

„Alle 30 Jahre werden KIs eine Million mal größer“

Jürgen Schmidhuber zählt zu den besten KI-Forschern der Welt. Er erklärt, wieso Computer ähnlich lernen wie wir, wo deutsche Mittelständler Chancen haben – und worin sich Leibniz irrte.

Professor Schmidhuber, mittlerweile ver geht kaum ein Tag, an dem Künstliche Intelligenz nicht irgendwo Schlagzeilen macht. Wo stehen wir denn heute, wie schlau sind die Computer?

Sie werden ständig schlauer. Und zwar nicht unbedingt deswegen, weil die Algorithmen immer raffinierter würden, deren Ursprünge teilweise in der Mitte des vergangenen Jahrhunderts liegen.

Sondern?

Weil nach wie vor gilt, dass alle fünf Jahre die Rechner ungefähr zehnmal billiger werden. Dieser Trend begann spätestens 1941, als Konrad Zuse den ersten Allzweckrechner baute, der ungefähr eine Operation pro Sekunde ausführen konnte. Inzwischen können wir Millionen von Milliarden Instruktionen ausführen zum selben Preis, den Zuse damals für eine Operation zahlen musste. Davon profitiert die Künstliche Intelligenz sehr, denn die neuronalen Netzwerke, die heutige Grundlage zahlreicher Anwendungen sind, werden bei gleichem Preis immer größer und leistungsfähiger.

Ein grundlegender Lernalgorithmus, der Sie und Ihren damaligen Studenten Sepp Hochreiter berühmt machte, verbirgt sich hinter dem Kürzel LSTM. Was kann der, was die Algorithmen davor nicht konnten?

LSTM kann viel besser als frühere Systeme mit Sequenzen arbeiten. Die ganze Welt ist ja sequenziell. Woran merken Sie das? Die Daten, die über Ihre Augen und Ohren hereinkommen, also über Ihre Kameras und Mikrofone, sind nicht statisch, nein, das sind Reihen von Daten, Videos oder Sprachsignalen. Lange konnte kein künstliches neuronales Netzwerk damit gut umgehen. Aber selbst LSTM musste warten, bis die Rechner billig genug waren, um in den 2000ern dann Durchbrüche zu erzielen – in der Sprachverarbeitung, Übersetzung, Videoerkennung, Textgenerierung oder dem Textverständnis. Viele Anwendungen der Plattform-Konzerne beruhen auf Methoden, die wir zu Beginn der Neunziger-Jahre an der TU München entwickelten.

Immer wieder fasziniert uns das Deep Learning auch deshalb, weil wir Menschen lernende Wesen sind. Aber ist das wirklich vergleichbar, nähern sich die KIs, an denen auch Sie arbeiten, tatsächlich dem an, wie wir Menschen sind?

Ja, meiner Ansicht nach ist das prinzipiell ganz nahe an dem, was Babys oder Kinder oder auch Erwachsene machen. Während wir interagieren mit unserer Umgebung, lernen wir, wie die Umgebung funktioniert. Wir lernen, vorherzusagen, was passiert, wenn wir uns auf die eine oder andere Weise verhalten. Und wir nutzen die Vorhersagen, um die Zukunft zu planen. Ein Baby lernt beispielsweise, wie sich die Video-Eingaben seiner „Kameras“, der Augen, ändern, wenn es auf seine Hand blickt und damit wackelt. Das muss es auch lernen, denn zunächst weiß es fast nichts, weil in seinen Genen nicht viel drinstecken kann.

Wie meinen Sie das?

Alle Informationen, die ein Baby von seinen Eltern mitbekommen hat, befinden sich in einer relativ kurzen Kette von Genaubsteinen, ein paar Milliarden Bits. Das ist fast nichts im Vergleich zu dem, was ein erwachsenes Gehirn enthält. Und diese wenigen Informationen der Gene sind noch dazu vor allem zuständig für wirklich wichtige Dinge, die mit Intelligenz gar nichts zu tun haben: für Magen und Leber und Immunsystem und den unglaublichen Selbstreplikations-Mechanismus. Na ja, und dann gibt es eben noch ein paar Millionen Bits, die uns von unseren menschenaffähnlichen Vorfahren unterscheiden. Die uns nach heutigem Verständnis intelligent machen. Diese wenigen Bits sind zwar fast nichts im Vergleich zu den Terrabytes an Daten, die sich ein Baby jeden Tag reinzieht. Doch sie kodieren den Lernalgorithmus, der dafür sorgt, dass das Baby aus den Daten lernt, wie die Welt funktioniert. Es baut daraus ein „Weltmodell“, mit dessen Hilfe es planen kann. Seit dem Jahr 1990 machen unsere KIs im Labor das ähnlich, zumindest ist mir persönlich nicht klar, worin ein fundamentaler Unterschied zu dem bestehen soll, was Babys machen.

Aber ist nicht das „Weltmodell“ selbst ein großer Unterschied? Ich kann meines gebrauchen, um ganz verschiedene Aufgaben zu bewältigen oder Entscheidungen zu treffen und umzusetzen. Viele erfolgreiche KIs sind heute hingegen hochspezialisiert: Sie verfügen über eine bestimmte Fertigkeit und sind darin wortwörtlich übermenschlich – aber eben nur in dieser einen Sache.

Ihr Weltmodell ist in der Tat wesentlich breiter als das unserer KIs, weil Sie in den Jahrzehnten seit Ihrer Geburt viel mehr Erfahrungen gesammelt haben als unsere KIs in wenigen Tagen der Interaktion mit unbekannten Umgebungen. Auch hat Ihr Hirn immer noch viel mehr adjustierbare Parameter als derzeitige KIs. Allerdings haben wir schon längst KIs, die nicht nur eine Aufgabe können, sondern zehn oder zwanzig verschiedene, und KIs, die neue Aufgaben schneller lernen, weil sie auf früher Gelerntem aufbauen können. Und bedenken Sie den oben erwähnten Trend: Alle 30 Jahre werden KIs eine Million mal größer.

Sie sind nicht nur Forscher, sondern auch Unternehmer. KI braucht Daten. Ihre Firma NNAISENSE ist indes winzig verglichen mit den führenden Tech-Konzernen, die über enorme Ressourcen verfügen. Haben Sie überhaupt eine Chance, langfristig eigenständig zu bleiben?

Klar! Tatsächlich sind unsere Kunden Industrieunternehmen. Sie können die Daten, welche die großen Konzerne im World Wide Web oder in Videospiele sammeln, überhaupt nicht brauchen. Nein, sie verfügen über eigene Daten aus den ganz eigenen industriellen Prozessen. Nehmen Sie einen unserer Investoren: das Unternehmen Schott, welches der Zeiss-Stiftung gehört und hochwertiges Glas herstellt. Die relevanten Daten liefern die Sensoren in der Schott-Fabrik, nicht Facebook. Schott hat schon mehr als 100 Jahre Erfahrung mit diesem schwierigen Prozess, schätzt aber dennoch die Vorteile, die KI in der Produktion bringt.

Was machen Sie denn da so mit Ihren Klienten und Projektpartnern?

Typischerweise lernt unser „tiefer digitaler Zwilling“ oder „deep digital twin“ aus den Daten unseres Partners, wie seine Fabrik samt den zugehörigen physikalischen Prozessen funktioniert. Dieses erlernte Weltmodell nutzen wir, um durch relativ billige „mentale Experimente“ anstatt teurer Experimente in der physikalischen Welt herauszufinden, wie die Fabrik anhand bestimmter Leistungs-Indikatoren verbessert werden kann, wie sie profitabler oder umweltfreundlicher werden kann. Einen „deep digital twin“ trainieren wir zum Beispiel mit unserem Partner EOS in Krailing, einem führenden Anbieter für 3-D-Druck in der Industrie. EOS baut tolle Maschinen, in denen zuckende Laserstrahlen Pulver schmelzen, das Schicht für Schicht aufgetragen wird, sodass über die Zeit ein kompliziertes Teil einer Flugzeugturbine entsteht oder ein anderes Produkt, das Sie mit herkömmlichen Verfahren gar nicht herstellen können. Tomographen in diesen Druckern liefern uns sehr viele Daten zum Trainieren des tiefen digitalen Zwilling.

In diesem Bereich ist der Zug noch nicht abgefahren für Deutschland?

Natürlich nicht. Viele deutsche Mittelständler sitzen auf wertvollen Datenschätzten, die sich von KIs auswerten lassen, um die entsprechenden industriellen Prozesse zu verbessern, um sie profitabler zu machen oder grüner oder vielseitiger. Das ist nach wie vor eine Riesenchance für Deutschland, denn hier gibt es ja sowieso schon zahlreiche Unternehmen, die Weltspitze sind in dem, was sie tun. Und die erkennen, dass KI ihre Produkte weiter verbessern kann – wobei diese KIs eben die Daten des jeweiligen Unternehmens brauchen, um angelernt zu werden.

Sie selbst wiederum stellen unseren Lernern auf dieser Digitec-Seite den Universalgelehrten Gottfried Wilhelm Leibniz vor. Warum ist es nützlich, sich jetzt mit solchen vor Jahrhunderten gewonnenen Einsichten zu beschäftigen?

Mein Spezialgebiet ist das maschinelle Lernen. Da geht es um adaptive Systeme, die Daten aus vergangenen Beobachtungen verwenden, um in der Zukunft mehr

Erfolg zu haben. Beim Menschen selbst geht es natürlich auch darum. Selbst bei ganzen Gesellschaften, die ihrerseits aus vielen kleinen Menschlein bestehen, geht es darum. Allgemein gilt: Aus der Vergangenheit lernt man in der Hoffnung, dass es einem später nützt. Dies erklärt das generelle Interesse an der Geschichte. Und da Geschichte vor allem durch technische Durchbrüche angetrieben wird, ist die Wissenschaftsgeschichte besonders ausschussreich für das Verständnis der großen Zusammenhänge. Deswegen interessiere ich mich sehr dafür.

Im Geschichtsunterricht kommt sie eher unterrepräsentiert vor.

Zu Unrecht. Sie ist viel bedeutender als die in der Schule dominierende Politikgeschichte, wo es vielleicht darum geht, welcher König in welchem Jahr was für eine Heerschar angeführt hat. Tatsächlich haben Könige, Kaiser, Politiker häufig nur reagiert auf technologische Neuerungen, die ihre Zeit prägten. In meinem Gebiet, der Informatik, ist es nun spannend herauszufinden, wer grundlegende Dinge erfunden hat, die heute in jedem Computer enthalten sind – und da kommt man ganz schnell auf den Leibniz.

Wenn ich mein Handy oder meinen Laptop anschalte, wo ist da der Leibniz drin?

Ihr Smartphone und Ihr Laptop haben einen internen Speicher. Die erste Maschine mit Speicher wurde 1673 von Leibniz gebaut. Ihr Smartphone und Ihr Laptop sind auch Binärcomputer. Da sind lauter kleine Schalter drin, die zwischen null und eins unterscheiden können. Leibniz war derjenige, der den ersten Binärcomputer beschrieb. Er dachte sich ein auf Lochkarten basierendes System aus, in dem Murmeln durch Anwesenheit oder Abwesenheit Nullen und Einsen kodieren. Und wie diese Murmeln interagieren können in einer Weise, die einer Rechnung entspricht.

Das klingt sehr veraltet oder?

Im Gegenteil. Nach diesem Prinzip funktionieren heutige Rechner mit elektronischen Transistoren. Wer mit dem Binärsystem rechnet, dem fällt es viel leichter, einen funktionstüchtigen Computer zu bauen, als jemandem, der das Dezimalsystem verwendet.

Hatte Leibniz geahnt, welches Potential er freisetzt?

Velleicht. Er war sehr ambitioniert. Er schuf eine „formale Algebra des Denkens“ und dachte, man könne jede mögliche sinnvolle Frage dadurch beantworten, dass man die Antwort einfach ausrechnet. Selbst Fragen der Philosophie. Hierzu brauchte es seiner Ansicht nach lediglich eine universelle Sprache und einen allgemeinen Kalkül, um Schlussfolgerungen zu ziehen.

Was sich als falsch herausgestellt hat.

Nun ja, im Jahr 1931 zeigte der österreichische Mathematiker Kurt Gödel die Schranken des Berechenbaren auf. Nicht jede vernünftige Frage lässt sich durch Rechnen bejahen oder verneinen. Aber eben doch sehr viele.

Das Gespräch führte Alexander Armbruster.



Der erste Informatiker

Wie Gottfried Wilhelm Leibniz den Computer erdachte / Von Jürgen Schmidhuber

Mancher, der den Namen Gottfried Wilhelm Leibniz hört, mag an einen herausragenden Juristen, Historiker, Philosophen und Diplomaten denken. Viele kennen ihn zudem als an einen der bedeutendsten Mathematiker aller Zeiten. Tatsächlich hat dieses Universalgenie, dessen Geburtstag sich in diesem Jahr zum 375. Mal jährt, noch mehr vollbracht: Er legte praktische und theoretische Grundlagen aller modernen Rechner.

Doch der Reihe nach: Die Konstruktion von Automaten begann bereits in der Antike. Das Antikythera-Getriebe (eine Art astronomischer Rechenmaschine) entstand vor mehr als 2000 Jahren. Die wohl erste programmierbare Maschine der Welt (das automatische Puppentheater des Heron von Alexandria) stammt aus dem 1. Jahrhundert. Ihre Energiequelle war ein fallendes Gewicht, das eine um die Stifte eines drehbaren Zylinders gewickelte Schnur zog. Komplexe Befehlssequenzen, die mehrere Minuten lang Türen und Puppen steuerten, wurden durch komplexe Umwicklungen kodiert.

Doch viele Aspekte der modernen Informatik lassen sich tatsächlich auf Leibniz zurückführen. Im Jahr 1673 erdachte er mit dem sogenannten Schrittzähler die erste Maschine, die alle vier Grundrechenarten ausführen konnte. Dieser wies hinaus über Wilhelm Schickards zahnradbasierte Rechenmaschine aus dem Jahr 1623 und Blaise Pascals überlegene Pascaline aus dem Jahr 1642. Das ist bedeutender, als es auf den ersten Blick vielleicht erscheinen mag, denn: Wer die Grundrechenarten beherrscht, vermag jede beliebige numerische Berechnung auszuführen – der österreichische Mathematiker Kurt Gödel kodierte ein Vierteljahrtau-

send später mithilfe der Grundrechenarten sogar beliebige formale Systeme und Rechenprozesse.

Der Schrittzähler war darüber hinaus der erste Rechner mit internem Speicher. Das sogenannte Leibniz-Rad speicherte den Multiplikanden einer Multiplikation. Es zählte während der Rechnung mit, wie viele Additionen zur Ausführung einer gegebenen Multiplikation bereits ausgeführt wurden. Solche internen Variablen sind heute natürlich unabdingbar in der Informatik.

Inspiriert vom antiken binären chinesischen I Ching dokumentierte Leibniz auch die binäre Arithmetik, die in praktisch allen modernen Rechenmaschinen Anwendung findet. Ein Rechner, der nur mit 0 und 1 als Elementarsymbolen arbeitet, ist leichter zu bauen als ein Dezimalrechner, der alle Ziffern von 0 bis 9 ver-



Unser Podcast zu
Digital- und
Technikthemen
www.faz.net/digitec-podcast

wendet. Es sollte jedoch erwähnt werden, dass Zahlendarstellungen im Dualsystem an sich sehr viel älter sind und bis ins alte Ägypten zurückreichen. Der algorithmische Aspekt des dualen Rechnens ist in des relativ jung.

Im Jahr 1679 beschrieb Leibniz tatsächlich die Grundprinzipien binärer Computer. Er stellte binäre Zahlen mithilfe von Murmeln dar, die durch Lochkarten gesteuert wurden. Dabei legte er



Deutscher Universalgelehrter: Gottfried Wilhelm Leibniz (1646–1716)

Foto ddp / Bearbeitung F.A.Z.

wesentlich das Funktionsprinzip elektronischen I Ching dokumentierte Leibniz auch die binäre Arithmetik, die in praktisch allen modernen Rechenmaschinen Anwendung findet. Ein Rechner, der nur mit 0 und 1 als Elementarsymbolen arbeitet, ist leichter zu bauen als ein Dezimalrechner, der alle Ziffern von 0 bis 9 ver-

höchst einflussreiche Ideen zu einer universellen Sprache und einem allgemeinen Kalkül für Schlussfolgerungen (Characteristica Universalis & Calculus Ratiocinator). Der KI-Pionier Norbert Wiener meinte einst: „In der Tat ist die allgemeine Idee einer Rechenmaschine nichts anderes als eine Mechanisierung des Leibniz'schen Kalküls Ratiocinator.“

Der Leibniz-Ausspruch „Calculemus!“ ist wiederum prägendes Zitat der Aufklärung: „Käme es zwischen Philosophen zur Kontroverse, so bräuchten sie nicht mehr zu streiten als Buchhalter. Sie müssten sich nur mit ihren Bleistiften und Schiefertafeln hinsetzen und zueinander sagen: Lasst uns rechnen!“ Da war also plötzlich einer, der die ganze Welt auf das Berechenbare reduzieren wollte. Heute gibt es viele von dieser Sorte, doch damals war das revolutionär.

Als ob seine Errungenschaften in der Informatik nicht ausreichten, um das Vermächtnis des Gottfried Wilhelm Leibniz als das eines der größten Wissenschaftler überhaupt zu zementieren, war er noch dazu der Erste, der die Integralrechnung veröffentlichte – im Jahr 1684. Damit erweiterte er sowohl die Pionierarbeiten

des Archimedes, der vor mehr als zwei Jahrtausenden die Infinitesimalrechnung einföhrte und schon Spezialfälle der Integralrechnung kannte, zum Beispiel für Kugeln und Parabelsegmente, als auch jüngere Durchbrüche der Integralrechnung, etwa durch Madhava von Sangamagrama und Kollegen im Indien des 14. Jahrhunderts. Dass Leibniz nicht nur zur Mathematik, Wahrscheinlichkeitstheorie und Technik sehr viel beitrug, sondern auch zur Linguistik, Biologie, Medizin, Geologie, Psychologie, Politik, Ethik, Theologie, Geschichte, Philologie und Philosophie, sei nur am Rande erwähnt.

Wie hat sich die Theorie des Rechnens nach Leibniz' Tod im Jahr 1716 nun weiterentwickelt? Über zwei Jahrhunderte später erweiterte Kurt Gödel Freges bereits erwähnte, von Leibniz inspirierte formale Sprache und führte schließlich die erste wirklich universelle Sprache zur Kodierung beliebiger formalisierbarer Prozesse ein (1931–34). Mit ihr zeigte er, dass es fundamentale Beschränkungen dessen gibt, was entscheidbar oder berechenbar ist. Damit versetzte er dem Leibniz'schen Projekt der universellen Pro-

blemlöser einen Schlag. Gödels bahnbrechende Arbeit aus dem Jahr 1931 legte die Grundlagen der modernen theoretischen Informatik und der Theorie der Künstlichen Intelligenz, insofern sie die Grenzen des Theorembeweisens, des Rechnens, der Logik und der Mathematik selbst aufzeigte. Gödel konstruierte formale Aussagen, die über die Berechnung anderer formaler Aussagen sprechen – insbesondere selbstreferentielle unentscheidbare Aussagen, die implizieren, dass ihr Wahrheitsgehalt nicht ermittelt werden kann durch einen Theorembeweiser, der systematisch alle möglichen Theoreme aus einer aufzählbaren Menge von Axiomen herleitet. Dies hatte enormen Einfluss auf die Wissenschaft und Philosophie des 20. Jahrhunderts. Einige missverstanden gar sein Ergebnis und dachten, er hätte gezeigt, dass der Mensch der KI überlegen sei.

Und wie hat sich die praktische Hardware nach Leibniz weiterentwickelt? Die ersten kommerziellen programmgesteuerten Maschinen (lochkartenbasierte Webstühle) entstanden in Frankreich um das Jahr 1800 durch Joseph-Marie Jacquard und andere – vielleicht die ersten „modernen“ Programmierer, die sozusagen industrielle Software schrieben. Sie inspirierten Ada Lovelace und ihren Mentor Charles Babbage. Er plante vergeblich, einen programmierbaren Universalrechner zu bauen. Im Jahr 1941 jedoch konstruierte Konrad Zuse die Z3, den ersten praktischen, funktionierenden, programmierbaren Allzweckrechner, basierend auf seiner Patentanmeldung aus dem Jahr 1936. Im Gegensatz zu Babbage verwendete Zuse das Leibniz'sche Binärrechnerprinzip anstelle der traditionellen dezimalen Arithmetik. Wie bereits erwähnt, vereinfachte dies die Hardware sehr.

Im Jahr 2021 feiern wir also nicht nur Leibniz' 375. Geburtstag, sondern auch den 90. Jahrestag von Gödels berühmter Arbeit von 1931 und das achtzigjährige Jubiläum des ersten funktionsfähigen programmgesteuerten Computers von Zuse (1941) auf der ganzen Welt.