

Ethik und KI

Beispiel autonomes Fahren

Cebit 2018

Michael Mörke

Integrata-Stiftung für humane Nutzung der IT

Was ist autonomes Fahren?

Fünf Stufen:

0. Fahrer macht alles selbst:
beschleunigen, bremsen, lenken, Umfeld beachten
1. Assistiertes Fahren:
Spurhalte-Assistent, Berganfahrhilfe,
2. Teilautomatisiert:
Einparkhilfe, Stauassistent, ...
3. Hochautomatisiert:
Autopilot, Spurwechsel, Verfolgung,
Geschwindigkeitsanpassung an Verkehrssituation
4. Vollautomatisiert:
Auto macht alles, nur nicht ungelernete Situationen
5. Vollständig autonom, fahrerlos
Nur Ziel und Startfreigabe werden extern vorgegeben



Chancen und Risiken

Chancen:

1. Globalisierung von Verkehrsregeln
2. Weniger Unfälle, weniger Verkehrstote
3. Andere Ziele der Aufmerksamkeit beim Fahren
4. Autos werden „geteilt / shared“ (kommunal)
5. Weniger Autos, weniger Parkplatzbedarf
6. LKW's fahren dichter, weniger Strassenplatzbedarf
7. Verbesserte Mobilität, auch in kleinen Kommunen
Bürgerbus durch autonome Autos ersetzen

Risiken:

1. Werden Lastwagenfahrer arbeitslos?
2. Wer prüft, ob alle Situationen beherrscht werden?
3. Unerwartete Unfälle in unerwarteten Situationen
4. Wer haftet bei Unfällen?

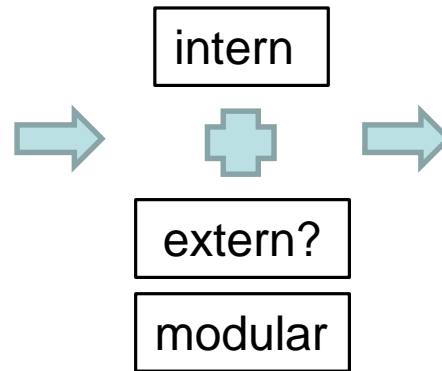
Wie funktioniert das?



Sensoren:

Kamera
Mikrofon
Abstandsradar
GPS
Geschwindigkeit
Querkräfte
Helligkeit
Belastung
Tankfüllstand
Bedientablet
Etc...

KI:



Aktoren:

Bremse
Gas
Lenkrad
Ansage
Lichtschalter
Blinker
Hupe
Scheibenwischer
Klimaanlage

Dilemma - 1

1. Situation:

Auto fährt auf ein Hindernis zu und hat nur die Wahl, entweder links die Oma oder rechts das Kind totzufahren.

2. Beurteilung:

Das ist Quatsch! (oder Spielzeug für Philosophen)

3. Begründung:

Ein autonomes Auto fährt nie ungebremst auf eine Wand zu. Nur unerwartete Ereignisse könnten es überfordern.

Beispiel: Felsbrocken stürzt plötzlich auf die Strasse. Wie oft?

4. Menschliche Ethik:

Leben der Insassen geht in plötzlichen Situationen vor!
Rest ist egal. War dann eben ein Unfall.

Dilemma - 2

1. Situation:
Auto steht vor Ampel und bemerkt hinter sich einen LKW,
der nicht mehr rechtzeitig bremsen kann.
Soll es nach vorne ausweichen und die Fußgänger überfahren?
Oder rechts auf den Radweg und dort Radfahrer gefährden?
2. Auch Quatsch!
3. Begründung:
Erst können LKW's autonom fahren. Dann siehe Fall 1.
Im gemischten Verkehr kommen solche Unfälle eben vor.
Wenn kein Fußgänger oder kein Radfahrer -> Bitte!
4. Ethik:
Unbeteiligte dürfen nicht geschädigt werden.
Es darf nicht Leben gegen Leben aufgerechnet werden
5. Beispiel (9/11): Auch ein von Terroristen entführtes Flugzeug
darf nicht abgeschossen werden!

Dilemma - 3

1. Situation:

Auto fährt Bergstrasse hinauf und vor ihm ein Radfahrer, der hin- und her-geigt, also zu wenig Platz zum Überholen. Soll es weiße Mittellinie überfahren an ungefährlicher Stelle?

2. Mensch würde eine Minute warten und dann genau das tun. Autonomes Auto wird es auch tun und Gesetz übertreten. Genau wie der Mensch!

3. Es hat also mindestens 2 Arten „Fahrstile“.
- mehr oder weniger risikoreich

4. Wie viele Fahrstile hat es wirklich?

Kann es (legale oder illegale) Autorennen machen?

Wer stellt den Fahrstil ein? Wie stellt man den ein?

Wie übt man einen Fahrstil mit seinem Auto?

Wer nimmt die Fahrprüfung ab? (Fahrlehrer?)

Wer übernimmt die Verantwortung? (Hersteller, Trainer? Halter?)

Ethik und Technik

Klassisch: Ethik oder Moralphilosophie:

Anwendung von Wertvorstellungen bei menschlichem Handeln

Traditionell: in Wirtschaft und in Technologie:

Ethischer Einsatz und Umgang mit Werkzeugen

Bewertung von technischem Handeln und Folgen des Einsatzes



Neu bei KI ist die Aufteilung in:

➤ Äußere Ethik (Technikethik):

Anwendung und Einsatz der Roboter durch Menschen

➤ Innere Ethik (Maschinenmoral):

Eingebautes Verhalten der KI gegen Menschen

Wir müssen uns mit beiden Formen von Ethik beschäftigen!

Neu: Wir müssen eine passende Ethik einbauen!

Ethik in KI einbauen

KI: Maschinen lernen,

1. Objekte und Situationen zu erkennen und
2. sich zu verhalten.

Zum Erkennen braucht es keine innere Ethik
– aber zum richtigen Handeln.

Ethik wird nur zum Handeln gegen Menschen benötigt.

Beim Lernen werden Lernziele vorgegeben.

Ein autonomes Auto verfolgt viele Ziele gleichzeitig.

Wenn das Auto sich moralisch richtig verhält,
hat es ethische Werte als Zielvorgabe gelernt.

Ethik in KI einbauen

Ethische Ziele lassen sich als menschliche Werte in Hierarchien darstellen, die – je nach Situation – wichtiger oder weniger wichtig sind.

(ähnlich wie Maslowsche Bedürfnispyramide)

KI muss Handlungsoptionen an Werten messen.

und zwar situationsbedingt!



**Maschinen moralisches Handeln beibringen
= Ethik in KI einbauen =
menschliche Werte als Ziele vorgeben.**

Ethik und Mobilität

Technikethik - Beispiel gemischter Verkehr:
Sollen alle Fahrzeuge miteinander kommunizieren können?
Müssen wir auch in Fahrräder Ethik einbauen?

Maschinenmoral - welche Ethik sollen wir einbauen?

1. Utilitaristischer Ansatz? (Nutzen-orientiert)
2. Deontologischer Ansatz? (Prinzipien-orientiert, fuzzy)
3. Tugend-Ethik? (Qualitäts-orientiert)

Vermutlich von allem etwas?

Können uns die Philosophen dabei helfen?

Müssen wir Philosophen den Ingenieuren an die Seite stellen?

Wieviel Zeit haben wir, Maschinenmoral zu entwickeln?

Zum Glück kommen autonome Autos in 5 Stufen.



**Die Ethik dazu müssen wir aber alle und als Gesellschaft machen!
Wir sollten es nicht nur Google überlassen!**

Ethik in KI einbauen



Offen:

1. Kann man die Ethik in der KI in Maschinen kapseln?
Die KI müsste dann bei Handlungen gegen Menschen immer vorher beim Ethikmodul über ein Interface anfragen, ähnlich dem ethischen Gewissen eines Menschen.
2. Welche Informationen müsste das Interface bereitstellen?
Vermutlich sind das viele: Wie strukturieren?
3. Kann man die gekapselte Ethik zertifizieren?
Wie gehen wir dann mit dem Weiter- Lernen um?
4. Gibt es dann für jeden Kulturkreis eine eigene Kapsel?

Machen Sie mit!

Wecken Sie andere auf!
Diskutieren Sie mit Ihren Bekannten!
Wenn möglich: arbeiten Sie mit an Normen!

Melden Sie sich bei uns dafür:
info@integrata-stiftung.de