

Was passiert bei Unfällen mit selbst fahrenden Autos

Seminar: Ist Künstliche Inteligenz gefährlich?

Jochen Clormann

Universität Heidelberg

18. Mai 2017

Agenda

- 1 Einleitung
- 2 Safe State
- 3 Kostenfunktion
- 4 Consequence Engine
- 5 Studie
- 6 Zusammenfassung & Einschätzung

Einleitung



Abbildung: Autonomer Tesla 2016 [7]

- Breites Forschungsgebiet von Industrie und Wissenschaft mit großen Zielen:
- + Verringert Anzahl an Unfällen
- + Erhöht die Mobilität
- + Verringert Umweltbelastung
- ± Weitere zahlreiche Möglichkeiten

Google WAYMO



Abbildung: Waymo by Google [8]

- *Early Rider* Programm in Phoenix, AZ
 - Studie: 94 % aller Unfälle durch menschliches Versagen
 - Ziel: Verringerung Verkehrsunfälle
 - 3M Meilen autonom gefahren
 - Lernfähige Autos mit Sicherheitsfahrer
 - Frequenz der „Eingriffe“ durch Fahrer
- 2015: 1.250 Meilen
2016: 5.000 Meilen

[6]

Google WAYMO



Abbildung: Google Waymo [10]

Einleitung



Abbildung: Unvorsichtiger Fußgänger [9]

- Nicht alle Verkehrsteilnehmer sind autonom
- Unvorhersehbare Ereignisse
- Gefahrensituationen
- Entscheidung zwischen:
 - Verkehrsregeln brechen
 - Sachschaden verursachen
 - Menschen verletzen

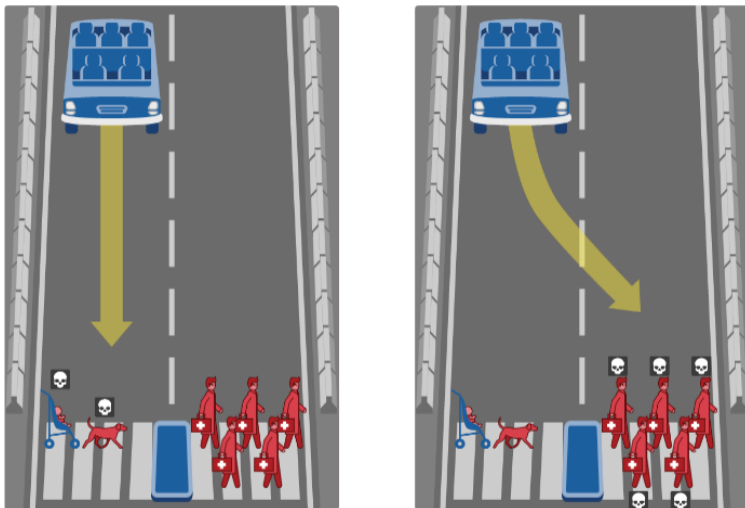


Abbildung: Töte Kleinkind und Hund **oder** 5 Ärzte [11]

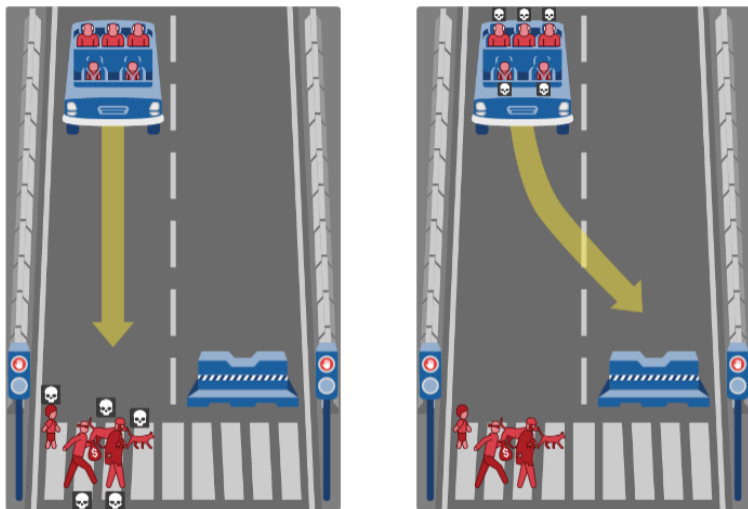


Abbildung: Töte Kind, Verbrecher, Obdachloser & 2 Tiere **oder** 2 Kleinkinder & 3 Erwachsene [11]

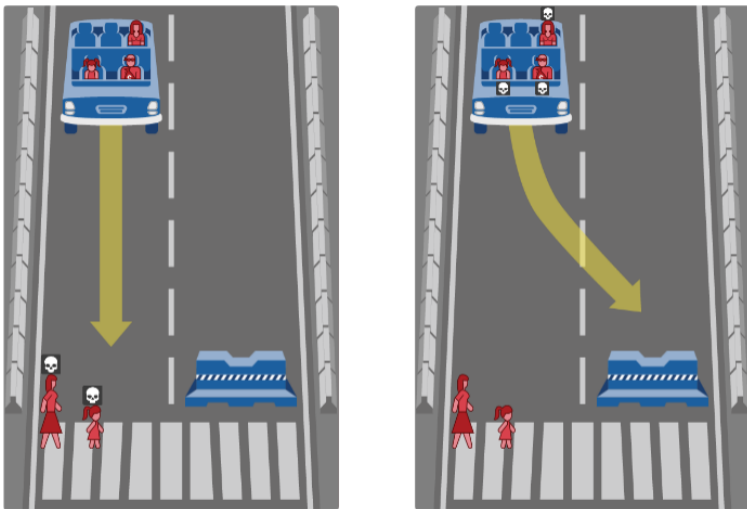


Abbildung: Töte Frau mit Kind **oder** Verbrecher mit Geiseln (Frau & Kind) [11]

Leitsatz für Gefahrensituationen

*...Alle Personen gut zu behandeln und **keine Diskriminierung** anhand von Rasse, Religion, Geschlecht, Behinderung, Alter, Nationalität, Sexualität, Geschlechtsidentität oder Geschlechtsausdruck **zu vollziehen**...*

Übersetzt aus [3]

Safe State

Definition

Das Verhüten von Unfällen darf nicht als eine Vorschrift des Gesetzes aufgefasst werden, sondern als ein Gebot menschlicher Verpflichtung und wirtschaftlicher Vernunft.

(Werner von Siemens, 1880)

- Sicherheitsnorm ISO 26262:2011
- Safe State ist erreicht wenn das eingehende und ausgehende Risiko im aktuellen und zukünftigen Zustand unter einer gewissen Grenze bleibt.

Risiko als Funktion

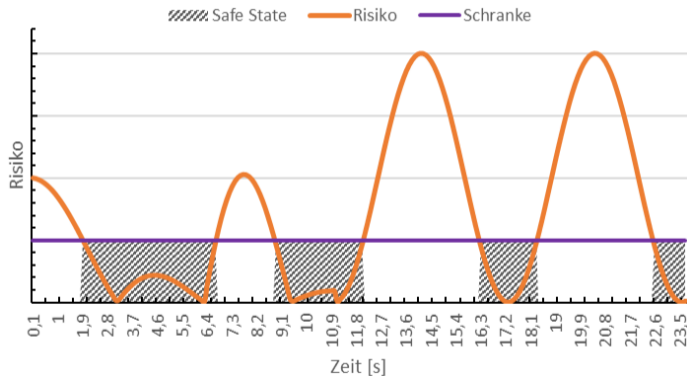


Abbildung: Zeit-Risiko Diagramm

Beinhaltet:

- Alle Objekte der Umgebung
- Bewegungsrichtungen
- Regeln & Gesetze
- Eigene „Mission“
- Verfügbare Leistung

Performance



Risk Level



Functional Limits



Safe State



[12, 13, 14, 22]

Projekt LEONIE



Abbildung: LEONIE auf öffentlicher Straße [23]

- Gut dokumentiertes Projekt der TU Braunschweig
- Radar und Lasertechnik
- Im Notfall übergabe an den Sicherheitsfahrer
- Forschungsbereiche:
 - Umgebungsüberwachung
 - (IT) Leistung
 - Wirkung

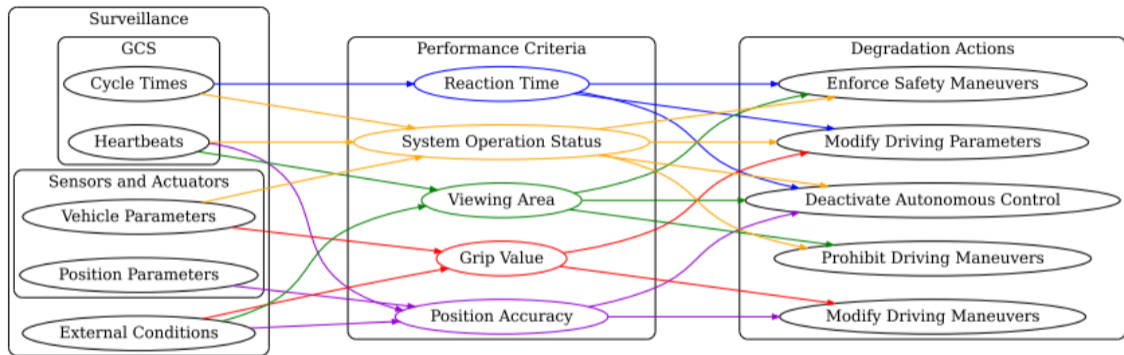


Abbildung: Beziehung zwischen Überwachung, Verarbeitung und Wirkung [5]

Kostenfunktion

Definition

Kostenfunktionen für **verschiedene Szenarien** betrachtet:

- Handlungswirkungen
 - „Eigener“ Schaden
 - Sachschaden
 - Tod / Verletzung
- Risiken
 - Unvorhergesehene Auswirkungen
- Eigene „Fähigkeiten“
 - „Bestes“ Szenario wählbar?
 - Unerreichbares „bestes“ Ergebnis
 - Erreichbares „schlechtestes“ Ergebnis

Anwendung bei Pfadabweichung

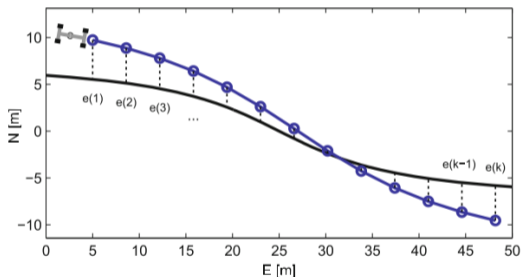


Abbildung: Aktueller und optimaler Pfad [4]

- Berechnung über SSE:

$$J = C_1 \sum_{i=1}^N e(i)^2 \quad (1)$$

- Entspricht der Pfadabweichung
- Einbeziehung weiterer Faktoren:

$$J = C_1 \sum_{i=1}^N e(i)^2 + \underbrace{C_2}_{\text{Gewicht}} \sum_{j=1}^{N-1} \underbrace{|\delta(j+1) - \delta(j)|}_{\text{Lenkwinkel}} \quad (2)$$

Weitere (externe) Parameter:

Parameter	Gewicht in Gefahr	Normalgewicht
Verkehrsregeln	gering ↔ hoch	hoch
Zeit zum Ziel	gering	hoch
Distanz zu anderen	gering	hoch
Anzahl Verletzte	hoch	hoch
Anzahl Tote	hoch	hoch

Tabelle: Fahrparameter mit Gewichtung

⇒ $C(\text{Unbekannte Parameter}) \gg C(\text{Bekannte Parameter})$

Ziel: Verringerung des Gesamtschadens

Verringerung des Gesamtschadens heißt:

- Minimierung der Kostenfunktion
- durch Anpassung der Gewichte
- Resultate:
 - Selbstopferung
 - Auswählen und „Anvisieren“ des robustesten Ziels
- 2 Beispiele..

Gesamtschaden verringern



Abbildung: Kradfahrer mit Schutzausrüstung[16]



Abbildung: Kradfahrer ohne Schutzausrüstung[15]



Abbildung: Situation: Plötzlich entgegenkommendes Fahrzeug [17]

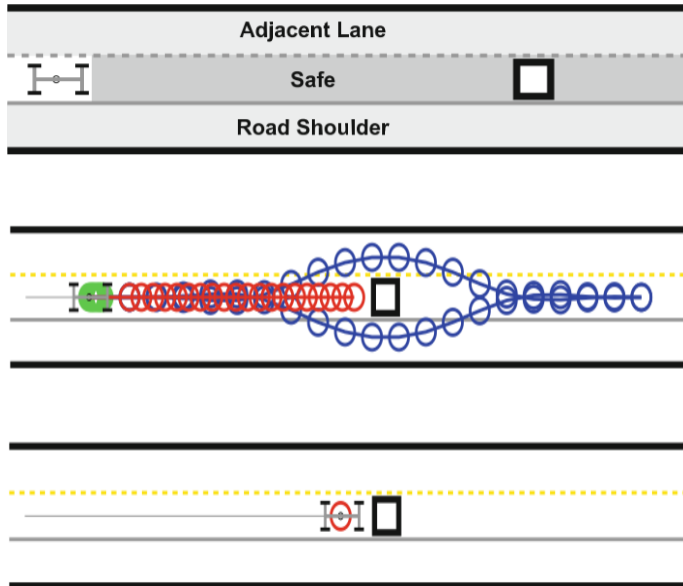
Möglichkeiten:

- Auffahrunfall mit Verletzten ggf. Toten
- Selbstopferung

Einschränkungen:

- Kein Verstoß gegen Verkehrsregeln
- Außer Personen sind in Gefahr

[4]



Zusammenfassung - Kostenfunktion

- Gewichtung verschiedener Parameter
- Abweichung zwischen optimalem und aktuellem Wert
- Fähigkeit „Ziele“ zu bestimmen
- Selbstopferung als Option
- Deadlock Problem bei zu strengen Einschränkungen

Consequence Engine

Definition

- Steuerungseinheit von autonomen Maschinen
- Bewertet einzelne Ereignisse mit numerischen Werten
- Klassifiziert nach ethisch, programmierten Prinzipien

```
if      (for all robot actions ,  
        the human is equally safe)  
then  
    output safe actions  
else  
    output action(s) for least  
    unsafe human outcome(s)
```

Listing 1: Basis Algorithmus

[2]

Algorithmus & Logik

```
:name: ethical_g  
:agent: robot
```

:Outcome Scores:

```
safe = 0  
collision = 4  
hole = 10
```

:Ethical Precedence:

```
human > robot
```

Listing 2: Schadens und Ethik Definition

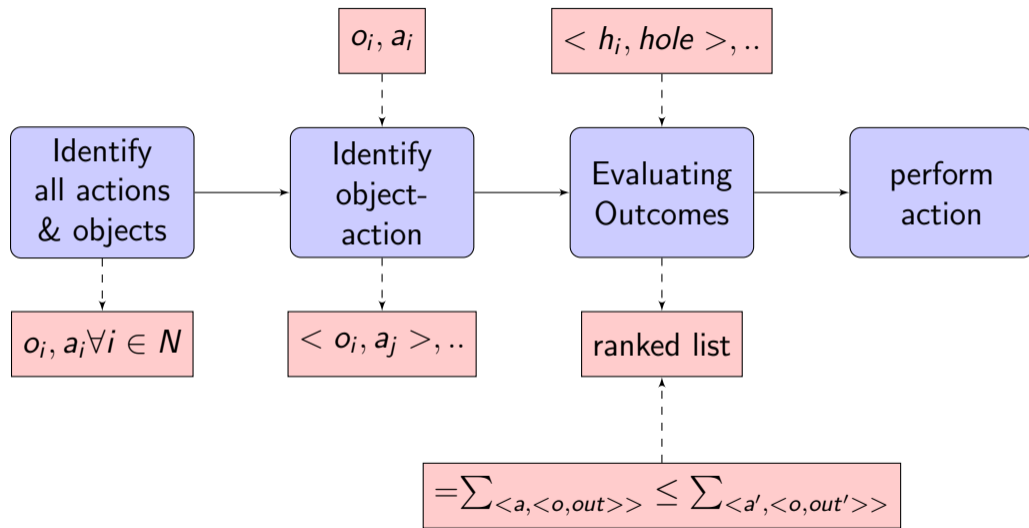
Implementierbares Modell

Gegeben:

- Objekte
 - Landschaft
 - Menschen
 - Gefahrenquellen
- Bewegungsrichtungen

Gesucht:

- Meist gefährdetes Objekt (MgO)
- Aktion um MgO zu beschützen
- „Beste“ Aktion



Versuch mit Mini Robotern

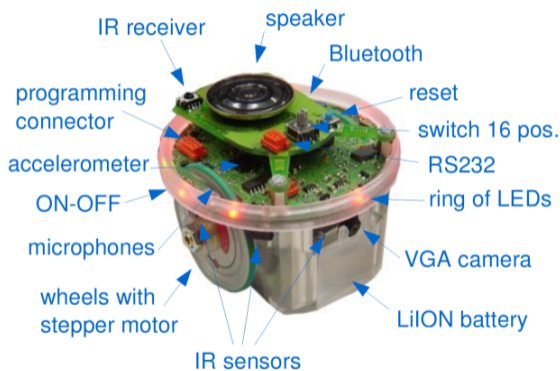
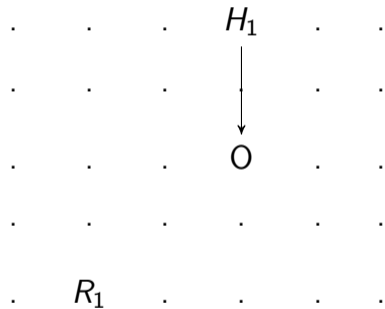


Abbildung: e-pukt Roboter [19]

Versuchsaufbau:

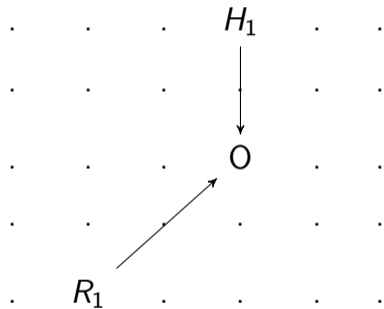
- Mensch bewegt sich auf Gefahrenstelle zu
- Roboter soll dies erkennen
- Roboter soll Menschen beschützen
- ggf. sich selbst der Gefahr aussetzen

Versuchsdurchführung



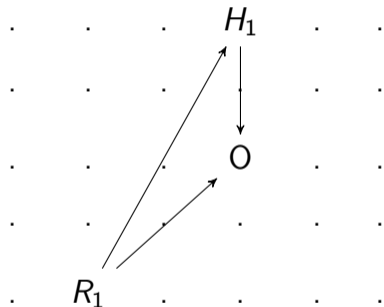
- 1 Mensch H_1 läuft auf Grube O zu

Versuchsdurchführung



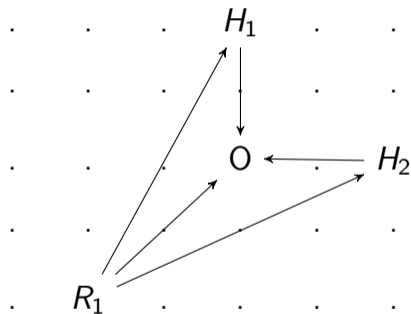
- 1 Mensch H_1 läuft auf Grube O zu
- 2 Roboter R_1 kann sich selbst in die Grube fahren

Versuchsdurchführung



- 1 Mensch H_1 läuft auf Grube O zu
- 2 Roboter R_1 kann sich selbst in die Grube fahren
- 3 Roboter R_1 kann Mensch H_1 aufhalten

Versuchsdurchführung II



- 1 Mensch H_1, H_2 läuft auf Grube O zu
- 2 Roboter R_1 kann sich selbst in die Grube fahren
- 3 Roboter R_1 kann Mensch H_1 oder H_2 aufhalten
- 4 Roboter muss sich zwischen zwei Menschen entscheiden

Ergebnisse

Ein „Mensch“:

- Mensch konnte immer gerettet werden

Zwei „Menschen“ :

- 9% beide Menschen gerettet
- 59% min. Ein Mensch gerettet
- 32% kein Mensch gerettet

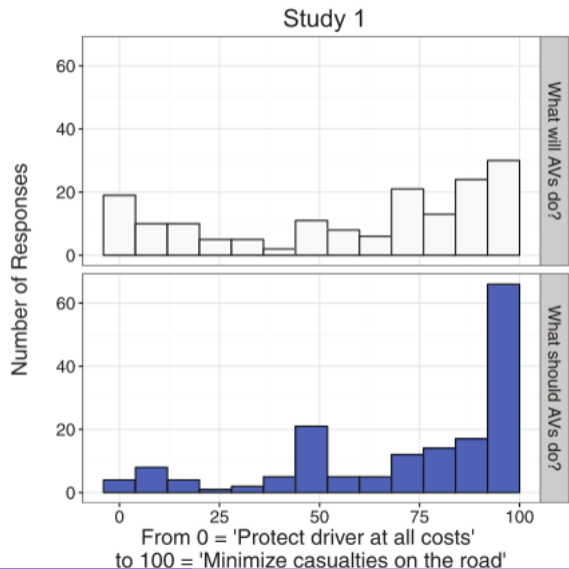
Einschätzung:

32% im Labor zu viel

Studie

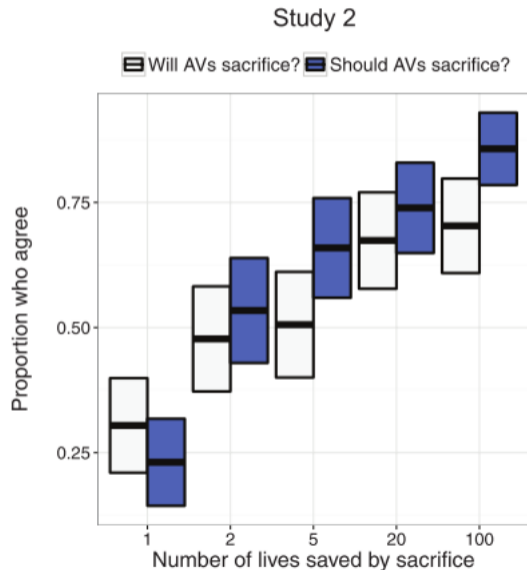
Studie 1:

- n=182
- Vergleich zwischen:
 - Wie werden AF programmiert
 - Wie sollten AF programmiert werden



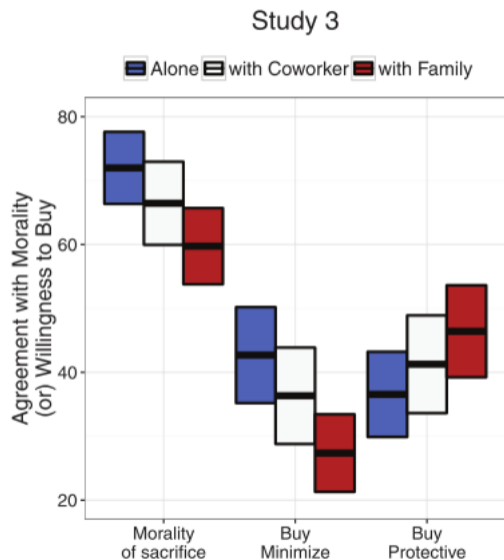
Studie 2:

- n=451
- Anzahl der Personen für die Fahrzeuginsassen geopfert werden sollten



Studie 3

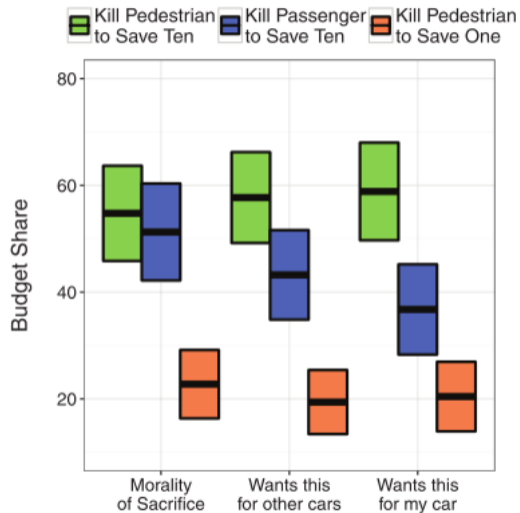
- n = 259
- Gerechte Selbstopferung
- Kaufe ein Auto das:
 - Schaden minimiert
 - Den Fahrer beschützt
- ..unter Berücksichtigung der Beziehung zu Passagieren



Studie 4:

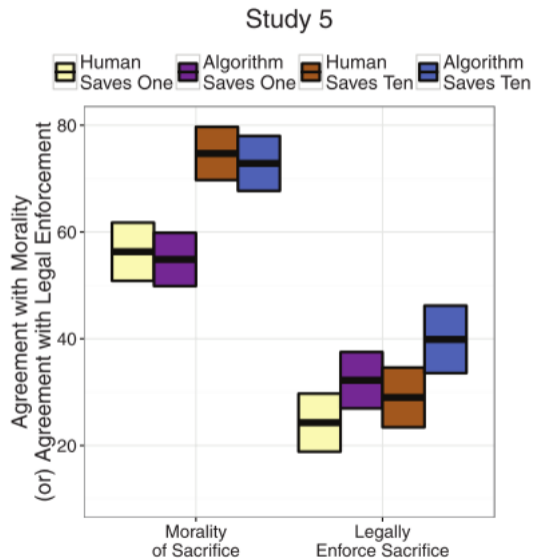
- n = 267
- Verteilung von 100 Budget Points auf:
 - ① Töte 1 Fußgänger, verschone 10 Fußgänger
 - ② Töte Passagier, verschone 10 Fußgänger
 - ③ Töte 1 Fußgänger, verschone 1 Fußgänger
- In den Bereichen:
 - Gerechte Selbstopferung
 - Für andere Autos
 - Für mein Auto

Study 4



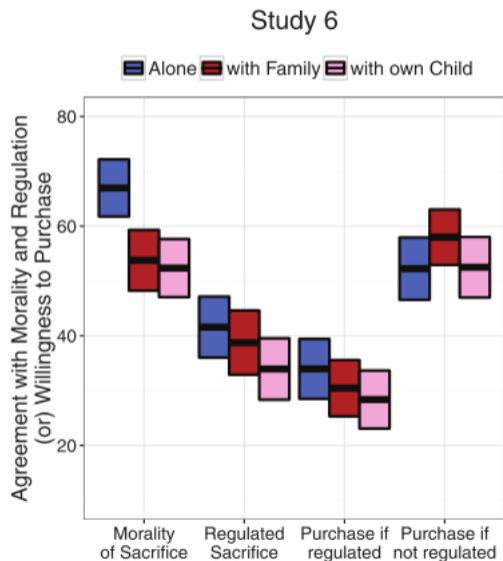
Studie 5:

- $n = 376$
- Vergleich zwischen:
 - Mensch opfert sich, verschont 1
 - Algorithmus opfert Passagier für 1
 - Mensch opfert sich, verschont 10
 - Algorithmus opfert Passagier für 10
- Gerechte Entscheidung
- Gesetzlich verpflichtend



Studie 6

- n = 393
- Kaufbereitschaft bei gesetzlich kontrollierten Algorithmen
- Fahren mit Familie oder Kind



Zusammenfassung Studienergebnisse

- Gesetzliche Regulierung nötig aber auch kontraproduktiv
- Zustimmung zur Minimierung des Gesamtschadens
- Abhängigkeit der Beifahrer Beziehung
- geringe Kaufbereitschaft für Selbst opfernde Autos

Zusammenfassung & Einschätzung

Zusammenfassung

- Safe State
 - Definition von Sicherheitsbedingungen
 - Sicherheitsanforderungen in Leistung, Risiko und Funktionalität
- Kostenfunktion
 - Berücksichtigung von Handlungsauswirkungen und Risiken
 - Selbstopferung
 - Deadlock
- Consequence Engine
 - Betrachtung einzelner Aktionen
- Studie
 - Befragte sind „einverstanden“ mit Selbstopferung
 - geringe Kaufbereitschaft

Einschätzung

Unfallursachen durch Fehlverhalten:

Ursache	Anzahl
Alkoholeinfluss	12.660
Falsche Straßennutzung	24.763
Geschwindigkeit	47.024
Abstand	50.667
Überholfehler	13.445
Vorfahrt & Manöverfehler	110.336
Fehlverhalten geg. Fußgänger	16.629
Σ	275.524

Weitere Unfallursachen:

Ursache	Anzahl
Glätte durch Regen	6.362
Glätte durch Schnee, Eis	4.974
Sichtbehinderung (Nebel)	456
Sichtbehinderung (Sonne)	3.116
Ungesicherte Baustellen	125
Wild	2.580
Tiere auf der Fahrbahn	1.076
Σ	18.689

Unfälle mit Todesfolge: 3.459

Entspricht ca. 792 Unfälle mit Personenschaden & 9 Tote pro Tag

[20, 21]

Einschätzung

Autonome Fahrzeuge können:

- Keine Fehler machen die Personen machen
- In Gefahrensituationen den Überblick behalten

Autonome Fahrzeuge können (noch) nicht:

- Entscheidungen treffen die..
 - ..moralisch vertretbar sind,
 - gesellschaftlich anerkannt sind,
 - keinen Autobauer benachteiligen.

Außerdem..

- Wer ist der Angeklagte??
- Wer sichert wie gegen Hacker??
- Wettbewerbsvorteil durch bessere Technik??

Quellen I

- [1] Bonnefon, J.-F., Shariff, A., & Rahwan, I. (2016). **The social dilemma of autonomous vehicles**. In *Science*, 352(6293), 1573–1576.
<http://doi.org/10.1126/science.aaf2654>
- [2] Dennis, L. A., Fisher, M., & Winfield, A. F. T. (2015). **Towards Verifiably Ethical Robot Behaviour**. In *CoRR*, abs/1504.0, 1–11. Retrieved from
<http://arxiv.org/abs/1504.03592>
- [3] Lin, P. (2016). **Why Ethics Matters for Autonomous Cars**. In *Autonomous Driving* (pp. 69–85). Berlin, Heidelberg: Springer Berlin Heidelberg.
http://doi.org/10.1007/978-3-662-48847-8_4

Quellen II

- [4] Gerdes, J. C., Thornton, S. M., Gerdes, J. C., Thornton, Á. S. M., & Thornton, S. M. (n.d.). **5 Implementable Ethics for Autonomous Vehicles**.
http://doi.org/10.1007/978-3-662-48847-8_5
- [5] Reschka, A., Böhmer, J. R., Nothdurft, T., Hecker, P., Lichte, B., & Maurer, M. (2012). **A surveillance and safety system based on performance criteria and functional degradation for an autonomous vehicle**. In *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, pp. 237–242.
<http://doi.org/10.1109/ITSC.2012.6338682>
- [6] <https://waymo.com/>
- [7] https://en.wikipedia.org/wiki/Autonomous_car#/media/File:Tesla_Model_S_Hwy_40.jpg

Quellen III

- [8] <https://www.google.com/selfdrivingcar/>
- [9] <http://blog.audi.de/2017/03/24/audi-mit-lte-v-auf-der-cebit/>
- [10] <https://www.youtube.com/watch?v=fbWeKhAPMig>
- [11] <http://moralmachine.mit.edu/hl/de>
- [12] [http://imagenes.km77.com/fotos/bbtcontent/clipping/
KM7KPH20090616_0113/3.jpg](http://imagenes.km77.com/fotos/bbtcontent/clipping/KM7KPH20090616_0113/3.jpg)
- [13] <http://www.gebrauchtwagen.net/auto-lexikon/seitenstreifen>
- [14] [https://www.automobil-produktion.de/maerkte/
bosch-ist-teil-der-global-ncap-initiative-106.html](https://www.automobil-produktion.de/maerkte/bosch-ist-teil-der-global-ncap-initiative-106.html)
- [15] <http://www.tagesspiegel.de/images/roller/11576590/3-format43.jpg>

Quellen IV

- [16] <http://cdn.salzburg.com/nachrichten/uploads/pics/2016-05/orginal/bischofshofen-motorradfahrer-bei-unfall-getoetet-41-64401963.jpg>
- [17] <https://www.mottekundbommel.de/wp-content/uploads/2016/02/Mittagspause-mit-den-Kollegen.gif>
- [18] <http://www.activistpost.com/wp-content/uploads/2014/08/0825.sdt-news.jpg>
- [19] https://www.generationrobots.com/img/cms/Epuck_devices.png
- [20] <https://www.destatis.de/DE/ZahlenFakten/Wirtschaftsbereiche/TransportVerkehr/Verkehrsunfaelle/Tabellen/FehlverhaltenFahrzeugfuehrer.html>

Quellen V

- [21] <https://www.destatis.de/DE/ZahlenFakten/Wirtschaftsbereiche/TransportVerkehr/Verkehrsunfaelle/Tabellen/UrsachenPersonenschaden2.html>
- [22] http://www.tagesspiegel.de/images/verkehr_dpa/6751762/3-format6001.jpg
- [23] https://www.tu-braunschweig.de/Medien-DB/presse/leonie/leonie3_tu-braunschweig.jpg