

Rationalisierung von Moral im Einsatz von autonomen Personenkraftwagen

Bilal Karim Reffas `bilal.karim.reffas@mni.thm.de`

25. August 2018

Technische Hochschule Mittelhessen,
Fachbereich Mathematik, Naturwissenschaften und Informatik,
Wiesenstraße 14, 35390 Gießen

Zusammenfassung. Autonome Fahrzeuge sollen das Leben vieler Menschen erleichtert und dabei noch die Todesfälle im Straßenverkehr drastisch senken. Es stellt sich jedoch die Frage, wie gehen die rational denkenden Maschinen mit Extremsituationen um? In dieser Hausarbeit werden philosophische Perspektiven angewendet, um solche Extremsituationen analysieren zu können.

Schlüsselwörter: Rationalisierung · Moral · Autonomes Fahren · Künstliche Intelligenz · THM · Moral Maschine

1 Einleitung

Seit Jahren wird vermehrt im öffentlichen Diskurs über autonome Autos berichtet. Statistiken zufolge sollen im Jahr 2030 27 Millionen autonome Pkws in Europa zugelassene sein [2].

Befürworter der autonomen Pkws sind der Meinung, dass die 3177 Verkehrstoten, die jährlich im Deutschen Straßenverkehr umkommen, gesenkt werden [3]. Außerdem sollen sie die Bequemlichkeit während der Reise erhöhen. Kritiker der neuen Technologie hingegen befürchten, dass die Technologie, vorallem in Extremsituationen, zu moralisch problematischen Ereignissen führen kann.

Der Wunsch von autonomen Pkws liegt in der Geschichte der Menschheit schon weit zurück. Schon Leonardo da Vinci hat bereits im Jahr 1500 Skizzen für den Bau eines autonomen Wagens entworfen [1]. Der erste revolutionäre Schritt kam mit der Erfindung des Geschwindigkeits-Tempomats von Ralph Teetor im Jahr 1945 [1]. Durch die rasante Weiterentwicklung von autonomen Pkws seit 2010 werden vermehrt mehr Diskussionen geführt über die Rationalisierung von Moral im Zusammenhang von autonomen Pkws, denn durch Technologien wie das Maschinelle Lernen und immer größere Datensätze, die auch verarbeitet werden können durch schnellere Rechner, kann die Technologie zur Realität werden.

Der größten Anbieter dieser Technologie, gemessen an den gefahrenen Kilometern, ist Alphabet-Tochter Waymo [5]. Alphabet-Tochter Waymo, die im Jahre 2009 gegründet worden ist, führt die Liste der „Selbstfahrenden-Kilometer“ mit einem großen Vorsprung an, mit einer Reichweite von 567.366 km, wobei nur 63

Eingriffe von Menschen notwendig waren [5].

Es stellt sich die Frage, wie handelt die Maschine moralisch, sobald Menschen nicht mehr in das System eingreifen. Ist es überhaupt möglich einer rationale Maschine, wie zum Beispiel dem Computer, Moral beizubringen? Ist es verwerflich, unter den Gesichtspunkten des heutigen Fortschritts der Künstlichen Intelligenz, autonome Pkws zuzulassen? Sollte es einen Führerschein für die Software dieser autonomen Pkws geben oder gar TÜV Prüfungen? Mit diesen Fragestellungen wird sich diese Hausarbeit beschäftigen. Im folgenden wird aus Gründen der besseren Lesbarkeit auf die gleichzeitige Verwendung männlicher und weiblicher Sprachformen verzichtet. Sämtliche Personenbezeichnungen gelten gleichermaßen für beiderlei Geschlecht.

1.1 Level des autonomen Fahrens

Ein autonomer PKW verfügt über ein Level, das die Fortschrittlichkeit des Systems klassifiziert.

- Level 1 (Menschlicher Fahrer)
Die erste Stufe kommt weitestgehend ohne Automatisierung aus. Der Fahrer hat die volle Kontrolle über die Geschwindigkeit, das Gaspedal, die Bremsen und das Lenkrad. Dennoch kann das Fahrzeug Sicherheitssysteme besitzen, wie zum Beispiel ein Abstandshalter.
- Level 2 (Teilautomatisiertes Fahren)
Pkw des zweiten Levels übernehmen die Technik, die Längs- und die Querverführung zum Beispiel beim Einparken oder beim Spureneinhalten. Hierbei handelt es sich um teilautomatisierte Fahrzeuge, die jederzeit vom Fahrer persönlich überwacht werden müssen. Deutsche Premiumhersteller setzen diese Technologie bei Stauassistenten ein. Tesla bewegt sich mit seiner Technologie zwischen dem Level 2 und 3 [6].
- Level 3 (Hochautomatisiertes Fahren)
Beim Level 3 kann der Fahrer sein Fahrzeug auch über längere Zeit unbewacht lassen. Das System erkennt, sobald es ohne den Fahrer nicht mehr in der Lage ist, den Pkw zu steuern. Dann erfolgt eine Warnung des Systems an den Fahrer, der anschließend die Kontrolle übernehmen muss.
- Level 4 (Vollautomatisiertes Fahren)
Ausgereifte Systeme brauchen kein Eingreifen des Fahrers, außer bei klar definierten Fahrsituationen. Momentan gibt es noch keine Fahrzeuge dieses Level zu kaufen. Die einzige Firma, die bereits straßentaugliche Pkws diesen Levels entwickelt hat, ist Google [6].
- Level 5 (Fahrerloses Fahren)
Ab Level 5 sind Fahrer überflüssig und das System kann mit jeder Situation umgehen. Im Moment gibt es noch kein Fahrzeug, das über diese Technologie verfügt [6].

2 Rationalisierung von Moral

Damit ein Subjekt Verantwortung übernehmen kann, muss dieses Subjekt mindestens handlungsfrei sein, Einfluss auf das Handlungsergebnis haben und Zurechnungsfähig sein. Das alles trifft theoretisch bei den Rechenmaschinen je nach Situation zu. Immanuel Kant unterscheidet die Zurechnungsfähigkeit je nach der Beziehung des Subjektes zum Geschehen. Jede Abstufung der Zurechnungsfähigkeit ist bei autonomen Pkws gegeben. Da rationale Systeme nicht empfinden können, ist es ziemlich eindeutig beurteilen zu können, dass diese eine Weltanschauung mit persönlichen Werten, Vorstellungen und Sichtweisen nicht besitzen. Daraus folgt, dass Handlungsmuster, Handlungskonventionen, Handlungsregeln und Handlungsprinzipien diesen Systemen beigebracht werden müssen, damit kritische Situationen abgewegt werden können. Da Moral sich auf das Verhalten von Menschen gegenüber Menschen bezieht und nicht auf das Verhalten von Tieren oder Rechenmaschinen, ist eine Moralbildung für die Rechenmaschine ohne Einflüsse von außen nicht möglich. Eine Implementierung des kategorischen Imperativ oder hypothetischen Imperativ sollte allerdings technisch keine Probleme darstellen, dass die Rechenmaschine keine Lebewesen überfahren soll, kann ihr beigebracht werden. Das würde aber auch bedeuten, dass autonome Fahrzeuge je nach Situation anders kritisch reagieren würden, weil Menschen je nach soziokulturellem Hintergrund auch andere Wertvorstellungen besitzen. Doch wie würde eine Moralbildung von autonomen Pkws konkret aussehen? Hierbei könnte man auf die großen Fortschritte der Künstlichen Intelligenz in den letzten Jahren verweisen. Es ist durchaus möglich eine Online Befragung über die ethischen Dilemma durchzuführen, die das Massachusetts Institute of Technology schon durchgeführt hat. Hierbei wurden 2000 Probanden gebeten, bei klassischen Dilemmata eine eindeutige Entscheidung zu treffen. Eine der Fragestellungen war: Ein autonomes Fahrzeug rast auf eine Menschenmenge zu. Entweder fährt es mitten hinein und tötet die Passanten oder es weicht aus, prallt gegen eine Barriere und tötet seinen Insassen. Das Ergebnis der Befragung ist eindeutig: Rund drei Viertel der Befragten entschieden sich dafür, lieber den einen Menschen opfern als die ganze Menschenmenge. Das nahezu gleiche Ergebnis gab es, als die Befragten annehmen sollten, sie selbst oder Mitglieder ihrer Familie säßen im Fahrzeug [7]. Auch die Kognitionswissenschaftler der Universität Osnabrück beschäftigen sich in ihrer Studie mit diesem Thema [8]. Ein zentrales Ergebnis ist, dass Computer moralische Entscheidungen von Menschen lernen können. Damit widerlegen sie eine Aussage der Ethik-Kommission der Bundesregierung [9]. Es zeigt sich außerdem, dass Unternehmen oft eine andere Meinung im Vergleich zu den Probanden haben. Der Daimler-Sicherheitschef ist zum Beispiel der Meinung, dass Selbstfahrende Pkws zuerst die Insassen retten sollten [10]. Damit die Maschine die Moral erlernen kann, lassen sich Machine Learning Modelle anfertigen, die in Extremsituationen eine Entscheidung treffen. Natürlich darf dabei nicht vergessen werden, dass das Machine Learning Modell nur so gut sein kann wie die Anzahl der trainierten Fragen und Antworten. Es könnten also auch kritische Situationen entstehen, die für die Maschine eine komplett neue Herausforderung darstellt. Da letztendlich die Autoindustrie die Rationalisierung

von Moral implementieren muss, ist es interessant, nach welchen Regeln dies geschehen wird. Abzuwarten bleibt, ob es eventuell noch Gesetze geben wird, die die Pläne, wie die des Daimler-Sicherheitschefs, unterbinden werden.

3 Philosophische Perspektiven

3.1 Utilitarismus

Das Grundprinzip des Utilitarismus ist, dass man bei der Durchführung einer Handlung die größtmögliche Summe an Nutzen für alle Betroffenen erreicht. Deshalb sind die meisten Menschen der Meinung, dass utilitaristische Regeln zur Minimierung von Opfern zu bevorzugen sind, wie in der Studie der Kognitionswissenschaftler der Universität Osnabrück bestätigt wird [9]. Dennoch würde die Mehrheit der Teilnehmer keinen Pkw mit solchen Algorithmen kaufen wollen. Im Sinne der utilitaristischen Moral würde die Alternative gewählt werden, bei der die Anzahl der Opfer minimiert würde. Um das Problem genauer zu definieren, haben die Wissenschaftler im Sinne einer experimentellen Ethik sechs Online-Umfragen durchgeführt. Versuchspersonen waren US-Amerikaner auf "Amazon Mechanical Turk", die für jeweils 25 Cent mitmachten [11]. Die Teilnehmer dieses Experiments mussten ihr Alter und das Geschlecht angeben, wie auch den Grad ihrer Begeisterung für autonome Fahrzeuge. Anschließend mussten die Probanden beantworten, ob sie die Frage verstanden haben. Daraufhin wurden zehn Prozent von der Auswertung ausgeschlossen. Durchweg waren die Versuchsteilnehmer mehrheitlich für die quantitative utilitaristische Lösung, dass es moralischer für autonome Fahrzeuge sei, die Mitfahrer zu opfern, wenn dadurch mehr Leben gerettet würden. In einer weiteren Umfrage gaben etwa 76 Prozent an, es sei moralischer einen Mitfahrer zu opfern als zehn Fußgänger. Ebenfalls wird bei einer weiteren Umfrage deutlich, dass bei einer steigenden Differenz zwischen den möglichen Opfergruppen, die Zustimmung zur utilitaristischen Lösung steigt. Selbst wenn ein Familienmitglied im Fahrzeug sitzt, ist noch eine knappe Mehrheit für das utilitaristische Opfern des Passagiers. Die Mehrheit wäre also für eine utilitaristische Programmierung, wobei nur 67 Prozent der Probanden glauben, dass die Fahrzeuge so programmiert würden. Es zeigt sich außerdem, dass die Menschen nicht sonderlich verängstigt wären, wenn Fahrzeuge mit einer utilitaristischen Ethik programmiert wären [11]. Bei einer weiteren Umfrage wird ersichtlich, dass weniger Opfer bevorzugt werden, solange es nicht die Angehörigen oder einen selbst trifft. Hierbei mussten die Teilnehmer auf einer Skala von 1-100 angeben, wie wahrscheinlich es wäre, dass sie ein Fahrzeug mit programmierter Opferminimierung beziehungsweise eines kaufen würden, dass vornehmlich die Mitfahrer schützt, auch wenn 10 oder 20 Menschen deswegen ums Leben kommen. So sehr die Befragten moralisch utilitaristisch entscheidende Fahrzeuge schätzen, so wenig wären sie bereit, so ein Fahrzeug zu kaufen und damit im Fall eines Dilemmas ihr eigenes Leben oder das eines Familienangehörigen für die Moral zu opfern. Für die Wissenschaftler ist das ein klassischer Fall des sozialen Dilemmas, wo jeder dazu neigt, die eigenen Kosten zu reduzieren, anstatt sich so zu verhalten, dass es zu einem optimalen Ergebnis für alle kommt. So würden die Menschen eben

utilitaristische, sich selbst und die Mitfahrer opfernde Algorithmen in autonomen Fahrzeugen begrüßen, ohne selbst ein solches Fahrzeug kaufen zu wollen. Auch wurde gefragt, ob die Teilnehmer solche autonomen Fahrzeuge kaufen würden, deren Algorithmen von der Regierung reguliert werden. Die Teilnehmer, die sich für das utilitaristische Opfern aussprachen, waren eher skeptisch, wenn dieses moralische Prinzip vom Staat verordnet werden würde. [11]. Insgesamt wären nur 40 Prozent damit einverstanden, um die 30 Prozent würden auch solche Pkws kaufen, wobei mehr als die Hälfte lieber solche kaufen würden, die nicht staatlich reguliert sind [11]. Die Wissenschaftler ziehen aus der Studie heraus, dass zwar gesetzliche Regulierungen notwendig sein könnten, um das Problem des sozialen Dilemmas zu lösen, sie wären aber eher unbeliebt in der Bevölkerung [11]. Die Mehrheit würde eine gesetzliche Vorschrift zur Einführung utilitaristisch entscheidender autonomer Fahrzeuge ablehnen und eine solche Regelung könne auch die Akzeptanz für autonome Pkws entscheidend verringern. Es lassen sich dennoch andere Alternativen beschreiben, die die Grundprinzipien des Utilitarismus wahren. Beispielsweise die persönlichen Daten von Krankenkassen sind öffentlich gespeichert, dass Pkwsysteme darauf zugreifen und die Daten nach bestimmten Mustern, zum Beispiel statistischer Lebenserwartung, auswerten könnten. Das würde ermöglichen, dass das Fahrzeug auf dieser Basis Entscheidungen treffen könnte.

3.2 Deontologie

Bei der Deontologie ist das Grundprinzip, dass die Software in einem autonomen Kraftfahrzeug so programmiert sein sollte, dass ein bestimmter, vorher festgelegter Regelsatz umgesetzt wird. Dadurch wird der moralische Status einer Handlung nicht nur anhand ihrer Konsequenzen bestimmt. Ein zu nennendes Beispiel wäre, dass Kinder die jünger als zehn Jahre alt sind, nicht überfahren werden dürften. Diese Regelsätze lassen sich technologisch in der Maschine umsetzen. Auch in einem deontologischen Ansatz könnte die Software, je nach Kultur, unterschiedlich moralische Vorstellungen besitzen.

4 Erster tödlicher Unfall

Ein autonomes Fahrzeug des 3 Levels (Beta Level 4) von Volvo, das mit der Software Uber ausgerüstet war, ist der erste bekannt gewordene Unfall mit einem autonomen Auto, bei dem eine Fußgängerin gestorben ist [12]. Eine 49-jährige Frau soll am 18. März 2018 gegen 22 Uhr in Tempe im US-Bundesstaat Arizona außerhalb eines Fußgängerübergangs die Fahrbahn überquert haben. Anschließend soll die Passantin direkt aus dem Schatten auf die Fahrbahn getreten sein. Daraufhin hat sie ein autonomes Pkws des 3-4 Levels erfasst. Die örtliche Polizei berichtet, dass dieser Zusammenstoß in jedem Modus, ob autonom oder manuell, schwer zu verhindern gewesen wäre [12]. Daraufhin hat der US Amerikanische Softwarehersteller Uber alle Testfahrten mit selbstfahrenden Pkws gestoppt [12].

5 Ausblick

Für die moralischen Dilemmata scheinen keine befriedigende Lösungen in Sicht zu sein. Es kann keine Lösungen von autonomen Pkws geben zu Situationen, die auch für Menschen sehr schwierig und komplex zu lösen sind. Die Rechenmaschinen können mindestens nur so gut handeln wie wir Menschen. Je mehr autonome Pkws auf den Straßen unterwegs sind, desto dringender brauchen wir klare Regeln und Antworten. Dies wird vorrausichtlich mit der rasanten Entwicklung der Künstlichen Intelligenz auch geschehen müssen. Es ist durchaus nicht einfach Algorithmen zu entwickeln, die moralische Werte und das Selbstinteresse ausgleichen, zumal dies in verschiedenen Kulturen anders ausfallen könnte, wie im Punkt zwei schon beschrieben. Mit den autonomen Pkws könnten sicher viele Unfälle vermieden werden, aber definitiv nicht alle und es werden neue Unfallursachen auftreten, für die es wenn es soweit ist andere Lösungen bedarf. Gegebenfalls muss die Infrastruktur den Gegebenheiten angepasst werden, auch ist die Interaktion mit der Umwelt wichtig für autonome Pkws, um Informationen nach außen ersichtlich zu machen. Auch könnten autonome Pkws nur in Gebieten unterwegs sein, wo Menschen sich nicht aufhalten, wie es mit Zügen der Fall ist, dass würde die Probleme moralischen Gedankenexperimente von alleine lösen. Flugzeuge sind auch oft autonom unterwegs, dennoch kann ein Unfall passieren. Hierfür gibt es die Blackbox, die alle wichtigen Metha-Informationen aufzeichnet. Dieses Modell könnte auch für autonome Pkws einsetzen werden. Eins wird wahrscheinlich klar, es werden sich die autonomen Pkws durchsetzen, die am sichersten sind, wo auch die Kosten einhergehen, denn die Versicherungsunternehmen werden die Kunden zur Kasse beten, die ein Fahrzeug besitzen von Herstellern mit eher nicht zuverlässiger Software. Man könnte auch die Gefahr minimieren, indem jedes Softwareupdate der Hersteller von autonomen Pkws geprüft wird. Erst nach bestandener Prüfung darf der Pkw auf die Straße. Eine Art TÜV Prüfung oder Führerscheinprüfung, also für autonome Pkws. Da Machine Learning Modelle nur die Mehrheit der Antworten widerspiegeln und dementsprechend reagieren, lässt sich bestreiten, ob es sich um eine zufriedenstellende Moral handelt. Eins muss aber klar sein, es stellt sich nicht die Frage, ob autonome Pkws Einzug in unser Leben finden werden, sondern eher wann.

Literatur

1. autonomes-fahren.de *Geschichte des Autonomen Fahrens* <https://www.autonomes-fahren.de/geschichte-des-autonomen-fahrens/>, Abruf: Juli 2018
2. Statista *SPrognose der Anzahl zugelassener Personenkraftwagen (Pkw) in Europa nach Art der Fahrzeugnutzung im Zeitraum der Jahre 2018 bis 2030* <https://de.statista.com/statistik/daten/studie/875082/umfrage/prognostizierter-pkw-bestand-in-europa-nach-art-der-pkw-nutzung/>, Abruf: Juli 2018
3. Statista *Statistisches Bundesamt. n.d. Anzahl der Verkehrstoten im Straßenverkehr in Deutschland von 1991 bis 2017* <https://de.statista.com/statistik/daten/studie/185/umfrage/todesfaelle-im-strassenverkehr/>, Abruf: Juli 2018
4. Frankfurter Allgemeine *Roboterautos ersparen Deutschland Milliarden* <http://www.faz.net/aktuell/wirtschaft/mehr-wirtschaft/dihk-selbstfahrende-autos-bringen-deutschland-milliardeneinsparungen-15691942.html>, Abruf: Juli 2018
5. David Fickling *Self-Driving Cars Need to Slow Down After Uber Crash* <https://www.bloomberg.com/gadfly/articles/2018-03-20/uber-crash-shows-need-for-collaboration-in-self-driving-cars>, Abruf: Juli 2018
6. Handelsblatt jur *Das sind die fünf Stufen des autonomen Fahrens* <https://www.handelsblatt.com/unternehmen/industrie/technik-hinter-dem-uber-unfall-das-sind-die-fuenf-stufen-des-autonomen-fahrens/21100770.html?ticket=ST-4599069-hVSfzUxF3TFWJvolUoz6-ap6>, Abruf: Juli 2018
7. Adrian Lobe *Wie die Moral in die Maschine kommt* <https://www.spektrum.de/kolumne/wie-die-moral-in-die-maschine-kommt/1492775>, Abruf: Juli 2018
8. Leon R. Sütfeld, Richard Gast, Peter König and Gordon Pipa *Using Virtual Reality to Assess Ethical Decisions in Road Traffic Scenarios: Applicability of Value-of-Life-Based Models and Influences of Time Pressure* <https://www.frontiersin.org/articles/10.3389/fnbeh.2017.00122/full>, Abruf: Juli 2018
9. Josephine Lütke *Autonomes Fahren: Computer mit menschlicher Moral?* <https://bit.ly/2Lf2cDw>, Abruf: Juli 2018
10. Friedhelm Greis *Selbstfahrende Autos sollen zuerst Insassen retten* <https://www.golem.de/news/daimler-sicherheitschef-selbstfahrende-autos-sollen-zuerst-insassen-retten-1610-123736.html>, Abruf: Juli 2018
11. Florian Rötzer *Moralisches Dilemma für autonome Fahrzeuge und ihre Nutzer* <https://www.heise.de/tp/features/Moralisches-Dilemma-fuer-autonome-Fahrzeuge-und-ihre-Nutzer-3380341.html>, Abruf: Juli 2018
12. Andreas Wilkens *Polizei zum ersten Todesfall mit autonomem Auto: Unfall war schwer zu verhindern* <https://www.heise.de/newsticker/meldung/Polizei-zum-ersten-Todesfall-mit-autonodem-Auto-Unfall-war-schwer-zu-verhindern-3999229.html>, Abruf: Juli 2018