

Wo stehen wir?

Wechselspiel zwischen menschlicher und künstlicher Intelligenz

geschrieben von Catrin Misselhorn

Die Chat-Software GPT gilt als Durchbruch der Künstlichen Intelligenz (KI). Um zu verstehen, warum das so ist, lohnt sich ein kurzer Blick auf die Geschichte. Der Begriff der KI fällt erstmals 1956 im Kontext einer Konferenz am Dartmouth College, deren Ziel es war, Maschinen zu konstruieren, die sich auf eine Art und Weise verhalten, die man bei Menschen als intelligent bezeichnen würde. Diese Definition ist grundsätzlich bis heute leitend für die KI geblieben. Wie daran deutlich wird, ist die KI von Anfang an von einem Wechselspiel zwischen menschlicher und Künstlicher Intelligenz geprägt. Auf der einen Seite war die Nachbildung oder zumindest Simulation menschlicher Intelligenz Ziel der Entwicklung von KI, auf der anderen Seite versprach sie auch ein besseres Verständnis der menschlichen Intelligenz. Das zeigt sich an den drei Paradigmen der KI-Forschung, die jeweils eine bestimmte philosophisch motivierte Vorstellung dessen, was Intelligenz ausmacht, mit bestimmten Ansätzen verbinden, wie man sie in künstlichen Systemen technisch realisieren kann.

Drei Paradigmen Künstlicher Intelligenz

Der erste Ansatz führt intelligentes Verhalten auf Symbolverarbeitung zurück. Menschliche und Künstliche Intelligenz lassen sich demnach gleichermaßen als Symbolverarbeitungsprozesse verstehen, die auf der Manipulation von physisch realisierten Symbolen auf der Grundlage vorgegebener Regeln basieren.

Dem zweiten Ansatz zufolge ist es ein Fehler, bei der Entwicklung von KI gänzlich von der Beschaffenheit des menschlichen Gehirns zu abstrahieren. Stattdessen gilt es, sich an der Verknüpfung der Neuronen im Gehirn zu orientieren und deren Funktionsweise mithilfe mathematischer Modelle nachzubilden, weshalb dieser Ansatz auch als Konnektionismus bezeichnet wird. Diese künstlichen neuronalen Netze sind zwar inspiriert vom Gehirn, aber sie sind weder strukturell noch funktional damit gleichzusetzen. Obwohl der Symbolverarbeitungsansatz später mit dem Label »Good Old-Fashioned AI« oder GOFAI versehen wurde, entstanden beide Ansätze ungefähr gleichzeitig um die Mitte des 20. Jahrhunderts.

Der dritte Ansatz hält die ersten beiden für unzulänglich, weil sie den Körper und seine Umwelt als wesentliche Bezugspunkte intelligenten Verhaltens ignorieren. Dieses Forschungsprogramm entstand in den 1980er Jahren und firmiert unter den Stichworten »Nouvelle AI«, »Embodied AI« oder verhaltensbasierte KI. Im Vordergrund stand die Idee, dass intelligentes Verhalten als physische Interaktion mit der realen Welt verstanden werden muss.

In der Folge wandte man sich von reinen Softwaremodellen dem Bau von Robotern zu, die über einen Körper, Sensoren und Aktoren verfügen. In den Vordergrund traten Fähigkeiten wie Wahrnehmung, Fortbewegung oder physische Manipulation der Umwelt, die zwar für ein kleines Kind scheinbar einfach zu erwerben sind, aber künstliche Systeme vor größere Herausforderungen stellen als anspruchsvolle, aber abstrakte kognitive Leistungen wie Schachspielen.

Deep Learning bringt den Durchbruch

Die Aussicht, intelligente Maschinen herzustellen, verdankt sich wesentlich dem Entstehen der Computertechnologie. Lange Zeit wurde der Begriff der KI für alle Ansätze computationaler Problemlösung gebraucht. Heute wird der Begriff zumeist enger zur Bezeichnung unterschiedlicher Formen maschinellen Lernens verwendet. Insbesondere das Deep Learning, eine Fortentwicklung des Konnektionismus, die sich neuronaler Netze mit mehreren Schichten zur Informationsverarbeitung bedient, brachte große Fortschritte. So konnte mithilfe tiefer neuronaler Netze vor gut zehn Jahren eine massive Steigerung der Erkennungsrate von Objekten auf Bildern erreicht werden.

Diese Verbesserung machte Bilderkennung für eine große Bandbreite von Anwendungsgebieten in Industrie und im klinischen Bereich geeignet. Jenseits der reinen Objekterkennung geht der Trend

dahin, Systeme zu schaffen, die Bilder aus Text oder anderen Bildern generieren können.

Große Sprachmodelle verändern die Welt

Auch die großen Sprachmodelle, zu denen die verschiedenen Versionen von ChatGPT & Co. gehören – seit Ende 2022 als Testversionen der Öffentlichkeit zugänglich –, bedienen sich tiefer neuronaler Netze. Diese Sprachmodelle werden in mehreren Schritten mithilfe einer Unmenge von Texten darauf trainiert, Eingaben in natürlicher Sprache zu erkennen und darauf in natürlicher Sprache zu antworten. Die Besonderheit daran ist, dass man mit diesen Systemen über eine riesige Bandbreite von Themen wie mit einem menschlichen Gesprächspartner kommunizieren kann. Die Bots kombinieren dabei nicht vorgegebene Textbauteile, sondern bringen auf der Grundlage der Trainingsdaten eigene Formulierungen hervor.

Die neueste Version GPT 4 kann auch Bilder verarbeiten. Die Aufgaben, die diese Systeme erfüllen können, sind schier unerschöpflich. Sie können nicht nur beliebige Gebrauchstexte schreiben, sondern Prosa, Gedichte oder Bilder verschiedenster Stilrichtungen erstellen, Programmieren, Drehbücher verfassen, Fragen des alltäglichen Lebens ebenso wie Prüfungsfragen aller Lernniveaus bearbeiten, Musik komponieren, Verträge rechtlich prüfen etc.

Der Turingtest als Maßstab Künstlicher Intelligenz?

Diese Chatbots haben auch deshalb so viel Furore gemacht, weil sie als die ersten Systeme gelten, die den Turingtest ohne Zuhilfenahme von Tricks bestanden haben. Alan Turing schlug als Maßstab für echte Künstliche Intelligenz ein Imitationsspiel vor, das auf der Idee beruht, dass einem Computer Denken zuzusprechen ist, wenn es ihm gelingt, einen menschlichen Interaktionspartner in einem Frage-Antwort-Dialog davon zu überzeugen, er sei ein Mensch.

Nun versucht zumindest ChatGPT nicht über seine maschinelle Natur hinwegzutäuschen, sondern betont auf direkte Anfrage, er sei kein Mensch. Gleichwohl unterscheidet sich die überwältigende Mehrheit der Ausgaben nicht grundlegend von menschlichen Äußerungen. Dennoch ist unstrittig, dass diese Chatbots nicht wirklich denken können.

Denken erfordert ein Operieren mit symbolischen Repräsentationen, die in logische Beziehungen miteinander gebracht werden mit dem Ziel, die Wahrheit einer Behauptung zu erweisen oder eine Handlung zu begründen. Diese Vorstellung des Denkens motivierte den Symbolverarbeitungsansatz. Derartige Ziele kennen Chatbots nicht, die Wörter rein nach der Wahrscheinlichkeit ihres gemeinsamen Auftretens aneinanderreihen. Deshalb werden sie auch als »statistische Papageien« kritisiert. Die Tatsache, dass die Chatbots den Turingtest bestehen, spricht also nicht dafür, dass sie denken können, sondern dass der Test unzureichend ist.

Auch der Übergang vom Denken zum Handeln in einer physischen Umwelt, den die Nouvelle AI als zentralen Aspekt intelligenten Verhaltens auffasste, wird durch die neuen Chatbots nicht geleistet. Doch es ist durchaus möglich, mit ihrer Hilfe robotische Systeme zu steuern, erste Versuche hat Microsoft bereits mit Drohnen unternommen. Bleibt die Frage, wie Menschen in einem substanziellen Sinn Verantwortung für derartige Systeme übernehmen können.

Der Beginn einer neuen industriellen Revolution?

Sascha Lobo verglich jüngst in seiner Spiegelkolumne die Größenordnung der durch diese Systeme bevorstehenden gesellschaftlichen, technologischen und kulturellen Umwälzungen mit der Industrialisierung. Die Technologie könnte einen enormen Automatisierungsdruck ausüben. Dabei ist der Kapitalaufwand im Unterschied zu früheren Automatisierungswellen zwar für die Nutzerinnen und Nutzer verhältnismäßig gering. Aber es kommt auf der Seite der Firmen, die diese Software zur Verfügung stellen, zu einer starken wirtschaftlichen Konzentration. Denn nur wenige Anbieter werden über die erforderlichen Server- und Rechenkapazitäten verfügen und Zugriff auf die benötigte immense Menge an Trainingsdaten haben. Diese bestehen in von Menschen erarbeiteten Inhalten, die – ohne diese Arbeit zu honorieren oder die Urheberinnen und Urheber auch nur zu erwähnen – von kommerziellen Modellen wie GPT 4 gewinnbringend genutzt werden. Noam Chomsky bezeichnet diese Form der Arbeitsteilung deshalb als »Hightech-Plagiat«.

Fraglich ist auch, zu welchen Auswirkungen auf den Arbeitsmarkt es kommt. Denn diese Form der Automatisierung scheint nicht in erster Linie einfache, ungelernte Tätigkeiten zu betreffen, sondern

solche Branchen, in denen früher neue Arbeitsplätze entstanden, darunter Forschung und Entwicklung in Naturwissenschaft und Technik, Bildung, Pflege, Rechtswesen und den Kreativbereich.

Auch die Konsequenzen für die Demokratie könnten bedrohlich sein, wenn die simulierte Kommunikation überhandnimmt. Es steht zu befürchten, dass die bereits mit den sozialen Medien einsetzende Erosion der demokratischen Öffentlichkeit dramatisch fortschreitet, wenn die Welt überschwemmt wird mit von Chatbots generierten Inhalten, während die zugrunde liegenden Algorithmen sowie die Datenbasis intransparent bleiben.

Keine Frage, die neue Technologie bietet vielfältige Möglichkeiten für die Erweiterung des menschlichen kreativen Potenzials. Doch die Kreativität, die nötig ist, um daraus einen positiven Entwurf für Wirtschaft, Gesellschaft und Kultur zu machen, bleibt bis auf Weiteres uns Menschen vorbehalten.

Dieser Text ist zuerst erschienen in Politik & Kultur [04/2023](#).