohl allgemein gehalten, sich nicht auf größere und schwierigere Problembeispiele anwenden ließen. Die Alternative zu den „schwachen Methoden“ wäre mächtigeres und spezifischeres Wissen zu benutzen, das umfassendere beweisführende Schritte erlaubt und leichter mit typischerweise auftretenden Fällen umgehen kann.

Seit den 80er Jahren hat sich die KI immer mehr auch zu einem Industriezweig entwickelt. Fast jedes größere U.S. Unternehmen finanzierte seine eigene KI-Forschungsgruppe und erforschte oder benutzte solche Expertensysteme. Das erste erfolgreiche kommerzielle Expertensystem „R 1“ lief ab 1982 bei der Digital Equipment Corporation unter der Leitung von Dean McDermott. Dieses Programm half bei der Konfiguration von Bestellungen von neuen Computersystemen und brachte dem Unternehmen bis 1986 eine Ersparnis von geschätzten 40 Millionen Dollar pro Jahr. 1981 riefen die Japaner das sogenannte „Fifth Generation“-Projekt ins Leben - ein 10-Jahres-Plan, um intelligente Computer zu bauen, auf denen PROLOG laufen sollte. Als Gegenantwort dazu gründeten die USA die „Microelectronics and Computer Technology Corporation“ (MCC), ein Forschungskonsortium, um die nationale Wettbewerbsfähigkeit zu sichern. Leider haben die KI-Komponenten der MCC und des „Fifth Generation“-Projekts ihre ambitionierten Zielsetzungen niemals erreicht. Dem großen Boom der KI-Industrie - angefangen von ein paar Millionen Dollar im Jahre 1981 bis zu mehreren Milliarden von Dollar 1988 - folgte alsbald eine Periode, die als „AI Winter“ (dt.: KI Winter) bezeichnet wurde. In dieser Zeit litten viele Unternehmen unter fehlgeschlagenen Projekten und damit verbundenen unerfüllbaren Versprechungen.

Obwohl sich die Computer-Wissenschaften in den 70er Jahren weitestgehend vom Gebiet der Neuronalen Netzwerke entfernt hatten, ging die Arbeit auf anderen Gebieten weiter. Physiker wie John Hopfield¹¹ benutzten Techniken aus statistischen Mechanismen, um die Speicher- und Optimierungseigenschaften von Netzwerken zu analysieren. Psychologen wie David Rumelhart und Geoff Hinton setzten das Studium der Neuronalen Netze bezüglich der Speichervorgänge fort. Der wirkliche Impuls kam Mitte der 80er Jahre als vier Gruppen unabhängig voneinander den sogenannten „Back Propagation Learning“-Algorithmus wiedererfanden, der erstmals 1969 von Bryson und Ho

¹⁰ M. L. Minsky, S. Papert, *Perceptrons: An Introduction to Computational Geometry* MIT Press, [7]

¹¹ amerikanischer Physiker, entwickelte 1982 ein Modell eines asynchronen Neuronalen Netzes: *Hopfield-Netze* sind einschichtige Neuronale Netze, die ausschließlich indirekte Rückkopplungen zwischen je zwei verschiedenen Knoten des Netzes besitzen aber keine direkte Rückkopplung zum gleichen Knoten [4]

entwickelt worden ist.¹² Dieser Algorithmus, der auf mehrschichtigen Perzeptronen basiert, wurde auf viele Lern-Probleme innerhalb der Computer-Wissenschaften als auch in der Psychologie angewendet.

Die vergangenen Jahre sind geprägt durch Revolutionen - sowohl inhaltlich als auch in der Methodik der Arbeiten in der KI. Es ist nun üblich, auf bestehenden Theorien aufzubauen, anstatt neue aufzustellen. Die Ansprüche basieren eher auf strikten Theorien oder handfesten experimentellen Beweisen als auf Intuition. Außerdem ist die Anwendbarkeit auf die reale Welt wichtiger geworden als die Demonstration an „spielzeughafte“ Beispielen. Einige bezeichnen diese Veränderung als Sieg der Ordnung und zwar die, die der Überzeugung sind, dass sich KI-Theorien auf mathematischer Striktheit gründen sollten. Zu den „Lotter-Lieschen“ (engl.: scruffies) dieses Fachgebietes zählen sie die, die lieber ein paar Dutzend Ideen ausprobieren, einige Programme schreiben und danach abschätzen, was am besten funktioniert. Beide Ansätze sind wichtig. Eine Umstellung auf Ordentlichkeit deutet darauf hin, dass das Gebiet einen gewissen Level an Stabilität und Reife erreicht hat. Inwiefern diese Stabilität durch eine neue „verlotterte“ Idee (engl.: scruffy idea) gestört würde, sei eine andere Frage. [7]

Bezüglich der Methodik hat sich die KI letztendlich unter den wissenschaftlichen Methoden einen festen Platz gesichert. Um akzeptiert zu werden, müssen sich die Hypothesen strikten empirischen Experimenten unterwerfen und die Ergebnisse müssen statistisch auf ihre Bedeutung analysiert werden. Durch die Nutzung des Internet und der gemeinsame Zugriff auf Sammlungen von Testdaten und Code ist es nunmehr möglich, Experimente nachzubilden.

Durch den Fortschritt im Lösen von Teilproblemen der KI ermutigt, haben Forscher wieder begonnen sich auf das „whole agent“ Problem (dt.: vollständiger Agent) zu konzentrieren. Die Arbeit von Allen Newell, John Laird und Paul Rosenbloom an „SOAR“ ist das bekannteste Beispiel für eine komplette Agenten-Architektur.¹³ Sie zielt darauf ab, die Arbeitsweise eines Agenten zu verstehen, der in eine reale Umgebung eingebettet ist und kontinuierlich Sensoreingaben erhält. Eine der wichtigsten Umgebungen für intelligente Agenten ist das Internet. KI-Systeme sind in webbasierten Applikationen alltäglich geworden. KI-Technologien sind die Grundlage vieler Internet-Werkzeuge, wie zum Beispiel Suchmaschinen oder Webseiten-Konstruktionssysteme. [7]

3 Autoren über KI

3.1 Hubert L. Dreyfus, Stuart E. Dreyfus, *Mind over Machine*

Im Kapitel „Logic Machines and Their Limits“ werden Computer als logische Maschinen dem menschlichen Denk-System gegenübergestellt. Menschen wiesen Flexibilität, Urteilsvermögen und Intuition auf, die jeglicher Zerlegung in Spezifikationen und Schlußfolgerungen widerstehe und damit genauso schwierig sei, einer logischen Maschine nahezubringen. Das Wesen eines Computers läge darin, präzisen Regeln zu folgen. Diese Regeln „lernt“ er sofort und kann sie direkt anwenden - wir Menschen hingegen müssen neu Erlerntes zuerst trainieren, um es sicher anwenden zu können.

¹² veröffentlicht in *Applied Optimal Control*, A. E. Bryson und Y.-C. Ho [7]

¹³ *SOAR: An architecture for general intelligence*, 1987, *Artificial Intelligence* 33(1), 1-64; [7]

Die Autoren stellen sich die Frage, welchen Level an Fähigkeiten wir von einer Maschine erwarten können zu erreichen - unter Berücksichtigung der Möglichkeiten, die die jetzige und zukünftige Programmierung bietet. Sie glauben, dass der Weg hin zur verbesserten Performanz einhergeht mit der Abkehr vom logischen Verarbeiten atomarer Fakten hin zur *Erkennung* ohne in isolierbare Elemente zu flüchten, d.h. die Gleichheit zwischen einer üblichen Situation und einer gespeicherten bildähnlichen Repräsentation einer vorhergehenden ähnlichen Situation zu erkennen. [5]

3.2 Hubert L. Dreyfus, *What Computers Still Can't Do*

Hier stellt der Autor als großes Problem die mangelnden Adaptionsfähigkeiten von menschlicher Denkweise bzw. den Strukturen und Entscheidungsvorgängen durch die Maschine dar. Der Unterschied zwischen menschlicher Zielsetzung und maschinellen Endzuständen wird mittels eines Zitats von dem japanischen Wissenschaftler Satosi Watanabe verdeutlicht:

Der Mensch macht Abschätzungen und Bewertungen anhand eines sehr unspezifischen und quasi gefühlsbehafteten Wertesystems. Währenddessen ist für einen Roboter eine Bewertung nur anhand einer Tabelle/Liste oder eines spezifischen Kriteriums möglich.[...]Dieser Unterschied ist fein aber tiefgreifend. Man könnte nun sagen, der Mensch hat *Werte* und die Maschine *Tatsachen* bzw. *Objekte*. Natürlich sind Menschen auch objektiv, aber es wird von einem Bewertungssystem abgeleitet und ist nicht der letztendliche Schiedsrichter seiner Handlungen, wie es für einen Roboter sein würde. Sobald die Zielsetzung festgelegt wird, kann die Maschine es genau wie ein Mensch auch verfolgen. Menschliches auf Nützlichkeit ausgerichtetes Verhalten kann ganz einfach durch eine Maschine simuliert werden, wenn die quantitative Nützlichkeit und die Wahrscheinlichkeit jedes alternativen Ereignisses festgehalten und der Maschine übergeben wird. Aber eine Maschine kann niemals an den Ursprung kommen, von wo aus diese Nützlichkeit abgeleitet wurde.

Watanabe ist überzeugt, dass genau diese Werte wesentlich und unentbehrlich für intelligentes Verhalten sind. In diesem Buch werden auch Herbert Simon¹⁴ und Walter Reitman erwähnt, die zwar richtig festgestellt haben, dass Gefühle und Motivation eine Rolle beim intelligenten Verhalten spielen, aber ihre Art, diese Rolle zu simulieren, sei, Programme zu schreiben, in denen „Gefühle“ bzw. Emotionen die Arbeit an einem Problem unterbrechen und fremde Faktoren oder Arbeiten an einem anderen Problem einführen. Sie scheinen nicht zu sehen, dass Gefühle und Bedenken unser kognitives (= auf Erkenntnis beruhendes) Verhalten begleiten und leiten. Das sieht der Autor als einen der typischen Fälle an, in denen man nicht erkennt, was man nicht in der Lage sei, zu programmieren bzw. programmiertechnisch zu realisieren. Er kommt in seinem Buch zu dem Schluß, dass diese alternative Vorstellung vom Menschen und seiner Fähigkeit, intelligent zu handeln, in Wirklichkeit eine Analyse der Art und Weise ist, auf die der Mensch durch seine Aktivitäten die Welt um sich herum erzeugt. Und es sei

¹⁴ (1916 - 2001), einer der einflussreichsten Sozialwissenschaftler des 20. Jahrhunderts, erhielt 1978 den Nobelpreis für Wirtschaftswissenschaften [8]

diese Welt, die die Bedingungen aufstellt, unter denen spezifische Fakten für den Menschen zugänglich werden. Die Fakten sind ursprünglich betreffend unserer Bedürfnisse organisiert. Das versetzt uns in die Lage, den grundlegenden Unterschied zwischen menschlicher und maschineller Intelligenz zu sehen. Künstliche Intelligenz müsse auf dem Level der Objektivität und Rationalität beginnen, wo die Fakten bereits erzeugt worden sind. Es fasst alle auftretenden Fakten aus einer Situation zusammen und versucht, die Ergebnisse zu verwenden, um intelligentes Verhalten zu simulieren. Aber diese Fakten, die aus dem Zusammenhang heraus entnommen werden, seien eine unhandliche Masse neutraler Daten mit denen KI-Forscher weitestgehend nicht umgehen können. Die Programme stürzen ab, sobald sich die Menge an Daten und Information erhöht. Wenn man davon ausgehe, dass die bisherigen Beschreibungen der menschlichen Intelligenz korrekt sind, dann gäbe es im Prinzip schon Gründe, warum Künstliche Intelligenz niemals vollständig realisiert werden könne. Neben dem technischen Problem, eine riesige Menge von Bits an Daten speichern zu müssen, gäbe es nach der letzten Analyse keine wirklich festgelegten Fakten. Außerdem handele es sich um eine ständig durch den Menschen veränderte Menge an Fakten, die ständig aktualisiert werden müsste. Selbst wenn der Philosoph oder KI-Forscher vorschlägt, diesen Einwänden zu begegnen, indem man die menschlichen Bedürfnisse formalisiert, würde man stets mit den gleichen ursprünglichen Problemen konfrontiert werden. Unbestimmte Ziele und Bedürfnisse sowie die Erfahrung von Befriedigung und Erfüllung könnten nicht von einer digitalen Maschine simuliert werden, dessen einziger Existenzmodus auf einer Reihe von bestimmten Zuständen basiert. [2]

3.3 A.K. Dewdney, *Yes, We Have No Neutrons*

Was unterscheidet „gute“ Wissenschaft von „schlechter“? Der Autor beantwortet diese Frage durch das Ziehen von klaren Grenzen zwischen seriösen und weniger fundierten Methoden und zeigt an speziellen Beispielen, was in der sogenannten Wissenschaft alles schief gehen kann. Vor allem aber macht er seine Vorstellung vom Ablauf einer begründeten und sachlich korrekten Methodik innerhalb einer seriösen Wissenschaft durch folgendes Schema (Abb. 1) deutlich:

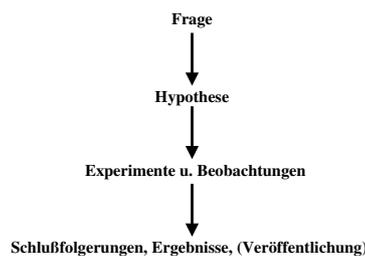


Abbildung 1. frei nach [1]

Schlechte Wissenschaft weicht demnach von diesem Schema ab und kann somit nur zu falschen Resultaten und stark verzerrten Versionen der Wahrheit führen. Die wirkliche Macht des Wissenschaftlers beruhe nicht auf einem Haufen Atome oder einem rekonstruierten Gen sondern in dem tief verwurzelten Verständnis des Universums und seinen Strukturen. Vieles von dem, was wir als „Talent“ bezeichneten, entspringe dieser Achtsamkeit. „Gute“ Wissenschaft sei letztendlich echte Magie. Es könnte einen ins Wanken bringen, zu erkennen, wie viele physikalische Phänomene mit einer ungeschmückten Genauigkeit Theorien und Formeln folgen, die rein garnichts mit unseren Wünschen oder kreativen Impulsen zu tun haben, außer mit der Realität selbst. Genauso verblüffend sei, herauszufinden, wenn Phänomene, die durch Theorien und Formeln vorausgesagt wurden, dann später tatsächlich existierten. Aber „wissenschaftliche Magie“ geschehe nur dann, wenn Wissenschaftler erlaubten, die Logik von sorgsam durchgeführten Untersuchungen über persönliche Hoffnungen oder Ängste über mögliche Ergebnisse zu stellen. Wenn es doch die Gedanken an Ruhm gäbe, würden wahre Wissenschaftler diese für sich im Hintergrund bewahren. [1]

4 Schlusswort

Hat man sich einen Überblick über die bisherige Entwicklung und die entscheidenden Fortschritte aber auch über die noch zu überwindenden Hürden und Fallstricke der KI verschafft, sollte sich ein objektives Bild einer Wissenschaft darstellen, die sich Mitteln und Methoden bedient, die bereits in anderen anerkannten Fachgebieten wie der Physik, Chemie oder Biologie allgemein etabliert sind. Die KI baut auf vielen verschiedenen Theorien auf und bezieht sich häufig auf andere Wissenschaften wie zum Beispiel die Psychologie oder die Sozialwissenschaften. Sie gründet sich auf Mathematik und Logik und verfeinert sich durch ihre psychologischen als auch philosophischen Anteile. Die Frage nach der Seriösität dieser Wissenschaft ergibt sich vordergründig aus der Bezeichnung *Künstliche Intelligenz* selbst, sowie der ungeklärten Definition, was genau unter KI zu verstehen ist. Ein weiteres Problem stellt die mangelnde Realisierbarkeit der bestehenden Ansätze und Algorithmen dar, denn selbst wenn man es geschafft hat, das Wissen um die Welt in Formeln zu vereinen, so fehlt es letztendlich an Rechenkapazität, um die Ideen umzusetzen und zu beweisen. Nicht zu vergessen sind typische Einzelfälle, in denen unhaltbare Versprechungen und Zielsetzungen ohne fundierten wissenschaftlichen Hintergrund die Glaubwürdigkeit und damit Seriösität angreifen. Mit seinen negativen Beispielen und sogenannten „schwarzen Schafen“ unterscheidet sich die KI jedoch nicht von anderen Fachgebieten. Hier gilt es abzuwägen, ob einzelne wissenschaftliche Fehlritte zur kompletten Ablehnung eines vielschichtigen Forschungsgebietes führen sollte, in dem die Anwendung von fundierten und anerkannten Methoden ganz offensichtlich überwiegt.

Literatur

1. A. K. Dewdney. *Yes, We Have No Neutrons*. John Wiley Sons, Inc., New York, 1997.
2. Hubert L. Dreyfus. *What Computers Still Can't Do - A Critique of Artificial Reason*. The MIT Press, Cambridge, Massachusetts, 1999.
3. Universität Heidelberg „Ein Blick über den Tellerrand“. 2004. URL: <http://www.cl.uni-heidelberg.de/kurs/ws02/prog2/html/page004.html>.
4. Universität Münster „Hopfield Netze“. 2002. URL: <http://wwwmath.uni-muenster.de/SoftComputing/lehre/material/nnscrip/script/Kapitel6.pdf>.
5. Stuart E. Dreyfus Hubert L. Dreyfus. *Mind Over Machine*. The Free Press, New York, 1986.
6. [lexi-tv.de/lexikon](http://www.lexi-tv.de/lexikon) Roboter. 2004. URL: <http://www.lexi-tv.de/lexikon>.
7. Peter Norvig Stuart Russell. *Artificial Intelligence - A Modern Approach*. Prentice Hall, New Jersey, 2003. Second Edition.
8. Wikipedia. Die freie Enzyklopädie im Internet, 2004. URL: <http://de.wikipedia.org>.