

Lotterie des Sterbens

Verkehr Bei selbstfahrenden Autos entscheidet der Rechner über Leben und Tod. Darf er notfalls auch die Insassen opfern? Über die Abgründe einer neuen Maschinenethik.

Eines Tages wird es geschehen, so oder ähnlich: Ein selbstfahrendes Auto saust übers Land, der Computer lenkt. Der Fahrer hat es gemütlich, er liest Zeitung. Da hüpfen drei Kinder auf die Straße, links und rechts stehen Bäume.

In diesem Augenblick muss der Computer entscheiden. Wird er das Richtige tun? Drei Menschenleben hängen davon ab.

Nein, vier.

Der Computer könnte ja auch zu dem Schluss kommen, es sei unter diesen Umständen am besten, den Fahrer zu opfern. Dann setzt er den Wagen an den Baum.

Und warum auch nicht? Sollen lieber drei Kinder sterben?

Kein schönes Ende für unseren Fahrer. Die Visionäre vom selbstfahrenden Auto dagegen machen es sich leicht: Der Mensch rollt entspannt dahin, sagen sie, bequem und sicher wie in der Eisenbahn. Aber es wird immer Unfälle geben, die nicht zu vermeiden sind: Kinder aus dem Nichts, jäh kreuzende Motorradfahrer oder ein schleudernder Bruchpilot auf der Gegenpur.

Was dann?

Bis der Mensch aus der Zeitung auftaucht und die Lage peilt, ist es zu spät. Die Maschine allein muss in so einem Szenario ausrechnen, was zu tun ist; sie ist hoffentlich darauf programmiert, den Schaden möglichst gering zu halten. Das heißt aber auch: Damit urteilt sie über Leib und Leben der Beteiligten.

Ob die Gesellschaft das erträgt? Die Frage stellt sich eher früher als später.

Erstaunlich schnell kam die Technik voran; bald könnten die ersten Autos einsatzreif sein, die ganz ohne Eingriffe des Fahrers auskommen – auf den Autobahnen schon 2020, wenn es nach den Optimisten geht. Ein paar Jahre später wäre es dann im gesamten Straßennetz so weit.

Viele Hersteller, von Daimler bis Tesla, haben selbstfahrende Autos in Arbeit; allen voraus eilt Google. Der Suchmaschinenkonzern wagt auch die weitestgehende Prognose. Chris Urmson, Projektleiter des Google-Autos, sagte kürzlich, sein älterer Sohn dürfe in vier Jahren den Führerschein machen: „Wir wollen, dass er dann gar keinen mehr braucht.“

Der Computer wäre sicherlich der bessere Fahrer – nie abgelenkt, reaktionsstark und stets im Bilde über das Verkehrsge-

schehen ringsum. Sollte er das Steuer übernehmen, dürfte die Zahl der Unfälle stark sinken. Dafür wird es andere Unfälle geben, neuartige und verstörende.

Denn künftig wird so manches Opfer vorher ausgesucht. Das Auto wählt, wen es verletzt und wen es verschont – im Grenzfall auch: wer weiterleben darf und wer nicht. Dagegen nimmt sich der gute alte Verkehrsunfall fast gemütlich aus.

Ein menschlicher Fahrer hat kaum eine Wahl, wenn es passiert: Er steigt in die Bremse, reißt das Steuer herum, lenkt hierhin oder dorthin, streift Passanten, kracht in ein anderes Auto oder gegen einen Baum. Egal, was er tut: Wer wollte ihn für die Reflexe einer Schrecksekunde haftbar machen? Selbst wenn der Unglücksfahrer sieben Fußgänger an der Ampel überrollt, weil er einem Kind ausweichen wollte, das ihm vor die Stoßstange lief: Schicksal!

Niemand trifft überlegte Entscheidungen im Augenblick höchster Not. Da walten nur noch die uralten Paniksysteme eines Lebewesens, das sich retten will.

Ganz anders der Computer. Er kennt kein Entsetzen, er bleibt kalt und jederzeit zurechnungsfähig. Im Nu analysiert er die Lage, tausendmal schneller als der Mensch. Er lenkt zudem ein Auto, das vollgepackt ist mit Sensoren und optischen Abstandsmessern. Kameras sorgen für stete Rundumsicht. Bald kann er, mittels Gesichtserkennung, womöglich auch Alter und Geschlecht der Unfallbeteiligten einschätzen.

Es ist, als könnte ein Fahrer mitten im Unfallgeschehen gleichsam den Film anhalten: Dann sähe er vor sich die Kinder, mitten auf der Straße – aber zugleich auch schon das voll besetzte Auto, das von hinten viel zu schnell herangerast kommt.

Und was nun? Vollbremsung oder nicht?

Der Computer hat die Wahl. Selbst in extremer Zeitnot kann er annähernd kalkulieren, was geschieht, wenn er so oder so handelt. Er ist imstande, rational zu entscheiden, und deshalb muss er das auch. Er kann nicht einfach panisch bremsen oder drei Kinder überfahren, wenn es einen weniger schlimmen Ausgang gibt.

Nur: Ist der Tod des unschuldigen Fahrers weniger schlimm?

Solche Fragen wirft die Zukunft auf, und Philosophen sind bereits zur Stelle.



Mit leicht sadistischer Einfallsfreude denken sie sich künftige Unfallszenarien aus, eines abgründiger als das andere. Die Aufgabe jeweils: Berechnen Sie den geringstmöglichen Schaden.

Vergleichsweise einfach ist noch die Wahl zwischen zwei Fußgängern. Ein Kind läuft auf die Straße, das Auto könnte nur noch auf den Gehweg ausweichen. Dort schiebt eine Greisin ihren Rollator. Dem Kind stehen, statistisch gesehen, noch viele Lebensjahre bevor. Die Greisin hätte viel weniger zu verlieren. Dennoch: Darf ein Auto zum Zwecke der Schadensminderung die arglose Dame töten?



JON FRICKEY / DER SPIEGEL

Heikler wird es, wenn indirekte Personenschäden zu kalkulieren sind. Angenommen, zwei Autos kommen entgegengebraust, links ein Volvo, rechts ein Kleinwagen, der halsbrecherisch überholt. Ein Zusammenstoß ist unausweichlich. Der computergelenkte Wagen kann den dicken Volvo rammen, der wohl mit ein paar Beulen in der Panzerung davonkäme. Oder aber er zerknüllt den Kleinwagen und nimmt den Tod der Insassen in Kauf.

Ein Computer, der auf geringstmöglichen Schaden programmiert ist, müsste den Kleinwagen verschonen. Aber ist das gerecht? Zumindest Volvo-Fahrer werden

Einwände haben. Wer will schon in einem rollenden Rammbock sitzen, der im Umkreis von Kollisionen wie magisch die autonomen Autos auf sich zieht?

Im Beispielfall kommt die Schuldfrage hinzu; schließlich war der Kleinwagen auf der falschen Spur. Sollte ihm der Computer dafür nicht einen Abschlag in Rechnung stellen?

Das sind nur Gedankenspiele zum Eingewöhnen. Aber es wird zu Unfällen kommen, die ihnen ähneln – und jeder einzelne dürfte Aufsehen erregen.

Natürlich ist es nicht der Computer, der in der Schrecksekunde entscheidet. Er ist

nur eine Maschine, er braucht eindeutige Regeln, eine Software, ein Programm. Wie sich das Auto in allen erdenklichen Unfallszenarien verhalten soll, muss geklärt sein, noch bevor es aus der Fabrik rollt.

Damit verwandelt sich, was bisher Schicksal war, in ein Denkproblem. Die Gesellschaft muss sich eine Moral für ihre Autos überlegen. Das Gute: Sie hat alle Zeit der Welt dafür. Das Schlechte: Eine befriedigende Lösung ist kaum vorstellbar.

Bei allen Entscheidungen sind Folgen zu bedenken. Soll man den Autos erlauben, mit der Rücksichtslosigkeit einer Kampfmaschine stets das Überleben der

Insassen zu retten? Auch wenn dafür drei Kinder umkommen?

Die Alternative ist auch nicht verlockend: ein Wagen, der den Fahrer killt, sobald es ihm angemessen dünkt. Wer würde so ein Gefährt kaufen? Die Maschine, die über das Weiterleben entscheidet, zumal das eigene, ist einfach gar zu unheimlich. Hinzu kommt, dass keine Elektronik ganz ohne Fehler ist. Was, wenn der Computer sich einmal vertut und gegen den Baum rast, nur weil ein Kaninchen über die Straße hoppelt?

Das Auto, wie wir es kennen, war dem Fahrer total zu Diensten; das ist sein Zauberei. Es gab aber auch schon früh den Grusel vor der Maschine, die zum Eigenleben erwacht. Man denke an Stephen Kings Killermobil „Christine“: Ein blutroter 58er Plymouth Fury rollt da durch die Straßen auf der Suche nach Opfern.

Es gibt noch eine Gattung von Maschinen, die über Leben und Tod entscheiden könnten: Killerdrohnen finden vielleicht bald ohne Bodenpiloten ihr Ziel. Auch das selbstfahrende Auto braucht, wie die Drohne, einen Algorithmus, der Zielpersonen aufspürt – nur dass es im Straßenverkehr um einen vorgeblich höheren Zweck geht: den geringstmöglichen Schaden an Leib und Leben.

Gibt es Auswege aus dem Dilemma der Autonomie?

Fachleute diskutieren sogar schon den Losentscheid. In Grenzfällen könnte ein Zufallsgenerator in der Steuerungssoftware auslösen, ob es das Kind auf der Straße oder die betagte Passantin auf dem Gehweg trifft. Aber eine Lotterie der Opfer klingt erst recht unmenschlich. Außerdem: Was nützt der Computer am Steuer, wenn er am Ende doch wieder kopflos wie ein Mensch reagieren soll?

Bleibt als Lösung, die leidige Autonomie des Autos zu beschränken. Dann wäre der Fahrer verpflichtet, sich ständig einsatzbereit zu halten. Aber selbst dann müsste die Maschine für den Fall gerüstet sein, dass er versagt – oder nicht schnell genug die Lage erfasst. Sie muss handeln, weil sie handeln kann.

Ein autonomes Auto, da hilft wohl nichts, ist immer im Dienst – erst recht, bevor es kracht: Was wäre nun der glimpflichste Verlauf? Das geht nur mit Rechnen. Für jeden Personenschaden braucht der Computer einen Zahlenwert, genauer: einen Preis. Ist das Leben eines Kindes mehr wert als das eines Erwachsenen? Wie ist das bei zwei Erwachsenen? Bei fünf?

Wie ist der Schuldfaktor zu gewichten? Fallen Kinder im Auto des Unfallgegners als Unschuldige ins Gewicht, auch wenn die

ser bei Rot noch über die Kreuzung jagte? All das muss in der Software festgelegt sein. Programmierer sprechen von Kostenfunktionen. Das sind Gleichungen, mit denen sich widerstreitende Interessen verrechnen lassen. Am Ende wählt das Auto die Lösung mit der geringstmöglichen Schadenssumme – mit einem Wort: den optimalen Unfall.

Aber sind Glück und Leid überhaupt berechenbar? Philosophen denken schon lange darüber nach. Der englische Sozialreformer Jeremy Bentham entwarf im späten 18. Jahrhundert sogar ein Formelwerk dafür. Sein Algorithmus sollte für jede

denschaftslos wie eine Maschine, notfalls ohne Rücksicht auf das Individuum.

Benthams Nutzenkalkül galt lange als närrisch, obendrein kaum praktikabel. Aber nun wird der alte Utilitarist für das selbstfahrende Auto wieder interessant: Lasse sich nicht das heikle Entscheidungsproblem nach seinen Prinzipien regeln?

Mit dem bloßen Verrechnen von Personenschäden ist es freilich nicht getan. Autos, die sich darauf beschränkten, wären eine stete Bedrohung fürs Gemeinwesen. Jederzeit könnten sie etwa die unbeteiligte Greisin auf dem Fußweg umfahren, um einem Kind auszuweichen, das plötzlich auf der Straße erschien. Das ist, als müsste ein Besucher im Krankenhaus Angst vor den Ärzten haben – hinter der nächsten Ecke könnten sie ihn packen und ausweiden, um mit seinen Organen fünf Todkranke zu retten.

Philosophen suchen bereits nach einer Lösung. Ein Ansatz: Jedes Nutzenkalkül braucht einen Rahmen; nicht alle Opfer sind gleichwertig. Zum Beispiel erwarten Verkehrsteilnehmer mit Recht, dass die anderen sich konform verhalten. Wer sich im Straßenraum bewegt, weiß, dass er aufpassen muss. Auf dem Gehweg dagegen kommen normalerweise keine Automobile dahergeschossen; hier darf der Fußgänger unbesorgt trödeln. Das Auto muss diese geschützte Sphäre achten – auch wenn es das Leben eines Kindes kostet. Die arglose Passantin hätte demnach die höheren Rechte.

Wie auch immer die neuen Regeln des Handelns sich herausbilden: Sie müssen mathematisch formuliert sein, in Zahlen und Gleichungen, damit der Computer sie versteht. Das ist der Preis. Je autonomer die Maschinen, desto maschineller die Moral.

Eine breite Debatte darüber ist unabdingbar. Gut möglich,

dass daran der Traum vom selbstfahrenden Auto schon scheitert, bevor er wahr wird – weil es eine Gesellschaft heraufbeschwören würde, in der kaum jemand leben will.

Dann macht der Fortschritt eben mal halt, nicht wahr? Und der Mensch bleibt bis auf Weiteres selbst am Steuer.

Auch das wäre keine unschuldige Entscheidung mehr. Bis zu 90 Prozent der Verkehrsunfälle gehen, je nach Schätzung, auf menschliches Fehlverhalten zurück. Über eine Million Menschen sterben weltweit jedes Jahr auf den Straßen – solange der Computer nicht übernehmen darf, wird sich daran wenig ändern. Ist es das wert?

Manfred Dworschak

Mail: manfred.dworschak@spiegel.de



Handlung ermitteln, wie viel Befriedigung oder Schmerz sie nach sich zieht. Intensität, Dauer, Zahl der Betroffenen – alles ging in die Gleichung mit ein.

Auf Bentham geht die Denkschule der Utilitaristen zurück. Ihre Anhänger sagen: Gut ist, was den größtmöglichen Nutzen (Englisch: utility) hervorbringt – und zwar umgerechnet auf alle Beteiligten. Das Prinzip der Menschenrechte hielt Bentham für „Unsinn auf Stelzen“. Eine Verfassung wie die deutsche hätte er verlacht: Sie verbietet es, einen Menschen zu töten (oder auch, nur zu foltern), selbst wenn das viele Leben retten könnte. Für Bentham war das gar keine Frage: Er träumte von einer Gesellschaft, die streng wissenschaftlich ein Maximum an Gemeinwohl produziert, lei-