Abo **KI der nächsten Generation**

An der ETH entwickeln sie eine Maschine, die tickt wie wir

Forschende arbeiten an einer menschenähnlichen künstlichen Intelligenz. Denn obwohl lernfähige Algorithmen erstaunliche Fähigkeiten besitzen, kann der Mensch sie um Längen schlagen – noch.

Joachim Laukenmann

Publiziert: 15.11.2022, 17:51

Wer ist klüger: Mensch oder Künstliche Intelligenz? Ein Intelligenztest liefert die Antwort.

Illustration: Künstliche Intelligenz (Dall-E)

In mehrfacher Hinsicht hat künstliche Intelligenz (KI) die Benchmark der menschlichen Fähigkeiten geradezu pulverisiert. Bei strategischen Spielen wie Schach oder Go ist der Mensch längst chancenlos. Auf Röntgenbildern erkennt KI Lungenleiden weit zuverlässiger als Radiologen und Radiologinnen.

Auch bei der Spracherkennung und -verarbeitung ist KI beeindruckend: Systeme wie GPT-3 von Open AI können bereits heute Texte generieren, die von handgeschriebenen Artikeln kaum zu unterscheiden sind.

All das sind Beispiele, wo eine KI anhand von grossen Datensätzen – etwa Schachpartien oder Röntgenaufnahmen – für ihre spezifische Aufgabe sehr umfassend trainiert werden kann.

Es kommt aber auch vor, dass sich Sprachsysteme wie GPT-3 übertölpeln lassen und sinnlose Inhalte formulieren. Das deutet auf eine gewisse Schwäche heutiger KI hin: Sie kann

nur das, wofür sie trainiert wurde und was sie kennt. Sobald ein Sprachsystem einer Kombination von Begriffen gegenübersteht, dem es so noch nie begegnet ist, fabuliert es Unsinn.

Ein Intelligenztest für künstliche Intelligenz

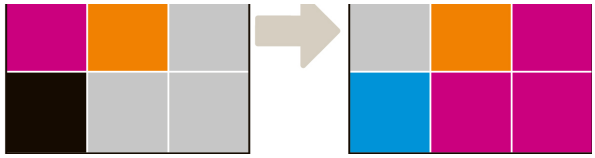
Am deutlichsten zeigt sich die Schwäche heutiger KI bei einer Herausforderung – einer Art Intelligenztest –, die der Mensch geradezu spielend erledigt, während jede noch so ausgefeilte KI kläglich scheitert. Das Scheitern der KI bei diesem Intelligenztest ist zugleich der Ausgangspunkt für Forschende der ETH Zürich, der Universität Zürich und des Lab42 ↗, um die KI der nächsten Generation zu entwickeln: eine, die wirklich tickt wie das menschliche Gehirn.

Worum es geht, zeigen die Muster in der Grafik. Das erste Muster wird nach einer gewissen Regel in das daneben gezeigte Muster verändert. Nach exakt der gleichen Regel wird das zweite Beispielmuster transformiert. Nun gilt es, das Konzept, nach dem die ersten beiden Muster verändert wurden, auf ein drittes Beispiel zu übertragen.

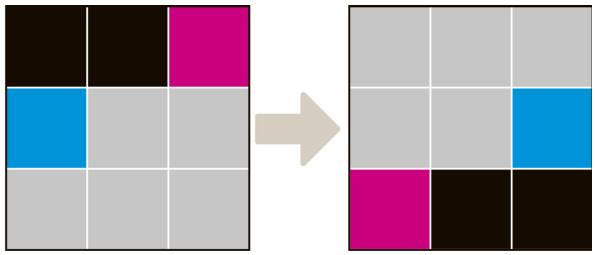
Können Sie den IQ-Test für Algorithmen lösen?

Zwei Beispiele zeigen Ihnen, wie eine Grafik in eine andere umgewandelt wird. Erkennen Sie die Gesetzmässigkeit? Dann können Sie die Aufgabe in der dritten Zeile bestimmt lösen.

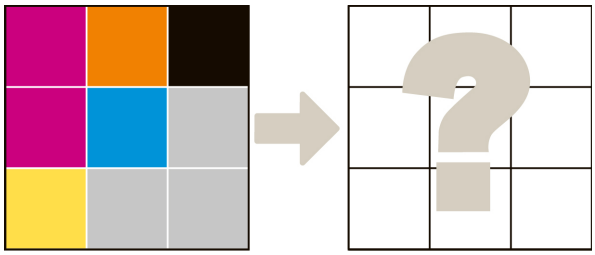
Beispiel 1
Links sehen Sie die Anfangsgrafik, rechts die Endgrafik



Beispiel 2
Links sehen Sie die Anfangsgrafik, rechts die Endgrafik



Aufgabe
Wie sieht die passende Endgrafik aus?



Die Auflösung finden Sie am Ende des Artikels

Tabelle: db; Quelle: arc-editor.lab42.global

Von Tausenden Aufgaben dieser Art können Menschen im Schnitt 85 Prozent lösen, die besten schaffen sogar mehr als 95 Prozent. Die Fachwelt spricht von der Abstraction and Reasoning Corpus Challenge, kurz: ARC-Challenge.

Die Herausforderung dieser Bilderrätsel besteht darin, anhand von wenigen Beispielen das Prinzip der Transformation zu erfassen und auf ein weiteres Beispiel anzuwenden. «Man braucht ein starkes Abstraktionsvermögen, um diese Aufgaben lösen zu können», sagt der Neuroinformatiker Benjamin Grewe, der eine Doppelprofessur an der ETH Zürich

und der Universität Zürich innehat und sich mit der Entwicklung einer menschenähnlichen KI beschäftigt.

KI-Forscher Benjamin Grewe mit einer Skulptur von HR Giger.
Foto: Michele Limina

Eine heutige KI scheitert an der ARC-Challenge, weil sie zu wenig Beispiele zum Üben hat – ganz anders als beim Schach oder bei der Erkennung von Tumoren auf einem Röntgenbild. Sie scheitert auch, weil sie nicht in der Lage ist, von den Spezialfällen, auf die sie trainiert wurde, zu generalisieren.

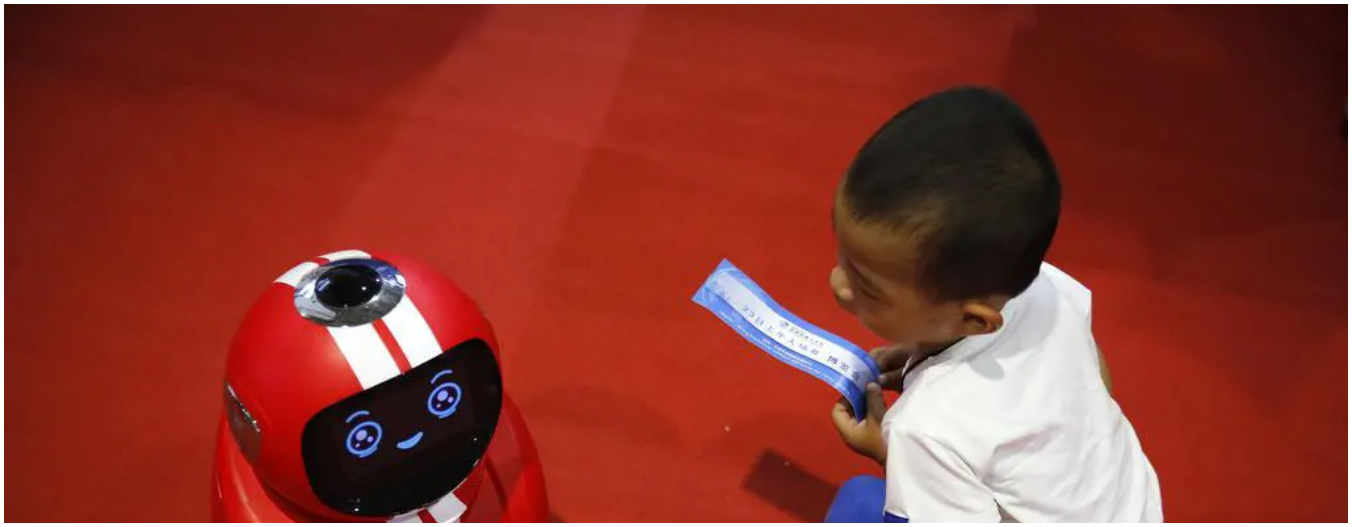
Heutige KI sind in dem Sinne Schmalspurintelligenzen. Sie sehen nicht über ihren engen Tellerrand hinaus. Begegnet der Autopilot einer Strassenszene, auf die er nicht trainiert wurde, weil die Bandbreite an möglichen Strassenszenen einfach zu gross ist, dann kann ein Unfall resultieren.

Der Mensch kann Erlerntes auf neue Situationen übertragen

Der Mensch indes besitzt die Fähigkeit, ein erlerntes Konzept auf eine neue Situation zu übertragen. Daher genügen ihm bei einer ARC-Challenge zwei bis fünf Beispiele, um ein Rätsel zu lösen. Entsprechend muss der Mensch nur wenige Fahrstunden nehmen, dann kann er auch in einer Stadt Auto fahren, die er nie zuvor gesehen hat, und Fahrsituationen meistern, denen er in der Fahrschule nicht begegnet ist.

Diese Fähigkeit zu abstrahieren ist im realen Leben essenziell. Denn sie erlaubt es, in einer sich ständig wandelnden Umwelt klarzukommen. «Wir wollen einen Algorithmus, der diese ARC-Challenge so wie wir sofort lösen kann, ohne sich die ARC-Beispiele jemals zuvor angeschaut zu haben», sagt Grewe.

Die zentrale Frage lautet also: Wie bringt man einer KI die Fähigkeit der Abstraktion und Übertragung bei? Die einfache Antwort lautet: indem man sie eine Umwelt erkunden lässt wie ein kleines Kind.



Künstliche Intelligenz muss lernen wie ein Kind: Ein Kind beobachtet einen intelligenten Roboter auf der World Robot Conference 2017 in Peking.

Foto: Andy Wong (Keystone)

«Kleine Kinder führen Aktionen aus und schauen, wie das ihre Umwelt verändert», sagt Grewe. Sie lernen, verschiedenfarbige Bauklötze hin und her zu schieben und zu stapeln. Dann verwenden sie das Erlernte, um mit Bauklötzen ein Haus zu bauen. «Auch der Algorithmus muss die Welt erforschen und schauen, was passiert.»

Im Grunde zerlegt der Mensch seine Umwelt in Einzelteile und Konzepte: Autos, Velos, Fussgängerstreifen, Ampeln, Farben, Bewegung nach links, Rotation nach rechts. Dann generiert er eine Vorstellung vom Zielzustand: Bei der ARC-Challenge zum Beispiel «rotes Quadrat nach oben schieben» oder auf der Strasse: «Auto um die Baustelle navigieren». So können wir Situationen, denen wir so genau noch nie begegnet sind, aus uns bekannten Konzepten zusammensetzen.

KI soll lernen, wie wir abstrakte Konzepte verarbeiten: Auto, Ampel, Velo, Schild.

Foto: Keystone

Entsprechend muss die KI lernen, ihre Umwelt nicht einfach als Anordnung von Millionen farbigen Pixeln zu sehen wie heute üblich. Sie muss die Pixelwelt in elementare Bausteine und Konzepte zerlegen, die Umwelt in sinnvolle Objekte einteilen und mit diesen Objekten Aktionen durchführen.

Das Team von Grewe an der ETH baut derzeit in Zusammenarbeit mit dem Lab42 eine kleine virtuelle Welt auf, in der ein virtueller Agent Klötze herumschieben, drehen oder andere Aktionen damit ausführen kann. «Der Agent führt zufällige Aktionen aus wie ein Baby und schaut, wie sich der Raum verändert, in dem er lebt.»

Trainingswelten sind wichtiger Schlüssel für menschenähnliche KI

Die Entwicklung solcher Trainingswelten, in denen KI lernen können wie Kinder, ist laut Grewe einer der wichtigsten

Schlüssel für die Entwicklung einer KI, die tickt wie das menschliche Gehirn. «Wenn ich eine KI entwickeln möchte mit gewissen Fähigkeiten, dann muss ich herausfinden, in welcher Trainingswelt sie aufwachsen muss.»

Kann KI mehr sein als eine Schmalspur-Intelligenz?
Illustration: Künstliche Intelligenz (Dall-E)

Daneben möchte Grewe besser verstehen, wie abstrakte Konzepte – Auto, Ampel, Farbwechsel, Quadrat – im Gehirn verarbeitet werden. Im Keller des Instituts auf dem Campus Irchel in Zürich zeigt Grewe einen kleinen schwarz ausgekleideten Mäusekäfig mit Touchscreen und Futterklappe für die Belohnung. Hier lösen Mäuse sehr einfache ARC-Challenges.

Diese sind so aufgebaut, dass es einfacher ist, Aufgabe 2 zu lösen, wenn die Mäuse das Konzept der Aufgabe 1 auf Aufgabe 2 übertragen.

Bis hinunter auf die Ebene einzelner Neuronen möchte Grewe verstehen, wie das Gehirn Konzepte bildet, und regt daher solche Hirnzellen unter einem Mikroskop an. «Das möchte ich dann auf die Algorithmen übertragen.»

Es braucht Kollaborationen mit der Industrie

Wenn alles gut läuft, meint Grewe, könnte er die grundlegenden Algorithmen für seine menschenähnliche KI in drei bis fünf Jahren bereit haben. Zunächst sollen die Trainingswelten deutlich kleiner sein als die komplexe Realität, in der wir Menschen aufwachsen und lernen. Etwa so wie die Trainingswelt für die ARC-Challenge. Denn wirklich grosse Trainingswelten seien sehr aufwendig und teuer. «Das kann die Forschung allein nicht leisten. Dafür braucht es Kollaborationen mit der Industrie.»

Zurzeit bereitet Grewe mit dem ursprünglichen Erfinder der ARC-Challenge, dem bei Google arbeitenden Franzosen François Chollet, und dem Lab42 eine ARC-Challenge mit rund 5000 Bilderrätseln vor, bei dem Hunderte Forschergruppen aus aller Welt den IQ ihrer KI testen sollen. «Wenn wir die ARC-Challenge lösen, können wir das eins zu eins auf autonome Autos, autonome Drohnen, Roboternavigation oder was auch immer anwenden», sagt Grewe. «Wenn dieser Ansatz durchschlägt, dann denke ich, dass wir das nächste Level von KI haben, nämlich Algorithmen, die Verhalten generieren und die ihre Umwelt wirklich verstehen.»

Zahlreiche Beispiele für die ARC-Challenge kann man hier probeweise lösen: arc-editor.lab42.global/playground ↗

Auflösung IQ-Test für Algorithmen

Die Grafik muss um 180 Grad gedreht werden

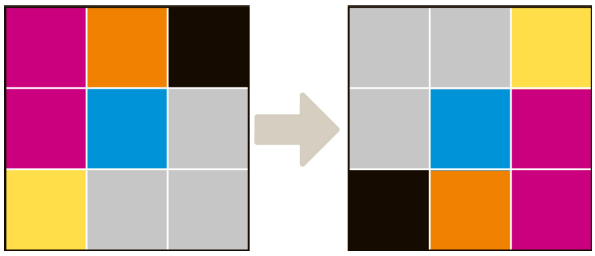


Tabelle: db; Quelle: arc-editor.lab42.global

Joachim Laukenmann ist Redaktor im Team Wissen. Seine Schwerpunkte sind Physik, Astronomie, Mobilität, Energie und Klimawandel. Er hat Physik studiert und in Kosmologie promoviert. 2008 erhielt er den Alstom Journalistenpreis. Er hat mehr als 20 Jahre Erfahrung im Wissenschaftsjournalismus. Mehr Infos

@JoLauki

Fehler gefunden? Jetzt melden.

26 Kommentare