

## Wie selbstfahrende Autos lernen, in Notsituationen ethisch zu handeln

Für wen entscheidet sich der Algorithmus, wenn er eines von zwei Menschenleben gefährden muss? Ein neuer Ansatz basiert auf Risikoverteilung, um dem Dilemma aus dem Weg zu gehen.

Herbie Schmidt

09.02.2023, 05.30 Uhr



In Brandenburg gibt es seit 2019 bereits die Möglichkeit, autonome Fahrzeuge im Strassenverkehr zu testen.

Janine Schmitz / Imago

Die Tür zum autonomen Fahren steht schon offen. Derzeit entwickeln Programmierer weltweit Systeme, die aus den Erkenntnissen von Sensoren und Kameras die richtigen Schlüsse ziehen. Algorithmen entstehen, die es den Systemen erlauben, blitzschnell Fahrentscheidungen zu treffen.

Immer wieder stellen sich nicht nur Kritiker die Frage, wie das autonom fahrende Auto entscheiden würde, wenn die Bremsen versagen und ein Anhalten unmöglich wird. Schützt das Auto den Fahrer oder die Menschenmenge? Und wenn die Wahl zwischen zwei Menschen besteht, für wen entscheidet der Algorithmus? Lässt sich Ethik überhaupt programmieren?

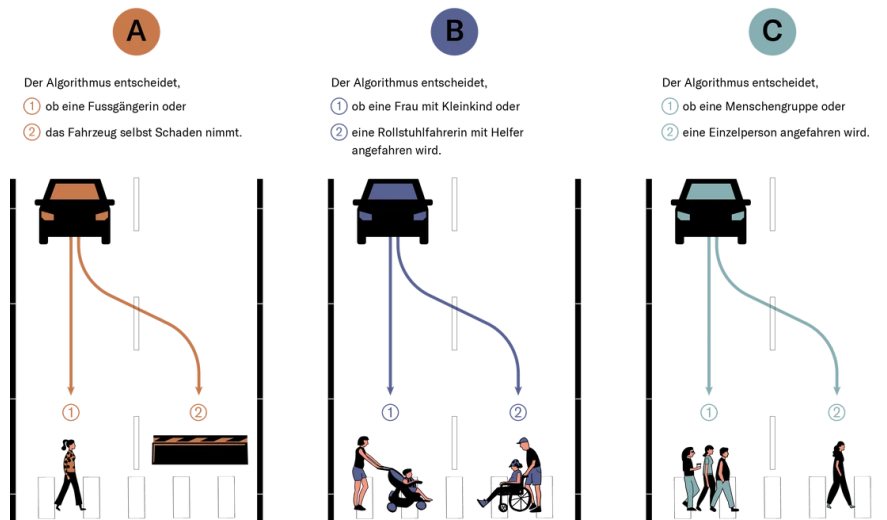
Man spricht in der Wissenschaft oft vom «Trolley-Problem»: Eine Strassenbahn gerät ausser Kontrolle und droht fünf Personen zu überrollen. Durch Umstellen einer Weiche kann das Tram auf ein anderes Gleis umgeleitet werden. Unglücklicherweise befindet sich dort eine weitere Person. Darf durch das Umstellen der Weiche der Tod einer Person in Kauf genommen werden, um das Leben von fünf Personen zu retten?

### **Eine Frage der Wahrscheinlichkeit**

Das auch als «Weichenstellerfall» bezeichnete Problem wird gerne auf autonom fahrende Autos übertragen, auch wenn sich das Szenario bis heute so nicht ereignet hat. Oft lautet die Antwort, auf eine solche Dilemma-Situation stosse man beim autonomen Fahren nie, der Fall sei zumindest höchst unwahrscheinlich. Die Aussage ist zu ausweichend. Es lässt sich nicht völlig ausschliessen, dass eine Entscheidung zwischen zwei schlechten Alternativen einmal für ein autonomes Fahrzeug entsteht.

## Welche Entscheidung soll das autonome Auto treffen?

Mögliche Szenarien mit einem Fahrzeug ohne Bremsfunktion



Quelle: Moral Machine

NZZ / eck.

Der im Jahr 2017 erschienene Bericht der deutschen Ethikkommission zum automatisierten und vernetzten Fahren beschäftigt sich mit diesen Dilemma-Situationen. Sie kommt darin zum Schluss, dass generelle Regelungen wie «Sachschaden vor Personenschaden» rechtlich nicht umgesetzt werden können. Die verschiedenen denkbaren Szenarien seien aufgrund ihrer Vielfalt und Komplexität für eine Normierung ungeeignet.

Immerhin hat sich die EU-Kommission im Jahr 2021 dazu durchgerungen, zwanzig Ethikempfehlungen herauszugeben. Diese sollen Programmierer bei der Entwicklung des autonomen Fahrens befolgen. Doch so einfach lassen sich Algorithmen nicht ethisch programmieren, es gibt eine riesige Zahl möglicher Fälle. Hinzu kommen Systeme, die sich durch eigenes Lernen während der Anwendung weiterentwickeln – nach dem Prinzip der künstlichen Intelligenz (KI).

Einen neuen Vorschlag zur Ethik beim autonomen Fahren präsentiert eine Forschergruppe der TU München in der kommenden Ausgabe der Fachzeitschrift «Nature Machine Intelligence». Das Team, bestehend aus Maximilian Geisslinger, Franziska Poszler und Markus Lienkamp, hat einen Algorithmus entwickelt, der mehr kann, als nur nach vorgegebener Route von A nach B zu fahren.

Das Neue bei dem System ist eine einprogrammierte Berechnung, welche Gefahren auf der Fahrt für andere Verkehrsteilnehmer bestehen könnten. Der Algorithmus hat dann die Aufgabe, die verschiedenen Risiken zu bewerten. Das Wissenschafterteam nutzt bei der Bewertung der verschiedenen Risiken fünf ethische Regeln:

- Maximal akzeptables Risiko, bei dem ein Fahrmanöver durchgeführt werden darf
- Besonderer Schutz derjenigen, die bei einem Unfall am schlimmsten betroffen wären
- Gleichbehandlung aller Menschen (unabhängig von Alter, Geschlecht, etc.)
- Minimierung des Gesamtrisikos
- Eigenverantwortung der Verkehrsteilnehmer

Nach diesen Kriterien spielten die Forscher rund zweitausend Fahrsituationen durch und entwickelten anschliessend den Algorithmus. Das Team weist darauf hin, dass seine fünf Kriterien zur Risikobewertung nur ein Beispiel darstellten. Es komme bei der Risikoabwägung auf die gesellschaftlichen Werte an, die je nach Anwendung gälten. Entscheidend sei

zudem dabei die künftige Gewichtung der genannten Ethikprinzipien.

## **Weg vom Schaden, hin zur Risikoverteilung**

Der Vorschlag aus Deutschland klingt interessant. Der Fokus liegt nun nicht mehr auf den Dilemma-Situationen, sondern auf der Risikoverteilung. Auf diese Weise werden Velofahrer und Fussgänger besser geschützt als mit anderen Algorithmen, weil ihre Vulnerabilität speziell berücksichtigt wird. Als besonders schwierig dürfte sich allerdings die Gewichtung ethischer Prinzipien erweisen.

Für Claudia Brändle vom Karlsruher Institut für Technologie ist dieser ethische Algorithmus «sehr spannend». Gegenüber diesem neuen ganzheitlichen Ansatz der Risikoverteilung «scheinen die klassischeren Konzepte für einen Umgang mit automatisierten Entscheidungen an Attraktivität zu verlieren».

Brändle sieht den Grund dafür bei der bisher starken Fokussierung auf Trolley-Probleme. «Sie machen vor allem die Grenzen der genannten ethischen Konzepte in der realen Anwendung deutlich.» Statt zum Thema zu machen, wo die moralische Verantwortung für Schäden liege, fokussiere der neue Ansatz auf die Verteilung des Risikos. Der vorgeschlagene Algorithmus könne ein guter Kandidat bei der tatsächlichen Umsetzung autonomen Fahrens sein. «Man muss aber wissen, auch dieser Algorithmus kann – genau wie andere – Risiken für Schäden lediglich minimieren, aber nicht ausschliessen.»

Kritischer sieht Ortwin Renn die Situation. Er ist wissenschaftlicher Direktor am Forschungsinstitut für Nachhaltigkeit in Potsdam. «Wenn man davon ausgeht, dass die vorgeschlagene KI-Software auf die Lösung ethischer Dilemmata ausgerichtet ist, gibt es eine Reihe von Argumenten, die diesen Anspruch zumindest relativieren», stellt er fest.

Renn moniert, dass zwischen den fünf genannten Ethikgrundsätzen Widersprüche auftreten könnten. «Man stelle sich vor: Links stehen zwei junge kraftstrotzende Männer und rechts eine gebrechliche ältere Frau. Zielt der Fahrer auf die ältere Frau, nimmt er in Kauf, dass diese Frau bei einem Aufprall wesentlich mehr Verletzungen davontragen wird als die beiden Männer.»

Getreu der neuen Studie müsste der Algorithmus dem zweiten Ethikprinzip Folge leisten: Die Priorität erhalte die ältere Frau, da sie höchst verwundbar ist. «Gleichzeitig würde aber dadurch das Gesamtrisiko verringert, weil jetzt eine Person statt zwei zu Schaden kommen würde. Damit wäre das erste Ethikkriterium erfüllt. Welches der beiden Kriterien hat dann Vorrang?»

## **Wie bewertet künstliche Intelligenz die Risiken?**

Einen zweiten Fallstrick sieht Renn in einigen der Kriterien wie «Verantwortung» oder «maximal vertretbares» oder «akzeptables» Risiko. Hier sei der Algorithmus auf das Ermessen des Programmierers oder der KI angewiesen. «Wann ist denn ein Risiko akzeptabel?», fragt der Potsdamer Professor. «Dafür gibt es keinen allgemeingültigen und von

allem prinzipiell akzeptierten Massstab. Im schlimmsten Fall macht das die KI nach Gutdünken und gibt vor, einen <objektiven> Massstab zu besitzen.»

Neben der Kritik am Algorithmus der Risikoverteilung verwendet er auch das alte Argument der Unwahrscheinlichkeit: «Sosehr sich die Community zur ethischen Verantwortbarkeit von autonomem Fahren mit diesen Dilemmata beschäftigt und händeringend nach Lösungen sucht, so unwahrscheinlich ist es, dass es je zu einem dieser Dilemmata kommt.» Es sei bei Versicherungen noch kein Fall gemeldet worden, bei dem es tatsächlich zu einer der oben geschilderten Dilemma-Situationen gekommen sei.

Was Renn und andere Experten, die sich zu Trolley-Problemen äussern, vergessen: Die Entwicklung hin zum autonomen Fahren auf höchster Stufe ist in vollem Gange. Das Auftreten von Dilemmata könnte immer wahrscheinlicher werden.

Der an der TU München entwickelte Algorithmus wurde bisher in den rund zweitausend Simulationen validiert. Mit dem Forschungsfahrzeug «Edgar» wird die Software künftig auch im Strassenverkehr getestet, jedoch vermutlich jeweils mit Sicherheitsfahrern an Bord. Der Algorithmus, in den die laufenden Erkenntnisse der Forschungsarbeit fliessen, steht als sogenannter Open-Source-Code öffentlich zur Verfügung.

## Passend zum Artikel

INTERVIEW

### **Autonomes Fahren: «Wir haben keine Zuverlässigkeitsprobleme»**

10.11.2021



---

### **Die Konkurrenz holt auf, das autonome Fahren stockt: Tesla ist in der Bewertungshölle angekommen**

20.10.2022



---

### **Die Zweifel am autonomen Fahren wachsen nach schwerem Crash**

17.08.2022



---

**Mehr von Herbie Schmidt (hdt) >**

---

### **Alpine A110 R: ein Sportgerät, das keinen Rückspiegel benötigt**

01.02.2023



---

### **Katar schickt den Autosalon Genf in die Wüste**

31.01.2023





## **Mazda CX-60 im Test**

24.01.2023



---

## **Windräder statt Schafzucht: Porsche hofft auf grünen Sprit aus Patagonien**

19.01.2023



---

## **Bentley Flying Spur Hybrid im Test: einer der letzten Alibi-Hybride mit kleiner Reichweite**

13.01.2023



---

**Mehr zum Thema Autonomes Fahren >**

## **Autonomes Fahren ist bisher ein unerfülltes Versprechen. Das könnte sich auch dank China gerade ändern**

22.10.2022



## Mit dem Roboter Optimus will Tesla wieder einmal die Welt retten

01.10.2022



### KOMMENTAR

## Elon Musk hört nicht auf, Dinge zu versprechen, die Tesla-Fahrer in Gefahr bringen

16.09.2022



## Ein Tesla-Besitzer verklagt den Autohersteller wegen falscher Versprechen zum automatisierten Fahren

15.09.2022



## Schanghai will Chinas Zentrum für intelligentes Fahren werden

12.09.2022



## Wer hat Angst vor Autos ohne Fahrer?

08.08.2022



---

Für Sie empfohlen >

## **Kommt es in Baselland zur Sensation? SVP-Regierungskandidatin Sollberger ist überraschend in der Defensive**

vor 3 Stunden



---

### **Die Airbnb-Debatte geht wieder los: Luzern stimmt über eine Regulierung ab**

vor 3 Stunden



---

### **In Zürcher Apotheken und Spitälern fehlen Medikamente. Ein Arzt sagt: «Wir haben immer wieder verzweifelte Eltern, die Schmerzmittel für ihr Kind suchen»**

vor 4 Stunden



---

### **SP und Grüne beginnen sich in der Debatte um Waffenexporte zu unterscheiden: «Wir sind nicht die Zwillinge, für die uns alle halten»**

vor 3 Stunden



---

### **Das Aussehen ist der häufigste Grund für Hassrede auf Social Media: Eine Studie zeigt die Verrohung der politischen Debatte**

vor 3 Stunden



---

### **Links-Grün will, dass die Stadt Zürich bis 2030 auf 10 Prozent Solarstrom kommt. Der Stadtrat sagt: Das ist unmöglich**

vor 3 Stunden



---

## **Ein kleiner Investor bringt Gautam Adanis Milliardenimperium ins Wanken: Wie Leerverkäufer mit gezielten Kampagnen Geld verdienen**

08.02.2023



---

## **Zürcher Stadtparlament möchte «goldene Fallschirme» für Amtspersonen bekämpfen – aber nicht so, wie es die SVP will**

08.02.2023



---

Copyright © Neue Zürcher Zeitung AG. Alle Rechte vorbehalten. Eine Weiterverarbeitung, Wiederveröffentlichung oder dauerhafte Speicherung zu gewerblichen oder anderen Zwecken ohne vorherige ausdrückliche Erlaubnis von Neue Zürcher Zeitung ist nicht gestattet.