Ethical decomposition as a new method to analyse moral dilemmata

Findings on mad trolleys and self-driving cars

Christian Thielscher[1], Bianca Krol[2], Stefan Heinemann[3], Michael Schlander[4]

Abstract
Introduction: Since P. Foot studied the trolley problem in 1967, it has been extensively discussed in ethics, decision theory, medicine, and other disciplines. With the invention of autonomous vehicles, it has become an important and urgent practical question.
Methods: Three well-known, one new and one slightly changed versions of the trolley problem are arranged in a specific order. Kantian and utilitarian solutions to the problems are discussed. Respondents' decisions were empirically tested and aligned with ethical theories.
Results: Both Kantian and utilitarian ethics provide rules for decision. However, both are incomplete and differ from each other sometimes. In one case, both recommend to "act", in another, "not to act". In these two cases, almost all respondents follow the mutual advice. In other cases, ethical theories as well as responses differ.
Discussion: Respondents did not behave irrationally; rather, they considered ethical theories in a sensible way. Disaggregating Kantian and utilitarian decisions helps to identify situations in which autonomous vehicles can be programmed for better adaption to user preferences.

Keywords: Self-Driving Cars, Moral Dilemmata, Kantian Ethics, Utilitarian Ethics, Experimental Ethical Research.

1 Introduction

Since P. Foot studied the trolley problem in 1967 [Fo67], it has been extensively discussed in ethics, decision theory, medicine, and other disciplines [Th17]. In particular, the question of why people decide differently in the trolley case vs. the fat man case (see below) has received a lot of attention [Ed15]. With the invention of autonomous vehicles – especially level 5 (SAE International's standard J3016) autonomous cars – it has become an important and urgent practical question [BSR16].

In Germany, for example, the "Ethics Commission on Automated Driving" recently recommended (amongst others) [BM18]:

[1] FOM University of Applied Sciences, Competence Center for Medical Economics, Leimkugelstr. 6, D-45141 Essen, christian.thielscher@fom.de
[2] FOM University of Applied Sciences, Institute for Empirical Research & Statistics, Leimkugelstr. 6, D-45141 Essen, bianca.krol@fom.de
[3] FOM University of Applied Sciences, Leimkugelstr. 6, D-45141 Essen/University Medicine Essen, stefan-heinemann@gmx.de
[4] German Cancer Research Center (DKFZ), Department of Health Economics, Im Neuenheimer Feld 581, D-69120 Heidelberg, michael.schlander@dkfz.de

doi:10.18420/inf2019_07

- Automated and connected driving is an ethical imperative if the systems cause fewer accidents than human drivers (positive balance of risk).
- Damage to property must take precedence over personal injury. In hazardous situations, the protection of human life must always have top priority.
- In the event of unavoidable accident situations, any distinction between individuals based on personal features (age, gender, physical or mental constitution) is impermissible.

The trolley problem addresses a thought experiment where a decision-maker is faced with an ethical dilemma. Which is the morally correct choice in a scenario where a runaway trolley on railway tracks is headed for five people who are unable to move:

1. Do nothing and the trolley kills five people?

2. Pull a lever and divert the trolley onto another track where it will kill only one person?

There are many versions of this "sacrifice-one-to-rescue-five"-dilemma including various narratives and numbers of options. The forerunner of the autonomous trolley problem was J. J. Thomsons' paper from 1985 [Th85], whereas in Foots' original version from 1967, the trolley was manned.[5] Foots' text reads as follows [Fo67]:

"Suppose that a judge or magistrate is faced with rioters demanding that a culprit be found for a certain crime and threatening otherwise to take their own bloody revenge on a particular section of the community. The real culprit being unknown, the judge sees himself as able to prevent the bloodshed only by framing some innocent person and having him executed. Beside this example is placed another in which a pilot whose aeroplane is about to crash is deciding whether to steer from a more to a less inhabited area. To make the parallel as close as possible it may rather be supposed that he is the driver of a runaway tram which he can only steer from one narrow track on to another; five men are working on one track and one man on the other; anyone on the track he enters is bound to be killed. In the case of the riots the mob has five hostages, so that in both the exchange is supposed to be one man's life for the lives of five. The question is why we should say, without hesitation, that the driver should steer for the less occupied track, while most of us would be appalled at the idea that the innocent man could be framed."

P. Foot also invented the problem which was later called "transplant":

"Why…do we not feel justified in killing people to obtain…spare parts for grafting on to those who need them?"

This question is at the heart of the trolley problem. P. Foot proposes that there are "positive" and "negative" duties. She writes [Fo67]:

[5] There have been earlier versions, e g. [We51].

"The steering driver faces a conflict of negative duties, since it is his duty to avoid injuring five men and also his duty to avoid injuring one. In the circumstances, he is not able to avoid both, and it seems clear that he should do the least injury he can. The judge, however, is weighing the duty of not inflicting injury against the duty of bringing aid. He wants to rescue the innocent people threatened with death but can do so only by inflicting injury himself. Since one does not in general have the same duty to help people as to refrain from injuring them, it is not possible to argue to a conclusion about what he should do from the steering driver case. It is interesting that, even where the strictest duty of positive aid exists, this still does not weigh as if a negative duty were involved. It is not, for instance, permissible to commit a murder to bring one's starving children food. If the choice is between inflicting injury on one or many there seems only one rational course of action; if the choice is between aid to some at the cost of injury to others, and refusing to inflict the injury to bring the aid, the whole matter is open to dispute. So, it is not inconsistent of us to think that the driver must steer for the road on which only one man stands while the judge (or his equivalent) may not kill the innocent person in order to stop the riots."

According to this logic, we may not kill an uninvolved person in order to rescue five others with his organs (as in the "transplant" case); however, the pilot and the trolley driver may kill one because otherwise they would kill five – here two negative duties collide and in this case, we kill as few as possible.

Thomson however, objects to Foots solution; she disagrees with the sentence "killing one is worse than letting five die". In order to prove its falsity, she introduced both the "bystander at the switch" as well the "fat man problem" (Foot had a different fat man problem in her paper) [Th85, p. 1397]:

"Let us begin by looking at a case that is in some ways like Mrs. Foot's story of the trolley driver. I will call her case Trolley Driver; let us now consider a case I will call Bystander at the Switch. In that case, you have been strolling by the trolley track, and you can see the situation at a glance: The driver saw the five on the track ahead, he stamped on the brakes, the brakes failed, so he fainted. What to do? Well, here is the switch, which you can throw, thereby turning the trolley yourself. Of course, you will kill one if you do. But I should think you may turn it all the same."

(The authors of this paper disagree with Thompson on the bystanders' recommended behavior as we will explain in more detail later.)

Thompson also introduced the "fat man" problem, and the problem which it creates [Th85, p. 1409].

"Consider a case - which I shall call 'Fat Man' - in which you are standing on a footbridge over the trolley track. You can see a trolley hurtling down the track, out of control. You turn around to see where the trolley is headed, and there are five workmen on the track where it exits from under the footbridge. What to do? Being an expert on trolleys, you know of one certain way to stop an out-of-control trolley: Drop a really heavy

weight in its path. But where to find one? It just so happens that standing next to you on the footbridge is a fat man, a really fat man. He is leaning over the railing, watching the trolley; all you have to do is to give him a little shove, and over the railing he will go, onto the track in the path of the trolley. Would it be permissible for you to do this? Everybody to whom I have put this case says it would not be. But why?"

Again, the key question is: why would most people throw the switch but not shove the fat man? And should an autonomous vehicle in a comparable situation – for example if it has to decide to either kill five children playing on a street or run into a pedestrian and kill him – decide like in a "fat-man-problem" or in the trolley case?

Thomson gives several potential explanations. First, she says that shoving is a direct infringement of the fat man's rights whereas throwing the switch isn't. However, as she herself notices, this explanation does not work. Imagine that you do not have to shove the fat man but just need to wobble the handrail: you still would not – although "wobbling" the handrail and "throwing" the switch are pretty close (at least morally). Thomson finally explains [Th85, p. 1410]:

"So the means by which the agent in 'Fat Man' gets the trolley to threaten one instead of five include toppling the fat man off the footbridge; and doing that is itself an infringement of a right of the fat man's. By contrast, the means by which the agent in 'Bystander at the Switch' gets the trolley to threaten one instead of five include no more than getting the trolley off the straight track onto the right-hand track; and doing that is not itself an infringement of a right of anybody's."

The authors of this paper do not accept this explanation. If one throws the switch he definitely kills the one person on the other track; it is hard to see how this is "no infringement" of his right to live. We also do not accept other solutions that were proposed, e.g., that wobbling the handrail requires the decider to be close to the victim of his doing (the fat man); whereas the switch thrower is distant from the killed person. It is easy to construct a specific case where a fat man stands on a trap door and a trigger can be pulled from somewhere to open it (however, still in this case, most people would not kill the fat man but would throw the switch). Finally, we do not think that people behaving differently in these situations are simply inconsistent in their behavior.

We believe that both P. Foot as well as Thomson point into the right direction. Our key hypothesis is the idea that "negative duties overrule positive duties" is a Kantian rule (see below) whereas the central utilitarian rule is "kill as few as possible". Respondents simply use both rules, if they agree, and, if they disagree, prefer one of these over the other, depending on the specific details of the story. If they agree, autonomous vehicles should as well.

By ethical decomposition of five test cases, we prove that

- Kantian and utilitarian decision rules agree that it is better to act – to throw the switch etc. – in one case (and, therefore, the majority of respondents does act),

- agree not to act in another case (so that respondents don't act), and
- disagree in the other cases (so that some respondents act while others don't). We finally test the hypothesis empirically.

2 Methods

In light of the various possibilities and potential issues and solutions relating to the trolley narrative, this paper aims to explore the decision of people in different variations of the trolley case. To do so we devised a set of cases - by developing one new case, changing another one and rearranging them together with three known cases - which serve to

- create a continuum of changing decisions so that almost all responders would "act" (throw the switch, shove the fat man, etc.) in the first case and almost none would act in the last case; and
- differentiate between Kantian and utilitarian perspectives. We used these two because they are seen as the most influential non-religious ethical theories [Sa10]. Of course, ethical decomposition works with other theories as well.

Finally, we empirically tested the cases with respondents. Responses were analyzed with chi square test.

3 Results

There are five cases in our continuum:

1.: In the "dynamite wagon" case the decision-maker is standing near the trolley track seeing an unmanned trolley hurtling down the track, out of control. There are three tracks – one with one person chained to the track, one with five persons chained, and one with the dynamite wagon. What to do? If the decision-maker does nothing, the unmanned trolley will run into the dynamite wagon loaded with explosives and the detonation will kill both the one person on one track, as well as the group of five people on the other track. However, the decision-maker can choose to throw the switch so that the trolley cart is diverted from the dynamite - but then kills either one or five.

2. and 3.: In addition to the "dynamite wagon", we propose two different versions of the trolley case: in the first one, the switch is loose so that it can move from one side to the other; if the trolley arrives it will cause the switch to point in one or the other direction with a 50/50 chance (this is what we call the „trolley, floppy switch" case). The traditional case where the switch points to the five is called "trolley, fixed switch".

4. and 5.: Finally, we use the "fat man" and "transplant" cases as described above.

The following table presents an overview of cases (table 1 – we will explain the column "expected response" in a second):

Problem	**Short description**	**Utilitarian decision**	**Kantian decision**	**Expected response**
1. Dynamite wagon	Trolley runs into dynamite wagon, thus killing all; bystander may throw switch in other direction, thus killing one or five	Throw switch	Throw switch	> 95%
2. Trolley, floppy switch	A trolley will kill either one or five; unclear position of switch; bystander may throw switch in either direction, thus killing one or five	Throw switch	Throw switch (?)	> 50%
3.Trolley, fixed switch	A trolley will kill five because switch points in their direction; bystander may throw switch in other direction, thus killing one	Throw switch	Do nothing	~ 50%
4. Fat man	A trolley will kill five unless you shove a fat man over bridge	Shove fat man	Do nothing	< 50%
5. Transplant	Five persons will die from disease unless you kill another person, transplanting his organs	Do nothing (?)	Do nothing	~ 0%

Tab. 1: Description of cases

We now discuss these cases in order to differentiate between the utilitarian and Kantian decision-making approach. By the "utilitarian" approach we mean Benthams' basic version where the sum of happiness is maximized [Be70]. With "Kantian" we refer to Kants' famous categorical imperative. Kant used different formulations which he thought are interchangeable. We employ two of these phrases [Ka61]:

1. “Handle nur nach derjenigen Maxime, durch die du zugleich wollen kannst, dass sie ein allgemeines Gesetz werde.” This can be loosely translated as: act on principles that are well-suited as foundation for a law.

2. “Handle so, daß du die Menschheit sowohl in deiner Person, als in der Person eines jeden andern jederzeit zugleich als Zweck, niemals bloß als Mittel brauchest.” That is, behave at all times such that you treat humanity both in the form of your person, as well as all others, as an end, but never merely as a means.

From the Kantian perspective, it is forbidden to kill an uninvolved person in order to save five because this would treat the individual as a means rather than as an end. This interpretation is close to Foots’ sentence: “killing one is worse than letting five die.”

In “dynamite wagon” both utilitarians, as well as Kantians will throw the switch. However, there is a problem for Kantians: in which direction? While it seems natural to kill one rather than five, it is not easy to see the principle which would allow for a law. For example, the principle “it is better to kill fewer people than more” cannot always be correct. Imagine you have to either rescue two Texan criminals sentenced to death and waiting for their penalty or one seven-year old girl.

“Everything else being equal if you have to kill either one or five then choose the one” seems to be a good candidate but still is far from being perfect. Imagine, for example, that a sadistic police officer in a totalitarian country forces you to decide: either you kill one of five innocent prisoners or he kills all five. Are you entitled to kill one? Which one? It seems as if the Kantian rule is indecisive in these cases – which is not very surprising since Kant did not want to develop a metric to decide in every dilemma; rather he wanted to identify ethical rules that are always correct (independent of the situation).

In “floppy switch”, utilitarians will clearly throw the switch as well, whereas Kantian logic is again open for discussion. One could argue that a Kantian might be hesitant to throw the switch, being afraid of using the one person as a means only. However, if he does not, he lets coincidence decide; in this case he uses the one or the five killed as fortuities’ means to save the other(s). Before the floppy switch becomes fixed (by chance), both the one as well as the five persons are “involved” by being threatened. If the Kantian throws the switch, he has to treat either the one or the five as a means and the other(s) as an end. In this conflict, some Kantians will feel entitled to throw the switch while others do not. It is close to the situation where a pilot has to decide where to crash his broken plane: in an inhabited area or redirect it into a less peopled region. Kantians willing to change the planes’ direction will point at the fact that people living close to aeroplane routes know that they might get into the way of busted machines, that pilots will try to kill as few as possible; and therefore, implicitly accept the pilots’ behavior.

In “trolley, fixed switch”, Kantians would NOT throw the switch whereas utilitarians would. In this case, the one person is uninvolved (not threatened by coincidence) and therefore, Kantian logic would forbid to kill him.

This also makes clear that the ethical watershed runs between “floppy switch” and “fixed switch” (and not between “fixed switch” and “fat man”).

The authors think that this is an important finding. We believe that some listeners who prefer Kantian ethics may mix “trolley, floppy switch” with “trolley, fixed switch” and treat the latter as if it was the former: it is easy to confound the two. Kantians who think about “fixed switch” may misunderstand that the switch indeed is fixed; in this case, they may think they are in “floppy switch” and decide to throw the switch – although they should not. By clarifying the situation, Kantians who threw the switch in “fixed switch” erroneously (because they thought it was a “floppy case”) will change their mind. Our empirical results point into this direction.

In fat man, pure utilitarians would shove or wobble, and even in transplant it is difficult for them to find a way not to kill the one person: they could, e.g., say that if they did people might not enter hospitals anymore (because they expect to get killed and eviscerated there), thus causing trouble and loss of happiness for the society. However, this introduction of indirect societal costs and benefits (that is, costs that are not directly linked to the case as such but rather potential consequences of it) creates significant uncertainties: which indirect costs are to consider and how are they valued? Where to start and stop with calculation of potential consequences? E.g., would people use bridges (like in the “fat man problem”) if they think they get shoved?

Therefore, neither Kantian nor utilitarian rules provide clear decisions on all of the cases. In addition, sometimes they agree whereas sometimes they disagree depending on the respective variation of the basic conflict (kill one to rescue five).

We empirically tested this intuition with responses to our cases. In July 2018, we discussed the cases from “dynamite wagon” to “transplant” with 50 students. All of them studied medicine at the university hospital of Heidelberg, all being in their 5th semester. First, we introduced the cases to them and asked them whether they would “act” (i.e., depending on the specific case, throw the switch, shove the fat man, or kill the transplant patient).

More specifically, we explained the “trolley, fixed switch” as follows: “An unmanned, out-of-control trolley runs towards five people who are chained to the track; it will kill all five. You are standing close to a switch. If you throw it, the trolley changes direction and runs into one person who is chained to the track, killing this one but saving the other five.” We also showed a little drawing (Fig. 1). We then asked the students whether they were familiar with the case or not.

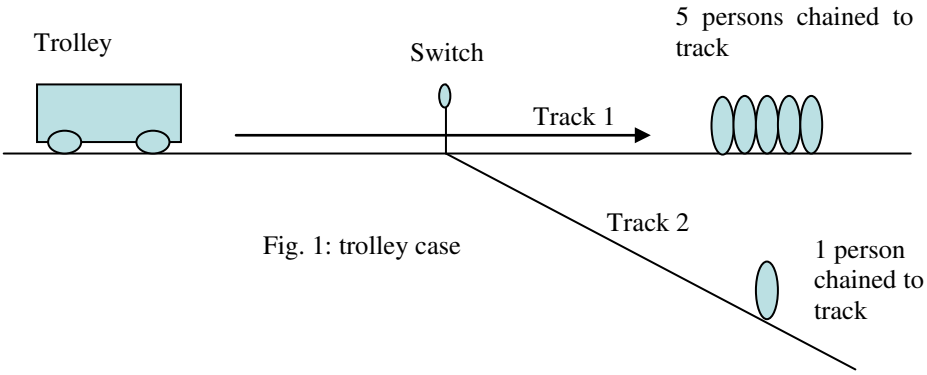

Fig. 1: trolley case

We continued by presenting the other cases (from "dynamite wagon" to "transplant").

Following this, the two groups of students (familiar / new to the problem) were asked to vote whether they would act (throw the switch, shove the fat man, etc.).

Afterwards, utilitarian vs. Kantian ethics and their perspective on the cases were discussed.

Finally, we had the same votes as before.

These are our findings (table 2):

	New to problem (n=30)		Familiar with problem (n=20)	
	before discussion	after discussion	before discussion	after discussion
Dynamite wagon***	21:3	23:1	16:0	16:0
Trolley, floppy switch***	21:9	23:4	12:6	12:6
Trolley, fixed switch***	23:5***	9:9***	17:2***	11:9***
Fat man***	7:18	8:17	4:12	3:15
Transplant***	0:25	0:25	0:19	0:19

Tab. 2: Results

The numbers in the boxes denote the number of students who act versus the number of students who do not. For example, in the group of students who were new to the problem, before discussion 21 would throw the switch in "dynamite wagon" whereas three would not. The numbers do not always add to the total number of students because some did not decide.

Results are statistically highly significant at the 0,001 level:

- All lines differ significantly from all other lines (by comparing the sum per line) – with one exception: trolley, floppy switch vs. trolley, fixed switch. However, in the group that was new to the problem, after discussion, those two cases did differ significantly.
- For both groups, the discussion did not change the responses significantly – again with one exception: both groups changed their responses significantly in trolley, fixed switch.

4 Discussion

The purpose of our study was to better understand ethical intuition and empirical responses to dilemma situations and to deduct recommendations for autonomous vehicles.

The overall pattern of our students' response is in line with expectations. With both Kantians and utilitarian ethicists throwing the switch in "dynamite wagon" (we didn't ask in which direction), we expected students to do the same – and the vast majority did. In "transplant", Kantians and (most) utilitarians would not act, and therefore, most respondents do not act as well. We conclude that in these cases, automatic control systems should do the same; guidelines for software developers might help to align engineering with unambiguous moral decisions. "Rolling dice" seems not to be the best strategy in these cases.

In the cases from "trolley" to "fat man", however, students need to decide which of two conflicting ethical rules to follow. Therefore, it was expected that the voting is in between the two ends of our continuum.

A bit of a surprise is the fact that before discussion, less students did not act in „fixed switch" than in "floppy switch"; this may be due to some uncertainty. However, after discussion, this effect disappeared.

Another interesting finding is that after discussion, more students did not act in "fixed switch". It seems that the comparison of the cases and ethical background fosters Kantian behavior.

Until now, the fact that people decide differently in comparable cases as we modelled here (all being versions of "kill one to rescue five") is often seen as a case of inconsistent behavior. An often-cited example for irrationalities is N. Taleb who highlighted incon-

sistencies in the thought processes of subjects when forecasting / estimating the likelihood of various scenarios [Ta08]. He calls this "pathologies in decision-making", and he also talks about "blindness" and the "paradox of perception", as well as "the pull of the sensational" and "the use of narrative to get attention". Taleb maps and separates general attributes of thinking and reasoning into what he calls System 1 and System 2. System 1 is experiential. It is associated with intuition that produces short cuts called "heuristics" that allow rapid action. System 2 is the cognitive one that we normally call "thinking" that is effortful, reasoned, logical, self-aware, etc. – the stuff that apparently lies at the basis of research and goes on in the classroom. This idea is also at the core of D. Kahneman's work on fast and slow thinking [Ka11].

However, there is debate whether this is indeed a pathology. P. Railton recently argued that System 1, being based on prior experiences and having condensed them to simple rules of thumb is wrong only if for some reason the current situation does not fit to former learnings [Ra15].

Another case often cited as a typical example for "irrational" behavior is the fact that people tend to pay higher attention to known lives as opposed to unknown. If miners are trapped in Chile, much effort is provided to rescue them; although this money would have saved much more people if used for feeding the hungry. However, there might be a rational explanation for this as well; e.g., if respondents use two rules of thumb as the following: "try to help were you can" and "you cannot save everybody in the world so focus on what you can do (and know about)".

After ethical decomposition, the responses given by the students are not irrational but rather in line with ethical theories. Even if there is some irrational part to decisions (as is probably always the case) it is much less clear whether all of the changes (from act to don't act) are due to inconsistencies. There might be good reasons to act "Kantian" in one situation and "utilitarian" in another.

We therefore think that just leaving it open to the artificial intelligence engineers how their car behaves may be not the best solution [GM17]. At least it should be possible to decompose "Kantian" versus "utilitarian" decisions in many potential situations; and if they agree, engineers should agree as well. Of course, this will work for some but not all situations. For example, in situations involving humans versus properties the German Ethics Commission on Automated Driving demands (in a Kantian way) that damage to property must take precedence over personal injury [BM18]. However, this is not undisputed. R. A. Posner, for example, says it is better to kill a child than 100.000 sheep [Po79].

On a more practical note, we reason that in addition to those of utilitarians, the interests of Kantian deciders need to be reflected in creating unmanned cars as well: a utilitarian self-driving vehicle will kill as few as possible, whereas a Kantian car is simply not entitled to kill anybody (if avoidable). For example, it is acceptable to a utilitarian if sometimes the vehicle software crashes and kills people – given that the pleasure created by autonomous vehicles is higher than the loss resulting from the crashes and that the

costs of building the vehicle in a safe way are higher than the costs of the pleasure lost by accidents. It is, therefore, much more difficult to construct a Kantian car because it requires fail-safe IT, well-developed Artificial Intelligence systems and extensive testing and planning. (Kant would probably also note that voluntarily creating vehicles which have the capacity to kill people are not in line with ethical demands and that the engineers who constructed them might want to read his books.) Thus if we do not know whether at least some people prefer Kantian cars, we should not risk killing them by accidence.

Of course, trolley type cases are not the only problem software developers face; rather, there is a multitude of challenges, e.g., of technology, cost-benefit-assessments, etc. [Ro19]. The trolley case tells something about human perspectives on ethics, and if we understand them better, we can transfer them to automatic systems.

There are two key limitations to our results.

First our experimental design did not control for group effects (e.g. that some respondents wanted to decide in the same way as the majority did).

Second, we cannot match voting yet with Kantian or utilitarian perspective; that is, we do not know which of our students see themselves as Kantian or utilitarian (or else). We can only conclude this from their respective voting. We will design further studies to find out.

5 Conclusion

Moral dilemmata can be analyzed both from utilitarian as well as Kantian perspective. Sometimes they agree in their decisions, sometimes they disagree. In cases where they agree, respondents decide almost unambiguously, and automatic control systems should do the same. Here, “rolling dice” seems not to be the best strategy.

Our experiments support the hypothesis that humans use both utilitarian and Kantian decision criteria. In some cases, they prefer utilitarian over Kantian rules and vice versa. This may explain behavior that looks “irrational” at first glance.

Ethical decomposition of conflicts can help with the difficult undertaking to tell clear cases (where utilitarians and Kantians agree) from unclear ones. In addition, it supports analysis of our moral intuition. More research is needed to experimentally scrutinize moral behavior of humans, and to translate it into guidelines for software development.

Acknowledgements: The authors would like to thank Linda O’Riordan, Julia Richenhagen, and Piotr Zmuda for valuable input.

Bibliography

[Be70] Bentham, J.: An Introduction to the Principles of Morals and Legislation. Edited by Burns, J. H. and Hart, H. L. A. Oxford University Press, Oxford, 1970

[BM18] BMVI Press Release: Ethics Commission on Automated Driving presents report. https://www.bmvi.de/SharedDocs/EN/PressRelease/2017/084-ethic-commission-report-automated-driving.html, accessed: July 20th, 2018

[BSR16] Bonnefon, J. F.; Shariff, A.; Rahwan, I.: The social dilemma of autonomous vehicles. Science, 352(6293), 2016, p. 1573-1576.

[Ed15] Edmonds, D.: Would You Kill the Fat Man? The Trolley Problem and What Your Answer Tells Us about Right and Wrong. Princeton University Press, Princeton, 2015

[Fo67] Foot, P.: The Problem of Abortion and the Doctrine of the Double Effect. Oxford Review, no. 5, 1967, p. 5-15

[GM17] Gogoll, J.; Müller, J.: Autonomous Cars: In Favor of a Mandatory Ethics Setting. Sci Eng Ethics (2017) 23, p. 691-700

[Ka11] Kahneman, D.: Thinking, Fast and Slow. Farrar, New York, 2011

[Ka61] Kant, I.: Grundlegung zur Metaphysik der Sitten. Edited by Valentiner, T. Reclam Stuttgart, 1961

[Po79] Posner, R. A.: Utilitarianism, Economics, and Legal Theory. The Journal of Legal Studies, 1979 8:1, p. 132f.

[Sa10] Sandel, M. E.: Justice. Farrar, New York, 2010

[Ra15] Railton, P.: "Dual-Process" Models of the Mind and the "Statistical Victim Effect". In: Cohen, I . G. et al.: Identified versus Statistical Lives An Interdisciplinary Perspective. Oxford University Press, Oxford, 2015, p. 24-42

[Ro19] Roff, H.: The folley of trolleys: Ethical challenges and autonomous vehicles. https://www.brookings.edu/research/the-folly-of-trolleys-ethical-challenges-and-autonomous-vehicles/, accessed: June 13th, 2019

[Ta08] Taleb, N.: The Black Swan. Penguin, London, 2008

[Th17] The Moral Sense Test. http://www.moralsensetest.com/, accessed: Dec 2nd, 2017

[Th85] Thomson, J. J.: The Trolley Problem. The Yale Law Journal, Vol. 94, No. 6 (May, 1985), pp. 1395-1415

[We51] Welzel, H.: Zum Notstandsproblem. ZStW Zeitschrift für die gesamte Strafrechtswissenschaft 63 [1951, 1], p. 47-56.