On Making Autonomous Vehicles Respect Traffic Law: a Case Study for Dutch Law

Henry Prakken
Utrecht University, Department of Information and Computing Sciences and University of Groningen, Faculty of Law,
The Netherlands
h.prakken@uu.nl

ABSTRACT

Among the problems that still need to be solved before autonomous vehicles can fully autonomously participate in traffic is the one of making them respect the traffic laws. This paper discusses this problem by way of a case study of Dutch traffic law. First it is discussed to what extent Dutch traffic law exhibits features that are traditionally said to pose challenges for AI & Law models, such as exceptions, open texture and vagueness and the need for commonsense knowledge. Then three approaches to the design of law-respecting AV are evaluated in light of the challenges posed by Dutch traffic law.

CCS CONCEPTS

• **Computing methodologies → Mobile agents**; • **Applied computing → Law**;

KEYWORDS

Autonomous vehicles, Traffic law

ACM Reference format:
Henry Prakken. 2017. On Making Autonomous Vehicles Respect Traffic Law: a Case Study for Dutch Law. In *Proceedings of ICAIL '17, London, United Kingdom, June 12-16, 2017,* 4 pages.
https://doi.org/10.1145/3086512.3086542

1 INTRODUCTION

Autonomous vehicles are one of the most spectacular recent developments of Artificial Intelligence. While currently allowed technology is limited to features such as adaptive cruise control, parking assistance with automated steering and lane keeping assistance, fully autonomous vehicles, which can drive to any location where it is legal to drive and make their own decisions without human intervention, may well be able to take part in ordinary traffic within the next decades [1]. Among the problems that need to be solved is the one of making autonomous vehicles (AV) respect the traffic laws. Solutions to this problem may well profit from computational models of legal reasoning but so far the field of AI & Law has hardly addressed this issue. The present paper aims to put this topic on the AI & Law research agenda by way of a case study of Dutch traffic law and its implications for the design of fully autonomous self-driving cars. In the literature on AV design there have to the best of my knowledge so far not been any systematic studies of the problem of respecting traffic law (as also observed by [7]). For example, in [8] the problem is not even mentioned while yet this paper discusses the DARPA Urban challenge, in which the AV had to obey California traffic rules. Therefore, the present study, while still a conceptual one, fills an important gap in the literature.

The problem of making AV respect traffic law is a special case of the more general problem of making intelligent autonomous systems respect the relevant laws. Note that this problem arises irrespectively of the legal question whether machines can be assigned responsibility in a legal sense. Even if a human remains legally responsible for the actions of the machine, the human faces the problem of ensuring that the machine behaves in such a way that the responsible human complies with the law. Currently, this kind of problem is mainly studied under the heading of 'machine ethics' [2]. While this may be the appropriate field for studying the related problem of making intelligent autonomous systems behave ethically responsibly, the problem of making them respect the law arguably belongs to AI & Law.

2 THE CLASSIC AI & LAW PROBLEMS VS THE NEW CHALLENGE

The task of making intelligent autonomous systems respect the law has some similarities and differences with more traditional tasks modelled in AI & Law. Just as in, for instance, legal decision or argumentation support, the task is to apply norms to facts in order to legally classify behaviour. Therefore, the task of making autonomous systems respect the law faces some of the same challenges as any task in which behaviour has to be legally classified, in particular the possibility of rule conflicts, ambiguities in the formulation of legal rules, the open-textured and vague nature of many legal concepts and the possibility of unforeseen exceptions on the basis of purpose or principle.

However, there are also differences. First, while AI & Law research has traditionally focused on support tools for humans carrying out legal tasks, with autonomous systems this is different: they do not support humans in their legal tasks (although they may support humans in other tasks) but they have to decide about the legal status of their own actions. In many cases it will be impossible for humans to check or override the system's decision.

Moreover, while tasks supported by traditional AI & Law tools often concern the application of the law to past cases, to determine the legal status of some past behaviour or some existing state of affairs, autonomous systems have to determine the legal status of their future actions. Among other things, this means that autonomous

ICAIL '17, June 12-16, 2017, London, United Kingdom

ACM ISBN 978-1-4503-4891-1/17/06…$15.00
https://doi.org/10.1145/3086512.3086542

systems do not face evidential problems in the legal sense. Even when traditional AI & law supports legal tasks with an eye to the future, such as deciding on benefit applications, drafting regulations or contracts or designing tax constructions, there are differences with autonomous systems. While traditionally supported future-oriented tasks concern behaviour in the non-immediate future and often contain classes of actions (as with contract drafting or regulation design), autonomous systems have to 'run-time' consider individual actions in the immediate future.

Another difference is that while the legal tasks traditionally modelled in AI & law require explanation and justification of decisions, with autonomous systems there is less need for this, since the primary problem is to generate legally acceptable behaviour. Therefore, the black-box nature of data-mining or machine-learning techniques will, unlike with traditional legal tasks, be less of a problem for the task of making autonomous systems respect the law.

Next, while much AI & Law research studies legal tasks in an adversarial setting (primarily a legal proceeding), with the task of letting autonomous systems respect the law there are no adversaries: all that counts is to let the system do what it has to do within the bounds of the law. Thus there will be less need for argumentation than as usual in AI & Law applications.

Yet another difference is that one may expect that the bulk of the cases encountered by an autonomous system will from a legal point of view be standard, mundane cases. For example, autonomous cars will not have to determine the legal responsibility for car accidents but will have to decide about driving from A to B in a way that respects the traffic regulations. While processing legislation in public administration also usually concerns standard cases, in the court room this is different.

Finally, the tasks traditionally modelled in AI & Law are usually strictly legal while autonomous systems have to balance legal considerations against other considerations. Autonomous systems are not designed to obey the law but for other purposes, such as driving from A to B. Sometimes behaviour fulfilling this purpose is from a legal point of view illegal but still socially acceptable; for example, slightly speeding in a queue of cars that all drive a few miles above the maximum speed, or passing a vehicle that is standing still where changing lanes is forbidden but there is no approaching traffic. This means that the behaviour of autonomous systems should not be seen as rule-governed but as rule-guided in that legal rules are just one factor influencing socially optimal or permissible behaviour. Other factors are e.g. social conventions, individual or social goals or simply common sense. And sometimes these other factors override the legal factors. Having said so, even rule-guided models of autonomous systems will have to specify what the law requires, and this is the problem discussed in this paper.

3 DUTCH TRAFFIC LAW

The main Dutch traffic regulations are the Road Traffic Act 1994 (Wegenverkeerswet 1994, WVW) and the Traffic Rules and Signs Regulations 1990 (Reglement Verkeersregels en Verkeerstekens 1990, RVV). The WVW hierarchically precedes the RVV. The Dutch traffic regulations are (like presumably in all jurisdictions) designed to promote two purposes: safe and efficient traffic. These purposes are codified in Article 5 WVV, which states that 'It is forbidden to behave in such a way that danger on the road is or may be caused or that the road traffic is or may be impeded'.

Traffic regulations apply to a relatively simple, closed and predictable world, at least compared to may other legal domains. This is one reason why the Dutch traffic regulations are rather precise and concrete (but with some exceptions to be discussed below). Another reason is the need for humans to safely and efficiently coordinate their actions in traffic, which requires clear and precise rules.

Yet Dutch traffic law contains a considerable number of vague and open-textured terms, which is an obvious potential obstacle for making AV respect traffic law (cf. [7]). For example, Article 5 WVW is is clearly vague and open-textured with concepts like causing danger on the road, impeding traffic, guilt, serious bodily harm, and causation of harm or damage. However, its vague notion of causing danger is to a large extent made precise in the RVV, since its rules are meant to enforce safe traffic behaviour. Article 5 WVW thus has two roles: as a fall-back option in case dangerous behaviour is not forbidden by specific RVV rules, and as a general implicit exception to specific RVV rules in case otherwise permitted or obliged behaviour would cause danger to or impediment of traffic.

An interesting case in this respect is that of a self-driving Google car which was stopped by the California police for driving too slowly. Google had for safety reasons set the car's maximum speed for roads in a 35mph zone at 25mph and one of its cars was causing a big queue of traffic while driving 24mph.[1] This behaviour was held by the police to violate the following rule from the Californian traffic regulations:

> No person shall drive upon a highway at such a slow speed as to impede or block the normal and reasonable movement of traffic, unless the reduced speed is necessary for safe operation, because of a grade, or in compliance with law.

Under Dutch law this would be a violation of Article 5 WVW.

The RVV also contains several vague or open-textured terms. For example, Article 3 para 1 WVW says 'Drivers shall keep to the right as much as possible'. This should arguably not be read in a physical sense but in the sense of what is reasonably right as much as possible given the traffic context. This reading makes the term open-textured. Some other examples are Article 28 RVV: sound or light signals are only allowed to avert a *threat of danger* and Article 43(3) WVW: on motorways, use of the shoulder is only allowed *in case of emergency*.

Other concepts in the RVV are clear for humans but require some recognition or judgement by AV, on the recognition of objects (such as tunnels, bicycle lanes, crossings and roundabouts) or of spatial relations (on or just before, at a short distance, blocking, making free), or of behaviours of humans (indicating to turn, moving with difficulty) or on observability (serious inhibition of visibility).

The main function of **civil liability rules** is to determine liability once something has happened. Currently, there is much discussion among legal scholars on who should be held liable when an AV causes an accident [1]. However, for present purposes traffic

[1] http://www.bbc.com/news/technology-34808105, accessed 21 December 2016.

liability rules are relevant in a different way: they can yield additional cues for driving behaviour in cases where no specific traffic rule is violated. The Dutch traffic liability rules are meant to protect the more vulnerable road user against the stronger. Accordingly, motor vehicle drivers are almost always held liable towards non-motorised traffic participants. Even if two motor vehicles collide, the norms are stringent in that only perfect drivers can fully escape liability. The underlying principle is that cars are inherently dangerous and the consequences of even minor traffic mistakes can be very serious, so people who decide to use cars should be encouraged to drive according to the highest possible standards.

4 AUTONOMOUS VEHICLES AND DUTCH TRAFFIC LAW

In this section I discuss some challenges that Dutch traffic law presents for autonomous vehicles by considering three different approaches to the design of law-respecting AV and evaluating each of them in light of the challenges posed by Dutch traffic law.

4.1 Three Approaches to Achieving Norm Compliance

One solution to the problem of making AV respect traffic law is to design the system in a way that guarantees that the system will not exhibit unwanted behaviour. This is the conventional solution when non-autonomous machines, tools or systems are used, sometimes called *regimentation*. A similar approach has been proposed for autonomous systems by [5], who proposes to verify the behaviour of systems off-line with so-called model-checking techniques. Note that in both cases good practice in AI & Law ([3]) requires that the regimented AV design must be linked to the relevant law for purposes of validation and maintenance. In this respect, [7] observe that current designs of AV do not have an explicit traffic law model but that a usual approach to test AV is with 'play books', which contain scenarios the AV must be able to handle. They warn that the traffic law model implicit in such scenarios may be incomplete or oversimplified, while testing whether this is the case is difficult since the traffic model remains implicit.

A limitation of the regimentation approach is that when systems are increasingly autonomous and have to operate in increasingly complex environments, their input and behaviour cannot be fully predicted, so that regimentation or advance off-line model checking is impossible or of limited value. How can norm compliance then be ensured? The question then arises whether an autonomous system should be designed to *reason* about how to behave lawfully or whether it can be *trained* to do so with machine-learning techniques applied to a large number of training cases. In the first approach there is the obvious need for explicit representation of legal information in the system and for giving the system explicit reasoning and decision making capabilities. This is still somewhat similar to the traditional AI & law systems for supporting human decision making, except that the human is taken out of the loop. An important issue then is whether the mundane nature of cases faced by the autonomous system can reduce the complexity of the classification and interpretation problems to such an extent that the machine can fully take over. On the other hand, the reasoning can, unlike in the traditional settings, be opaque in that there is less need for explaining or justifying why the behaviour is legally correct. Incidentally, the latter combined with the run-time and forward-oriented setting with mundane cases, makes that current AI & Law research on evidential legal reasoning and sophisticated legal argument will likely be less relevant here.

The other approach is that the ability to behave legally correctly is acquired implicitly by training. The currently usual approaches to designing AV to a large extent rely on machine learning approaches. For example, the March 2016 edition of the Google Self-Driving Car Project Monthly Report[2] says

> . . . rather than teaching the car to handle very specific things, we give the car fundamental capabilities for detecting other road users or unfamiliar objects, and then we give it lots of practice in a wide range of situations.

This approach is similar to regimentation in that it aims to equip the AV with law-respecting behaviours without giving it explicit normative reasoning capabilities. However, it differs from regimentation in that it does not aim to fully *guarantee* correct behaviour. Accordingly, validation of correct behaviour is not done by formal means but by empirical testing. For very advanced autonomous systems, like robots operating in daily life, this approach might be equivalent to solving the notorious AI common-sense problem, but for more modest systems this approach might be more realistic. One interesting question is how autonomous vehicles classify on this scale. Below we will discuss some interpretation and classification problems in Dutch traffic law that are relatively easy for humans but seem very hard for the current generation of autonomous vehicles. Finally, note that like with the regimentation approach, the AV design should in agreement with good practice in AI & Law be linked to the relevant law.

4.2 Required abilities of AV

I now discuss which abilities an AV that can fully autonomously drive in all Dutch traffic situations should have.

As regards interpreting sensor data, current AV design mainly seems to be concerned with relatively 'low level' cognitive capabilities like determining the own and other road users' location, speed and direction, distinguishing drivable from non-drivable areas and recognising obstacles. According to [7] current AV technology does not yet allow for advanced object recognition, while yet this is required by Dutch traffic law. [7] discuss the example of Article 16 RVV, which states that 'Road users must not cut across military columns and motorised funeral processions'. This requires the ability to classify vehicles. Similarly, as we saw in Section 3, the AV needs to recognise other types of objects, such as tunnels or bus stops. For humans such classification tasks are straightforward; the (legal or commonsense) definitions of the various types of vehicles and other objects are generally precise. The added complexity compared to traditional AI & Law is that the AV needs to perform such classification from its sensor data. This is a major problem, since according to [1] making sense of sensor data is probably the hardest part of designing fully autonomous AV.

[2] https://static.googleusercontent.com/media/www.google.com/lt//selfdrivingcar/files/reports/report-0316.pdf (accessed January 5, 2017).

Important kinds of objects to be recognised are traffic signs, traffic lights and road lining. A usual approach here is letting the AV use maps on which these things are indicated. This has two obvious limitations: permanently changed traffic situations not yet incorporated in the maps, and temporary changes, such as the directions given by authorised officials on the basis of Article 12 WVW or Article 82 WVV. According to [6]:

> Google's cars can detect and respond to stop signs that aren't on its map, a feature that was introduced to deal with temporary signs used at construction sites.

According to [6], Google's reply to this is as follows:

> Google says that its cars can identify almost all unmapped stop signs, and would remain safe if they miss a sign because the vehicles are always looking out for traffic, pedestrians and other obstacles.

But this has the potential problem that the resulting behaviour does not comply with the second purpose of Dutch traffic law, namely, to promote efficient traffic. Overly cautious behaviour might result in a violation of Article 5 WVW's prohibition to impede other traffic. See also the case of the Google car driving too slowly discussed above in Section 3, and see [6], who remarks that

> But in a complex situation like at an unmapped four-way stop the car might fall back to slow, extra cautious driving to avoid making a mistake.

A step further than object classification and recognition of traffic signs and lights and road lining is situational awareness and interpretation. For example, an AV should be able to distinguish between ordinary pedestrians (merely to be avoided) and police officers giving directions. [6] discusses how the Google car currently deals with this:

> Pedestrians are detected simply as moving, column-shaped blurs of pixels – meaning, (...), that the car wouldn't be able to spot a police officer at the side of the road frantically waving for traffic to stop.

While the Google car may thus avoid colliding with a police officer, it may fail to obey the officer's directions.

Article 19 RVV, which states that 'A driver must at all times be able to bring his vehicle to a standstill within the distance that he can see to be clear', seems within the capabilities of current AV technology. However, the behavioural cues arising from the traffic liability rules require that an AV has sophisticated means to interpret a situation and its context from its sensor data. For example, an AV must be able to adjust its speed to the state of the road, the type of environment (country side, busy shopping street, and so on), to unusual weather conditions, to the presence of special features like bus stops, pedestrian crossings, objects besides the road that block or impair the car's view, and to the presence or absence of foot paths or bicycle paths. Moreover, knowledge of the local context may also be relevant. For example, in one Dutch case the court remarked that in a large city like Amsterdam it is not unusual for cyclists to ride their bicycle in the dark without light. If capabilities like object classification or recognising types of persons is still to a large extent beyond current AV technology, then the same can be inferred from this for these more advanced capabilities to interpret a situation from the AV's sensor data.

Finally, as remarked by [7], the purpose to promote both safe and efficient traffic requires that an AV has the ability to recognise and respect social cues. For instance, it should be able to interpret gestures by or eye contact with other drivers in complex traffic situations. This, too, seems a hard problem for AV.

5 CONCLUSION

This paper has studied the problem of making autonomous vehicles respect traffic law by way of a case study of Dutch traffic law. For AI & Law it has put a new topic on its research agenda, while for AV designers it has hopefully created a deeper awareness of the challenges that traffic law poses for AV.

As for related research, there is currently much attention for so-called 'moral algorithms' for letting AV deal with situations akin to moral dilemmas. For example, [4] study moral decision problems that AV could face by presenting people with several simple situations and asking them what the AV should do. In one such situation the AV has to choose between killing a pedestrian or killing the AV's passenger by driving into a wall. [4] conclude that it is important to study which moral algorithms should be programmed into AV. From our legal point of view, this importance can be challenged. As we saw above, In Dutch traffic liability law the main emphasis is not on what to choose in dilemmas such as the ones discussed by [4] but on anticipation, that is, on how much care has to be taken to avoid such situations. Therefore, instead of focusing on moral algorithms for situations of the kind studied by [4], the first priority should arguably be to study 'legal algorithms' for avoiding such situations, that is, for careful and prudent driving behaviour that yet does not make the traffic too slow to be efficient.

The present study has focused on fully autonomous self-driving cars. For less advanced stages of automation-assisted driving, some of the negative conclusions may not hold or may hold to a lesser extent while other conclusions may still hold with the same force, depending on the degree and kind of autonomy of the AV. Moreover, in the future traffic laws may be adapted to the presence of AV on the road. Nevertheless, the present paper has in any case laid the foundations for a systematic study of the problem of making autonomous vehicles respect the law.

REFERENCES

[1] J. M. Anderson, N. Kalra, K.D. Stanley, O. Sorensen, C. Samaras, and O.A. Oluwatola. 2016. *Autonomous Vehicle Technology. A Guide for Policy Makers.* RAND Corporation, Santa Monica, CA.

[2] M. Anderson and S.L. Anderson (Eds.). 2011. *Machine Ethics.* Cambridge University Press, Cambridge.

[3] T.J.M. Bench-Capon and F.P. Coenen. 1992. Isomorphism and legal knowledge based systems. *Artificial Intelligence and Law* 1 (1992), 65–86.

[4] J.-F. Bonnefon, A. Shariff, and I. Rahwan. 2016. The social dilemma of autonomous vehicles. *Science* 352 (2016), 1573–1576.

[5] J. Broersen. 2014. Responsible intelligent systems. The REINS Project. *Kuenstliche Intelligenz* 28 (2014), 209–214.

[6] Lee Gomes. 2014. Hidden obstacles for Google's self-driving cars. *MIT Technology Review* August 28 (2014). Available at https://www.technologyreview.com/s/530276/hidden-obstacles-for-googles-self-driving-cars/.

[7] R.E. Leenes and F. Lucivero. 2014. Laws on robots, laws by robots, laws in robots: regulating robot behaviour by design. *Law, Innovation and Technology* 6 (2014), 193–220.

[8] S. Thrun. 2010. Toward robotic cars. *Commun. ACM* 53 (2010), 99–106.