Eine neue Intelligenz, die sich selber nicht versteht

Autor(en): **Ulmi, Nic**

Objekttyp: **Article**

Zeitschrift: **Horizonte : Schweizer Forschungsmagazin**

Band (Jahr): **29 (2017)**

Heft 113

PDF erstellt am: **23.05.2024**

Persistenter Link: https://doi.org/10.5169/seals-821477

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.
Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.
Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

Ein Dienst der *ETH-Bibliothek*
ETH Zürich, Rämistrasse 101, 8092 Zürich, Schweiz, www.library.ethz.ch

http://www.e-periodica.ch

Eine neue Intelligenz, die sich selber nicht versteht

Was ist eine Entdeckung wert, die man nicht versteht? Die Entwicklung intelligenter Maschinen in Laboratorien stellt das eigentliche Wesen des Wissens infrage. *Von Nic Ulmi*

Hypothesen formulieren, Entdeckungen machen: Diese Begriffe bilden das Herzstück jedes wissenschaftlichen Ansatzes. Seit Anfang der 2000er Jahre sind Maschinen mit künstlicher Intelligenz daran, dieses Gebiet zu erobern. Sie produzieren auf neuen Wegen Ergebnisse und übernehmen Aufgaben, die einst Menschen vorbehalten schienen. Diese Maschinen, die immer leistungsfähiger werden und doch nicht in der Lage sind, ihre eigenen Erfolge zu verstehen, beschwören eine befremdliche Vision herauf: eine automatisierte, industrielle, von Robotern betriebene Forschung.

Drei Beispiele: An der Tufts-Universität in Massachusetts arbeitet ein Netz aus künstlichen Neuronen daran, ein altes Rätsel der Biologie zu lösen. Es entwickelt eine Hypothese zur Regeneration von Planarien-Süsswasserwürmern, die einen verlorenen Schwanz oder Kopf nachwachsen lassen können. An der Universität Adelaide in Australien wiederum hat ein Roboter das optimale Mittel gefunden, um ein Bose-Einstein-Kondensat herzustellen. Das Gebilde aus Bosonen, die beinahe bis auf den absoluten Nullpunkt gekühlt wurden, veranschaulicht makroskopische Quantenphänomene. Schliesslich haben die Maschinen des Start-ups Insilico Medicine an der Universität Johns Hopkins in Baltimore Moleküle entwickelt, die in der Onkologie nützlich sein könnten.

> «Das menschliche Gehirn erbringt im Gegensatz zu Maschinen Abstraktionsleistungen mühelos.»
>
> Roger Schank

Roger Schank, ein Pionier der künstlichen Intelligenz (KI), zeigt sich von diesen Beispielen wenig beeindruckt. «Bei allen handelt es sich absolut nicht um künstliche Intelligenz», protestiert er. Der Forscher, der an den Universitäten Yale, Northwestern und Carnegie Mellon tätig ist, befürchtet, dass sich ein altbekanntes Szenario wiederholen wird: Anfang der 1970er und Ende der 1980er Jahre gab es zweimal einen Hype um die künstliche Intelligenz, beide Male folgte aber eine «Dürreperiode». Die übersteigerten Erwartungen lösten sich auf, das öffentliche Interesse erlosch, die Geldquellen versiegten. «Im Moment ist die Presse begeistert, bald wird sie sich jedoch wieder anderem zuwenden. Das Problem ist, dass dann die Forschung über KI erneut zum Stiefkind wird. Das finde ich schade.»

Das Steak und der Haarschnitt

Worum handelt es sich dann, wenn nicht um Intelligenz? «Das sind Beispiele für ein Pattern Matching (Musterabgleich, A.d.R.)», antwortet Schank. «Facebook verwendet dieses Verfahren, um Ihr Gesicht auf Fotos zu identifizieren.» Mit anderen Worten: Die Maschinen generieren Motive – das Schema eines Moleküls, den Regenerationsplan eines Strudelwurms –, die sie dann mit Mustern in bestehenden Datenbanken vergleichen. «Wissenschaftler hingegen machen auf eine ganz andere Art Entdeckungen: Am Anfang steht die Ratlosigkeit. Dann formulieren sie eine Hypothese zum Phänomen, das sie nicht verstehen, und testen diese. Genau diesen Prozess nennt man Wissenschaft.»

Roger Schank erzählt gern eine Geschichte, um den Unterschied zwischen menschlichem Entdecken und automatischem Lernen zu veranschaulichen. Er nennt sie «Das Steak und der Haarschnitt»: «Ich diskutierte mit einem Kollegen in Yale und beschwerte mich, dass es mir nie gelingt, ein Steak zu bekommen, das wirklich ‹saignant› ist: Das Fleisch kommt immer zu lange gebraten auf den Tisch. ‹Weshalb bloss?›, fragte ich mich. ‹Ich habe früher in England gelebt, und dort schaffte es kein Coiffeur, mir einen Flattop zu schneiden›, meinte mein Kollege nur. Heureka! Jede der beiden Geschichten erklärt die andere. Auf einer höheren Abstraktionsebene sind sie identisch. Wir haben es in beiden Fällen mit Personen zu tun, die problemlos fähig wä-

«Bei Wissenschaftlern steht vor einer Entdeckung die Ratlosigkeit.»

Roger Schank

ren, die verlangte Leistung zu erbringen, die jedoch die Anweisung nicht ausführen, weil sie diese als zu extrem empfinden.» Und die Moral von der Geschichte? «Ein menschliches Gehirn erbringt diese Abstraktionsleistung mühelos. Es ist in der Lage, eine Sache so zu betrachten, als wäre es eine andere. Die Motivation schöpft es daraus, dass wir von Natur aus Ziele haben: unseren Appetit sättigen, unsere Neugierde stillen, die durch etwas Unbekanntes geweckt wurde.»

So funktioniert unser Gehirn. Doch die neue Welle der Forschungsarbeiten und Applikationen im Bereich der KI seit Ende der 1990er Jahre zielt nicht mehr darauf ab, das kognitive Modell des menschlichen Geistes nachzubilden. Diese Wende öffnet die Tür zu einer ganz neuen Wissenswelt, bei der das Know-how durch automatisches Lernen aus Megadaten entsteht. Diese Neuausrichtung hat die KI aus ihrer Lethargie befreit und bereits Früchte getragen: Vorschläge zum Weiterstöbern bei Amazon, Siri, den Sieg von AlphaGo über einen der besten Go-Spieler der Welt oder auch die Voraussage der Genexpression eines Bakteriums gemäss einer Studie der Universität Pennsylvania von 2016.

Zur Freude der Maschine

Doch kann diese Technologie wirklich wissenschaftliche Erkenntnisse hervorbringen? Oder braucht es dazu das existenzielle Bedürfnis, zu verstehen, den Drang, zu wissen, eine «libido sciendi», wie der Heilige Augustinus es formuliert hätte?

Wie sieht das Jürgen Schmidhuber, Co-Direktor des Instituts Dalle Molle für Forschung im Bereich künstliche Intelligenz (Idsia) in Lugano? Über ihn schrieb die New York Times im November 2016: «Wenn die KI erwachsen geworden ist, wird sie ihn vielleicht ‹Papa› nennen.» Er entwickelte 1997 den Ansatz des «long short-term memory», den heute Programme zur mündlichen Spracherkennung nutzen. Er ist der Ansicht, dass im Gehirn von Wissenschaftlern, Künstlern und Babys, aber auch in künstlichen neuronalen Netzen vergleichbare Belohnungssysteme wirken können. Diese verschiedenen Formen von Intelligenz können Genugtuung empfinden, wenn sie im Chaos der Welt Wiederholungen und Regelmässigkeiten entdecken, die ihnen zuvor unbekannt waren.

«Stellen Sie sich ein Programm vor, das eine Bilderreihe modellieren soll, die hundert fallende Äpfel zeigt», erklärt der Forscher. «Ohne die Wirkung der Schwerkraft zu kennen, braucht es sehr viele Bits, um die Daten zu codieren. Sobald jedoch diese Regelmässigkeit erkannt wurde, kann das Programm sie verwenden, um Voraussagen zu machen. Und so benötigt es weniger Bits, um die Information zu speichern. Dieser Unterschied zwischen vorher und nachher, das heisst diese Kompression der Daten, misst die Tiefe der neuen Kenntnisse, die das Netz erworben hat. Dies löst ein Signal der Belohnung aus, ein Moment innerer Freude für das Netz, sozusagen.»

Entdeckungen werden belohnt

Jürgen Schmidhuber hat diesen Belohnungsmechanismus in seiner formellen Fun- und Kreativitätstheorie formalisiert. Er vergleicht ihn mit der Erfahrung, die ein Musiker macht, wenn dieser eine Harmonie entdeckt, und auch mit dem Humor: «Wenn die Pointe eines Witzes eine unerwartete Wendung in der Geschichte bringt, können plötzlich Daten komprimiert werden. Dann lachen wir. Dieses Phänomen ist zentral, um Maschinen zu programmieren, die über eine künstliche Neugierde verfügen und Entdeckungen machen. Dafür muss das Netz als Duo funktionieren», erklärt der Forscher. «Auf der einen Seite steht der Generator, der Aktionen und Experimente durchführt, die Daten generieren; seine Motivation schöpft er aus dem Bestreben nach einer Maximierung der Belohnungen. Auf der anderen Seite steht das Modell, das jedes Mal eine Belohnung schickt, wenn es eine neue Regelmässigkeit entdeckt und dadurch Daten komprimieren kann. Ein solches System muss man einrichten, um einen künstlichen Wissenschaftler zu schaffen.»

«Das neuronale Netz von Insilico Medicine verwendet ein solches duales System», sagt Polina Mamoshina, Genetikerin und Informatikerin, die am Projekt mitarbeitet. «Der Datengenerator ist so programmiert, dass er virtuell und zufällig molekulare Strukturen schafft. Der Diskriminator trainiert mithilfe von Datenbanken, Moleküle zu erkennen, die fähig sind, das Wachstum von Tumoren zu hemmen. Am Anfang besteht das Ziel des Generators darin, den Diskriminator in die Irre zu führen, damit dieser falsche Identifikationen vornimmt und daraus lernen kann.

«Bei der Pointe eines Witzes können wir plötzlich Daten komprimieren. Dann lachen wir.»

Jürgen Schmidhuber

Von den 60 Molekülen, die der Generator gefunden und der Diskriminator validiert hat, wurden einige bereits als Wirkstoffe gegen Krebs patentiert. «Das ist ein ermutigendes Zeichen für die Präzision des Systems», ergänzt Mamoshina. «Wir werden nun den Prozess der Validierung der übrigen Moleküle *in vitro* und später *in vivo* in Angriff nehmen.» Für die Forscherin revolutioniert dieser Ansatz die Disziplin: Anstatt blind eine grosse Menge von Verbindungen zu durchsieben, besteht das Ziel darin, Medikamente mit massgeschneiderten Eigenschaften zu schaffen.

Während wir auf neugierige Maschinen von Jürgen Schmidhuber warten und auf Systeme, welche die Parallelen zwischen den Geschichten vom Steak und vom Haarschnitt von Roger Schank erkennen, werden das automatische Lernen und die Megadaten die wissenschaftliche Praxis revolutionieren. An der Universität Bristol ruft Nello Cristianini, Professor für künstliche Intelligenz, dazu auf, diese neuen Instrumente positiv aufzunehmen, den Anwendungsbereich aber genau abzustecken:

«Ich arbeite seit 20 Jahren mit maschinellem Lernen. Ich kann sagen, dass es funktioniert. Die Maschine ist in dem Sinne lernfähig, als sie ihre Leistung mit zunehmender Erfahrung verbessert.»

Bereits heute können diese Ansätze einen beträchtlichen Teil der Gewinne von Amazon generieren, wenn die richtigen Bücher den richtigen Leuten empfohlen werden. «Es ist bemerkenswert, dass diese Algorithmen weder ein psychologisches Modell jedes Anwenders enthalten noch ein Modell mit einer Literaturkritik für jedes einzelne Buch», sagt Cristianini weiter. «Sie führen rein statistische Überlegungen durch: Leute mit bestimmten Merkmalen und einem bestimmten Verhalten haben bestimmte Bücher gekauft. Daraus resultiert eine wichtige Erkenntnis: Es ist möglich, eine Prognose zu machen, ohne eine Theorie zu haben.»

> «Ich arbeite seit 20 Jahren mit maschinellem Lernen und kann bestätigen, dass es funktioniert.»
>
> Nello Cristianini

Doch lässt sich dieses Modell wirklich auf die Wissenschaft übertragen? «Es gibt keinen philosophischen Grund, der dagegen spricht», antwortet der Forscher. «Ein Computer könnte Molekülmodelle erstellen und ihre Toxizität voraussagen. Was damit gewonnen wäre? Dass ein Medikament *in silico* entwickelt werden könnte, ohne alle möglichen Moleküle herstellen und sie an Mäusen testen zu müssen. Was damit verloren wäre? Dass man nicht weiss, weshalb das Medikament funktioniert.»

Da man nicht weiss, welcher Logik die Maschine folgt, ist das automatische Lernen eine Black Box. Diese Intransparenz ist laut Cristianini insbesondere ausserhalb der akademischen Welt heikel: «Wenn die Algorithmen den Zugang zu Rechten regeln, zum Beispiel wer an eine Schule aufgenommen wird oder eine Versicherung abschliessen kann oder wer auf Bewährung entlassen wird.»

Das Ende der Theorie?

Kann diese Art des Lernens noch als Wissenschaft bezeichnet werden? «Man darf sich nichts vormachen», fährt Nello Cristianini fort. «Falls man mit maschinellem Lernen dank korrekter Prognosen Geld machen kann, werden wir wohl schrittweise neu definieren, was wir unter Wissenschaft verstehen. Die Gelder für die Forschung werden in erfolgversprechende praktische Anwendungen fliessen, während es schwierig werden dürfte, andere Ansätze zu finanzieren.» Ist zu befürchten, dass «das Ende der Theorie» kommt, wie es Chris Anderson ausrief? 2008 prophezeite der Verantwortliche des Magazins Wired, dass die Datenflut die wissenschaftliche Methode überflüssig machen werde.

«Man muss sich fragen, worin das Ziel von Theorien besteht», antwortet Nello Cristianini. «Für mich ist die Antwort glasklar. Eine schöne Theorie der Mechanik oder der Thermodynamik zu entwickeln, ein Stück unseres Universums zu verstehen – das ist ein unermesslicher kultureller Wert. Wir müssen die Funktionsweise der Welt und von uns selber kennen.» Eine gute Theorie hat einen konkreten Wert. «Eine Black-Box-Prognose reicht nicht aus, wenn viel auf dem Spiel steht, zum Beispiel, wenn eine Sonde auf den Mars geschickt oder eine chirurgische Operation geplant wird. In diesen Fällen will man genau wissen, was passieren würde, wenn man diesen oder jenen Parameter ändert. Dies erfordert kontrafaktische Überlegungen, die nur mit Theorien realisierbar sind.»

Auf der einen Seite probabilistische Maschinen, auf der andern Seite menschliche Wesen, die vom Wunsch geleitet werden, zu verstehen. Die in ihren biologischen Funktionen verwurzelt sind und Theorien aufstellen. Werden die Maschinen die Menschen ersetzen? «Genau festgelegte Aufgaben können Maschinen übernehmen», bestätigt der Forscher Cristianini. Beispiele? «Ein Protein zeichnen, von dem die Aminosäuresequenz bekannt ist. Ausgehend von einem vollständigen menschlichen Genom mit drei Milliarden Buchstaben die 20 000 Gene finden, aus denen es besteht … Doch die Bedeutung und Tragweite der Entdeckung zu erkennen wird dem Menschen vorbehalten bleiben.»

In den Sozialwissenschaften ist eine analoge Arbeitsteilung denkbar. Cristianini arbeitet mit Historikern zusammen, um die historischen Transitionen in der englischen und italienischen Geschichte anhand der systematischen Lektüre von Zeitungen zu messen. «Niemand kann 500 Millionen Artikel lesen. Deshalb geht die Maschine sie für uns durch. Weshalb dieses oder jenes Resultat wichtig ist, erklärt dann aber der Historiker.»

Ergänzung: Auch Jürgen Schmidhuber ist von diesem Konzept überzeugt. «In den wissenschaftlichen Laboratorien und anderswo werden Maschinen monotone Arbeiten erledigen, die für Menschen eigentlich unattraktiv sind. Dadurch gehen sicher Arbeitsplätze verloren. Wir werden deshalb als ganze Gesellschaft dafür zu sorgen haben, dass die Gewinne verteilt werden. Dies wird über ein bedingungsloses Grundeinkommen, Robotersteuern oder andere Mechanismen geschehen müssen.» Die durch die Automatisierung der Wissenschaft aufgeworfenen Fragen bleiben ebenso offen wie die epistemologischen. Ist eine automatische Vorhersage wissenschaftliches Wissen? Nello Cristianini stellt sich dieser Herausforderung: «Gerade habe ich zwei Wissenschaftsphilosophen eingestellt, die über diese Frage nachzudenken beginnen.»

Nic Ulmi ist freier Journalist in Genf.

He starts to shake.

H (CONT'D)
It may never be forgiven, but that is just too bad. I have to leave, but I'm not free of the world.

C
Yes. Perhaps I should take it from here. I'm not going to do something.

H
You can't afford to take this anywhere. It's not a dream. But I've got a good time to stay there.

C
Well, I think you can still be back on the table.

H
Mmm. It's a damn thing scared to say. Nothing is going to be a thing but I was the one that got on this rock with a child and then I left the other two.

He is standing in the stars and sitting on the floor.

Benjamin ist der erste automatische Drehbuchautor der Welt. Konzipiert wurde er von Ross Goodwin und dem Filmemacher Oscar Sharp. Die Szene aus seinem ersten Film Sunspring gibt eine Ahnung davon, wie lang er wohl noch üben wird, bis seine Dialoge und Geschichten auch für menschliche Betrachter Sinn ergeben.
Bild: thereforefilms.com/sunspring