

Ulrike Barthelmeß
Ulrich Furbach

Künstliche Intelligenz aus ungewohnten Perspektiven

Ein Rundgang mit Bergson,
Proust und Nabokov

 Springer Vieweg

Die blaue Stunde der Informatik

Die blaue Stunde – die Zeit am Morgen zwischen Nacht und Tag, die Zeit am Abend ehe die Nacht anbricht. Wenn alles möglich scheint, die Gedanken schweifen, wenn Zeit für anregende Gespräche ist und Neugier auf Zukünftiges wächst, auf alles, was der nächste Tag bringt.

Genau hier setzt diese Buchreihe rund um Themen der Informatik an: Was war, was ist, was wird sein, was könnte sein?

Von lesenswerten Biographien über historische Betrachtungen bis hin zu aktuellen Themen umfasst diese Buchreihe alle Perspektiven der Informatik – und geht noch darüber hinaus. Mal sachlich, mal nachdenklich und mal mit einem Augenzwinkern lädt die Reihe zum Weiter- und Querdenken ein. Für alle, die die bunte Welt der Technik entdecken möchten.

Weitere Bände in der Reihe <http://www.springer.com/series/15985>

Ulrike Barthelmeß · Ulrich Furbach

Künstliche Intelligenz aus ungewohnten Perspektiven

Ein Rundgang mit Bergson, Proust und
Nabokov

 Springer Vieweg

Ulrike Barthelmeß
Koblenz, Deutschland

Ulrich Furbach
Koblenz, Deutschland

Die blaue Stunde der Informatik
ISBN 978-3-658-24569-6 ISBN 978-3-658-24570-2 (eBook)
<https://doi.org/10.1007/978-3-658-24570-2>

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

Springer Vieweg

© Springer Fachmedien Wiesbaden GmbH, ein Teil von Springer Nature 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 allgemein beschreibenden Bezeichnungen, Marken, Unternehmensnamen etc. in diesem Werk bedeutet nicht, dass diese frei durch jedermann benutzt werden dürfen. Die Berechtigung zur Benutzung unterliegt, auch ohne gesonderten Hinweis hierzu, den Regeln des Markenrechts. Die Rechte des jeweiligen Zeicheninhabers sind zu beachten.

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.

Einbandabbildung: designed by eStudio Calamar S.L.

Springer Vieweg ist ein Imprint der eingetragenen Gesellschaft Springer Fachmedien Wiesbaden GmbH und ist ein Teil von Springer Nature

Die Anschrift der Gesellschaft ist: Abraham-Lincoln-Str. 46, 65189 Wiesbaden, Germany

Danksagung

Manche Teile dieses Buches wurden durch Diskussionen in Vorlesungen und Vorträgen zum Thema künstliche Intelligenz motiviert. Der Gedanke, die Behandlung von Erinnerungen in der KI und in der Literatur zu verknüpfen, wurde angeregt durch das Buch von Jonah Lehrer *Prousts Madeleine* [1]. Danken möchten wir Isabelle Barthelmeß, Beate Körner, Claudia Schon und Holger Schultheis für das Lesen und Kommentieren von Teilen des Buches. Besonderer Dank gilt Frau Dr. Sabine Kathke, die als Lektorin äußerst hilfreich und wertvoll für uns war.

Literatur

1. Lehrer J (2010) Prousts Madeleine. Piper, München

Inhaltsverzeichnis

1	Einleitung	1
2	Künstliche Intelligenz	7
3	Körper und Geist	25
4	Zeit und Erinnern	39
5	Repräsentation von Wissen	49
6	Mnemotechnik	69
7	Die Kunst des Sehens	77
8	Freier Wille	95
9	Erinnern – ein kreativer Akt	107
10	Rahmen und Struktur	135
11	Bewusstsein	153
12	Sprache	167
13	Nachwort	187

Abbildungsverzeichnis

Abb. 2.1	Imitationsspiel. Links: Computer und Mann, beide über der Linie, geben sich als Frau aus. Der Fragesteller unter der Linie soll herausfinden, welcher von beiden ein Mann bzw. eine Frau ist. Rechts: Computer und Mensch geben sich als Mensch aus, der Fragesteller unter der Linie soll herausfinden, welcher von beiden ein Mensch ist	10
Abb. 5.1	Semantisches Netz mit den Konzepten Lebewesen, Vogel, Säugetier und Mensch. Sokrates ist eine Instanz von Mensch, und Säugetiere haben Lunge und sind sterblich	55
Abb. 5.2	Aktive Nervenzelle	62
Abb. 5.3	Künstliches neuronales Netz. An den roten Knoten liegen Eingabesignale an, die über die gewichteten Kanten weitergeleitet werden. Die grünen Knoten sind Ausgabeknoten, die schließlich die Ausgabesignale liefern. Zahlen an den Kanten bedeuten, dass die Signale entlang der Kanten mit diesem Faktor multipliziert, also verstärkt oder abgeschwächt, werden	63
Abb. 5.4	Perceptron mit drei Eingabe- (rot) und einem Ausgabeknoten (grün)	65

X Abbildungsverzeichnis

Abb. 5.5	Boolesche Funktionen. Die obere Kante mit dem festen Eingangswert 1 wird als Bias bezeichnet. Mit dem Gewicht an dieser Kante kann der Schwellwert des Neurons modifiziert werden, sodass er bei allen Neuronen gleichmäßig, hier > 0 , gesetzt werden kann. Links ein Ein-Neuron-Netz zur Berechnung der Booleschen Funktion Oder. Wenn x oder y den Wert 1 haben, liefert das Neuron den Wert 1, sonst 0. Wenn z. B. $x = 0$ und $y = 1$ ist, erreicht das Neuron den Wert $(1 * -0,5) + (1 * 0) + (1 * 1)$, also $+0,5$. Da $0,5 > 0$ gilt, feuert das Neuron und gibt den Wert 1 aus. Rechts ein Ein-Neuron-Netz zur Berechnung der Booleschen Funktion Und. Wenn x und y den Wert 1 haben, liefert das Neuron den Wert 1, sonst 0	65
Abb. 5.6	Neuronales Netz mit einem inneren verdeckten Knoten	66
Abb. 7.1	Gustav Klimt: Adele Bloch-Bauer I	81
Abb. 7.2	False positives aus	90
Abb. 7.3	Erzeugte Bilder und reale Objekte aus	90
Abb. 7.4	Die Wason-Selection-Task	93



1

Einleitung

Zusammenfassung Die Bedeutung der Informatik und der künstlichen Intelligenz wird in diesem einleitenden Kapitel erläutert. Es wird dafür plädiert, damit zusammenhängende Themen nicht nur aus der Sicht des Naturwissenschaftlers oder Technikers zu diskutieren, sondern auch Philosophen, Literaten und Künstler zur Sprache kommen zu lassen.

Technische Entwicklungen haben seit jeher unser Leben beeinflusst und verändert. Am deutlichsten wird das wohl am Beispiel der Dampfmaschine und der darauf folgenden industriellen Revolution im Europa des 19. und 20. Jahrhunderts. Die Mechanisierung der industriellen Fertigung hatte nicht nur die Arbeits-, sondern auch die Lebensbedingungen der Menschen verändert. Die gesellschaftlichen Strukturen wandelten sich, es setzte eine Landflucht ein, und die Städte in den Industriegebieten boten, zumindest für die Arbeiterklasse, ungesunde und menschenunwürdige Lebensumstände.

Die Informatik, international zumeist Computer Science genannt, hat in ähnlich drastischer Weise unsere Lebens- und Arbeitswelt geprägt. Als ich Anfang der 1970er-Jahre begann, Informatik zu studieren, musste ich der Frage, was ich studiere, immer auch eine Erklärung hinterherschicken: Informatik, das heißt, Computer und so . . . Damals wurden Computer höchstens in großen Unternehmen eingesetzt, die sich die immensen Anschaffungs- oder Leasingkosten der sogenannten Mainframe-Computer leisten konnten. Der erste PC kam ja erst in den 1980er-Jahren auf den Markt, und damit wandelte sich auch allmählich das Bild der Informatik in der Gesellschaft. Man konnte plötzlich mithilfe von Computern kommunizieren, die Ticketverkäufe von Fluggesellschaften wurden über Computernetzwerke abgewickelt,

Datenbanksysteme durchzogen die öffentliche Verwaltung, und wir lernten langsam, mit dieser Technik umzugehen. In Betrieben mussten sich Angestellte mit dem „Kollegen Computer“ anfreunden, Mikroprozessoren hielten Einzug in unsere Automobile und verbesserten den Komfort. Nicht ganz so von der Öffentlichkeit wahrgenommen, wurden Computer gleichzeitig auch immer wichtiger für die Rüstungsindustrie und damit auch für unsere Sicherheit. Mittlerweile durchdringt die Informatik unsere gesamte Arbeits- und Lebenswelt. Ganz langsam, relativ unspektakulär, aber nicht minder drastisch werden wir in so ziemlich allen Lebenslagen von Computersystemen und Netzen abhängig. Das Internet bildet sozusagen die Lebensadern unserer Gesellschaft. Lernen, Wissenserwerb, Finanzwesen oder Konsum sind ohne das weltweite Internet nicht mehr vorstellbar. Gleichzeitig, und je intensiver wir diese Netze nutzen, sind wir auch transparenter geworden – wir sprechen mittlerweile vom gläsernen Bürger. Die Suchmaschinen-, die sozialen Netzwerk- oder die Versandplattform-Betreiber wissen immer mehr von uns, über unsere Kauf- und Lebensgewohnheiten. Daten von Benutzern sind ein wichtiges Wirtschaftsgut geworden; Unternehmen, die Zugriff auf große Mengen von Benutzerdaten haben, erreichen riesige Marktwerte, ganze Geschäftsmodelle basieren auf diesen Kundendaten. Interessant sind diese Daten, weil sie mit Techniken, die im Bereich der Künstlichen-Intelligenz(KI)-Forschung entwickelt wurden, ausgewertet werden können, sodass vielfältige Rückschlüsse auf Personen, die diese Daten erzeugen, möglich sind. Natürlich können diese Informationen auch dazu dienen, uns das Leben angenehmer zu machen: Mein Online-Versandhändler weiß, welche Produkte ich bevorzuge, und lässt mir entsprechende Sonderangebote zukommen, mein persönlicher Assistent im Smartphone oder in der Lautsprecherbox im Wohnzimmer weiß, welchen Musikgeschmack ich habe, und macht mir entsprechende Vorschläge.

Diese KI-Techniken unterscheiden sich von den Algorithmen, die wir in der Informatik einsetzen. Natürlich handelt es sich hierbei auch um Algorithmen, also Computerprogramme, die von Menschen entwickelt worden sind, allerdings verfügen diese Programme über die Fähigkeit zu lernen. Dies bedeutet, dass sie während ihres Einsatzes sich selbst weiterentwickeln und modifizieren – in welche Richtung dies geht, hängt von den Daten ab, auf die die Verfahren während ihres Einsatzes treffen. Bei diesen sogenannten Machine-Learning(ML)-Verfahren lässt sich nicht von vorne herein im Detail absehen, wie sie sich weiterentwickeln werden. Gerade die ML-Verfahren haben in letzter Zeit der KI-Forschung und -Entwicklung einen immensen Aufschwung beschert: KI-Systeme gewinnen in Quizshows, sie besiegen in höchst komplexen Spielen, wie zum Beispiel Go, Weltklassenspieler, sie helfen unsere Fahrzeuge sicherer zu machen, indem sie den Fahrer in vielfältiger Weise

unterstützen, und sie beginnen gar Fahrzeuge autonom fahren zu lassen. KI-Systeme werden auch dazu benutzt, militärische Waffensysteme intelligent und womöglich sogar autonom zu machen – die Folgen für den Einsatz und die Auswirkungen auf militärische und politische Entscheidungen werden von vielen KI-Forschern als äußerst kritisch und gefährlich bewertet. Es gibt dazu einen Aufruf, autonome tödliche Waffen weltweit zu ächten, den bereits Tausende von Wissenschaftlern und namhafte Unternehmen unterzeichnet haben.¹

Wir leben und arbeiten mit Computersystemen, die sich intelligent verhalten oder gar intelligent sind. Diese Systeme können selbstständig Entscheidungen treffen und sich weiterentwickeln. Manche Forscher prognostizieren sogar KI-Systeme, die intelligenter und mächtiger als Menschen werden und womöglich die Menschheit unterdrücken könnten. Dies sind keineswegs nur einige Science-Fiction-Fans, sondern ernsthafte Wissenschaftler und Personen des öffentlichen Lebens. Was immer man darüber denkt und in welche Richtung die Diskussion geht – wir haben den Eindruck, dass so manche Argumente eher emotional als technisch fundiert sind, wenn Fragen wie „Können Maschinen intelligent sein, können sie Emotionen oder Bewusstsein haben, können sie uns Menschen überflügeln?“ diskutiert werden.

Sind dies alles Fragen, die nur mithilfe einer fundierten technisch naturwissenschaftlichen Bildung beantwortet werden können? Sollen wir die Diskussion über diese gesellschaftlich relevanten Fragestellungen den Naturwissenschaftlern alleine überlassen?

Wir meinen, dass wir alle uns mit diesen Fragen und mit den Folgeerscheinungen der KI auseinandersetzen sollten. Im Gegensatz zu Schwanitz in [1] möchten wir Bildung nicht nur auf Geisteswissenschaften und Kunst beschränken. Zur Bildung gehört sicherlich auch ein gewisses Verständnis naturwissenschaftlicher Zusammenhänge. Nur dann können wir wirklich verstehen, welche Entwicklungen unsere Gesellschaft betreffen und verändern. Dies ist eigentlich offensichtlich – seit jeher verwenden und diskutieren Geisteswissenschaftler, Philosophen und Literaten, aber auch bildende Künstler neue naturwissenschaftliche Erkenntnisse in ihrer Arbeit und denken sie weiter. Wir werden darauf eingehen, dass Marcel Proust sehr ausführlich über das Gedächtnis und das Erinnern schreibt, Vladimir Nabokov sich mit Erinnern und Zeit beschäftigt und Henri Bergson über das Verhältnis von Körper, Geist und Bewusstsein arbeitet. Und wir verweisen auf den Neurowissenschaftler und Nobelpreisträger Eric Kandel, der auch die Einflüsse der Naturwissenschaften auf die Künstler und Maler im Wien des beginnenden 20. Jahrhunderts beschreibt.

¹<https://autonomousweapons.org>.

Wir möchten im Folgenden einige dieser Arbeiten aus den Geisteswissenschaften, der Literatur und der Philosophie besprechen und ihre Themen mit Fragestellungen und Methoden aus der KI-Forschung in Zusammenhang bringen. Durch dieses Verketteten von natur- und ingenieurwissenschaftlichen Methoden mit Diskussionen aus der Kunst und den Kulturwissenschaften verbinden wir die Hoffnung, ein wenig zum Verständnis der tiefen gesellschaftlichen Veränderung, die mit der Verbreitung von KI-Techniken einhergeht, beizutragen. Vielleicht können dadurch aber auch KI-Treibende sehen, dass die Fragestellungen, die sie untersuchen, auch in völlig anderen Bereichen relevant sind und dort untersucht werden.

Die Auswahl der hier präsentierten KI-Techniken ist dabei keineswegs erschöpfend. Wir haben uns auf Themen aus den eigenen Forschungsbereichen, wie z. B. logisches Schließen, Wissen und Erinnern, konzentriert. Ergänzt wurden Themen, die in letzter Zeit in der KI-Forschung besonders relevant geworden sind, wie z. B. maschinelles Lernen und künstliche neuronale Netze.

Ähnliches gilt für die Themen und Techniken aus der Kunst und Geisteswissenschaft. Wir haben uns mit Autoren und Wissenschaftlern beschäftigt, die den Menschen, seine Körperlichkeit und sein Bewusstsein zum Gegenstand haben. Im Vordergrund stand die Frage, was Gedächtnis und Denken ausmacht, welche Rolle unsere Emotionen spielen, wenn wir uns als Menschen im Leben, in der Welt bewegen, wenn wir als Künstler die Möglichkeiten ausschöpfen, um neue originäre Sichtweisen auf uns und die Welt entwerfen.

Die Versuchung, Science-Fiction-Literatur heranzuziehen, um Parallelen zu wissenschaftlichen Entwicklungen zu diskutieren, ist natürlich sehr groß. Und es wäre sicher spannend, entsprechenden wissenschaftlich-technischen Spekulationen nachzugehen. Wir haben aber darauf verzichtet, da dies den Rahmen unserer Untersuchung sprengen würde, denn es geht uns hier nicht darum, mögliche Folgen möglicher Entwicklungen in einer fernen Zukunft mit ihren sozialen und moralischen Implikationen zu erörtern. Uns interessiert, inwiefern so scheinbar entlegene Bereiche wie KI, Kunst und Philosophie eine Verwandtschaft offenbaren. Vielleicht kann die KI sogar von geisteswissenschaftlichen Erkenntnissen profitieren, ähnlich der modernen Technik, bei der man sich unter der Überschrift Bionik clevere Lösungen aus der Natur abschaut?

Aufbau des Buches Im nächsten Kapitel wird das Forschungsgebiet KI charakterisiert, sodass dann in den darauf folgenden Kapiteln einzelne Techniken oder Methoden der KI erläutert und mit Ansätzen aus Philosophie und Kunst in Zusammenhang gebracht werden können. Darüber hinaus behandeln wir in einzelnen Kapiteln Themen wie Kreativität, freier Wille, Bewusstsein und Sprache. Die Kapitel können, müssen aber nicht in chronologischer

Reihenfolge gelesen werden. Inhaltliche Bezüge werden durch entsprechende Kapitelverweise aufgezeigt. Allerdings wäre die Lektüre des folgenden Kapitels über künstliche Intelligenz für das Verständnis der folgenden Ausführungen hilfreich. Einige der Kapitel beginnen mit frei assoziierten Gedanken, die zur Einstimmung auf das Thema beitragen mögen – sie sind jeweils kursiv gedruckt.

Literatur

1. Schwanitz D (1999) Bildung. Eichborn Verlag, Frankfurt a. M



2

Künstliche Intelligenz

Zusammenfassung Eine allgemeingültige Definition von künstlicher Intelligenz existiert leider nicht. Wir benutzen einen Aufsatz des Computerpioniers Alan Turing, um den Begriff zu durchleuchten, und wir beschreiben verschiedene Tests, um künstliche Intelligenz zu bewerten. Schließlich wird auch der Stand der Kunst des Forschungsgebietes skizziert.

In diesem Abschnitt wollen wir das Forschungsgebiet der künstlichen Intelligenz (KI) charakterisieren. In verschiedenen Lehrbüchern findet man über das Thema unterschiedliche Beschreibungen bzw. eigentlich Umschreibungen des Gebietes. Das Problem ist ganz einfach, dass es eigenartig erscheint, künstliche Intelligenz zu definieren, wenn schon der Begriff Intelligenz nicht klar definiert ist. Nun ist eine solche Situation aber gar nicht so ungewöhnlich – wir operieren häufig mit Begriffen, für die wir keine klare und eindeutige Definition haben. Nehmen wir Begriffe wie Glück oder Bewusstsein, zu denen zahllose Untersuchungen und Diskussionen existieren; selbst ein Begriff wie Pornografie ist nicht klar definierbar, obwohl er sogar in der Gesetzgebung und Rechtsprechung eine Rolle spielt. Ähnlich ist es mit der Intelligenz – wir verwenden den Begriff, wir haben sogar verschiedenste Tests, die irgendetwas wie Intelligenz messen, aber eine trennscharfe Definition gelingt uns nicht.

2.1 Was ist künstliche Intelligenz?

Anstatt jetzt eine Übersicht über die vielfältigen Definitionsversuche der Begriffe Intelligenz und KI zu geben, greifen wir zurück auf einen Aufsatz

aus dem Jahr 1950, als die Computerentwicklung noch in den Kinderschuhen steckte. Dies mag seltsam erscheinen, wenn man bedenkt, was die Informatik und die KI-Forschung seitdem an immensen Fortschritten vorzuweisen hat.

Der Aufsatz, den wir hier diskutieren wollen, hat den Titel „Computing machinery and intelligence“ und ist von Alan Turing [1]. Mit bemerkenswertem Weitblick konnte dieser Pionier schon damals die Entwicklung der KI und ihre Auswirkungen voraussehen und diskutieren. Ich stolperte über diesen Titel aus dem Jahre 1950, als ich gerade auf der Suche nach einem Dissertationsthema war. Das war 1977 – ich hatte mich in theoretische Informatik für meine Dissertation eingearbeitet. Turing ist bei Informatikern weltweit aufgrund seiner herausragenden Beiträge zu theoretischen Grundlagen der Informatik und auch der Mathematik bekannt. Für die breite Öffentlichkeit, und besonders für die britische, steht er für die Decodierung der Chiffriermaschine Enigma, die von der Deutschen Wehrmacht im Zweiten Weltkrieg eingesetzt wurde. Mehr zur Person findet sich in folgender Box.¹

Alan Turing, 1912 in London geboren und 1954 in Wilmslow, Cheshire, gestorben. Turing studiert Mathematik in einer Zeit, als die Grundlagen verschiedener neuartiger Disziplinen in den Naturwissenschaften gelegt werden. Einstein veröffentlicht seine Arbeiten zur Relativitätstheorie, der Logiker Kurt Gödel hat gerade seine richtungsweisende Arbeit herausgebracht, welche David Hilberts Traum der vollständigen Formulierbarkeit der Mathematik zunichtemacht, und verschiedene Mathematiker beginnen über den Begriff der Berechenbarkeit nachzudenken. Turing legt mit seiner einfachen und doch mächtigen (Gedanken-)Maschine, die später nach ihm Turing-Maschine benannt wird, ein bis heute gebräuchliches Instrument für die Analyse von Algorithmen und Computerprogrammen vor (siehe Kap. 3).

Während des Zweiten Weltkriegs arbeitet Turing an der Entschlüsselung deutscher Funksprüche mit. In Bletchley Park ist er Mitglied einer militärischen Dienststelle, welche sich das Ziel gesetzt hat, die deutsche ENIGMA-Verschlüsselungsmaschine zu knacken. In der Tat trägt Turing maßgeblich zum Erfolg dieses Unternehmens bei.

Ab 1948 wendet sich Turing in verstärktem Maße der Entwicklung von programmierbaren digitalen Computern zu. Der Aufsatz über künstliche Intelligenz [1] entsteht während seiner Zeit an der University of Manchester, wo Turing an der Entwicklung der Manchester Mark I beteiligt ist. Turing verfasst auch das Programmierhandbuch für diese Maschine, das zahlreiche Programme enthält, zumeist arithmetische Funktionen, die in einer einfachen Programmiersprache geschrieben sind.

Vor seiner Arbeit an der Decodierung der Enigma hat Turing einige Jahre in Princeton verbracht, wo er bei Alonzo Church studierte und auch 1938 promovierte. Aus dieser Zeit finden sich Hinweise, dass Turing sich im Zusammenhang

¹Die Darstellung in diesem Abschnitt ist in großen Teilen einem Aufsatz entnommen, den einer der Autoren in einer Informatikfachzeitschrift anlässlich des Turing-Jahres veröffentlicht hat [2].

mit Gödels Unvollständigkeitsresultat mit Begriffen wie Kreativität und Intuition beschäftigt hat. Beispielsweise hat er bereits in seiner Dissertation [3] Logiken vorgestellt, welche Raum für die mathematische Intuition lassen; er diskutiert dort in einem gesonderten Abschnitt die Rolle der Intuition und der Erfindungsgabe (ingenuity) beim Suchen formaler Beweise. Wir werden diese Problematik ausführlich im Kap. 7 diskutieren; mittlerweile weiß man nämlich, dass auch Emotionen eine maßgebliche Rolle bei intelligentem Schließen spielen.

Die Arbeiten jener Zeit können als direkte Fortsetzung der Ideen der großen Philosophen der Aufklärung verstanden werden. In Kap. 3 gehen wir bei der Behandlung der Körper-Geist-Problematik auf die Sichtweise von Descartes ein. Im Zusammenhang mit dem Begriff der KI ist auch die Arbeit von Gottfried Wilhelm Leibniz aus dem 18. Jahrhundert von Bedeutung. Leibniz hatte unter anderem auch mechanische Rechenmaschinen entworfen und die Vorstellung entwickelt, zukünftige Maschinen bauen zu können, die Probleme der Mathematik und Philosophie selbstständig lösen könnten. In Zukunft würden Philosophen sich nicht mehr mit Argumentationen „aufhalten“, sie würden einfach sagen „calulemus“ – „lasset uns rechnen“. Eine ausführliche Diskussion der Straße, die von Leibniz über Frege, Cantor, Hilbert und Gödel bis hin zu Turing führt, ist in [4] zu finden. Vor diesem Hintergrund ist es gar nicht so erstaunlich, dass sich Turing bereits damals schon mit der Frage, ob Maschinen denken können, auseinandergesetzt hat.

Turing wird 1952 wegen Homosexualität zu einer Gefängnisstrafe verurteilt. Vor die Wahl gestellt, entscheidet er sich für die medizinische Behandlung mit Hormonen. Wahrscheinlich in Folge dieser Therapie erkrankt er an Depressionen und verstirbt 1954 durch Suizid.

Alan Turing gilt heute als ein herausragender Informatikpionier, dessen Arbeiten noch immer allgegenwärtig in der modernen Informatik sind. 2013 wird er offiziell von der britischen Regierung rehabilitiert.

Can machines think? – Diese Frage stellt Alan Turing im ersten Satz seines Artikels [1]. Für die meisten Informatiker in den 1970ern waren denkende Maschinen noch kein Thema – erstaunlich, dass jemand sich schon Jahre vor der berühmten Dartmouth-Konferenz (1956), die als „Wiege der künstlichen Intelligenz“ gilt, mit intelligenten Maschinen beschäftigt hatte. Wenn man sich den Stand der Kunst im Bereich der digitalen Computer aus dieser Zeit ansieht, ist der Aufsatz von Alan Turing unglaublich visionär.

Turing argumentiert in diesem Artikel für die Fähigkeit von Maschinen, intelligent zu sein, und, viel wichtiger, er schlägt darin eine Art Test vor, mit dessen Hilfe man überprüfen kann, ob sich eine Maschine intelligent verhält. Anhand dieses Aufsatzes soll im folgenden Abschnitt erläutert werden, was wir uns unter dem Forschungsgebiet künstliche Intelligenz vorstellen können. Wir lehnen uns eng an Turings Aufsatz an, indem wir uns weitgehend an seine Struktur halten, seine Argumente wiedergeben und diese mit modernen KI-Entwicklungen in Beziehung setzen.

2.2 Das Imitationsspiel

Turing schlägt ein Imitationsspiel vor, welches zeigen soll, ob ein Computer intelligent ist. Mittlerweile wird dieses Spiel als Turing-Test bezeichnet:

Stellen wir uns einen Raum mit zwei Personen vor, A ist männlich, und B ist weiblich. Ein Fragesteller beliebigen Geschlechts befindet sich in einem anderen Raum, von dem aus er schriftlich mit A und B kommunizieren kann, ohne sie zu sehen.

Das Ziel des Spieles für den Fragesteller ist es nun zu ermitteln, welche der beiden Personen der Mann und welche die Frau ist. Die Rolle der Person B, also der Frau, ist es, dem Fragesteller zu helfen; Turing schlägt vor, dass dies am besten möglich sei, wenn sie die Wahrheit sagt. Person A dagegen muss mit seinen Antworten versuchen, den Fragesteller zu täuschen. Turing ersetzt nun die Person A in diesem Spiel durch eine Maschine und stellt die Frage, ob in dieser Konfiguration des Spieles der Fragesteller genauso oft getäuscht wird wie in der Version ohne Maschine. Diese Frage versteht er als verfeinerte Version der ursprünglichen Frage „Können Maschinen denken?“

In der ersten Version des Imitationsspieles, in der noch alle Parts von Menschen übernommen werden, spielt das Geschlecht der beteiligten Personen eine maßgebliche Rolle. In der weiteren Diskussion kommt Turing schließlich zu einer Version des Imitationsspieles, in der Part A durch den Computer übernommen wird und Part B durch einen Mann und beide sich für eine Frau ausgeben müssen (Abb. 2.1, links). Mittlerweile wird in der Literatur die

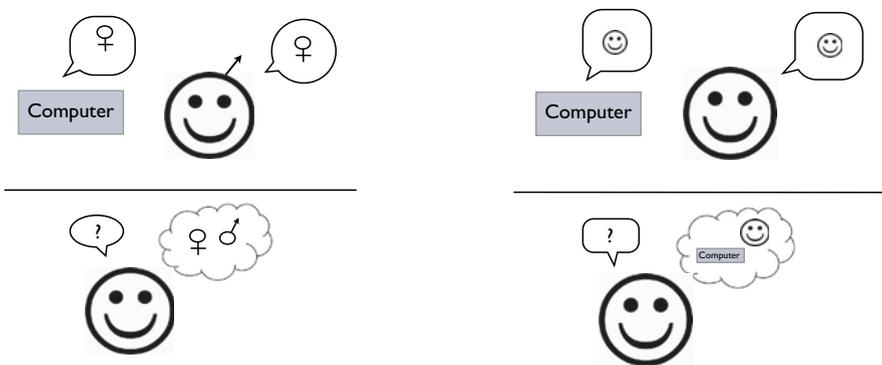


Abb. 2.1 Imitationsspiel. Links: Computer und Mann, beide über der Linie, geben sich als Frau aus. Der Fragesteller unter der Linie soll herausfinden, welcher von beiden ein Mann bzw. eine Frau ist. Rechts: Computer und Mensch geben sich als Mensch aus, der Fragesteller unter der Linie soll herausfinden, welcher von beiden ein Mensch ist

folgende Variante als „Turing-Test“ bezeichnet (Abb. 2.1, rechts). Das Spiel wird zwischen einer Maschine A, einem Menschen B und einem Fragesteller C gespielt. C soll herausfinden, welcher der beiden Konversationspartner A oder B ein Mensch ist.

Dieser Turing-Test geht also von der Hypothese aus, dass der Computer den Fragesteller so täuschen soll, dass dieser ihn fälschlicherweise für einen Menschen hält. Natürlich ist dies eine Sichtweise auf Intelligenz, die sehr stark an verbalem Verhalten orientiert ist. Dadurch bieten sich sofort einige Kritikpunkte an, auf die Turing auch unmittelbar eingeht – wir werden dies weiter unten aufgreifen.

2.3 Der Computer

In einem ausführlichen Abschnitt erklärt Turing, was er unter einem digitalen Computer, und insbesondere einem, der die Rolle der Maschine in dem Imitationsspiel übernehmen soll, versteht. Wenn man bedenkt, dass der übliche Sprachgebrauch zu dieser Zeit derart war, dass man beim Begriff Computer an einen Menschen dachte, der Berechnungen durchführt, ist es nicht verwunderlich, dass Turing gründlich erläutert, an welche Art von Computer er hier denkt. Er beschreibt die Architektur eines digitalen Computers, der aus Speicher, Rechenwerk und einer Kontrolleinheit besteht. Ausführlich geht er auf die Programmsteuerung eines solchen Computers ein und kommentiert dabei auch, dass die Kontrolle von Maschinen auch ein Zufallselement enthalten könnte, wodurch man solchen Maschinen sogar eine Art freien Willen zusprechen könnte (siehe Kap. 8).

In einem gesonderten Abschnitt erläutert Turing, dass der digitale Computer, wie er ihn beschrieben hat, eine „discrete state machine“ darstellt, die aber gleichzeitig universellen Charakter hat. Jede andere neue Maschine, die man sich ausdenkt, um bestimmte Berechnungen durchzuführen, ist (prinzipiell) unnötig; die gleiche Berechnung kann von einem geeignet programmierten digitalen Computer ausgeführt werden.

Damit formuliert Turing das Imitationsspiel genauer, indem er fragt: „Gibt es einen digitalen Computer, der im Imitationsspiel die Rolle der Maschine in befriedigender Weise übernehmen kann?“ Er beantwortet die Frage auch sofort positiv. Nach seiner Meinung wird es in 50 Jahren möglich sein, Computer so zu programmieren, dass sie das Spiel so beherrschen, dass ein durchschnittlicher Befrager mit nicht mehr als 70 % Wahrscheinlichkeit die Maschine nach fünf Minuten Fragezeit erkennen kann. Nun ist diese Vorhersage nicht eingetroffen – aber so ist das recht oft mit Prophezeiungen. Wichtiger erscheint

Turings Aussage, dass sich bis dahin die allgemeine Meinung so geändert haben wird, dass man dann keinen Widerspruch erwarten wird, wenn man von denkenden Maschinen spricht.

Nachdem Turing seine Meinung kurz und knapp formuliert hat, geht er ausführlich auf verschiedene Gegenargumente zu seiner These ein. Einige davon wollen wir im Folgenden diskutieren; vorher jedoch werfen wir noch einen Blick auf einen Wettbewerb, der sich eng an den Turing-Test anlehnt.

Im Jahre 1991 schrieb Hugh Gene Loebner ein Preisgeld von \$100.000 aus für das Programm, welches als erstes den Turing-Test besteht. Seitdem gibt es den jährlichen Loebner-Wettbewerb, bei dem Programme versuchen, genau wie von Turing gefordert, einen menschlichen Befrager zu täuschen. Mit \$2000 wird jedes Jahr das Programm prämiert, welches es schafft, die meisten Juroren zu täuschen. Hier wird allerdings zugelassen, dass das Thema der Unterhaltung stark beschränkt wird. Zweifelhaft ist, ob dieser Wettbewerb das Gebiet der KI voranbringt; sicher ist lediglich, dass sich Chatbot-Entwickler² immer neue Techniken einfallen lassen, um die Befrager des Wettbewerbs zu täuschen. Marvin Minsky, einer der Doyens der KI, sah im Loebner-Preis eine „abscheuliche und unproduktive jährliche Publicity-Kampagne“ und bot ein Preisgeld von \$100 für denjenigen, der Loebner davon abbringt, diesen Wettbewerb zu veranstalten. Loebner hat darauf reagiert, indem er argumentiert, dass der erste Gewinner des Wettbewerbs, also der \$100.000-Gewinner, den Wettbewerb beendet und damit die \$100 von Minsky zusätzlich erhalten muss. Folglich sei Minsky jetzt Co-Sponsor des Loebner-Preises.

2.4 Turings Einwände

Theologie und Kopf in den Sand Turing diskutiert und widerlegt in seinem Aufsatz nun einige der möglichen Einwände. Kurz geht er auf einen möglichen theologischen Einwand ein, wenn nämlich argumentiert wird, dass Gott dem Menschen eine unsterbliche Seele gegeben hat, aber nicht Tieren und Maschinen. Turing widerlegt das Argument, indem er schließt, dass Gott auch Elefanten eine Seele geben könnte, da er ja allmächtig sei; genauso gut könne er aber auch zulassen, dass Maschinen eine Seele haben. Und überhaupt, wenn man an Galileo denke, könne man den Verdacht bekommen, dass in der Vergangenheit diese Art Argumente zweifelhaft gewesen seien.

Ebenfalls kurz geht Turing auf das Kopf-in-den-Sand-Argument ein: Die Konsequenzen denkender Maschinen seien furchtbar. Deshalb müsse man hoffen und glauben, dass Maschinen nicht denken können. Obwohl Turing

²Chatbot: eine Kombination aus „to chat“ und „Roboter“.