Brake or Drive: On the Relation Between Morality and Traffic Rules when Driving Autonomously [1]

Astrid Rakow [2], Maike Schwammberger[3]

Abstract: For a safe and successful future with autonomous traffic agents (ATAs), these ATAs need to be enabled to understand and abide by traffic rules. However, purely formalising and analysing traffic rules is not enough to solve this task. In this paper, we discuss the role of moral for ATAs that follow traffic rules. In particular, moral values may enable an ATA to prioritise traffic rules, in case of conflicts. We outline an approach that uses formal verification to identify situations where traffic rules are in conflict with each other, with moral values or with specific goals of an ATA. We sketch how moral values and reasoning can help an ATA to resolve such conflicts autonomously.

Keywords: autonomous traffic agents, traffic rules, moral reasoning, conflicts, ethics

1 Introduction and Motivation

While more and more autonomous vehicles are driving on the roads, we identify an important issue: The autonomous traffic agents (ATAs) that steer these vehicles need to be able to understand and abide by traffic rules. In our sense, an ATA is an entity that acts autonomously in public road transport, meaning in environments that are regulated by traffic rules. We consider ATAs with SAE levels three to five, from driving autonomously only in specific situations to fully automated driving [SA18].

In general, „following traffic rules“ is implicitly implemented into ATAs in most cases. Traffic rules are not a major concern in the AV research community [ADF20; Pr17]. Without a formal representation of traffic rules, it is complicated for an ATA to reason about traffic rules. Due to this, it can be difficult to impossible to link an action of an AV to a traffic rule as its reason. However, for tasks like explainability and understandability, and with that ultimately trustworthiness of autonomous vehicles such a connection of observed actions and behaviours with their reasons is of importance, as is e. g. argued in [SK22].

In [AS22], we sketch our vision of building a *Digital Highway Code*: A fully formalised version of the natural language road traffic regulations that exist throughout the world (cf. UK Highway Code [De17] or the German „Straßenverkehrsordnung (StVO)“ [Bu13]. But a simple translation of the natural language sentences from these traffic regulations

[1] This research was supported by the German Research Council (DFG) in the PIRE Projects SD-SSCPS and ISCE-ACPS under grants no. DA 206/11-1, FR 2715/4-1, FR 2715/5-1.
[2] University of Oldenburg, Ammerländer Heerstraße 114-118, 26129 Oldenburg, Germany, a.rakow@uol.de
[3] University of Oldenburg, maike.schwammberger@informatik.uni-oldenburg.de

doi:10.18420/se2023-ws-12

is not sufficient. From [AS22], we identify the following tasks as a necessity for building a Digital Highway Code: `(t1)` deal with inconcise, culturally differing and ambiguous natural language formulations of the known traffic regulations, `(t2)` translate the traffic rules into a machine-readable format, `(t3)` identify conflicting traffic rules, `(t4)` prioritise traffic rules situation-dependently to solve conflicts, and `(t5)` verify that autonomous traffic agents can actually understand and follow the suggested Digital Highway Code.

In this vision paper, we discuss the role of moral values for ATAs abiding by traffic rules. We constitute that for tasks `(t1)` and `(t4)` moral values have to be taken into account to a certain degree, since ATAs need to apply a common sense interpretation that must be based on moral values. We observe that implementing moral ATAs is hence intertwined with high demands of informedness about their context, so that ATAs have to deal with uncertainties. After sketching our vision of how to equip an ATA with morality, we outline how the formalisation `(t2)` and verification `(t5)` of traffic rules can be used to identify conflicts between traffic rules `(t3)` and how to solve the identified conflicts `(t4)`.

Outline: In Sect. 2, we motivate *why* ATAs should be equipped with moral values. Following that, Sect. 3 sketches *how* we envision to equip ATAs with moral. In Sect. 4.2, we describe how we plan to use model checking to identify conflicts between traffic rules. We also sketch how these conflicts may be resolved by adding moral values. We conclude in Sect. 5.

2 The Need of Moral Values in ATAs

Here we motivate why ATAs need to take moral values into account. Many believe the more machines take control over important decisions, the more these machines have to act in morally appropriate ways [Pi97; To21]. But does an ATA also need to consider moral values despite being regulated by traffic rules? What scope for moral choices do traffic rules offer?

Traffic rules regulate the public road traffic. *They constrain the behaviour of traffic participants by telling them, when to stop, how fast to go at most, when to turn on lights and so on.* Traffic rules assume a certain infrastructure as given, and their rules are preconditioned on context information that is hence usually accessible to the drivers. *Traffic signs announce speed limits, and sirens inform traffic participants of approaching emergency vehicles.* Current traffic rules are natural language constructs that evolve over time. As a consequence, traffic rules are ambiguous and contain contradictions [Re21]. Traffic participants must hence apply a „common sense“ interpretation when abiding by traffic rules [De21]. The evolution of traffic rules is certainly influenced by moral values, *as e. g. the mission of the US Department of Transportation illustrates, that greatly shaped the US traffic laws. It is committed to „Serve the United States by ensuring a fast, safe, efficient, accessible and convenient transportation system that meets our vital national interests and enhances the quality of life of the American people, today and into the future“* [US22].

According to our everyday experience, traffic rules find a wide social acceptance and are followed by human drivers to a large degree. Nevertheless, human drivers often take the liberty to break (or „bend") some traffic rules, often without negative consequences [Re21]. Consequently, „driving through traffic requires that the vehicle conforms with societal expectations for roadway behaviour" that „have their foundation in core moral issues found in philosophy and ethics", as Thornton et al. remarked in [Th17, p. 1429]. ATAs need to know the *road culture*, that is a „set of knowledge, rules and norms of behaviour" „shared by all traffic users" [BID18, p. 82]. An ATA must be able to extrapolate the behaviour of others, which is influenced by moral values, and it has to choose behaviour that is morally acceptable and predictable by others and its user/ owner/ passenger. Consequently, an ATA must also take moral values into account when making decisions.

We argue that, even if a machine-interpretable, conflict-free version of traffic rules would exist, an ATA still would have to use some kind of moral common sense in rare, extreme situations. Think of scenarios like, *„Due to a heavy rainfall, the road rather suddenly gets flooded"* or *„A lorry's load is about to fall onto the road, endangering a following vehicle"*. In both cases, common sense – and hence moral values – have to be applied to resolve the situation safely. Moreover, these situations may occur so suddenly that even in shared control situations, a transfer of control to the driver might not be possible, so that the ATA has to make decisions autonomously.

On top of that, traffic rules provide a large degree of freedom for choosing what to do exactly. *A speed limit does not strictly define the speed to be chosen.* It has been observed in [CDF22] that several of the UK's traffic rules are „should"-rules, that is, they do not express an obligation but rather a strong recommendation. This freedom is often used by human drivers to their liking. *The decision whether fuel savings should be traded for the reduction of travel time is a moral decision. When buying a car, customers choose which hardware meets their needs – and hence also their moral values – best.* Since moral values of an owner or driver of an ATA also influence how the ATA should perform when adhering by traffic rules, an adaptation of the ATA to the moral values of the user seems to suggest itself.

We summarise our main points. Traffic rules alone do not sufficiently govern all situations arising in today's traffic and even within the realm of standard traffic situations, the ATA's manoeuvre choices should take moral values into account. In particular, in order to interact with other participants, an ATA needs to know about common/shared moral values. For user acceptance, the ATA has to take the user's individual moral values into account when fine-tuning manoeuvres. The necessary consequence is, that moral values should guide the design as well as behaviour of an ATA on the road.

3 Our Vision of Moral ATAs

In this section, we outline our vision of moral ATAs that is based on the current state-of-the art of ethical machines, i. e. machines that consider moral theories when taking decisions. Accompanying our outline, we introduce central notions of the field of machine ethics. „Ethics is the philosophical science which studies morality and is concerned with what is morally good and bad and morally right and wrong“ [Dr75; Si22]. *Machine Ethics* is „concerned with the behaviour of machines towards human users and other machines“ [AA07, p. 1]. In the field of computer science, one focus is the problem of automating moral reasoning [Sl22].

We think of ATAs basically as rational decision makers as studied in economy. Facing a decision problem, an ATA chooses the choice alternative that maximizes the expected utility based on the preference order of these alternatives. We follow Sen [Se74], envisioning morality to be an ordering over the agent's preference of alternatives. [Sl22]

We think, ATAs should be explicit ethical agents rather than implicit or fully ethical. This means, they contribute to evaluating whether a choice alternative is more or less ethical than another. Though, the degree to which they take moral decisions is limited by their abilities or the context. Their evaluation is determined by humans to a large extent. In contrast, fully ethical agents take fully autonomously moral decisions, and implicit moral agents take moral decisions, where the evaluation is completely preprogrammed. We take this position, because the ATAs have to function in contexts where other traffic participants apply common sense and are guided by moral values. As argued in the previous section, an ATAs needs to know about the road culture, in order to predict the behaviour of others, to be predictable and to foster acceptance. We agree with [Re21] that an ethical goal function should be developed by legislation, it should be agreed upon in society, may reflect cultural differences and be subject to ongoing review. Note, that we follow [To21] by focusing on ethical machines implementing some kind of moral theory, rather than on philosophical notions like „conciousness“, „intentionality“ or „agency“.

We envision that ATAs can be equipped with several moral theories. A *moral theory* is an explanation of what makes an action right or what makes an entity good. So, it establishes the righteousness of an action or entity [Va15, p. 78]. Most commonly considered theory types are consequentialism, deontological ethics, and virtue ethics. In *consequentialism*, one defines a choice as morally good if its consequences maximize the well-being or more abstractly, a utility. In a *deontological ethics* one defines moral rules or duties for an agent, while in a *virtue ethics* one considers virtues, e. g. friendliness and courage, as central concepts to distinguish whether a choice is morally good or bad. There is no consensus on which theory is best suited for any particular domain [Na20]. We envision that several stakeholders should be allowed to specify their own moral theory, which then is used according to its assigned priorities. *For instance, the legislation's moral theory of [Re21] would get the highest priority, while e. g. a manufacturer's would get a considerable lower priority.* That way, an ATA would be forced to adhere by moral values like „saving lives“, „avoidance of injuries“, „avoidance of damage“ and additionally, it could take into account the moral values of an individual such as „saving fuel“ or „optimize travel time“. We see it as future work to develop approaches to formally integrate multiple prioritised moral theories that may contradict each other in parts and reinforce each other in others. As Tolmeijer et al. noticed in their survey [To21], there is a need to further study the combination of different theories. They report that about a quarter of the approaches are *hybrid*, i. e. combining at least two classical moral theory types, and approximately half of those have a *hierarchical approach*, in which deontological features are standardly dominant over consequentialist ones. We expect that a formal analysis of a system of multiple prioritised moral theories might help to expose challenging ethical corner cases and thereby foster the discussion between stakeholders.

We think that an ATA should primarily implement a top-down approach to moral theories. Allen et al. [ASW05] proposed distinguishing between *top-down approaches*, that are theory-driven, and *bottom-up approaches*, in which a system learns moral behaviour from examples. We already noted that we agree with [Re21] that high priority moral values should be the result of a public discussion. To this end, a top-down approach seems appropriate. We imagine that a bottom-up approach can be appropriate for low-priority values, e. g., an ATA may learn moral values regarding *„fuel saving vs reduction of travel time"*.

Moral decision-making is intertwined with dealing with uncertainties. While traffic rules aim to be executable – the required contextual information is either provided by the infrastructure or is rather easily accessible to the driver through observation –, moral theories depend heavily on contextual information. *For example, the consequentialists objective „Maximize the number of saved lives." requires knowing these numbers. Not to mention moral theories that depend on information like intent of other traffic participants.* This immensely increases the need for information, which will usually not be fully satisfied. Moral ATAs must therefore develop strategies to make moral decisions despite the lack of information.

We envision an ATA to coarsely classify its moral choice alternatives according to their moral value and to the confidence the ATA has in its evaluation. When the ATA faces a decision problem it classifies its choice alternatives roughly *, e. g., whether the alternative is considered morally very good or slightly bad*. A confidence value captures how certain the ATA is that the choice alternative actually belongs to the assigned class. Uncertainties may arise due to perturbed, imprecise or missing sensory data, but also due to conflicting moral evaluations, *e. g. think of Robin Hood and the question „Is stealing for the poor okay?"*. We consider it future work, to formally define decision taking approaches of ATAs that are facing uncertainties and that are equipped with multiple prioritised moral theories. As related work we refer to [De21] where De Freitas et al. propose an approach to developing common sense tests for ATAs. They stress ATAs have to deal with uncertainties, limited information, and multiple metrics, among others.

We conjecture, that there are decisions that an ATA should not take unsupervised. Up to now, we argued that an ATA must take moral values into account. Especially in emergencies, where the time span is too short to get the human back into the loop, implementing no moral guidance would most likely lead to morally bad behaviour. We believe, that an ATA can often confidently choose an alternative that is morally good, or at least not morally slightly bad. There are however cases, where an ATA cannot take a decision with sufficient confidence. In such cases, we want to return the moral control to a stakeholder. We consider it as a design task to identify these scenarios and implement mechanisms to transfer control to the appropriate addressee within appropriate time frames.

Conclusion. We think that ATAs have to abide by moral values. We envision them to have multiple, prioritised, moral theories and their morally guided decision-making takes moral uncertainties into account. We believe that moral operational domains have to be identified, i. e., it must be defined in which situations an ATA is allowed to take what moral decisions, and it must be determined when an ATA is required to take certain moral decisions.

4 Detecting and Resolving Traffic Rule Conflicts

The previous sections make it clear that morality is important for ATAs if only because morally bad decisions must be avoided. We hope that moral theories will furthermore enable ATAs to interpret ambiguous traffic rules and resolve conflicts between traffic rules through common sense. As part of our efforts to formalise traffic rules for ATAs, we therefore want to investigate this hypothesis. To this end, we propose an approach in which we determine cases, where our formal traffic rules are in conflict, and we then examine whether adding moral can resolve the conflict. In the sequel, we describe this vision and its relation to our research agenda in more detail. We first outline our previous work and related work that lays the foundation for identifying conflicting traffic rules. We then sketch a stepwise approach to identifying conflicts among traffic rules. Finally, we discuss future work on the analysis of moral ATAs in situations where traffic rules are in conflict.

4.1 Preliminary Work

In [Sc18], the traffic logic *Urban Multi-lane Spatial Logic (UMLSL)* was introduced to formalise turn manoeuvres at urban intersections. Further on, an extension of timed automata, *Automotive-Controlling Timed Automata (ACTA)* was introduced for specifying crossing controllers that use UMLSL to reason about temporal aspects of traffic situations. The model-checking tool UPPAAL [BDL04] has been used in [BS19] to analyse properties, e. g. collision freedom, of these crossing controllers. For formalising road traffic rules, [SA21] extends UMLSL to *Urban Spatial Logic for Traffic Rules (USL-TR)*. With USL-TR spatial aspects of traffic rules (e. g. two cars driving one behind the other) can be formalised and with ACTA, temporal aspects of traffic rules (e. g. one event happening before another event) can be specified. For each traffic rule, a *traffic rule controller* would be specified, as is suggested in [AS22]. A connection of the USL-TR traffic rule controllers with *belief, desire, intention (BDI) agents* has also been proposed in [AS22]. For this, previous work on formalising and verifying temporal properties of traffic rules with BDI agents from [ADF21] will be taken as a basis. In this section, we argue that morality also needs to be taken into account for that extension.

4.2 Moral guidance when abiding by traffic rules

Reed et al. state in [Re21] that there are situations, where human drivers will break a traffic rule in order not to break another, more important traffic rule or moral rule. They mention that drivers often drive onto the sidewalk to let an emergency vehicle pass, for example, or to avoid a collision – even though the UK Highway Code states that a driver should not drive onto a sidewalk. Collenette et al. differentiate in [CDF22] between „must“ and „should“ rules in the UK Highway Code. They state that, if a „must“- and a „should“-rule are in conflict, the „must“-rule has to be followed, while, if two „should“-rules are in conflict, the ATA may choose which rule to follow. They do not further specify how this choice should be taken nor what to do if two „must“ rules are in conflict. Also, De Freitas et al. [De21] discuss that conflicts can exist, among traffic rules as well as between traffic rules and ethical principles. They mention that in some of these cases, rule hierarchies are already encoded in law. For instance, the advice of a police officer must be followed, even if this advice includes breaking a traffic rule.

To summarise, different works exist on conflicts between traffic rules. We identify from such related work and existing traffic rule books that two types of conflicts for traffic rules can exist: Situation-independent conflicts, where following one traffic rule is always more important than following another rule, and situation-dependent conflicts, where traffic rules are prioritised differently depending on different traffic situations and contexts. We propose to identify conflicting traffic rules first and then to resolve the conflict with the help of moral rules. As motivated in Sect. 3, we assume that the choice alternatives are divided into rough classes ranging from morally good to morally bad. However, we expect that many conflicts will occur whose choice alternatives lie in the grey area of the moral scale. In these moral dilemma situations, the ATAs cannot make a decision (cf. Moral Machine Dilemmas [Ma22]). For this reason, not all conflicts among traffic rules will be automatically solvable by incorporating moral into ATAs.

4.3 Analysing Traffic Rules – Fighting Complexity

To our best knowledge, there are no approaches to formally analyse traffic rules to find conflicts, nor approaches to explore how morality can be used to resolve these conflicts. This is discussed now for the approach introduced in Sect. 4.1.

Formal verification guarantees to cover all corner cases exhaustively, but its application is often hindered by the complexity of the examined model. As scenarios *„driving onto the sidewalk“* and *„police officer“* from Sect. 4.2 illustrate, the analysis of traffic rules requires a rich model, capturing various kinds of infrastructure, *e. g. sidewalks, intersections*, traffic participants, *e. g. emergency vehicles, police officers, pedestrians*, and characterising complex dynamics, *e. g. an imminent accident influenced by weather conditions*. We hence plan to analyse increasingly complex models. We summarise our approach very abstractly as follows. On encountering a conflict, we resolve the conflict and then increase the model complexity. The conflict resolution will introduce case distinctions. These cases will be considered separately. The partitioning will decrease the currently considered complexity. Note that in each of the cases, we just have to consider rule abiding behaviour.

Analysing sets of traffic rules. We start the analysis of traffic rules by ignoring their environmental context. *„In villages, you may drive 50km/h." becomes „You may drive 50km/h."* We hence do not consider any environment or agent models for now. Moreover, we verify increasingly larger sets of traffic rule automata in parallel. If formal verification detects a conflict at this stage, it might turn out to be spurious, because the context of the rules is ignored. Adding context, as described below, will resolve such conflicts. However, if the conflicting rules are valid in the same context, the conflict must be resolved. For this purpose, we envision that the rules can often be hierarchically ordered. Note that, in order to encode that „must"-rules have higher priority than „should"-rules, we determine the cases where these rules „conflict". We then encode the precedence of „must" over „should" explicitly for these cases. To determine these conflicts, we treat „should"-rules as „must"-rule and resolve the conflicts during the analysis process.

Adding Context. For adding context in analysis of traffic rules, we will introduce separate models for each context and examine whether the conflict remains. *In villages, you may drive 50km/h and outside of villages you may drive 100km/h.* Without this distinction, the rules are conflicting. We hence consider two separate models, m_village and a model m_outsideOfVillage in the further analysis. By adding rule context, we expect that some previously identified conflicts will be solved. Note that by introducing separate models for different contexts, the respective environment models can be kept simpler. *The m_outsideOfVillage model will not have shared routes for cars and trams.* If the analysis of traffic rules with an included context does not resolve a conflict, we can add an ATA model.

Adding ATA Models. If we add ATA models, we might be able to resolve a conflict or to discover that an agent never has to face the conflict. That is, we can formally establish whether an agent never gets into a situation where conflicting rules are applicable. *A small ATA delivering pizza will not have to decide whether giving way to the emergency vehicle is more important than respecting the maximal load capacity of the bridge ahead, because it is so light-weight.* Moreover, adding ATA models entails adding its beliefs, desires and intentions. This might resolve conflicts, that are otherwise possible. *The pizza robot might be programmed to plan its manoeuvres very cautiously, so that it always has enough time to get out of the way.*

However, adding an ATA model might also lead to new conflicts; e. g. conflicts that exist between the ATA's behaviour and traffic rules. So, adding (preliminary) ATA models will also be helpful for debugging or optimising these ATAs. We expect to detect cases where an agent simply breaks a traffic rule because it is not aware of the rule. These cases have to be fixed by making the ATA obey the rules, thus changing beliefs, desires or intentions of the agent. In other cases, the ATA's goals will morally justify braking traffic rules.

Adding Morality. *If an ATA is on its way to deliver a heavily injured person to the hospital, but a really slow vehicle and dense oncoming traffic hinders the ATA to proceed, then mounting an empty sidewalk to bypass the obstacle might very well be acceptable. For instance, it could be stated that mounting an empty sidewalk is okay, if lives can be saved and neither injuries nor damages are caused.*

In general, an ATA often has to achieve some progress or maintain liveness (i. e. its ability to proceed to function). These system properties can be of great importance, *e. g. saving an injured*. However, since the moral valuation often will not be that clear, a hierarchy resolving conflicts between traffic rules and ATA's intentions cannot always be defined. Interdisciplinary input is needed, to determine how moral guidance for such cases can be encoded in a general manner. As discussed in Sect. 2, including morality means a high burden for the agent by increasing its information need. We hence aim to only identify situations where the traffic rules need to be enhanced. We then add moral aspects so that an ATA is able to follow the rules adequately. In summary, we see adding of morality as a means of resolving conflicts between traffic rules. Further on, we conclude the analysis of formalised traffic rules can help with identifying cases where ATAs need to reason morally. Regarding our overall endeavour to formalise the rules from traffic rule books, we consider it as future interdisciplinary work to develop approaches to concretise ambiguous phrases.

Related Work. Our preliminary work from Sect. 4.1 is not the only approach that suggests to formalise and analyse traffic rules. We give an overview over other approaches that examine conflicting traffic rules and implement conflict resolution strategies. With the (R)ules (o)f (T)he (R)oad (A)dvisor (ROTRA), a translation of natural language traffic rules into Prolog code is introduced in [CDF22]. ROTRA can provide recommended and possible actions that are generated from a set of the UK Highway Code rules to an ATA. In this approach, some conflicting traffic rules have been identified, but it has not been answered how an ATA should handle such conflicts. A combination of this approach with our work on moral rules would be ideal. In another approach, the authors of [NR22] introduce „LTL sketches“, with which incomplete LTL formulae can be formulated. With this, an engineer could partially formalise a traffic rule, even if they do not know some technicalities, thereby reducing the ambiguity of its natural language version. Further on, safety distance related traffic rules from the Vienna Convention on Road Traffic have been verified in [RIA16] using Isabelle/HOL. In further work [Ri17], some authors of [RIA16] use a verification procedure using Isabelle/HOL, where they start from abstractly formalised traffic rules which they then concretise in a verified manner. In [DMR20], we describe an approach to generate a traffic scenario catalogue. This approach leads to a coverage of the possible traffic evolutions of an abstract world model by the generated abstract scenarios. This work is of interest for our analysis part, where we add context to our traffic rule analysis. The authors of [Gr22] extend this work by introducing run-time monitors that allow for the analysis of the traffic scenarios at run-time. With this, new conflicts that must be handled using moral rules could be identified at run-time. In this vision paper, we focus on the role of moral when reasoning about traffic rules. However, besides moral reasoning, legal reasoning is needed when analysing driving decisions in the conflict situations that we have discussed. For this, the authors of [We22] open up another version of a conflict by describing the *Congruence Problem*: They link a system's interpretation of a formalised traffic rule with the legal interpretation of that traffic rule. With legal interpretation, they mean the judgement of a juridical expert. It is of interest to use their approach to verify that adding morality to ATAs helps with minimising the Congruence Problem.

5 Summary/ Conclusion

In this paper, we argue that ATAs need to be equipped with moral values in order to interpret traffic rules appropriately. The approach that we propose detects when traffic rules are conflicting. We suggest to characterise such situations in terms of their context and to make an ATA preserve moral values in this context. Since a moral evaluation has a high demand on contextual information, we restrict the scope of moral decision-making. We believe that in many cases appropriate moral values can be encoded into an ATA, so that a common sense interpretation of traffic rules can be achieved, but there are also many cases where encoding morality will be a very challenging and an interdisciplinary endeavour.

Literatur

[AA07] Anderson, M.; Anderson, S. L.: The status of machine ethics: a report from the AAAI Symposium. Minds and Machines 17/1, S. 1–10, März 2007.

[ADF20] Alves, G. V.; Dennis, L.; Fisher, M.: Formalisation and Implementation of Road Junction Rules on an Autonomous Vehicle Modelled as an Agent. In: Formal Methods. FM 2019 Int. Ws. LNCS, Springer, S. 217–232, 2020.

[ADF21] Alves, G. V.; Dennis, L. A.; Fisher, M.: A Double-Level Model Checking Approach for an Agent-Based Autonomous Vehicle and Road Junction Regulations. J. Sens. Actuator Networks 10/3, S. 41, 2021.

[AS22] Alves, G. V.; Schwammberger, M.: Towards a Digital Highway Code using Formal Modelling and Verification of Timed Automata. In: 4th Ws. on Formal Methods for Autonomous Systems (FMAS/ASYDE@SEFM) 2022. Bd. 371. EPTCS, S. 77–85, 2022.

[ASW05] Allen, C.; Smit, I.; Wallach, W.: Artificial Morality: Top-down, Bottom-up, and Hybrid Approaches. Ethics and Information Technology 7/3, S. 149–155, Sep. 2005, ISSN: 1572-8439.

[BDL04] Behrmann, G.; David, A.; Larsen, K. G.: A Tutorial on Uppaal. In: Formal Methods for the Design of Real-Time Systems. Springer, S. 200–236, 2004.

[BID18] Babanoski, K.; Ilijevski, I.; Dimovski, Z.: The traffic morale, ethics and culture among the road traffic users. Horizons 22/, S. 77–85, Nov. 2018.

[BS19] Bischopink, C.; Schwammberger, M.: Verification of Fair Controllers for Urban Traffic Manoeuvres at Intersections. In: Formal Methods. FM 2019 Int. Ws. Bd. 12232. LNCS, Springer, S. 249–264, 2019.

[Bu13] Bundesrepublik Deutschland: Straßenverkehrsordnung (StVO), de, 2013, URL: www.gesetze-im-internet.de/stvo_2013/index.html, Stand: 01. 04. 2013.

[CDF22] Collenette, J.; Dennis, L. A.; Fisher, M.: Advising Autonomous Cars about the Rules of the Road. In: 4th Ws. on Formal Methods for Autonomous Systems (FMAS/ASYDE@SEFM) 2022. Bd. 371. EPTCS, S. 62–76, 2022.

[De17] Department for Transport: Using the road (159 to 203) - The Highway Code - Guidance - GOV.UK, 2017, URL: www.gov.uk/guidance/the-highway-code/using-the-road-159-to-203, Stand: 13. 04. 2018.

[De21] De Freitas, J.; Censi, A.; Walker Smith, B.; Di Lillo, L.; Anthony, S. E.; Frazzoli, E.: From driverless dilemmas to more practical commonsense tests for automated vehicles. Procs. of the National Academy of Science 118/11, März 2021.

[DMR20] Damm, W.; Möhlmann, E.; Rakow, A.: A Scenario Discovery Process Based on Traffic Sequence Charts. In: Validation and Verification of Automated Systems: Results of the ENABLE-S3 Project. Springer, S. 61–73, 2020.

[Dr75] Drobnickij, O.: Ethics. In: Themes in Soviet Marxist Philosophy: Selected Articles from the ‘Filosofskaja Enciklopedija’. Springer, S. 185–199, 1975.

[Gr22] Grundt, D.; Köhne, A.; Saxena, I.; Stemmer, R.; Westphal, B.; Möhlmann, E.: Towards Runtime Monitoring of Complex System Requirements for Autonomous Driving Functions. In: 4th W. on Formal Methods for Autonomous Systems (FMAS/ASYDE@SEFM) 2022. Bd. 371. EPTCS, S. 53–61, 2022.

[Ma22] Max Planck Society for the Advancement of Sciene e.V.(web site provider): The Moral Machine, 2022, URL: www.moralmachine.net, Stand: 15. 07. 2022.

[Na20] Nallur, V.: Landscape of Machine Implemented Ethics. Science and Engineering Ethics 26/5, S. 2381–2399, Okt. 2020.

[NR22] Neider, D.; Roy, R.: Expanding the Horizon of Linear Temporal Logic Inference for Explainability. In: 30th IEEE Int. Requirements Engineering Conf. Ws. (RE 2022). IEEE, S. 103–107, 2022.

[Pi97] Picard, R. W.: Affective Computing. MIT Press, 1997.

[Pr17] Prakken, H.: On the problem of making autonomous vehicles conform to traffic law. Artificial Intelligence and Law 25/3, S. 341–363, Sep. 2017.

[Re21] Reed, N.; Leiman, T.; Palade, P.; Martens, M.; Kester, L.: Ethics of automated vehicles: breaking traffic rules for road safety. Ethics and Information Technology 23/4, S. 777–789, Dez. 2021.

[Ri17] Rizaldi, A.; Keinholz, J.; Huber, M.; Feldle, J.; Immler, F.; Althoff, M.; Hilgendorf, E.; Nipkow, T.: Formalising and Monitoring Traffic Rules for Autonomous Vehicles in Isabelle/HOL. In: Procs. of int. Conf. on integrated Formal Methods (IFM) 2017. Bd. 10510. LNCS, Springer, S. 50–66, 2017.

[RIA16] Rizaldi, A.; Immler, F.; Althoff, M.: A Formally Verified Checker of the Safe Distance Traffic Rules for Autonomous Vehicles. In: Procs. of NASA Formal Methods Symp. NFM 2016. Bd. 9690. LNCS, Springer, S. 175–190, 2016.

[SA18] SAE Int.: J 3016: Surface vehicle recommended practice – (R) Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles, 2018.

[SA21] Schwammberger, M.; Alves, G. V.: Extending Urban Multi-Lane Spatial Logic to Formalise Road Junction Rules. In: 3rd Ws. on Formal Methods for Autonomous Systems (FMAS) 2021. Bd. 348. EPTCS, S. 1–19, 2021.

[Sc18] Schwammberger, M.: An abstract model for proving safety of autonomous urban traffic. Theoretical Computer Science 744/, S. 143–169, 2018.

[Se74] Sen, A.: Choice, Ordering and Morality. In: Practical Reason. Blackwell, 1974.

[Si22] Singer, P.: ethics, 2022, URL: `www.britannica.com/topic/ethics-philosophy`, Stand: 04. 12. 2022.

[SK22] Schwammberger, M.; Klös, V.: From Specification Models to Explanation Models: An Extraction and Refinement Process for Timed Automata. In: 4th Ws. on Formal Methods for Autonomous Systems (FMAS/ASYDE@SEFM) 2022. Bd. 371. EPTCS, S. 20–37, 2022.

[Sl22] Slavkovik, M.: Automating Moral Reasoning. In: Int. Research School in Artificial Intelligence in Bergen (AIB 2022). Bd. 99. OASIcs, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 6:1–6:13, 2022.

[Th17] Thornton, S. M.; Pan, S.; Erlien, S. M.; Gerdes, J. C.: Incorporating Ethical Considerations Into Automated Vehicle Control. IEEE Transactions on Intelligent Transportation Systems 18/6, S. 1429–1439, 2017.

[To21] Tolmeijer, S.; Kneer, M.; Sarasua, C.; Christen, M.; Bernstein, A.: Implementations in Machine Ethics: A Survey. ACM Comput. Surv. 53/6, 2021.

[US22] U.S. Department of Transportation (DOT): Mission and Vision, 2022, URL: `www.grants.gov/learn-grants/grant-making-agencies/department-of-transportation.html`, Stand: 02. 12. 2022.

[Va15] Vaughn, L. W.W. Norton & Company, 2015.

[We22] Westhofen, L.; Becker, J. S.; Hagemann, W.; Möhlmann, E.; Stierand, I.: Towards a Sound and Complete Interpretation of Traffic Rules for Automated Driving – Experiences and Challenges. In: Procs. of Methodologies for Translating Legal Norms into Formal Representations. to be published, 2022.