

Klaus Mainzer

Künstliche Intelligenz – Wann übernehmen die Maschinen?



1010 1000	0100 0110
0100 1010 1010	1010
0111 0100 0111	1010 1000 1010
1000 0100 1010	0110 0110
1010 1010 1000	0111 0100 1010
0111 0100	1010 1000 1010



Springer

Technik im Fokus

Weitere Bände zur Reihe finden Sie unter
<http://www.springer.com/series/8887>.

Klaus Mainzer

Künstliche Intelligenz – Wann übernehmen die Maschinen?

2., erweiterte Auflage

 Springer

Prof. em. Dr. Klaus Mainzer,
Emeritus of Excellence
Technische Universität München
München, Deutschland

„Konzeption der Energie-Bände in der Reihe Technik im Fokus: Prof. Dr.-Ing. Viktor Wesselak, Institut für Regenerative Energiesysteme, Hochschule Nordhausen“

ISSN 2194-0770

ISSN 2194-0789 (electronic)

Technik im Fokus

ISBN 978-3-662-58045-5

ISBN 978-3-662-58046-2 (eBook)

<https://doi.org/10.1007/978-3-662-58046-2>

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature 2016, 2019

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Der Verlag, die Autoren und die Herausgeber gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag noch die Autoren oder die Herausgeber übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen. Der Verlag bleibt im Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutionsadressen neutral.

Springer ist ein Imprint der eingetragenen Gesellschaft Springer-Verlag GmbH, DE und ist ein Teil von Springer Nature.

Die Anschrift der Gesellschaft ist: Heidelberger Platz 3, 14197 Berlin, Germany

Vorwort zur 2. Auflage

In der 2. Auflage wird auch das Thema Künstliche Intelligenz und Quantum Computing (Abschn. 10.3) berücksichtigt, nachdem diese Technologie absehbar ist. Neben der Innovation von Künstlicher Intelligenz rücken die Herausforderungen von Sicherheit und Verantwortung in den Vordergrund. Sie werden in der 2. Auflage durch die zusätzlichen Kap. 11 und 12 berücksichtigt: Im Machine Learning benötigen wir mehr Erklärung und Zurechnung von Ursachen und Wirkungen, um ethische und rechtliche Fragen der Verantwortung (z. B. beim autonomen Fahren oder in der Medizin) entscheiden zu können. Zudem ist Künstliche Intelligenz bereits eine Schlüsseltechnologie, die den globalen Wettstreit der Gesellschaftssysteme entscheiden wird. Wie sollen wir unsere individuellen Freiheitsrechte in der KI-Welt behaupten? Europa wird sich nicht nur als technischer KI-Standort, sondern auch mit seinem moralischen Wertesystem positionieren müssen.

München
im August 2018

Klaus Mainzer

Vorwort

Künstliche Intelligenz beherrscht längst unser Leben, ohne dass es vielen bewusst ist. Smartphones, die mit uns sprechen, Armbanduhren, die unsere Gesundheitsdaten aufzeichnen, Arbeitsabläufe, die sich automatisch organisieren, Autos, Flugzeuge und Drohnen, die sich selbst steuern, Verkehrs- und Energiesysteme mit autonomer Logistik oder Roboter, die ferne Planeten erkunden, sind technische Beispiele einer vernetzten Welt intelligenter Systeme. Sie zeigen uns, wie unser Alltag von KI-Funktionen bestimmt ist.

Auch biologische Organismen sind Beispiele von intelligenten Systemen, die wie der Mensch in der Evolution entstanden und mehr oder weniger selbstständig Probleme effizient lösen können. Gelegentlich ist die Natur Vorbild für technische Entwicklungen. Häufig finden Informatik und Ingenieurwissenschaften jedoch Lösungen, die anders und sogar besser und effizienter sind als in der Natur. Es gibt also nicht „die“ Intelligenz, sondern Grade effizienter und automatisierter Problemlösungen, die von technischen oder natürlichen Systemen realisiert werden können.

Dahinter steht die Welt lernfähiger Algorithmen, die mit exponentiell wachsender Rechenkapazität (nach dem Mooreschen Gesetz) immer leistungsfähiger werden. Sie steuern die Prozesse einer vernetzten Welt im Internet der Dinge. Ohne sie wäre die Datenflut nicht zu bewältigen, die durch Milliarden von Sensoren und vernetzten Geräten erzeugt werden. Auch Forschung und Medizin benutzen zunehmend intelligente Algorithmen, um in einer wachsenden Flut von Messdaten neue Gesetze und Erkenntnisse zu entdecken.

Seit ihrer Entstehung ist die KI-Forschung mit großen Visionen über die Zukunft der Menschheit verbunden. Löst die „künstliche Intelligenz“

den Menschen ab? Einige sprechen bereits von einer kommenden „Superintelligenz“, die Ängste und Hoffnungen auslöst. Dieses Buch ist ein Plädoyer für Technikgestaltung: KI muss sich als Dienstleistung in der Gesellschaft bewähren.

Seit meinem Studium war ich von den Algorithmen fasziniert, die künstliche Intelligenz erst möglich machen. Man muss ihre Grundlagen kennen, um ihre Leistungen und Grenzen abschätzen zu können. Erstaunlicherweise, das ist eine wesentliche Einsicht dieses Buchs, ändern noch so schnelle Supercomputer nichts an den logisch-mathematischen Grundlagen, die von menschlicher Intelligenz bewiesen wurden. Erst auf der Grundlage dieses Wissens lassen sich auch gesellschaftliche Auswirkungen bewerten. Zu diesem Zweck hatten wir bereits Ende der 1990er Jahre das Institut für Interdisziplinäre Informatik an der Universität Augsburg gegründet. An der Technischen Universität München kam die Leitung der Carl von Linde-Akademie und im Rahmen der Exzellenzinitiative 2012 die Gründung des Munich Center for Technology in Society (MCTS) hinzu. Im Themennetzwerk der Deutschen Akademie für Technikwissenschaften (acatech) steht ebenfalls „Technik im Fokus – Daten Fakten Hintergründe“, wie diese neue Buchreihe im Springer-Verlag heißt. Als langjähriger Autor im Springer-Verlag habe ich diesen Beitrag gerne geschrieben und danke für die bewährte Unterstützung.

München
im September 2015

Klaus Mainzer

Inhaltsverzeichnis

1	Einführung: Was ist KI?	1
	Literatur	6
2	Eine kurze Geschichte der KI	7
	2.1 Ein alter Menschheitstraum	7
	2.2 Turing-Test	10
	2.3 Vom „Allgemeinen Problemlöser“ zum Expertensystem	11
	Literatur	13
3	Logisches Denken wird automatisch	15
	3.1 Was heißt logisches Schließen?	15
	3.2 KI-Programmiersprache PROLOG	18
	3.3 KI-Programmiersprache LISP	20
	3.4 Automatisches Beweisen	27
	Literatur	42
4	Systeme werden zu Experten	43
	4.1 Architektur eines wissensbasierten Expertensystems ..	43
	4.2 Programmierung von Wissenspräsentationen	45
	4.3 Eingeschränktes, unsicheres und intuitives Wissen ..	50
	Literatur	54

5	Computer lernen sprechen	55
	5.1 ELIZA erkannte Zeichenmuster	55
	5.2 Automaten und Maschinen erkennen Sprachen	58
	5.3 Wann versteht mich mein Smartphone?	71
	Literatur	79
6	Algorithmen simulieren die Evolution	81
	6.1 Biologische und technische Schaltpläne	81
	6.2 Zelluläre Automaten	87
	6.3 Genetische und evolutionäre Algorithmen	92
	Literatur	96
7	Neuronale Netze simulieren Gehirne	99
	7.1 Gehirn und Kognition	99
	7.2 Neuronale Netze und Lernalgorithmen	104
	7.3 Emotionen und Bewusstsein	125
	Literatur	136
8	Roboter werden sozial	139
	8.1 Humanoide Roboter	139
	8.2 Kognitive und soziale Roboter	142
	8.3 Schwarmintelligenz von Robotern	149
	Literatur	154
9	Infrastrukturen werden intelligent	157
	9.1 Internet der Dinge und Big Data	157
	9.2 Vom autonomen Fahrzeug zum intelligenten Verkehrssystem	165
	9.3 Von Cyberphysical Systems zu intelligenten Infrastrukturen	169
	9.4 Industrie 4.0 und Arbeitswelt der Zukunft	174
	Literatur	181
10	Von der natürlichen über die künstliche zur Superintelligenz?	185
	10.1 Neuromorphe Computer und künstliche Intelligenz	185
	10.2 Natürliche und künstliche Intelligenz	200
	10.3 Quantencomputer und Künstliche Intelligenz	210

10.4 Singularität und Superintelligenz?	221
10.5 Technikgestaltung: KI als Dienstleistung	232
Literatur	241
11 Wie sicher ist Künstliche Intelligenz?	245
11.1 Neuronale Netze sind eine Black Box	245
11.2 Entscheiden unter unvollständiger Information	256
11.3 Wie sicher sind menschliche Institutionen?	259
Literatur	264
12 Künstliche Intelligenz und Verantwortung	267
12.1 Social Score und Neue Seidenstraße	267
12.2 Künstliche Intelligenz und globaler Wettstreit der Wertsysteme	273
Literatur	278
Weiterführende Literatur	281
Personenverzeichnis	285
Sachverzeichnis	287

Einführung: Was ist KI?

1

Nachdem mich das Klingeln meines Weckers etwas unsanft aufgeschreckt hat, wünscht mir die vertraute und freundliche Frauenstimme von Susanne einen guten Morgen und fragt, wie ich geschlafen habe. Etwas verschlafen erkundige ich mich nach meinen Terminen heute Morgen. Susanne erinnert mich an einen Termin in Frankfurt bei unserer Zweigstelle. Freundlich, aber bestimmt erinnert sie mich auch an das ärztlich verschriebene Bewegungstraining. Ich blicke auf meine Armbanduhr, die meinen aktuellen Blutdruck und die Blutwerte angibt. Susanne hat Recht. Ich müsste etwas tun. Susanne und der Wecker stecken in meinem Smartphone, das ich nach Dusche und Ankleiden in die Tasche stecke und zum Wagen eile. Zum Cockpit meines Wagens gewandt erkläre ich kurz mein Fahrziel. Nun habe ich Zeit für einen Kaffee und lese entspannt die Zeitung. Mein Wagen steuert selbstständig auf die Autobahn. Unterwegs weicht der Wagen einem Baufahrzeug aus. Er hält sich vorbildlich an die Verkehrsvorschriften und kommt dennoch besser voran als einige menschliche Fahrer, die gestresst mit überhöhter Geschwindigkeit, Lichthupe und zu kurzen Abständen schneller sein wollen. Menschen sind eben chaotische Systeme, denke ich noch. Dann bitte ich Susanne, mir Marktprofile unserer Produkte anzugeben, die sie blitzschnell mit Big Data Algorithmen ausfiltert. In der Frankfurter Zweigstelle angekommen lasse ich den Wagen selbstständig einparken. Die Halbleiterproduktion in unserem Betrieb ist weitgehend automatisch. Auch spezielle Kundenwünsche können Online im Verkauf eingegeben werden. Die Produktion richtet sich dann selbstständig auf diese speziellen Wünsche ein. Nächste Woche will ich nach Tokio und unseren japanischen Geschäftspartner treffen. Ich muss ihn noch bitten, mich nicht in eines der neuen Roboterhotels unterzubringen. Beim letzten Mal ging dort alles automatisch wie beim Einchecken auf dem Flughafen. Selbst im Empfang

saß eine freundliche Roboterdame. Mit menschlichem Service wird es zwar etwas teurer sein. Aber da bin europäisch „altmodisch“ und ziehe wenigstens im Privaten menschliche Zuwendung vor . . .

Das war kein Science Fiction Szenario. Das waren KI-Technologien, die technisch heute realisierbar sind und als Teilgebiet der Informatik und Ingenieurwissenschaften entwickelt werden. Traditionell wurde KI (Künstliche Intelligenz) als Simulation intelligenten menschlichen Denkens und Handelns aufgefasst. Diese Definition krankt daran, dass „intelligentes menschliches Denken“ und „Handeln“ nicht definiert sind. Ferner wird der Mensch zum Maßstab von Intelligenz gemacht, obwohl die Evolution viele Organismen mit unterschiedlichen Graden von „Intelligenz“ hervorgebracht hat. Zudem sind wir längst in der Technik von „intelligenten“ Systemen umgeben, die zwar selbstständig und effizient, aber häufig anders als Menschen unsere Zivilisation steuern.

Einstein hat auf die Frage, was „Zeit“ sei, kurz geantwortet: „Zeit ist, was eine Uhr misst“. Deshalb schlagen wir eine Arbeitsdefinition vor, die unabhängig vom Menschen ist und von messbaren Größen von Systemen abhängt [1]. Dazu betrachten wir Systeme, die mehr oder weniger selbstständig Probleme lösen können. Beispiele solcher Systeme können z. B. Organismen, Gehirne, Roboter, Automobile, Smartphones oder Accessoires sein, die wir am Körper tragen (Wearables). Systeme mit unterschiedlichem Grad von Intelligenz sind aber auch z. B. Fabrikanlagen (Industrie 4.0), Verkehrssysteme oder Energiesysteme (smart grids), die sich mehr oder weniger selbstständig steuern und zentrale Versorgungsprobleme lösen. Der Grad der Intelligenz solcher Systeme hängt vom Grad der Selbstständigkeit, von der Komplexität des zu lösenden Problems und der Effizienz des Problemlösungsverfahrens ab.

Es gibt danach also nicht „die“ Intelligenz, sondern Grade von Intelligenz. Komplexität und Effizienz sind in der Informatik und den Ingenieurwissenschaften messbare Größen. Ein autonomes Fahrzeug hat danach einen Grad von Intelligenz, der vom Grad seiner Fähigkeit abhängt, einen angegebenen Zielort selbstständig und effizient zu erreichen. Es gibt bereits mehr oder weniger autonome Fahrzeuge. Der Grad ihrer Selbstständigkeit ist technisch genau definiert. Die Fähigkeit unserer Smartphones, sich mit uns zu unterhalten, verändert sich ebenfalls. Jedenfalls deckt unsere Arbeitsdefinition intelligenter Systeme die For-

schung ab, die in Informatik und Technik unter dem Titel „Künstliche Intelligenz“ bereits seit vielen Jahren erfolgreich arbeitet und intelligente Systeme entwickelt [2].

► **Arbeitsdefinition** Ein System heißt intelligent, wenn es selbstständig und effizient Probleme lösen kann. Der Grad der Intelligenz hängt vom Grad der Selbstständigkeit, dem Grad der Komplexität des Problems und dem Grad der Effizienz des Problemlösungsverfahrens ab.

Es ist zwar richtig, dass intelligente technische Systeme, selbst wenn sie hohe Grade der selbstständigen und effizienten Problemlösung besitzen, letztlich von Menschen angestoßen wurden. Aber auch die menschliche Intelligenz ist nicht vom Himmel gefallen und hängt von Vorgaben und Einschränkungen ab. Der menschliche Organismus ist ein Produkt der Evolution, die voller molekular und neuronal kodierter Algorithmen steckt. Sie haben sich über Jahrtausende entwickelt und sind nur mehr oder weniger effizient. Häufig spielten Zufälle mit. Dabei hat sich ein hybrides System von Fähigkeiten ergeben, das keineswegs „die“ Intelligenz überhaupt repräsentiert. Einzelne Fähigkeiten haben KI und Technik längst überholt oder anders gelöst. Man denke an Schnelligkeit der Datenverarbeitung oder Speicherkapazitäten. Dazu war keineswegs „Bewusstsein“ wie bei Menschen notwendig. Organismen der Evolution wie Stabheuschrecken, Wölfe oder Menschen lösen ihre Probleme unterschiedlich. Zudem hängt Intelligenz in der Natur keineswegs von einzelnen Organismen ab. Die Schwarmintelligenz einer Tierpopulation entsteht durch das Zusammenwirken vieler Organismen ähnlich wie in den intelligenten Infrastrukturen, die uns bereits in Technik und Gesellschaft umgeben.

Die Neuroinformatik versucht, die Arbeitsweise von Nervensystemen und Gehirnen in mathematischen und technischen Modellen zu simulieren. In diesem Fall arbeiten KI-Forscher wie Naturwissenschaftler, die Modelle der Natur testen wollen. Das kann für die Technik interessant sein, muss aber nicht. KI-Forscher arbeiten häufig als Ingenieure, die effektive Problemlösungen unabhängig von der Naturvorlage finden. Das trifft auch für kognitive Fähigkeiten wie Sehen, Hören, Fühlen und Denken zu, wie das moderne Software-Engineering zeigt. Auch im Fall des Fliegens war die Technik erst dann erfolgreich, als sie die Gesetze der

Aerodynamik verstanden hatte und z. B. mit Düsenflugzeugen andere Lösungen als in der Evolution entwickelte.

In Kap. 2 beginnen wir mit einer „Kurzen Geschichte der KI“, die mit den großen Computerpionieren des 20. Jahrhunderts verbunden ist. Dem Computer wurde zunächst logisches Schließen beigebracht. Die dazu entwickelten Computersprachen werden bis heute in der KI benutzt. Logisch-mathematisches Schließen führt zu automatischen Beweisen, die Computerprogramme sichern helfen. Andererseits ist ihre Analyse mit tiefliegenden erkenntnistheoretischen Fragen der KI verbunden (Kap. 3). Um aber gezielt Probleme in unterschiedlichen Fachgebieten zu lösen, reichen allgemeine Methoden nicht aus. Wissensbasierte Expertensysteme simulierten erstmals Diagnosen von Medizinern und Analysen von Chemikern. Heute gehören Expertensysteme zum Alltag in Forschung und Beruf, ohne dass sie noch als „künstliche Intelligenz“ wahrgenommen werden (Kap. 4). Einer der spektakulärsten Durchbrüche der KI sind sprachverarbeitende Systeme, da Sprache traditionell als Domäne des Menschen gilt. Die dabei verwendeten Tools zeigen, wie unterschiedlich Technik und Evolution Probleme lösen können (Kap. 5).

Natürliche Intelligenz entstand in der Evolution. Es liegt daher nahe, die Evolution durch Algorithmen zu simulieren. Genetische und evolutionäre Algorithmen werden mittlerweile auch in der Technik angewendet (Kap. 6). Biologische Gehirne ermöglichen nicht nur erstaunliche kognitive Leistungen wie Sehen, Sprechen, Hören, Fühlen und Denken. Sie arbeiten auch wesentlich effizienter als energiefressende Supercomputer. Neuronale Netze und Lernalgorithmen sollen diese Fähigkeiten entschlüsseln (Kap. 7). Der nächste Schritt sind humanoide Roboter in menschenähnlicher Gestalt, die mit Menschen am Arbeitsplatz und im Alltag zusammen wirken. In einem stationären Industrieroboter werden Arbeitsschritte in einem Computerprogramm festgelegt. Soziale und kognitive Roboter müssen demgegenüber lernen, ihre Umwelt wahrzunehmen, selbstständig zu entscheiden und zu handeln. Dazu muss intelligente Software mit Sensortechnologie verbunden werden, um diese Art der sozialen Intelligenz zu realisieren (Kap. 8).

Automobile werden bereits als Computer auf vier Rädern bezeichnet. Als autonome Fahrzeuge erzeugen sie intelligentes Verhalten, das den Menschen als Fahrer mehr oder weniger vollständig ersetzen soll. Welche Anwendungsszenarien sind damit in Verkehrssystemen verbun-

den? Wie die Schwarmintelligenz in der Natur zeigt, ist Intelligenz nicht auf einzelne Organismen beschränkt. Im Internet der Dinge können Objekte und Geräte mit intelligenten Softwareschnittstellen und Sensoren versehen werden, um kollektiv Probleme zu lösen. Ein aktuelles Beispiel ist das industrielle Internet, in dem Produktion und Vertrieb sich weitgehend selbstständig organisieren. Eine Fabrik wird dann nach unserer Arbeitsdefinition intelligent. Allgemein spricht man mittlerweile von Cyberphysical Systems, die Smart Cities und Smart Grids ebenso erfassen (Kap. 9).

Seit ihrer Entstehung ist die KI-Forschung mit großen Visionen über die Zukunft der Menschheit verbunden. Wird es neuromorphe Computer geben, die das menschliche Gehirn vollständig simulieren können? Wie unterscheiden sich analoge Verfahren der Natur und digitale Technik? Konvergieren die Technologien des Künstlichen Lebens mit Künstlicher Intelligenz? Das Buch diskutiert neue Forschungsergebnisse über logisch-mathematische Grundlagen und technische Anwendungen von analogen und digitalen Verfahren.

Bei aller Nüchternheit der alltäglichen KI-Forschung motivieren und beeinflussen Hoffnungen und Ängste die Entwicklung hoch technisierter Gesellschaften. Gerade in den Hochburgen der amerikanischen Informations- und Computertechnologie wie z. B. Silicon Valley glaubt man, einen Zeitpunkt („Singularität“) voraussagen zu können, ab dem KI den Menschen ablösen wird. Die Rede ist bereits von einer kollektiven Superintelligenz.

Einerseits würde auch eine Superintelligenz, wie in diesem Buch gezeigt wird, den Gesetzen der Logik, Mathematik und Physik unterliegen. Wir benötigen daher fachübergreifende Grundlagenforschung, damit uns die Algorithmen nicht aus dem Ruder laufen. Andererseits fordern wir Technikgestaltung: Nach den Erfahrungen der Vergangenheit sollten wir zwar die Chancen erkennen, aber auch genau überlegen, zu welchem Zweck und Nutzen wir KI in Zukunft entwickeln sollten. KI muss sich als Dienstleistung in der Gesellschaft bewähren [2]. Das ist ihr ethischer Maßstab (Kap. 10).

Literatur

1. Mainzer K (2003) KI – Künstliche Intelligenz. Grundlagen intelligenter Systeme. Wissenschaftliche Buchgesellschaft, Darmstadt
2. DFKI (Deutsches Forschungszentrum für Künstliche Intelligenz). <http://www.dfki.de/web>. Zugegriffen: 8.1.2016

2.1 Ein alter Menschheitstraum

Ein Automat ist im antiken Sprachgebrauch ein Apparat, der selbstständig (autonom) agieren kann. Selbsttätigkeit charakterisiert nach antiker Auffassung lebende Organismen. Berichte über hydraulische und mechanische Automaten werden bereits in der antiken Literatur vor dem Hintergrund der damaligen Technik erwähnt. In der jüdischen Tradition wird Ende des Mittelalters der Golem als menschenähnliche Maschine beschrieben. Mit Buchstabenkombinationen des „Buchs der Schöpfung“ (hebr.: Sefer Jezira) kann der Golem programmiert werden – zum Schutz des jüdischen Volkes in Verfolgungszeiten.

Mit Beginn der Neuzeit wird Automation technisch-naturwissenschaftlich angegangen. Aus der Renaissance sind Leonardo da Vincis Konstruktionspläne für Automaten bekannt. Im Zeitalter des Barocks werden Spielautomaten auf der Grundlage der Uhrmachertechnik gebaut. P. Jaquet-Droz entwirft ein kompliziertes Uhrwerk, das in eine menschliche Puppe eingebaut war. Seine „Androiden“ spielen Klavier, zeichnen Bilder und schreiben Sätze. Der französische Arzt und Philosoph J. O. de Lamettrie bringt die Auffassung von Leben und Automaten im Zeitalter der Mechanik auf den Punkt: „Der menschliche Körper ist eine Maschine, die ihre (Antriebs-)Feder selbst spannt“ [1].

Der barocke Universalgelehrte A. Kircher (1602–1680) fördert bereits das Konzept einer Universalsprache, in der alles Wissen dargestellt werden soll. Hier schließt der Philosoph und Mathematiker G. W. Leibniz unmittelbar an und entwirft das folgenreiche Programm einer „Universalmathematik“ (mathesis universalis). Leibniz (1646–1716) will Denken und Wissen auf Rechnen zurückführen, um dann alle wissenschaftlichen Probleme durch Rechenkalküle lösen zu können. In seinem Zeitalter der Mechanik stellt man sich die Natur wie ein perfektes Uhrwerk vor, in dem jeder Zustand wie durch ineinandergreifende Zahnräder determiniert wird. Entsprechend führt eine mechanische Rechenmaschine jeden Rechenschritt einer Rechenfolge nacheinander aus. Leibnizens Dezimalmaschine für die vier Grundrechenarten ist die Hardware seiner Rechenkalküle. Grundlegend ist die Idee einer universellen symbolischen Sprache (lingua universalis), in der unser Wissen nach dem Vorbild der Arithmetik und Algebra repräsentiert werden kann. Gemeint ist ein Verfahren, mit dem „Wahrheiten der Vernunft wie in der Arithmetik und Algebra so auch in jedem anderem Bereich, in dem geschlossen wird, gewissermaßen durch einen Kalkül erreicht werden können“ [2].

Die weitere technische Entwicklung von den Dezimalrechenmaschinen für die vier Grundrechenarten zur programmgesteuerten Rechenmaschine fand nicht in der Gelehrtenstube, sondern in den Manufakturen des 18. Jahrhunderts statt. Dort wird das Weben von Stoffmustern zunächst mit Walzen nach dem Vorbild barocker Spielautomaten, dann durch hölzerne Lochkarten gesteuert. Diese Idee der Programmsteuerung wendet der britische Mathematiker und Ingenieur C. Babbage (1792–1871) auf Rechenmaschinen an. Seine „Analytical Engine“ sah neben einem vollautomatischen Rechenwerk aus Zahnrädern für die vier Grundrechenarten und einem Zahlenspeicher weiterhin eine Lochkartensteuerungseinheit, ein Dateneingabegerät für Zahlen und Rechenvorschriften und eine Datenausgabevorrichtung mit Druckwerk vor [3]. Wenn auch die technische Funktionstüchtigkeit beschränkt war, so wird die wissenschaftliche und wirtschaftliche Bedeutung sequentieller Programmsteuerung im Zeitalter der Industrialisierung richtig erkannt.

Babbage philosophiert auch über Analogien und Unterschiede seiner Rechenmaschinen zu lebenden Organismen und dem Menschen. Seine Mitstreiterin und Lebensgefährtin Lady Ada Lovelace, Tochter des romantischen Dichters Lord Byron, prophezeit bereits: „Die Analytical

Engine wird andere Dinge außer Zahlen bearbeiten. Wenn man Tonhöhen und Harmonien auf sich drehende Zylinder überträgt, dann könnte diese Maschine umfangreiche und auf wissenschaftliche Weise erzeugte Musikstücke jeder Komplexität und Länge komponieren. Allerdings kann sie nur das schaffen, was wir ihr zu befehlen wissen“ [4]. In der Geschichte der KI wird dieses Argument der Lady Lovelace immer wieder genannt, wenn es um die Kreativität von Computern geht.

Durch Elektrodynamik und elektrotechnische Industrie in der zweiten Hälfte des 19. Jahrhunderts werden neue technische Voraussetzungen zur Konstruktion von Rechnern gelegt. Während Holleriths Tabulierungs- und Zählmaschine zum Einsatz kommt, denkt der spanische Ingenieur Torres y Quevedo über Steuerungsprobleme für Torpedos und Boote nach und konstruiert 1911 den ersten Schachautomaten für einen Endkampf Turm-König gegen König.

Elektrizität, Licht und Strom inspirieren auch Literaten, Science-Fiction Autoren und die beginnende Filmindustrie. 1923 erdichtet der tschechische Schriftsteller Capek eine Familie von Robotern, mit denen die Menschheit von schwerer Arbeit befreit werden soll. Schließlich werden die Roboter wenigstens im Roman mit Emotionen versehen. Als Maschinenmenschen können sie ihr Sklavendasein nicht länger ertragen und proben den Aufstand gegen ihre menschlichen Herren. In den Kinos laufen Filme wie „Homunculus“ (1916), „Alraune“ (1918) und „Metropolis“ (1926).

In der Industrie- und Militärforschung werden in den 30er Jahren zwar erste Spezialrechner für begrenzte Rechenaufgaben gebaut. Grundlegend für die KI-Forschung wird aber die Entwicklung von universellen programmgesteuerten Computern, die für unterschiedliche Anwendungen programmiert werden können. Am 11. April 1936 meldet der deutsche Ingenieur K. Zuse (1910–1995) seine „Verfahren zur selbsttätigen Durchführung von Rechnungen mit Hilfe von Rechenmaschinen“ zum Patent an [5]. 1938 ist mit der Z1 eine erste mechanische Version fertig, die 1941 durch die Z3 mit elektromechanischen Relaischaltern ersetzt wird.

Was der Ingenieur Zuse technisch entwirft, wird von dem britischen Logiker und Mathematiker A. M. Turing (1912–1954) ebenfalls im Jahr 1936 als logisch-mathematischer Begriff definiert: Was ist überhaupt ein maschinelles Rechenverfahren, unabhängig von sei-

ner technischen Umsetzung? Turings ideale Rechenmaschine setzt neben einem unbegrenzten Speicher nur kleinste und einfachste Programmbeefehle voraus, auf die im Prinzip jedes noch so komplizierte Computerprogramm zurückgeföhrt werden kann [6].

2.2 Turing-Test

Als Geburtsjahr der KI-Forschung im engeren Sinn gilt 1950, als Turing seinen berühmten Aufsatz „Computing Machinery and Intelligence“ veröffentlicht [7]. Hier findet sich der sogenannte „Turing-Test“. Turing schlägt vor, einer Maschine genau dann „Künstliche Intelligenz“ zuzusprechen, wenn ein Beobachter nicht in der Lage ist zu unterscheiden, ob er es mit einem Menschen oder einem Computer zu tun hat. Beobachter und Testsystem (Mensch oder Computer) kommunizieren über ein Terminal (heute z. B. mit Keyboardtastatur und Bildschirm). Turing stellt in seiner Arbeit Musterfragen und Musterantworten aus unterschiedlichen Anwendungsbereichen vor wie z. B.:

Beispiel

F: Schreiben Sie mir bitte ein Gedicht über die Firth of Forth-Brücke.

A: In diesem Punkt muss ich passen. Ich könnte nie ein Gedicht schreiben.

F: Addieren Sie 34.957 zu 70.764.

A: (wartet ca. 30 Sekunden und gibt dann die Antwort) 105.721.

F: Spielen Sie Schach?

A: Ja.

F: Mein König steht auf e8; sonst habe ich keine Figuren mehr. Sie haben nur noch ihren König auf e6 und einen Turm auf h1. Sie sind am Zug. Wie ziehen Sie?

A: (nach einer Pause von 15 Sekunden) Th1-h8, matt.

Turing ist 1950 überzeugt: „Ich glaube, dass am Ende dieses Jahrhunderts die allgemeinen Ansichten der Gelehrten sich soweit geändert haben werden, dass man ohne Widerspruch von denkenden Maschinen reden können.“ Dass Computer heute schneller und genauer rechnen und besser Schach spielen, kann tatsächlich kaum noch bestritten

werden. Menschen irren aber auch, täuschen, sind ungenau und geben ungefähre Antworten. Das ist nicht nur ein Mangel, sondern zeichnet sie manchmal sogar aus, um sich in unklaren Situationen zu Recht zu finden. Jedenfalls müssten diese Reaktionen auch von einer Maschine realisiert werden können. Dass Turings Testsystem kein Gedicht schreiben wollte, also den Kreativitätstest von Lady Lovelace nicht bestanden hat, konnte Turing kaum erschüttern. Welcher Mensch ist schon kreativ und kann Gedichte schreiben?

2.3 Vom „Allgemeinen Problemlöser“ zum Expertensystem

Als sich 1956 führende Forscher wie J. McCarthy, A. Newell, H. Simon u. a. zur Dartmouth-Konferenz über Maschinenintelligenz trafen, waren sie von Turings Frage „Can machines think?“ inspiriert. Bezeichnend war die interdisziplinäre Zusammensetzung dieser Konferenz aus Computerwissenschaftlern, Mathematikern, Psychologen, Linguisten und Philosophen. So trat die Gruppe um den universell gebildeten H. Simon, dem späteren Nobelpreisträger für Wirtschaftswissenschaften, für ein psychologisches Forschungsprogramm ein, um kognitive Prozesse menschlicher Problem- und Entscheidungsfindung auf dem Computer zu simulieren.

Die erste Phase der KI-Forschung (etwa Mitte der 50er bis Mitte der 60er Jahre) ist noch von euphorischen Erwartungen bestimmt [8, 9]. Ähnlich wie in Leibnizens *Mathesis Universalis* sollen allgemeine Problemlösungsverfahren für Computer formuliert werden. Nachdem Newell, Shaw und Simon 1957 mit dem LOGICAL THEORIST ein Beweisprogramm für die ersten 38 Sätze aus Russells und Whiteheads Logikbuch „*Principia Mathematica*“ angegeben hatten, sollte das GPS (General-Problem-Solver)-Programm 1962 die heuristische Basis für menschliches Problemlösen überhaupt festlegen. Die Enttäuschung war groß angesichts der praktischen Ergebnisse. Erfolgreicher erwiesen sich erste spezialisierte Programme wie STUDENT zum Lösen von Algebra-Aufgaben oder ANALOGY zur Mustererkennung analoger Objekte. Es zeigte sich, dass erfolgreiche KI-Programme von geeigneten Wissensbasen („Datenbanken“) und schnellen Abrufverfahren abhängen.

In der zweiten Phase der KI (etwa Mitte der 60er bis Mitte der 70er Jahre) lässt sich eine verstärkte Hinwendung zum praktischen und spezialisierten Programmieren beobachten. Typisch ist die Konstruktion von spezialisierten Systemen, Methoden zur Wissensrepräsentation und ein Interesse an natürlichen Sprachen. Am MIT entwickelte J. Moser das Programm MACSYMAL, das eigentlich eine Sammlung von Spezialprogrammen zur Lösung von mathematischen Problemen in der üblichen mathematischen Symbolik darstellte. Weiterführende Programme dieser Art (z. B. zum Integrieren und Differenzieren) sind bis heute praktisch im Gebrauch.

T. Winograd stellte 1972 ein Robotikprogramm vor, um verschieden geformte und gefärbte Bauklötze mit einem Magnetarm zu manipulieren. Dazu wurden die Bauklötze mit ihren Eigenschaften und Ortsangaben in Datenstrukturen repräsentiert. Programmierung der Ortsangaben wurde mit dem Magnetarm durch Umstellung der Bauklötze ausgeführt.

In der dritten Phase der KI (ca. Mitte der 70er bis Mitte der 80er Jahre) rücken wissensbasierte Expertensysteme in den Vordergrund, die erste praktische Anwendungen versprachen. Abgegrenztes und überschaubares Spezialwissen menschlicher Experten wie z. B. von Ingenieuren und Ärzten sollte für den tagtäglichen Gebrauch zur Verfügung gestellt werden. Bei wissensbasierten Expertensystemen handelt es sich um KI-Programme, die Wissen über ein spezielles Gebiet speichern und aus dem Wissen automatisch Schlussfolgerungen ziehen, um konkrete Lösungen zu finden oder Diagnosen von Situationen bereitzustellen.

Im Unterschied zum menschlichen Experten ist das Wissen eines Expertensystems beschränkt. Es besitzt kein allgemeines Hintergrundwissen, keine Erinnerungen, Gefühle und Motivationen, die von Person zu Person trotz gemeinsamem Spezialwissen unterschiedlich sein können: Ein älterer Hausarzt, der eine Familie über Generationen kennt, wird anderes Hintergrundwissen in die Diagnose eines Familienmitgliedes einfließen lassen als der junge Spezialist, der soeben von der Universität kommt.

Wissen ist ein Schlüsselfaktor in der Darstellung eines Expertensystems. Wir unterscheiden zwei Arten von Wissen. Die eine Art des Wissens betrifft die Fakten des Anwendungsbereichs, die in Lehrbüchern und Zeitschriften festgehalten werden. Ebenso wichtig ist die Praxis im jeweiligen Anwendungsbereich als Wissen der zweiten Art. Es handelt

sich um heuristisches Wissen, auf dem Urteilsvermögen und jede erfolgreiche Problemlösungspraxis im Anwendungsbereich beruhen. Es ist Erfahrungswissen, die Kunst des erfolgreichen Vermutens, das ein menschlicher Experte nur in vielen Jahren Berufspraxis erwirbt.

E. A. Feigenbaum, einer der Pioniere dieser Entwicklung, verglich Mitte der 80er Jahre die Entwicklung wissensbasierter Expertensysteme mit der Geschichte der Automobilindustrie. In der Welt der KI wäre es sozusagen 1890, als die ersten Automobile auftraten. Sie waren handbetriebene pferdelose Wagen, aber bereits Automobile, d. h. selbstbewegliche Fahrzeuge. So wie seinerzeit Henry Ford die ersten Prototypen für die Massenproduktion weiterentwickelte, so würden nach Feigenbaum auch wissensbasierte Systeme in die Massenproduktion gehen. Wissensbasierte Systeme wurden also als „Automobile des Wissens“ verstanden [10].

Literatur

1. de La Mettrie JO (2009) *L'homme machine / Die Maschine Mensch*. Meiner, Hamburg (übers. u. hrsg. v. C. Becker)
2. Leibniz GW (1996) *Philos. Schr. VII*. Olms, Hildesheim, S 32 (hrsg. C.I. Gerhardt, repr.)
3. Babbage C (1975) On the mathematical powers of the calculating engine. In: Randell B (Hrsg) *The Origins of Digital Computers – Selected Papers*. Springer, Berlin, S 17–52 (unpublished manuscript, Dec. 1837)
4. Lovelace C (1842) Translator's notes to an article on Babbage's Analytical Engine. *Scientific Memoirs* 3:691–731 (hrsg. V. R. Taylor)
5. Zuse K (1936) Verfahren zur selbsttätigen Durchführung von Rechnungen mit Hilfe von Rechenmaschinen. Deutsche Patentanmeldung Z 23624 (11. April 1936)
6. Turing AM (1936–1937) On computable numbers, with an application to the Entscheidungsproblem. *Proc London Math Soc Ser 2*(42):230–265
7. Turing AM (1987) *Computing machinery and intelligence (1950)*. In: Turing AM (Hrsg) *Intelligence Service*. Schriften. Brinkmann u. Bose, Berlin, S 147–182
8. Feigenbaum EA, Feldman J (1963) *Computers and Thought*. McGraw-Hill, New York
9. Mainzer K (1985) Der Intelligenzbegriff in erkenntnis- und wissenschaftstheoretischer Sicht. In: Strombach W, Tauber MJ, Reusch B (Hrsg) *Der Intelligenzbegriff in den verschiedenen Wissenschaften*. Schriftenreihe der Österreichischen Computer Gesellschaft, Bd. 28. Oldenbourg, Wien, S 41–56

10. Weitere Quellen zu den historischen Wurzeln der KI in K. Mainzer (1994), Computer – Neue Flügel des Geistes? Die Evolution computergestützter Technik, Wissenschaft, Kultur und Philosophie. De Gruyter, Berlin/New York, S 103 ff. Die Entwicklung der KI in den 1990er Jahren ist dokumentiert in der deutschen KI-Zeitschrift „Künstliche Intelligenz“ des Fachbereichs 1 der Gesellschaft für Informatik e. V.

3.1 Was heißt logisches Schließen?

In der ersten Phase der KI-Forschung war die Suche nach allgemeinen Problemlösungsverfahren wenigstens in der formalen Logik erfolgreich. Dort wurde ein mechanisches Verfahren angegeben, um die logische Allgemeingültigkeit von Formeln zu beweisen. Das Verfahren konnte auch von einem Computerprogramm ausgeführt werden und leitete in der Informatik das automatische Beweisen ein.

Der Grundgedanke ist einfach zu verstehen. In der Algebra werden Buchstaben x, y, z, \dots durch Rechenoperationen wie z. B. Addieren (+) oder Subtrahieren (−) verbunden. Die Buchstaben dienen als Leerstellen (Variablen), um Zahlen einzusetzen. In der formalen Logik werden Aussagen durch Variablen A, B, C, \dots bezeichnet, die durch logische Junktoren wie z. B. „und“ (\wedge), „oder“ (\vee), „wenn-dann“ (\rightarrow), „nicht“ (\neg) verbunden werden. Die Aussagenvariablen dienen als Leerstellen, um Aussagen einzusetzen, die entweder wahr oder falsch sind. So ist z. B. $A \wedge B$ eine logische Formel, die durch Einsetzung der wahren Aussagen $1 + 3 = 4$ für A und $4 = 2 + 2$ für B in die wahre Aussage $1 + 3 = 4 \wedge 4 = 2 + 2$ übergeht. In der Arithmetik ergibt sich daraus der wahre Schluss $1 + 3 = 4 \wedge 4 = 2 + 2 \rightarrow 1 + 3 = 2 + 2$. Allgemein ist aber der Schluss $A \wedge B \rightarrow C$ nicht wahr. Demgegenüber ist der Schluss $A \wedge B \rightarrow A$ logisch allgemeingültig, da für die Einsetzung von beliebigen wahren

In einem entsprechenden Computerprogramm terminiert dieses Verfahren für die Aussagenlogik. Es zeigt also in endlicher Zeit, ob die vorgelegte logische Formel allgemeingültig ist. Allerdings wächst die Rechenzeit nach den bisher bekannten Verfahren exponentiell mit der Anzahl der Literale einer Formel. Was nun die „Künstliche Intelligenz“ betrifft, so können Computerprogramme mit der Resolutionsmethode automatisch über die Allgemeingültigkeit logischer Schlüsse wenigstens in der Aussagenlogik im Prinzip entscheiden. Menschen hätten große Schwierigkeiten, bei komplizierten und langen Schlüssen den Überblick zu bewahren. Zudem sind Menschen wesentlich langsamer. Mit steigender Rechenkapazität können Maschinen also wesentlich effizienter diese Aufgabe des logischen Schließens erledigen.

In der Prädikatenlogik werden Aussagen in Eigenschaften (Prädikate) zerlegt, die Objekten zu- oder abgesprochen werden. So wird in der Aussage $P(a)$ (z. B. „Anna ist Studentin“) einem Individuum namens „Anna“ (a) das Prädikat „Studentin“ (P) zugesprochen. Diese Aussage ist wieder entweder wahr oder falsch. In einer prädikativen Aussageform $P(x)$ werden Leerstellen (Individuenvariablen) x, y, z, \dots verwendet, für die Individuen a, b, c, \dots einer angenommenen Anwendungsdomäne (z. B. die Studierenden einer Universität) eingesetzt werden können. Neben den logischen Verknüpfungen der Aussagenlogik können nun auch Allquantoren $\bigwedge x$ („Für alle x “) und Existenzquantoren $\bigvee x$ („Es gibt ein x “) verwendet werden. So ist z. B. $\bigwedge x P(x) \rightarrow \bigvee x P(x)$ ein allgemeingültiger Schluss der Prädikatenlogik.

Für die Formeln der Prädikatenlogik lässt sich ebenfalls ein verallgemeinertes Resolutionsverfahren angeben, um wieder aus der Annahme der allgemeinen Ungültigkeit einer Formel einen Widerspruch abzuleiten. Dazu muss eine Formel der Prädikatenlogik in eine Normalform gebracht werden, aus deren Klauseln sich mechanisch ein Widerspruch ableiten lässt. Da aber in der Prädikatenlogik (im Unterschied zur Aussagenlogik) nicht allgemein die Allgemeingültigkeit einer Formel entschieden werden kann, kann es vorkommen, dass das Resolutionsverfahren kein Ende findet. Das Computerprogramm läuft dann unbegrenzt weiter. Es kommt dann darauf an, Teilklassen zu finden, in denen das Verfahren nicht nur effizient, sondern überhaupt terminiert. Maschinelle Intelligenz kann zwar die Effizienz von Entscheidungsprozessen steigern und sie

beschleunigen. Sie ist aber auch (wie menschliche Intelligenz) den prinzipiellen Grenzen logischer Entscheidbarkeit unterworfen.

3.2 KI-Programmiersprache PROLOG

Um ein Problem mit einem Computer zu lösen, muss das Problem in eine Programmiersprache übersetzt werden. So war FORTRAN eine der ersten Programmiersprachen, bei denen ein Programm aus einer Folge von Befehlen an den Computer besteht wie z. B. „springe an die Stelle z im Programm“, „schreibe in die Variable x den Wert a“. Im Zentrum stehen Variablen, also maschinentechnisch Register- bzw. Speicherzellen, in denen Eingabewerte gespeichert und verarbeitet werden. Wegen der eingegebenen Befehle spricht man auch von einer imperativen Programmiersprache.

In einer prädikativen Programmiersprache wird demgegenüber Programmierung als Beweisen in einem System von Tatsachen aufgefasst. Diese Wissensdarstellung ist aus der Logik wohlvertraut. Eine entsprechende Programmiersprache heißt „Programming in Logic“ (PROLOG), die seit Anfang der 1970er Jahre in Gebrauch ist [4, 5, 6]. Grundlage ist die Prädikatenlogik, die wir schon in Abschn. 3.1 kennengelernt haben. Wissen wird in der Prädikatenlogik als Menge von wahren Aussagen dargestellt. Um Wissensverarbeitung geht es in der KI-Forschung. Daher wurde PROLOG zu einer zentralen Programmiersprache der KI.

Hier wollen wir einige Bausteine von PROLOG einführen, um den Zusammenhang mit der Wissensverarbeitung zu verdeutlichen. Die logische Aussage „Die Objekte O_1, \dots, O_n stehen in der Relation R“ entspricht einem Sachverhalt bzw. Faktum, dem man in der Prädikatenlogik die allgemeine Form $R(O_1, \dots, O_n)$ gibt. In PROLOG schreibt man:

NAME(O_1, \dots, O_n),

wobei „NAME“ ein beliebiger Name einer Relation ist. Zeichenfolgen, die in der syntaktischen Form von Fakten wiedergegeben werden, heißen Literale.

Beispiel

Ein Beispiel eines Faktums bzw. Literals lautet:

verheiratet (sokrates, xantippe),
 verheiratet (abélard, eloise),
 ist lehrer (sokrates, platon),
 ist lehrer (abélard, eloise).

Behauptungen und Beweise über vorgegebene Fakten lassen sich nun in Frage-Antwort-Systemen darstellen. Dabei werden Fragen mit einem Fragezeichen und Ausgaben des Programms mit einem Stern bezeichnet:

? verheiratet (sokrates, xantippe),
 * ja,
 ? ist lehrer (sokrates, xantippe),
 * nein.

Fragen können sich auch gezielt auf Objekte beziehen, für die in diesem Fall Variablen eingesetzt werden. In Programmiersprachen werden dafür anschauliche Bezeichnungen wie z. B. „mann“ für einen beliebigen Mann bzw. „lehrer“ für einen beliebigen Lehrer verwendet:

? verheiratet (mann, xantippe),
 * mann = sokrates,
 ? ist lehrer (lehrer, xantippe),
 * lehrer = sokrates.

Allgemein lautet eine Frage in PROLOG „sind L_1 und L_2 und ... und L_n wahr?“ oder kurz:

? L_1, L_2, \dots, L_n

wobei L_1, L_2, \dots, L_n Literale sind. Logische Schlussregeln wie der direkte Schluss (modus ponens) lauten „wenn L_1 und L_2 und ... und L_n wahr ist, dann ist auch L wahr“ oder kurz:

$L :- L_1, L_2, \dots, L_n.$