Incorporating Ethical Considerations Into Automated Vehicle Control

Sarah M. Thornton, Selina Pan, Stephen M. Erlien, and J. Christian Gerdes

***Abstract*— Not only do automated vehicles need to meet specifications for technical performance, they also need to satisfy the societal expectations for behavior in traffic with humans. Societal expectations, such as accident avoidance and adherence to traffic laws, have their foundation in core moral issues found in philosophy and ethics. Thus, engineers designing control algorithms for automated vehicles can benefit from applying principles and frameworks from philosophy to drive design decisions. In particular, we use a set of ethical frameworks to map design decisions for a model predictive control problem to philosophical principles. Deontology, a rule-based ethical framework, motivates the development of constraints on the system. Consequentialism, a cost-based ethical framework, motivates the construction of the objective function. The choice of weights is guided by the concepts of virtue ethics and role morality to determine behavior for different types of vehicles. The strong link between ethical principles and actual vehicle behavior developed through this approach is demonstrated experimentally by implementing alternative design choices on a test vehicle in a simple driving scenario.**

***Index Terms*— Autonomous vehicles, vehicle control, ethics, model predictive control.**

I. Introduction

DESIGNING control algorithms for automated vehicles presents new challenges for engineers. Traditionally, control systems have desired specifications and performance measures to which programmers design the control algorithms. For fully automated vehicles, the ultimate desired performance outcome is the ability to drive safely and smoothly through traffic. Setting specifications for achieving this desired outcome is challenging because traffic situations involving humans are not easily quantified. Driving through traffic requires that the vehicle conform with societal expectations for roadway behavior. Expectations such as collision avoidance and observance of traffic laws go beyond mere technical specification to touch on moral issues that are long established and formally characterized in philosophy.

The objective of accident avoidance, for instance, is fundamentally motivated by the idea of caring for life and avoiding harm. Haidt described care (and its opposite, harm) as one of the foundational principles for moral reasoning [1]–[4]. The vehicle's compliance with traffic laws involves another moral foundation - the degree of respect for authority. Interactions with other road users should be furthermore based on fairness and reciprocity, yet another category in the five (or sometimes six) moral foundations found in Haidt's work. The fact that these societal expectations of automated vehicles map so cleanly to ethical principles in philosophy suggests that philosophy can be a useful resource for translating such expectations to specifications. In the same realm of ethics in engineering but with different applications, Mladenovic and McPherson [5] draw from social justice in the design of a traffic control framework for automated vehicles. Miller *et al.* [6] and Van den Hoven *et al.* [7] all advocate for the consideration of ethics throughout the entire engineering process. In particular, Miller et al. summarizes ethics models used in operations research and extends it to ethical decision-making machines. Lin also makes the case that ethical issues are central to the design of automated vehicles [8].

But as both Lin *et al.* [9] and Wallach and Allen [10] point out, a single philosophical framework is unlikely to be sufficient for programming autonomous systems. As a result, researchers have proposed solutions that combine different concepts from philosophy. Deontology, a rule-based ethical framework, and consequentialism, a cost-based ethical framework, both contribute structured guidelines for vehicle behavior. Goodall presents a three-tiered system for ethical decision-making in automated vehicles [11]. The first tier is a rational approach in which the vehicle follows the ethical principles of deontology and consequentialism. The other tiers involve artificial intelligence and a combined rational-artificial intelligence approach. Gerdes and Thornton present the two ethical frameworks of deontology and consequentialism as parallel to constraints and cost, respectively, in an optimal control problem [12]. Since many semi-autonomous and autonomous vehicles are already designed based on this type of control formulation (Gray *et al.* [13], [14]; Gao *et al.* [15]; Erlien *et al.* [16], [17]; Falcone *et al.* [18]), such an approach makes the links between philosophy and engineering quite direct. Thus, our work uses the notion of multiple ethical frameworks along with parallels between deontology, consequentialism and a constrained optimization problem as a starting point.

Our ultimate goal is to use ethical principles to make engineering decisions that result in reasonable, justifiable automated vehicle behavior. In particular, we use approaches in normative ethics, deontology and consequentialism, to cast driving goals as rules or costs as appropriate. These provide

Manuscript received April 15, 2016; revised July 8, 2016; accepted September 1, 2016. Date of publication October 11, 2016; date of current version May 29, 2017. This work was supported by the Daimler and Benz Foundation. The work of S. M. Thornton was supported by NSF through a Graduate Fellowship. The Associate Editor for this paper was J. Miller.
(Corresponding author: Sarah M. Thornton.)

The authors are with the Dynamic Design Laboratory, Department of Mechanical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: smthorn@stanford.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TITS.2016.2609339

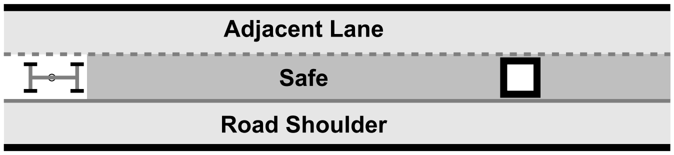

Fig. 1. The highlighted regions indicate safe driving regions. The safest region is the current lane of the vehicle minus any obstacles. Driving regions decrease in safety as the vehicle departs the lane.

constraints and cost functions for an optimization problem with a Model Predictive Control (MPC) structure that can be solved in real-time. This enables ethically motivated design decisions to be demonstrated and compared on a test vehicle for a simple traffic scenario. As with any optimization problem, there is a challenge in choosing appropriate weights for different objectives. To assist with this, we employ an ethical theory known as virtue ethics in the form of role morality. Virtue ethics and role morality provide an ethical framework based on alignment with character (Hursthouse [19], Harman [20]). As applied to vehicles, this framework guides the algorithm design to achieve desired behavior for different types of vehicles. To the knowledge of the authors, this is the first quantitative and in-vehicle experimental endeavor to incorporate ethics in the design of autonomous vehicle control.

This paper is structured as follows. Section II presents a motivating driving scenario based on a lane blocked by an obstacle. Section III describes how the philosophical concepts of deontology and consequentialism relate to engineering choices. These ideas combine in Section IV, which maps driving goals associated with the simple scenario to their respective philosophical concepts. The driving goals considered are path tracking, steering, obstacle avoidance and traffic laws. While many of these are straightforward, traffic laws can combine elements of cost and constraint. Section V formalizes the MPC formulation and Section VI shows in-vehicle experimental results, highlighting the impact of different formulations of traffic laws. This raises a larger question of virtue ethics and role morality as applied to automated vehicles, discussed with additional in-vehicle experiments in Section VII.

II. Scenario

To contextualize the relationship between ethics and engineering in autonomous vehicles, we construct a simple, realistic driving scenario that involves a variety of factors, including collision avoidance, mobility considerations and traffic laws. Figure 1 depicts the scenario of an autonomous vehicle traversing a two-lane roadway at constant speed. The current lane is obstructed by an obstacle ahead of the vehicle. This simple scenario prompts a range of possibilities for engineering decisions. Section V details how the various parameters from this scenario (collision avoidance, mobility, traffic laws, speed) fit into the controller.

One engineering design option is to program the vehicle to prioritize the ability to continue moving. This would mean entering the adjacent lane or the road shoulder to move around the obstacle to continue on its way. If the vehicle chooses the option of entering the adjacent lane, it could cause the vehicle to briefly violate a traffic law, for example, if the lane divider is a double solid yellow line. If the vehicle travels into the road shoulder to move around the obstruction, it obeys the double yellow line traffic law and continues to move, but the road shoulder would need to be accessible and safe, and is not meant for hosting long periods of travel. Clearly, competing demands of safety, legality and mobility must be weighed when considering theses options.

With a potential sacrifice in mobility, another approach is to program the automated vehicle to adhere strictly to traffic laws. Here, the underlying assumption is that traffic laws are implemented as hard rules. However, such strict obedience to traffic laws has consequences: the vehicle could be stopped indefinitely if trapped between the dual goals of avoiding the obstruction and obeying the double yellow line. This action could negatively impact the mobility or safety of surrounding vehicles.

The scope of engineering decisions that must be determined does not merely lie in the type of action to take; the degree of the action needs to be assessed too. For example, in the case of crossing the double yellow line to maneuver around the obstruction and encouraging smooth traffic flow, the amount of clearance between the vehicle and the obstruction is a design consideration. A narrow space between the vehicle and obstruction allows the vehicle to stay closer to its original designated lane should oncoming traffic emerge, but can increase the risk of brushing against the side of the obstruction. A wider berth ensures passage without hitting the obstruction, but the vehicle is then positioned further into the adjacent lane and would take longer to return to its original designated lane. The degree to which a vehicle is tuned to disobey a traffic law is therefore another engineering decision that requires careful ethical consideration. In addition to the type and degree of action taken, another layer of engineering design decisions involves the fact that different types of vehicles could be given different traffic law clearances depending on their expected role in society. We discuss this later on through the example of an ambulance and taxi, which exhibit very different behavior on roads due to their differing purposes.

Deconstructing this simple and common driving scenario highlights the many different vehicle behaviors that can result from making different engineering design decisions and underscores the need to make those design decisions in a systematic, reasoned and justifiable manner. The decisions should be explainable not only to engineers but also to other road users that will share the streets with the automated vehicle and to regulators who ensure traffic safety. Incorporating ethical frameworks from philosophy systematically into engineering design choices can help achieve these objectives.

III. Philosophical Frameworks

Ethical principles have been a topic of analysis among philosophers for centuries. We now examine those principles in the context of framing vehicle behavior. An autonomous vehicle is programmed by an engineer, and that programming adheres to a system of decision-making and control logic. Although control logic and ethics are clearly not equivalent, there do exist ethical frameworks that provide applicable motivation to mathematical frameworks, which we examine in this section.

Deontology is one of the major normative ethical theories. Deontological ethics follow a set of rules that determine the correct, ethical action, and these rules are to be followed with no exception. Isaac Asimov's Three Laws of Robotics [21] are an example of deontological ethics, which state:

1) A robot may not injure a human being, or, through inaction, allow a human being to come to harm.
2) A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
3) A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.

The Three Laws of Robotics state a clear architecture of behavioral rules for the robots in Asimov's stories to follow, effectively serving as constraints on the robots' behavior. As long as a robot remains within the conditionals of the Three Laws, it can operate as necessary. Of course, Asimov also demonstrated the limits of such an approach as his robot stories often involved strange behavior arising from conflicts between these laws.

Thus, deontology provides one type of motivating structure for the programming of automated vehicle algorithms: rules that can be defined and followed on the road. Such rules are analogous to conditionals and constraints used in decision-making and control algorithms to bound and manipulate system behavior (for example, a conditional for actuation saturation or a constraint in an optimization problem). For an autonomous vehicle, examples of such rules can be found in constraints designed to prevent the vehicle from causing harm to human beings, from inducing property damage to itself or to other objects, or from violating traffic laws. A key feature of a deontological framework is that rules can be hierarchical, thus setting clear priorities. From a programming perspective, the ability to weave together dependencies and hierarchies provides the advantage of clarity in reasoning for the development of the algorithm.

Another central normative ethical theory is *consequentialism*, which evaluates the moral validity of actions solely based on their consequences. We focus on a form of consequentialism known as *utilitarianism*, which analyzes the expected utility of a scenario and evaluates the consequences of actions based on what produces the most good [22]. The guiding principle is to always achieve the best outcome, i.e., "the ends justify the means."

Consequentialism, through its more specific form of utilitarianism, provides a basis for casting ethical decision making as an optimization problem. In control theory, optimal control uses an optimization problem to mathematically determine an optimal solution to be used as the control action. Specifically, the optimal control action (i.e., the ethically correct decision) is the feasible solution that minimizes the cost function (i.e., the desired outcome toward which one strives). An example of such a cost function for autonomous vehicles could be to minimize harm to the vehicle occupants. The optimal solution in consequentialist terms would be to maneuver the vehicle to achieve that goal of minimizing harm, no matter what. This approach also has some limitations, such as the difficulty in actually forming or evaluating the cost function (as is the case with definitions such as "harm" [12]) or making that cost function comprehensive (by, for instance, considering road users other than the occupants in this case).

IV. Design Choices

The frameworks of deontology and consequentialism are the result of extensive research in philosophy. Given their seminal presence in philosophy, we use these frameworks as tools to motivate and explain design decisions in the programming of automated vehicles. In this paper, we adopt an MPC formulation of the problem since the explicit consideration of constraints and costs in MPC maps well to the concepts of deontology and consequentialism. The following sections demonstrate how approaching the problem with these two philosophical frameworks in mind leads to a systematic treatment of different objectives in the problem. Specifically, we set constraints that direct the vehicle to avoid collisions, follow dynamical equations and stay within its steering capabilities. In the cost function, we specifically seek to direct the vehicle toward the desired outcomes of tracking a prescribed path and providing acceptable occupant comfort. In contrast with these other objectives, it is less clear whether traffic laws represent a cost or constraint so different representations are explored.

A. Path Tracking

A key objective of an autonomous vehicle is to follow a designated path. Following a path is a physical condition based on a measure of position difference, so a deontological constraint could be used to ensure path tracking. This would take the form of an equality constraint for the position of the vehicle to be equal to the desired position on the path. Upon further examination, however, following the path is not a strict requirement when it comes to maintaining safety; if an obstacle appears in the path, then the vehicle should have the option to deviate. Accounting for this line of reasoning, a cost function alternatively can serve as the instrument that accomplishes the goal of tracking a path via optimization, as depicted in Fig. 2. Thus, in the choice of framing path tracking in a consequentialist manner, we allow the vehicle freedom to deviate; if path tracking were denoted as a rule in a deontological framework, then a safety hazard arises in rule conflict and problem feasibility. Because the fundamental idea behind deontology requires that rules are to be followed without exception, we bypass it in favor of the more flexible principles from consequentialism for the specific goal of path tracking.

We now translate the objective of path tracking in a consequentialist form into a mathematical structure. In order to follow the path, the vehicle must achieve the two goals of minimizing lateral deviation from the path, e, and heading error, $\Delta\psi$, using a cost function, such as

$$J_x = \sum_k x^{(k)\top} Q^{(k)} x^{(k)} \tag{1}$$

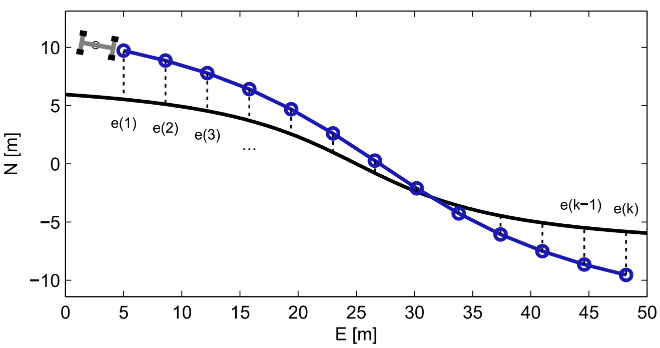

Fig. 2. Generating a cost from the difference between a desired path (black) and the vehicle's actual path (blue with dots).

where x is the vehicle state vector encompassing e and $\Delta\psi$ explained in further detail in Section V-A and the Appendix, k is the discrete time step in the prediction horizon, and the weight matrix Q only has nonzero entries corresponding to e and $\Delta\psi$.

B. Steering

The vehicle steering encompasses a few different design goals. The steering must operate within the actuator limits, should contribute toward path tracking and obstacle avoidance, and should perform smoothly. The first of those goals, operating within the actuator limits, can be cast in the form of a constraint on a maximum slew rate. The reasoning behind this design choice is that this cap is a physical limit on an actuator; physical limits must be highly prioritized in the control scheme and therefore are enforced as constraints in the deontological sense of strict rule obedience. A solution requiring control inputs that violate physical limits is simply not feasible. Thus, categorizing the slew rate limit in a deontological manner is most appropriate. We can represent this limit mathematically:

$$\left| F_{\text{yf}}{}^{(k)} - F_{\text{yf}}{}^{(k-1)} \right| \leq F_{\text{yf,max slew}} \tag{2}$$

where F_{yf} is the lateral front tire force.

When obeying the deontological constraint of maximum slew rate, additional design goals regarding the steering arise. One key goal is the smoothness of the steering, which is affected by the change in input from time step to time step. Steering smoothness is a reasonable consideration to include in the control algorithm because most occupants expect a level of comfort as they ride in a vehicle. Occupant comfort is a desirable feature, but, similarly to path tracking, if encoded as a hard and fast rule of matching a specific rate via an equality constraint or remaining under a rate via an inequality constraint, could result in safety compromises if an emergency maneuver is required. Specifically, if the vehicle needs to suddenly swerve to avoid an obstacle, it will be constrained by a requirement to maintain smooth steering and may not be able to steer and maneuver sufficiently to avoid a collision. However, when viewed from a consequentialist standpoint, we can include a cost associated with steering smoothness that the algorithm will minimize, but do so while obeying more highly prioritized rules associated with safety. Therefore, we cast steering smoothness for occupant comfort as a consequentialist cost. Occupant comfort level is incorporated into the objective function by associating a cost to the change in steering, or effectively, lateral front tire force:

$$J_{F_{\text{yf}}} = \gamma \sum_k \left|\left| F_{\text{yf}}^{(k)} - F_{\text{yf}}^{(k-1)} \right|\right|_2^2 \tag{3}$$

where γ is the associated cost. By minimizing the differential in lateral front tire force throughout the prediction horizon, the differential in front steer angle is also minimized. Thus, this term in the cost function results in smooth steering enforcement, which can factor into occupant comfort.

C. Obstacle Avoidance

Obstacle avoidance is a high priority in navigating roadways. As mentioned within the context of Sections IV-A and IV-B, the possibility of collisions and preserving the ability to avoid them are the basis for choosing to associate path tracking and steering smoothness with consequentialist costs rather than deontological rules. Now, we expound on the goal of collision avoidance. The choice of assigning collision avoidance as a deontological rule is natural; collision avoidance is arguably the highest priority of the automated vehicle.

The deontological rules that govern obstacle avoidance arise from dividing the environment into tubes in which the vehicle can safely travel and describing the envelope in which each tube lies. In the example scenario, the vehicle can choose from one of three tubes—pass the vehicle by entering the lane on the left, pass the vehicle by moving onto the shoulder to the right or stay in the lane. The boundaries of each tube in the environment can be constructed in reference to a nominal path which, in the scenario considered here, is simply the centerline of the lane in which the vehicle travels. These envelopes are descibed by a set of time-varying constraints on the maximum and minimum lateral offset from the nominal path, e, necessary to remain in the tube. The vehicle's trajectory over the prediction horizon is constrained to remain within this envelope to ensure the trajectory is collision-free. The environmental envelope may require the vehicle to deviate from the nominal path due to the fact that the nominal path need not be obstacle-free.

Figure 3 illustrates the methodology for the generation of the environmental envelope, beginning with the scenario illustrated in Fig. 3a as an example. The environment is sampled at discrete points along the nominal path based on the maintained longitudinal vehicle speed, U_{x}, and the assumption that the distance along the path is a function of only U_{x}. Figure 3b shows the future position of the vehicle k steps into the prediction horizon. In order to align with the discrete sampling, we extend the objects to match the sampling, shown in Fig. 3c. This extension allows for the identification of feasible gaps (defined as distances greater than vehicle width) between the objects. A graph search algorithm links adjacent feasible gaps, forming tubes like those shown in Fig. 3d. One tube must contain the full prediction horizon for the vehicle to avoid collisions. This tube methodology is similar to LaValle's

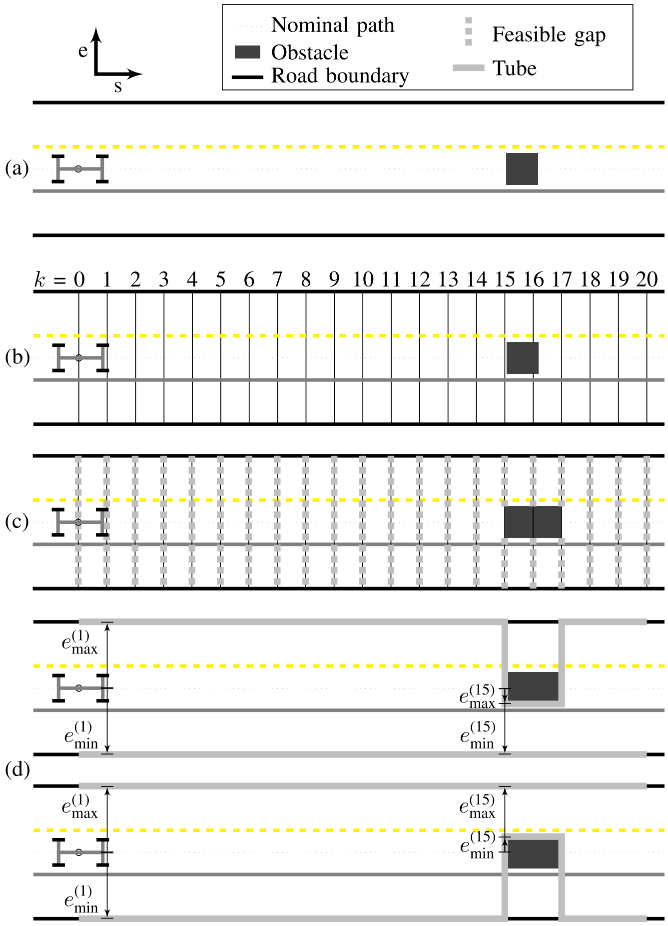

Fig. 3. Environmental envelope generation. The process consists of (a) starting with a set of obstacles along the nominal path, (b) discretization along the s direction, (c) extension of objects along that same s direction, which creates alignment with the discretization, and from which feasible gaps between objects are identified, and (d) connecting adjacent gaps into tubes which define maximum ($e_{max}^{(k)}$) and minimum ($e_{min}^{(k)}$) lateral deviation from the nominal path at each time step, k. Here, two tubes are given as examples.

vertical cell decomposition in [23], tube feasibility for robotic arm motion planning, as in Suh and Bishop [24], and driving corridors presented by Ziegler *et al.* [25].

The set of collision-free trajectories corresponding to a single tube is a convex set, due to the property that any linear combination of generated trajectories that is contained within a tube will also be contained within that same tube. This property opens up the ability to use fast optimization techniques for optimal trajectories to be identified quickly.

The following linear inequality represents the bound on e at each time step k represented by each tube:

$$H_{\text{env}} x^{(k)} \leq G_{\text{env}}^{(k)} \tag{4}$$

with

$$H_{\text{env}} = \begin{bmatrix} H_{\text{env,left}} \\ H_{\text{env,right}} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

$$G_{\text{env}} = \begin{bmatrix} G_{\text{env,left}} \\ G_{\text{env,right}} \end{bmatrix} = \begin{bmatrix} e_{\max}^{(k)} - \frac{1}{2}d - d_{\text{buffer}} \\ -e_{\min}^{(k)} - \frac{1}{2}d - d_{\text{buffer}} \end{bmatrix}$$

The environmental envelope is denoted with the subscript env, the lateral deviation bounds for time step k are given as $e_{\max}^{(k)}$ and $e_{\min}^{(k)}$, and the vehicle width is d. Occupant comfort can be further specified via d_{buffer}, which represents a preferred minimum distance between obstacles and the vehicle and can account for vehicle orientation changes in determining minimum gaps from obstacles as well.

D. Traffic Laws

Traffic laws present the most ambiguous choice between rule- and cost-based design. The scenario posed in Section II serves as a clear example of this conflict. Traffic laws by definition enforce structure and rules, and thus naturally can be framed as deontological. However, humans do not always treat traffic laws as deontological. Realistically, drivers in Section II's scenario often judge factors such as clearance from the obstacle, traffic in the adjacent lane, and speed for overtaking. The driver then makes a decision on whether or not to cross the double yellow line. Since humans will often opt to cross the line, particularly when overtaking a bicyclist, human compliance with traffic laws seems less deontological and more like a consequentialist weighting of safety, mobility and legality. Thus, in moving from human driver actions to programming automated driving, the decision to treat traffic laws as deontological or consequentialist is a fundamental design choice.

If defined as a rule, obedience of traffic laws can easily result in congestion, as described in Section II. If defined as a cost, the implication is that laws are *a priori* programmed to be broken. Thus, given the dilemma, we employ a "soft" constraint by incorporating a slack variable to encode traffic laws in the MPC formulation. The slack variable creates a scalable cost on constraint violation. When that cost is comparably lower than other objectives, the constraint is treated in a consequentialist manner because the slack variable augments the constraint to make the constraint less strict. In contrast, a very high weight on the slack variable causes the constraint to dominate all other objectives in a deontological manner because of the large penalty associated with making the slack variable value non-zero. Continuing with the two-lane roadway example and translating our ethically motivated design decisions into the algorithm, we include a cost on the slack variable corresponding to either crossing the road divider S_{left} or entering the road shoulder S_{right}. Treating traffic law obedience as deontological or consequentialist results in dramatically different vehicle behavior, which is demonstrated in Section VI. These different driving outcomes underscore the importance of matching societal expectations to programming decisions through these philosophical frameworks.

V. MPC Formulation

The control algorithm attempts to determine the optimal path for the vehicle in light of the costs and constraints placed on its motion. The vehicle divides the world into a number of feasible tubes, with each tube representing a separate convex optimization problem. The vehicle then calculates the optimal path in each tube as illustrated in Figure 4, choosing the

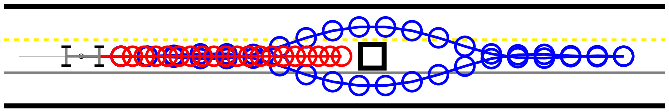

Fig. 4. The three tubes define the three possible maneuver options to avoid an obstacle. The left and right tubes are depicted in blue while stopping is depicted in red.

option with the lowest cost. The following sections describe the mathematics behind solving for the optimal trajectory in each tube and its associated cost.

A. Vehicle Model

The vehicle model used in the online MPC controller is a bicycle model with four states. There are two velocity states, vehicle sideslip (β) and yaw rate (r), and two position states, heading deviation ($\Delta\psi$) and lateral deviation (e), the details of which can be found in the Appendix. Thus, the vehicle state vector is

$$x = [\beta \quad r \quad \Delta\psi \quad e]^\top. \tag{5}$$

In this paper, the actuator is considered to be front steering, and the vehicle is assumed to be equipped with steer-by-wire technology. This enables the computer algorithm to command the desired front tire force. For simplicity, the results here consider a constant longitudinal speed maintained by a PD cruise controller unless the vehicle needs to brake, though this is not strictly necessary.

B. Optimization

For each tube in the environment, the optimal path and control inputs are the solution to the following optimization problem:

$$\begin{aligned}
\text{minimize} \quad & \sum_k x^{(k)\top} Q^{(k)} x^{(k)} && \text{(6a)}\\
& + \gamma \sum_k \left|\left| F_{\text{yf,opt}}^{(k)} - F_{\text{yf,opt}}^{(k-1)} \right|\right|_2^2 && \text{(6b)}\\
& + \sum_l [\sigma_{\text{env}} \quad \sigma_{\text{env}}]\, S_{\text{env,opt}}^{(l)} && \text{(6c)}\\
& + \sum_l [\sigma_{\text{tra}} \quad \sigma_{\text{tra}}]\, S_{\text{tra,opt}}^{(l)} && \text{(6d)}\\
\text{subject to} \quad & x^{(k+1)} = A_d^{(k)} x^{(k)} + B_d^{(k)} F_{\text{yf,opt}}^{(k)} + d_d^{(k)} && \text{(6e)}\\
& \left| F_{\text{yf,opt}}{}^{(k)} \right| \le F_{\text{yf,max}} && \text{(6f)}\\
& \quad k = 0 \ldots (T-1) && \\
& H_{\text{env}} x^{(l)} \le G_{\text{env}}^{(l)} + S_{\text{env,opt}}^{(l)} + S_{\text{tra,opt}}^{(l)} && \text{(6g)}\\
& \quad l = (T_{\text{split}} + 1) \ldots T && \\
& \left| F_{\text{yf,opt}}{}^{(i)} - F_{\text{yf,opt}}{}^{(i-1)} \right| \le F_{\text{yf,max slew}} && \text{(6h)}\\
& \quad i = 0 \ldots T &&
\end{aligned}$$

where $T_{\text{split}} + 1$ is when the time steps are longer for the environmental envelope as detailed by Erlien *et al.* [17]. The variables to be optimized are the front lateral tire forces $F_{\text{yf,opt}}$, the slack variable on the environmental deontological constraint ($S_{\text{env,opt}}$) and the slack variable on the traffic laws ($S_{\text{tra,opt}}$). The tunable parameters in this optimization problem are the costs on the vehicle states ($Q^{(k)}$), the cost on the change between inputs (γ), and the costs on the slack variables (σ_{env}, σ_{tra}).

Fig. 5. X1, an all electric, steer- and drive-by-wire research testbed.

TABLE I

X1 VEHICLE PARAMETERS

Parameter	Symbol	Value	Units
Vehicle mass	m	2009	kg
Yaw moment of inertia	I_{zz}	3000	$\text{kg} \cdot \text{m}^2$
Distance from front axle to CG	a	1.53	m
Distance from rear axle to CG	b	1.23	m
Track width	d	1.63	m
Front cornering stiffness	$C_{\alpha f}$	140	$\text{kN} \cdot \text{rad}^{-1}$
Rear cornering stiffness	$C_{\alpha r}$	170	$\text{kN} \cdot \text{rad}^{-1}$

The slack variables are not only used to represent traffic laws but also to set a hierarchy of deontological constraints and ensure that the problem always returns a feasible solution. Hence, the slack variables are unbounded. Higher weights on a slack variable give it a higher precedence in a deontological framework. Here the slack variables are used to weight the constraint version of traffic laws below that of obstacle avoidance. In a more comprehensive version of a vehicle control system, they could also be used to incorporate vehicle stability constraints such as those in Beal and Gerdes [26] and Bobier and Gerdes [27]. The slack variable weights for such constraints should be placed below that of collision avoidance in a deontological sense as demonstrated in Funke *et al.* [28]. Additional cost terms could also be incorporated in the cost function such as those presented by Wei *et al.* [29].

Optimization problem (6i) is a quadratic program with a significantly sparse structure that can be leveraged with an efficient solver for real-time implementation [30]. For this work, CVXGEN, developed by Mattingley and Boyd [31], is used to solve for the vector of optimal front lateral tire forces $F_{\text{yf,opt}}$. The control input to the autonomous vehicle at the next time step is the first solution $F_{\text{yf,opt}}^{(0)}$ which, in turn, is translated to a desired steering angle as described in the Appendix.

VI. EXPERIMENTAL RESULTS

The example from Section II sets the stage for demonstrating ethically motivated algorithm design in real vehicle experiments. Depending upon the driving situation and the

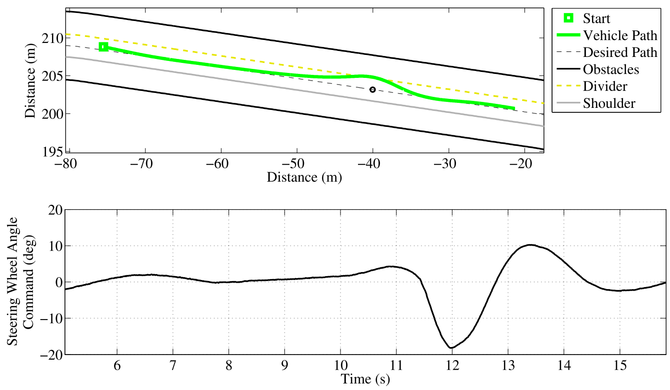

Fig. 6. The left tube is chosen because the traffic lane divider is considered safe to cross.

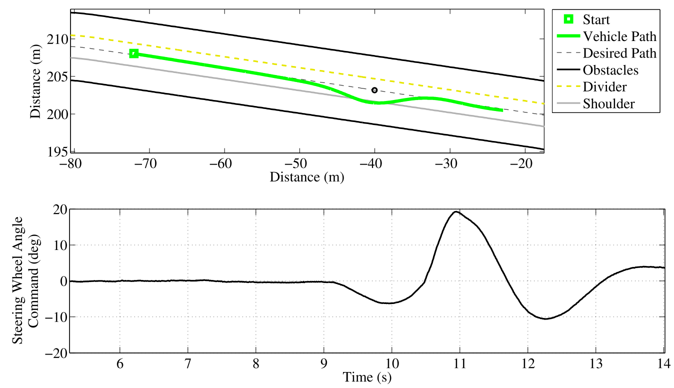

Fig. 7. The right tube is chosen because evaluation of the scenario determined it is safer to pass around the obstacle via the road shoulder.

TABLE II
WEIGHTS RESULTING IN A PASS ON THE LEFT

Parameter	Symbol	Value	Units
Lateral error	Q_e	0.7	m^{-1}
Heading error	$Q_{\Delta\psi}$	0.5	rad^{-1}
Smoothness	γ	0.1	kN^{-1}
Environmental slack	σ_{env}	500	m^{-1}
Road divider slack	σ_{left}	10	m^{-1}
Road shoulder slack	σ_{right}	150	m^{-1}

TABLE III
WEIGHTS RESULTING IN A PASS ON THE RIGHT

Parameter	Symbol	Value	Units
Lateral error	Q_e	0.7	m^{-1}
Heading error	$Q_{\Delta\psi}$	0.5	rad^{-1}
Smoothness	γ	0.1	kN^{-1}
Environmental slack	σ_{env}	500	m^{-1}
Road divider slack	σ_{left}	150	m^{-1}
Road shoulder slack	σ_{right}	10	m^{-1}

engineering design choices made in the algorithms, the vehicle chooses a different tube or a different trajectory within that tube. In these experiments, the weights are changed to reflect different interpretations of the traffic laws. The philosophical treatment of different objectives can therefore be translated through mathematics to the actual motion of the vehicle. The weights were first chosen in simulation and validated in-vehicle for the experiments. The actual numerical values for the weights are not crucial; instead, the relative values affect the solution to the optimization problem. In practice, they can be derived as a function of geography or determined by perception algorithms.

The test vehicle used for these experiments is X1, an all-electric, drive- and steer-by-wire vehicle shown in Fig. 5. The parameters for this vehicle are specified in Table I. The vehicle is equipped with an integrated GPS/INS system that provides real-time estimates of the vehicle states. In these experiments, obstacles and road boundary locations are assumed to be known. All of the following experiments took place on a cement surface which the controller models with constant friction.

A. Traffic Laws as Consequentialist Costs

In Section IV, we discussed the design choice to implement traffic laws as slack variables on the environmental constraints in the MPC formulation. Changing the weights on the slack variables can influence the vehicle's behavior as it navigates the scenario described in Section II. A consequentialist approach allows for flexibility in weighting the road bounds to reflect the severity of costs associated with crossing them. Table II shows a set of weights where the shoulder is treated effectively as a hard constraint by placing a high cost on the slack variable (representing, for instance, a curb on the side of the road) while the double yellow lane divider is treated more as a cost with lower weight. As seen in Fig. 6, the vehicle maneuvers to the left of the obstacle and crosses the divider. Because of the relatively high weight put on the divider relative to the path tracking weight, the vehicle does not cross far into the adjacent lane but rather stays fairly close to the obstacle.

Upon interchanging the costs on road shoulder and road divider (representing a stricter adherence to the law but an open space adjacent to the lane), Fig. 7 shows that the vehicle maneuvers to the right of the obstacle. Table III shows the weights chosen for this particular situation. Since the weights describe essentially a mirror image of the previous experiment, both the trajectory and the resulting steering angle are reflections of the original case.

B. Traffic Laws as Deontological Constraints

Another philosophical approach to accounting for traffic laws is to describe them deontologically as rigid rules. As the slack variable weights increase relative to the other weights, the road boundary constraints begin to resemble hard or deontological constraints. With this choice of weights, shown in Table IV, the tubes to the left and right of the vehicle have unacceptably high costs and therefore the vehicle must remain in the tube corresponding to the lane. Since that tube is blocked by the obstacle, the vehicle must brake to a complete stop. Figure 8 shows the trajectory as an independent longitudinal PD controller commands a brake force to bring the vehicle to a stop before the end of the tube.

The MPC formulation is flexible enough to incorporate different treatments of the environmental boundaries and the

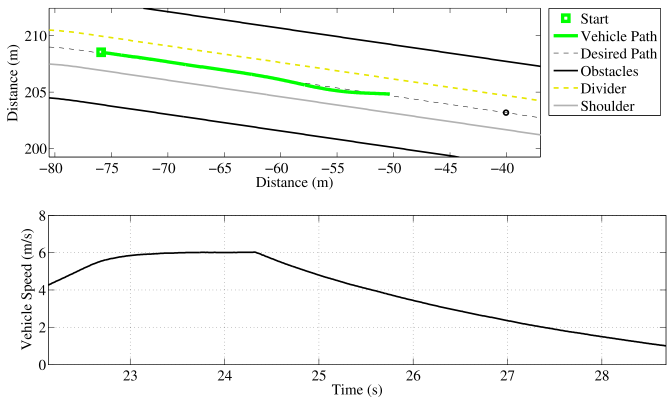

Fig. 8. Since left and right path options are weighted equivalently, the tube is considered to be blocked and the vehicle brakes to a complete stop.

TABLE IV
WEIGHTS RESULTING IN A FULL STOP

Parameter	Symbol	Value	Units
Lateral error	Q_e	10	m^{-1}
Heading error	$Q_{\Delta\psi}$	1	rad^{-1}
Smoothness	γ	0.1	kN^{-1}
Environmental slack	σ_{env}	500	m^{-1}
Road divider slack	σ_{left}	150	m^{-1}
Road shoulder slack	σ_{right}	150	m^{-1}

corresponding traffic laws. As this simple example demonstrates, however, treating the double yellow lane boundary as a hard, deontological constraint removes much of the flexibility from the vehicle path and fails to reflect the way humans drive. This suggests that traffic laws may have to be modified to give programmers the same flexibility demonstrated by their human driver counterparts. Without that option, programmers need to choose how much to weight the traffic laws in a consequentialist approach. In this MPC formulation, the relative values of the weights chosen for the traffic slack variables determine how the vehicle maneuvers around obstacles. These values can be a function of geography such as when a lane divider is a single, dashed white line or when it is a double, solid yellow line. Additionally, the weights can be determined by a perception algorithm using information from lidars and cameras that determine the safety of the terrain in each tube. These results demonstrate the flexibility of the MPC formulation to account for responsible decision-making.

VII. Vehicle Character

Thus far, the variations in the MPC formulation have been motivated by traffic situations for one particular vehicle. The experimental results indicate the principles behind the ethical theories of consequentialism and deontology at work, through the costs and constraints of the MPC formulation. However, these theories do not exclusively guide the choice of relative numerical values for the weights, which have a huge impact on design goals beyond safety for automated vehicles. Thus, we motivate this section by introducing a third normative ethical theory known as *virtue ethics*. Virtue ethics points the focus of ethical behavior toward character rather than correct actions or outcomes as emphasized by deontology and consequentialism. Under the framework of virtue ethics, a decision is ethical as long as it adheres to the disposition of the moral being. In other words, a person operates virtuously provided they always perform the correct action in the correct situation in alignment with their character [19].

The idea of the character of an agent naturally leads to a more specific concept which we will utilize in this work called *role morality*. Role morality is the idea that behavior within the context of a specific professional role and situation is acceptable but may not be acceptable outside that setting [32]. Role morality is cited in fields such as law and medicine to justify behavior that might otherwise be deemed unacceptable were the behavior to take place outside a professional situation specific to that field. An extreme example is Sanson, the executioner of Paris, as noted by Applbaum [33]. A less extreme example is that of a physician prescribing medication to someone who is not officially their patient. While writing prescriptions is acceptable within the professional bounds of a doctor-patient relationship, it is illegal outside of this role. These acceptable roles and codes of conduct are based upon the societal expectation of the service provided by professionals in that particular field; therefore, role morality is derived not from any individual in charge of making the rules but from a collective decision on what is best for society [32].

An important issue in the development of automated vehicles is therefore the type of role or character different vehicles should have. Specifically, the role of a vehicle affects the level of strictness to which a vehicle should adhere to a traffic law. Previous experiments in Figures 6, 7, and 8 showcase the ability of an autonomous vehicle to choose whether or not to violate a traffic law of obeying the double yellow line boundary due to safety considerations. The level of fidelity of either adherence or violation begs for a guiding principle, which can be found in the idea of role morality. Different types of vehicles have different roles to play in society. An ambulance running a red light while carrying a passenger experiencing life-threatening conditions to the hospital can acceptably break traffic laws: the role of an ambulance in society is to transport people to the hospital as quickly as possible in order to save lives. Conversely, a taxi may not run a red light to save time because its societal role does not merit that behavior.

Thus, while deontology and consequentialism enable vehicle goals to be described as either constraints or costs, role morality can help determine the strength of the applied rules and costs for different vehicles. For example, role morality establishes the context for why an ambulance would be programmed to consider breaking traffic laws more strongly than a taxi.

Using the MPC formulation presented in Section V, the character of a vehicle that can acceptably break laws, such as an ambulance, can be modeled by varying the weights on different objectives. As an example, the weights on tracking error can be relatively reduced to imitate the desired behavior of an emergency response vehicle. The relatively lower costs on lateral and heading error allow the vehicle more freedom to deviate from the path. The experimental results using the

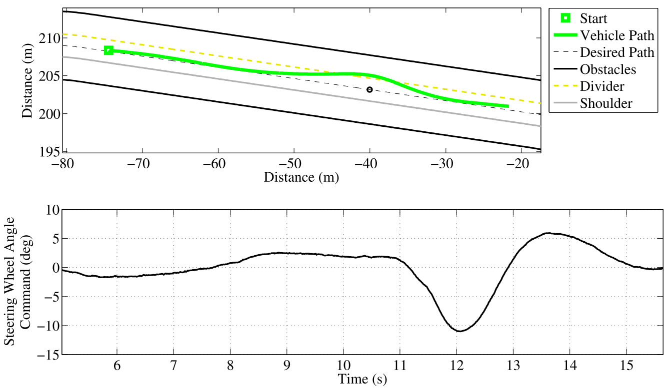

Fig. 9. Reduction of the cost on path following allows the vehicle behavior to model emergency response vehicle character.

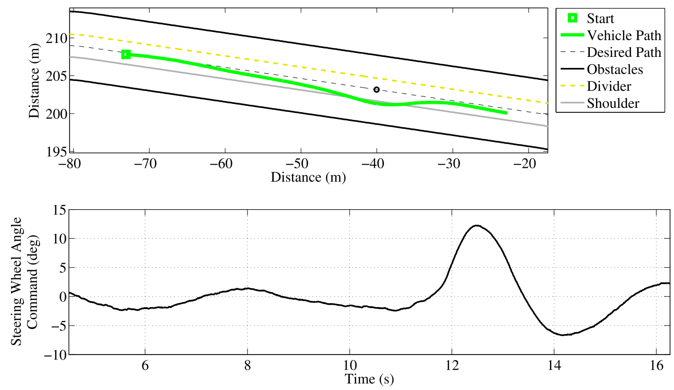

Fig. 10. Reduction of the cost on path following allows the vehicle behavior to model emergency response vehicle character.

TABLE V

WEIGHTS CAUSING AN AMBULANCE VEHICLE CHARACTER TO PASS ON THE LEFT

Parameter	Symbol	Value	Units
Lateral error	Q_e	0.3	m^{-1}
Heading error	$Q_{\Delta\psi}$	0.25	rad^{-1}
Smoothness	γ	0.1	kN^{-1}
Environmental slack	σ_{env}	500	m^{-1}
Road divider slack	σ_{left}	5	m^{-1}
Road shoulder slack	σ_{right}	200	m^{-1}

TABLE VI

WEIGHTS CAUSING AN AMBULANCE VEHICLE CHARACTER TO PASS ON THE RIGHT

Parameter	Symbol	Value	Units
Lateral error	Q_e	0.3	m^{-1}
Heading error	$Q_{\Delta\psi}$	0.25	rad^{-1}
Smoothness	γ	0.1	kN^{-1}
Environmental slack	σ_{env}	500	m^{-1}
Road divider slack	σ_{left}	200	m^{-1}
Road shoulder slack	σ_{right}	5	m^{-1}

weights in Tables V and VI are shown in Fig. 9 and 10, respectively. These relative weights were determined using the process described in Section VI. Because the lower relative costs on lateral and heading error with respect to smoothness allow for greater path deviation, the emergency vehicle begins executing the maneuver earlier, as shown in Fig. 9 and 10. In this particular scenario, the smoothness of the maneuver is potentially advantageous for an injured passenger.

While a vehicle can indeed be programmed to weight traffic laws and maneuver objectives according to its societal role, developing an engineered system which may not always observe the law is an uncomfortable proposition. This raises another question of role morality: should the engineers designing such a system be responsible for setting weights on the laws? Or, is a better solution to adapt the law to the programming realities of automated systems? While not producing a simple answer, the translation of ethical considerations to engineering costs and constraints does help to raise the appropriate questions that must be resolved to field automated vehicles.

VIII. CONCLUSION

The normative ethical theories of deontology and consequentialism provide guiding principles for responsible programming of autonomous vehicles. In particular, these concepts map well to an MPC framework which minimizes the consequential costs subject to deontological constraints. Making these connections can enable engineers working at the deepest levels of programming automated vehicles to connect their design choices with broader issues of societal acceptance. This paper has examined how to incorporate objectives such as path tracking, vehicle occupant comfort, and traffic laws as priorities in the cost function together while obstacle avoidance and vehicle slew rate limits enter as constraints. The concept of role morality provides a further basis for different weighting schemes within the control formulation depending on the vehicle type and function.

More complex scenarios will obviously require the consideration of additional objectives for the vehicle. Nevertheless, the simple obstacle avoidance maneuver in this paper already points out key challenges for balancing a respect for the law with the desire to integrate smoothly with human drivers in traffic. Realizing the benefits of automated vehicles will require further integration of legal and ethical considerations with the underlying control code.

ACKNOWLEDGMENT

The authors would like to thank Patrick Lin from California Polytechnic State University, San Luis Obispo for his expertise on robot ethics. They would also thank the students in the Dynamic Design Lab for their invaluable help with discussions regarding this research.

APPENDIX

A. Velocity States

The vehicle sideslip angle, β, and yaw rate, r, are the velocity states in the vehicle model. The sideslip angle is:

$$\beta = \tan\left(\frac{U_{\mathrm{y}}}{U_{\mathrm{x}}}\right) \quad (7)$$

$$\approx \frac{U_{\mathrm{y}}}{U_{\mathrm{x}}} \quad (8)$$

where the lateral and longitudinal velocities are denoted in the body-fixed frame as U_y and U_x, respectively. As simplifications, we assume $U_x \gg U_y$ and that U_x is constant.

The equations of motion for the sideslip and yaw rate are:

$$\dot{\beta} = \frac{F_{yf} + F_{yr}}{mU_x} - r \tag{9}$$

$$\dot{r} = \frac{aF_{yf} - bF_{yr}}{I_{zz}} \tag{10}$$

Here, the lateral tire force of the [front, rear] axle is denoted as $F_{y[f,r]}$, the vehicle mass is denoted as m, the yaw inertia is denoted as I_{zz}, and the distances from the center of gravity of the vehicle to the front and rear axles are denoted a and b, respectively.

The front tire slip angle, α_f, and rear tire slip angle, α_r, can be expressed as:

$$\begin{aligned} \alpha_f &= \tan^{-1}\left(\beta + \frac{ar}{U_x}\right) - \delta \\ &\approx \beta + \frac{ar}{U_x} - \delta \end{aligned} \tag{11}$$

$$\begin{aligned} \alpha_r &= \tan^{-1}\left(\beta - \frac{br}{U_x}\right) \\ &\approx \beta - \frac{br}{U_x} \end{aligned} \tag{12}$$

The linear expressions result from small angle approximations.

The brush tire model proposed by Fiala and presented in the following form by Pacejka provides a model of the relationship between the lateral tire forces and tire slip angles:

$$\begin{aligned} F_y &= \begin{cases} -C_\alpha \tan\alpha + \frac{C_\alpha^2}{3\mu F_z}|\tan\alpha|\tan\alpha & \\ \quad -\frac{C_\alpha^3}{27\mu^2 F_z^2}\tan^3\alpha, & |\alpha| < \arctan\left(\frac{3\mu F_z}{C_\alpha}\right) \\ -\mu F_z \operatorname{sgn}\alpha, & \text{otherwise} \end{cases} \\ &= f_{\text{tire}}(\alpha) \end{aligned} \tag{13}$$

Here, the surface coefficient of friction is given as μ, the normal load is given as $F_{z[f,r]}$, and the tire cornering stiffness is given as C_α.

The vehicle model used by the online MPC controller utilizes the front tire force to keep the problem linear with regards to the input. The resulting steer angle, δ, follows from (11) and (13):

$$\delta = \beta + \frac{ar}{U_x} - f_{\text{tire}}^{-1}(F_{yf}) \tag{14}$$

To address the nonlinearity of the rear tires, the brush tire model is linearized at a given rear tire slip angle ($\bar{\alpha}_r$), and the rear tire force (F_{yr}) is thus modeled as an affine function of α_r:

$$F_{yr} = \bar{F}_{yr} - \bar{C}_{\bar{\alpha}_r}(\alpha_r - \bar{\alpha}_r) \tag{15}$$

where $\bar{F}_{yr} = f_{\text{tire}}(\bar{\alpha}_r)$ and $\bar{C}_{\bar{\alpha}_r}$ is the equivalent cornering stiffness at $\bar{\alpha}_r$. In the initial time steps of the prediction horizon, the current rear slip angle, α_r, is chosen to be $\bar{\alpha}_r$. This allows the MPC controller to explicitly consider rear tire saturation in the near term prediction [26].

The equations of motion of the velocity states can now be formulated as affine functions of the states and input, F_{yf}:

$$\dot{\beta} = \frac{F_{yf} + \bar{F}_{yr} - \bar{C}_{\bar{\alpha}_r}\left(\beta - \frac{br}{U_x} - \bar{\alpha}_r\right)}{mU_x} - r \tag{16}$$

$$\dot{r} = \frac{aF_{yf} - b\left[\bar{F}_{yr} - \bar{C}_{\bar{\alpha}_r}\left(\beta - \frac{br}{U_x} - \bar{\alpha}_r\right)\right]}{I_{zz}}. \tag{17}$$

B. Position States

The position states of the vehicle, the heading deviation ($\Delta\psi$) and lateral deviation (e), are in reference to a nominal path that need not be obstacle-free.

The equations of motion of the heading deviation and lateral deviation are:

$$\dot{\Delta\psi} = r \tag{18}$$

$$\dot{e} = U_x \sin(\Delta\psi) + U_y \cos(\Delta\psi) \tag{19}$$

To approximate the above nonlinear equations as linear functions of the vehicle states, small angle assumptions for $\Delta\psi$ and β are employed to yield:

$$\dot{e} \approx U_x \Delta\psi + U_x \beta \tag{20}$$

Thus, (16), (17), (18), and (20) combine to produce a continuous state-space representation of the vehicle model:

$$\dot{x} = A_c(\bar{\alpha}_r)x + B_c F_{yf} + d_c(\bar{\alpha}_r) \tag{21}$$

with

$$x = \begin{bmatrix}\beta & r & \Delta\psi & e\end{bmatrix}^\top$$

$$A_c(\bar{\alpha}_r) = \begin{bmatrix} -\frac{\bar{C}_{\bar{\alpha}_r}}{mU_x} & \frac{b\bar{C}_{\bar{\alpha}_r}}{m{U_x}^2} - 1 & 0 & 0 \\ \frac{b\bar{C}_{\bar{\alpha}_r}}{I_{zz}} & -\frac{b^2\bar{C}_{\bar{\alpha}_r}}{I_{zz}U_x} & 0 & 0 \\ 0 & 1 & 0 & 0 \\ U_x & 0 & U_x & 0 \end{bmatrix}$$

$$B_c = \begin{bmatrix}\frac{1}{mU_x} & \frac{a}{I_{zz}} & 0 & 0\end{bmatrix}^\top$$

$$d_c(\bar{\alpha}_r) = \begin{bmatrix}\frac{\bar{F}_{yf} - \bar{\alpha}_r\bar{C}_{\bar{\alpha}_r}}{mU_x} & -\frac{b(\bar{F}_{yf} - \bar{\alpha}_r\bar{C}_{\bar{\alpha}_r})}{I_{zz}} & 0 & 0\end{bmatrix}^\top$$

Here, c denotes a continuous-time model. $A_c(\bar{\alpha}_r)$ indicates a linearization of A_c about $\bar{\alpha}_r$.

References

[1] J. Haidt and C. Joseph, "Intuitive ethics: How innately prepared intuitions generate culturally variable virtues," *Daedalus*, vol. 133, no. 4, pp. 55–66, 2004.

[2] J. Haidt, J. Graham, and C. Joseph, "Above and below left–right: Ideological narratives and moral foundations," *Psychol. Inquiry*, vol. 20, nos. 2–3, pp. 110–119, 2009.

[3] A. M. Brandt and P. Rozin, Eds., *Morality and Health*. New York, NY, USA: Routledge, 1997.

[4] J. Graham, J. Haidt, and B. A. Nosek, "Liberals and conservatives rely on different sets of moral foundations," *J. Pers. Social Psychol.*, vol. 96, no. 5, pp. 1029–1046, May 2009.

[5] M. N. Mladenovic and T. McPherson, "Engineering social justice into traffic control for self-driving vehicles?" *Sci. Eng. Ethics*, vol. 22, no. 4, pp. 1131–1149, Aug. 2015. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/26231407

[6] K. W. Miller, M. J. Wolf, and F. Grodzinsky, "This 'ethical trap' is for roboticists, not robots: On the issue of artificial agent ethical decision-making," *Sci. Eng. Ethics*, pp. 1–13, Apr. 2016. [Online]. Available: http://link.springer.com/10.1007/s11948-016-9785-y
[7] J. Van den Hoven, G. J. Lokhorst, and I. Van de Poel, "Engineering and the problem of moral overload," *Sci. Eng. Ethics*, vol. 18, no. 1, pp. 143–155, Mar. 2012.
[8] P. Lin, "Why ethics matters for autonomous cars," in *Autonomes Fahren*. Ladenburg, Germany: Springer, 2015, pp. 69–85.
[9] P. Lin, K. Abney, and G. A. Bekey, *Robot Ethics: The Ethical and Social Implications of Robotics*. Cambridge, MA, USA: MIT Press, 2011.
[10] W. Wallach and C. Allen, *Moral Machines: Teaching Robots Right From Wrong*. London, U.K.: Oxford Univ. Press, 2008.
[11] N. J. Goodall, "Ethical decision making during automated vehicle crashes," in *Proc. TRB Annu. Meeting*, Washington, DC, USA, 2014, pp. 58–65.
[12] J. C. Gerdes and S. M. Thornton, "Implementable ethics for autonomous vehicles," in *Autonomes Fahren*. Ladenburg, Germany: Springer, 2015, pp. 87–102.
[13] A. Gray, Y. Gao, T. Lin, J. K. Hedrick, H. E. Tseng, and F. Borrelli, "Predictive control for agile semi-autonomous ground vehicles using motion primitives," in *Proc. Amer. Control Conf.*, Jun. 2012, pp. 4239–4244.
[14] A. Gray, Y. Gao, T. Lin, J. K. Hedrick, and F. Borrelli, "Stochastic predictive control for semi-autonomous vehicles with an uncertain driver model," in *Proc. 16th Int. IEEE Conf. Intell. Transp. Syst.*, Oct. 2013, pp. 2329–2334.
[15] Y. Gao *et al.*, "Spatial predictive control for agile semi-autonomous ground vehicles," in *Proc. 11th Int. Symp. Adv. Vehicle Control*, 2012.
[16] S. Erlien, S. Fujita, and J. C. Gerdes, "Safe driving envelopes for shared control of ground vehicles," *Adv. Automotive Control*, vol. 46, no. 21, pp. 831–836, 2013.
[17] S. M. Erlien, S. Fujita, and J. C. Gerdes, "Shared steering control using safe envelopes for obstacle avoidance and vehicle stability," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 2, pp. 441–451, Feb. 2016.
[18] P. Falcone, F. Borrelli, J. Asgari, H. E. Tseng, and D. Hrovat, "Predictive active steering control for autonomous vehicle systems," *IEEE Trans. Control Syst. Technol.*, vol. 15, no. 3, pp. 566–580, May 2007.
[19] R. Hursthouse, "Virtue ethics," in *The Stanford Encyclopedia Philosophy*, E. N. Zalta, Ed., 2013. [Online]. Available: http://plato.stanford.edu/entries/ethics-virtue/
[20] G. Harman, "Skepticism about character traits," *J. Ethics*, vol. 13, no. 2, pp. 235–242, Sep. 2009.
[21] I. Asimov, *I, Robot*. New York, NY, USA: Gnome Press, 1950.
[22] J. Driver, "The history of utilitarianism," in *The Stanford Encyclopedia Philosophy*, E. N. Zalta, Ed., 2014. [Online]. Available: http://plato.stanford.edu/entries/utilitarianism-history/
[23] S. M. LaValle, *Planning Algorithms*. Cambridge, U.K.: Cambridge Univ. Press, 2006.
[24] S.-H. Suh and A. B. Bishop, "Collision-avoidance trajectory planning using tube concept: Analysis and simulation," *J. Robot. Syst.*, vol. 5, no. 6, pp. 497–525, Dec. 1988.
[25] J. Ziegler, P. Bender, T. Dang, and C. Stiller, "Trajectory planning for bertha—A local, continuous method," in *Proc. IEEE Intell. Veh. Symp.*, Jun. 2014, pp. 450–457.
[26] C. E. Beal and J. C. Gerdes, "Model predictive control for vehicle stabilization at the limits of handling," *IEEE Trans. Control Syst. Technol.*, vol. 21, no. 4, pp. 1258–1269, Jul. 2012.
[27] C. G. Bobier and J. C. Gerdes, "Staying within the nullcline boundary for vehicle envelope control using a sliding surface," *Veh. Syst. Dyn.*, vol. 51, no. 2, pp. 199–217, 2013.
[28] J. Funke, M. Brown, S. M. Erlien, and J. C. Gerdes, "Prioritizing collision avoidance and vehicle stabilization for autonomous vehicles," in *Proc. Intell. Veh. Symp.*, Jun./Jul. 2015, pp. 1134–1139.
[29] J. Wei, J. M. Dolan, and B. Litkouhi, "A prediction- and cost function-based algorithm for robust autonomous freeway driving," in *Proc. IEEE Intell. Veh. Symp.*, Jun. 2010, pp. 512–517.
[30] J. Mattingley, Y. Wang, and S. Boyd, "Code generation for receding horizon control," in *Proc. IEEE Int. Symp. Comput.-Aided Control Syst. Design (CACSD)*, Sep. 2010, pp. 985–992.
[31] J. Mattingley and S. Boyd, "CVXGEN: A code generator for embedded convex optimization," *Optim. Eng.*, vol. 13, no. 1, pp. 1–27, Mar. 2012.
[32] R. R. Radtke, "Role morality in the accounting profession—How do we compare to physicians and attorneys?" *J. Bus. Ethics*, vol. 79, no. 3, pp. 279–297, May 2008.
[33] A. I. Applbaum, *Ethics for Adversaries: The Morality Roles Public Professional Life*. Princeton, NJ, USA: Princeton Univ. Press, 2000.

Sarah M. Thornton received the B.Sc. degree in mechanical engineering from UC Berkeley in 2011 and the M.Sc. degree in mechanical engineering from the Massachusetts Institute of Technology in 2013, with a focus on developing an adaptive shift control algorithm for automatic transmissions. She is currently pursuing the Ph.D. degree in mechanical engineering from Stanford University. Her current research interests include the area of ethical and uncertain decision making for automated vehicles.

Selina Pan received the B.S.E. degree in aerospace engineering from the University of Michigan in 2008, and the M.Sc. and Ph.D. degrees in mechanical engineering from UC Berkeley in 2009 and 2014, respectively. She is currently a Post-Doctoral Researcher with Stanford University. Her current research interests include ethics, driver adaptation, and integrated path planning and tracking in autonomous vehicles. Her previous research was in unmanned aircraft and automotive engine control.

Stephen M. Erlien received the Ph.D. degree from Stanford University, Stanford, CA, USA, in 2015. He is currently the Controls Lead with Peloton-Technology, Mountain View, CA, where he applies shared control to heavy trucks to improve safety and efficiency. His current research interests include human–computer shared control of dynamic systems and constraint-based control. He was a recipient of the National Science Foundation Graduate Research Fellowship Award in 2009.

J. Christian Gerdes received the Ph.D. degree from UC Berkeley in 1996. He was the Project Leader for virtual proving grounds development with the Vehicle Systems Technology Center, Daimler-Benz Research and Technology, Portland, OR. He is currently a Professor with the Department of Mechanical Engineering and the Department of Aeronautics and Astronautics, Stanford University, Stanford, CA. He is also the Director of the Center for Automotive Research, Stanford University.