HAL
open science

Ethical decision making for autonomous vehicles

Nelson de Moura, Raja Chatila, Katherine Evans, Stéphane Chauvier, Ebru Dogan

To cite this version:

Nelson de Moura, Raja Chatila, Katherine Evans, Stéphane Chauvier, Ebru Dogan. Ethical decision making for autonomous vehicles. IEEE Symposium on Intelligent Vehicle, Oct 2020, Las Vegas (virtual), United States. hal-03022605

HAL Id: hal-03022605

https://hal.sorbonne-universite.fr/hal-03022605v1

Submitted on 24 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Ethical decision making for autonomous vehicles

Nelson de Moura[1,3] Raja Chatila[1] Katherine Evans[2,3] Stéphane Chauvier[2] Ebru Dogan[3]

***Abstract*— We address ethical dilemma situations that may arise during autonomous driving. To evaluate how this deliberation could work, we propose a decision-making algorithm based on a Markov Decision Process (MDP) which controls the vehicle in normal conditions. When a dilemma situation is detected, the collision severity is determined by an evaluation of the harm incurred by different types of road users. Then, to illustrate different moral approaches, three different policies are proposed: one based on ralwsian contractarianism, another on utilitarianism, and finally on egalitarianism. Each policy supports a different view of the concept of fairness, potentially producing different behaviors for the same dilemma situation.**

I. INTRODUCTION

One of the main motivations for research in autonomous vehicles (AV) is the possibility of traffic accident reduction. The U.S. 2015 road accident data indicates that 94% of all accidents are caused by human error [1]. Although such a decrease is expected by most experts, the deployment of AVs in the real world will create new types of accidents, due to the AV's sensorial limits and the interaction between autonomous systems and road users, among other reasons. These issues present a risk for the AV's acceptability, since without a careful treatment, these situations would be solved using criteria defined for common driving, which could result in morally unjustifiable results. Thus, even if situations like the ones described remain scarce, public disapproval would increase with time, threatening the use of not only AVs, but the full range of autonomous solutions that may be deployed in future years.

Ethical dilemma situations raise concerns for vehicle manufacturers, the general public, and policy makers. Given that every measurement and action taken are probabilistic, and the enormity of interactions that can happen in an urban context, ethical dilemma situations, regardless of the AV's impressive capabilities, can and will occur [2]. As such, the AV's decision-making must be implemented in such a way as to incorporate ethical considerations. The main contribution of this article is the proposition of an algorithm to represent the AV's normal decision-making, an idea of how to measure action consequences, and three ethical dilemma policies which provide different deliberative methods from which the ideal action can be chosen. The purpose is to simulate different situations and to stimulate reflections on the possibility of ethical AV decisions.

This article is organized as follows: section II presents relevant research on decision-making for AVs and the implementation of ethical decision-making in autonomous systems. Section III details the decision-making algorithm used to guide the AV, while section IV introduces a harm quantifier for AV actions and proposes three different policies to deliberate in dilemma situations. Then, section V shows the results obtained in the AV simulation and finally, section VI provides a summary of the obtained results and next steps.

II. RELATED WORK

According to the driver's problem solving model proposed by [3], and adapted for the autonomous vehicles context by [4], the AV's planning problem structure is popularly defined by three levels: strategic, tactical and operational. The expression *decision making* refers to the tactical level of the planning problem, when a global route has been traced by the strategic level but the AV must decide how it should be executed, given the movements of other road users and any infrastructure limitations that may exist.

A. Decision making for AVs

Building on the deterministic decision-making models used in the 2007 DARPA Challenge [5], probabilistic decision making is often modelled as a Markovian decision process (MDP) [6]. In [7], the state space is considered as continuous, actions are represented as behaviors by a Dynamic Bayesian Network (DBN) which is realized as Bézier parameters, where Monte Carlo Sampling (MCS) is used to approximate the transition model. Uncertainties related to perception and the behaviour of surrounding vehicles were introduced into the reward function calculation using MCS in [8], with a continuous state space and a discrete action set.

Adding state observation uncertainties can be achieved by expanding the MDP to a Partially Observable Markov Decision Process (POMDP), as is done in [4]. Limiting itself to a lane change operation, this model uses a variant of the real-time belief space tree (RTBSS) to calculate the resulting policy in real-time. In another implementation of POMDP, [9] proposes the *generalized policy graph* to be used as a solution instead of a policy. Each node is a high-level action for the AV (stop, go, etc.), while edges are represented by the observation, guiding the Bellman backup through the graph. Also adopting the POMDP model, [10] returns to the implementation in [7], this time using particle filtering to calculate α-vectors and a state space

[1] Institut des Systèmes Intelligents et de Robotique, Sorbonne Université, 4 Place Jussieu, 75005 Paris, France. (gomes@isir.upmc.fr, raja.chatila@isir.upmc.fr)

[2] Sciences, Normes, Démocratie, Sorbonne Université, 1, rue Victor Cousin, 75005 Paris, France (katherine.evans@vedecom.fr, stephane.chauvier@sorbonne-universite.fr)

[3] Institut VEDECOM, 23bis Allée de Marronniers, 78000 Versailles, France (ebru.dogan@vedecom.fr)

representation in the form of an offline decision tree, which it then executes online. With intersections as a use-case, [11] defined a POMDP to control the AV's acceleration using an Adaptive Belief Tree (ABT) to sample the most likely future states, creating a belief representation. Bayesian change-point detection is used in [12] to classify observed states from other vehicles based on predefined action-policies (formed by low-level commands), which are used as AV actions. To calculate the solution, these policies are used in the value calculation, creating a sort of 'single stage' MDP.

Another way to calculate a POMDP problem solution is to use reinforcement learning. Similar to the action-policies used in [12], [13] proposes a deep reinforcement learning (DRL) network capable of producing high-level decisions and low-level actions, which are represented by option policies for a four-way intersection environment. For the lane change operation, [14] tries to tackle the usual slow learning rate of DRL algorithms using only thirteen features to model the state: using information about distances, relative velocities in respect to other users, and the AV's absolute velocity. It also proposes an action verification using a highway code formalization [15].

As was done in [14], including risk assessment is a common theme in POMDP models. Using GIDAS (German in-depth accident study) data to build a car-pedestrian benchmark, [16] trained a hybrid POMDP planner and DRL algorithm to generate collision-free trajectories for AVs. Returning to an MDP implementation, [17] assess the risk for highway environments using common parameters, such as time to collision (TTC), time-headway (TH) and time-to-front (TTF) for future states, while the action state sequences are realized by a dynamic model predictive control (MPC).

B. Ethical decision making

Since the advent of autonomous systems and our expectations for their deployment in the real world, certain ethical concerns related to these systems have become a well known problem. In [18], two implementations of the casuistry approach were used to compare similarities and differences between ethical dilemma cases and to retrieve other cases that may be relevant to the one considered. Ethical constraints for autonomous weapons were addressed by [19] through an ethical governor capable of evidence-based reasoning, which transforms evidence into logical assertions to be used by mandatory constraints.

Formal methods were used by [20] to model dilemma situations, making it possible to apply an ethical framework which classifies actions as acceptable or unacceptable. Beyond this, [21] proposes a formal language to model a generic ethical dilemma, allowing the artificial agent to reason about an ethical choice during its occurrence using the information at hand. However, both methods do not include a probabilistic component in their decision-making model.

To deliberate in an AV dilemma scenario, [22] advocates for an AV's mixed strategy, the result of the Nash equilibrium. Using another approach, [23] applies the Maximin principle from rawlsian contractualism, outlining an algorithm to implement such a principle while using a "veil of ignorance" to conceal information that could create a bias.

But one can only go so far with hypothetical situations. Parameters such as dynamic interactions between agents and uncertain action consequences must be taken into account to produce fair, realistic, and acceptable ethical deliberation. Focusing on a situation when an AV must decide between a collision with pedestrians or a physical barrier, [24] used an estimation of the AV's deformation against the barrier and statistical data to define a risk velocity for the pedestrian collision. Addressing low level decision-making, [25] defined normal control constraints, such as steering wheel limits, traffic rules and smooth driving, as deontological and consequentialist constraints in an MPC controller for an overtaking operation.

III. MODELING AV DECISION MAKING

Five components define a finite horizon MDP [6]: state space (S), action set (A), transition probability function (P), reward function (R) and the discount factor (γ). Unlike many MDP (or POMDP) implementations that consider the state of the other road users in the state definition, the proposed method uses only the AV's configuration, composed by the AV's middle rear-axis point coordinates (x,y), direction θ, scalar velocity v and steering angle ϕ, given by 1, as state variables.

$$s_t = [x_t, y_t, \theta_t, v_t, \phi_t] \tag{1}$$

The behavior prediction of other road users is determined outside the MDP process and accounted for in the reward function. A policy is the final result, dictating which action should be executed at each state to achieve the best trajectory, then passed to the AV's controller. Up to this point no ethical considerations are used in the decision making process. This is addressed in section IV.

A. State space discovery

Starting from the AV's initial position, all $s_i \in S$ are discovered by propagating the position using an action through the vehicle model. Actions are defined *a priori*, consisting of linear acceleration and steering angle velocity pairs, $(\dot{v}, \dot{\phi})$, which are applied over a period of time Δt_{pol}. For the AV's motion model, the single track vehicle model with front wheel drive and without slippage is used. Its analytical model is given by equation 2.

$$\begin{cases} \dot{x}_{t+1} &= v_t \cos\theta_t \cos\phi_t \\ \dot{y}_{t+1} &= v_t \sin\theta_t \cos\phi_t \\ \dot{\theta}_{t+1} &= \frac{v_t}{l} \sin\phi_t \end{cases} \tag{2}$$

The available actions are combinations of $\dot{v} = \{-c, 0, c\}\,(m/s^2)$ and $\dot{\phi} = \{-40, -20, 0, 20, 40\}\,(^\circ/s)$, c being a constant.

B. Transition probability definition

Given that $\dot{\phi}$ controls the AV's direction (and thus which maneuver is executed), it defines the transition probability. As a hypothesis, if a positive acceleration is chosen, then the AV velocity will increase (or respectively decrease for a negative one). Such simplification is sufficient to analyze possible dilemma situations. The position of other road users are considered to be known.

In each set of actions with equal acceleration, the chosen action will have a probability of 0.8 of being successful, meaning that for state $s_{0,0}$, in figure 1, the action a_3 takes AV to $s_{1,3}$ with probability of 0.8 and to $s_{1,2}$ or $s_{1,4}$ with probability of 0.1 each. If only one neighbor exists, then it becomes 0.9/0.1 (action a_0).

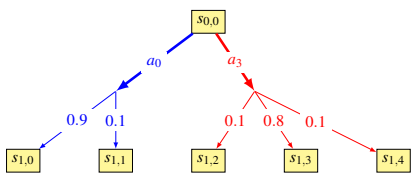

Fig. 1: State transition uncertainty for a_0 and a_3

C. Reward function

The reward function indicates to the system which actions are preferable to achieve a certain goal. It is based on two parameters: performance (s_{perf}) and action consequences (s_{conseq}), as given by equation 3.

$$R(s_t, a_t, s_{t+1}) = s_{\text{perf}} + s_{\text{conseq}} \tag{3}$$

Performance for the vehicle is measured as distance to the strategic path and its endpoint. Equation 4 calculates the reward using three variables: q_{lat} represents the lateral distance to the trajectory, $\Delta\theta$ is the offset angle between the direction of the AV and the trajectory and q_{eta} represents the estimated time of arrival (ETA).

$$s_{\text{perf}} = w_{\text{lat}} \cdot q_{\text{lat}} + w_{\text{dir}} \cdot \Delta\theta + w_{\text{eta}} \cdot q_{\text{eta}} \tag{4}$$

If the AV has a velocity equal to zero, then equation 4 is replaced by a fixed cost c_{eta}. The weights w_{lat}, w_{dir}, w_{eta} are calibrated heuristically, enabling the AV to arrive at its goal following the trajectory (disregarding other road users).

Action consequences are measured by two parameters, adherence to the traffic code (s_{traf}) and proximity to other road users (s_{prox}), as equation 5 indicates. The values s_{traf} and s_{prox} can assume negative values or 0, with the latter approximately 10 times lower than the former (s_{perf} can assume positive or negative values). Such measure avoid trade-offs between traffic code violations and possible collisions. Traffic code violations are not considered to be constitutive of a dilemma situation event, since the objective of this paper is to focus in unavoidable collisions.

$$s_{\text{conseq}} = s_{\text{traf}} + s_{\text{prox}} \tag{5}$$

A cascade of *if-else* emulates some traffic rules into the AV's decision process by adding costs when necessary.

- Velocity above limit: add the cost c_{vel} to reward.
- AV in opposite lane: add the cost c_{oplane} to reward.
- AV in the sidewalk: add the cost c_{sidewalk} to reward.

The evaluation of s_{prox} considers the proximity to other road users and the occurrence of a collision. Figure 2 defines the security zones close to the vehicle, which are the regions between the vehicle's body and the red lines. These frontiers represent an implicit risk measure for the reward function, since their role is to verify if breaking while maintaining the same direction is still a valid collision avoidance policy and if it is not, the action's reward is negatively impacted.

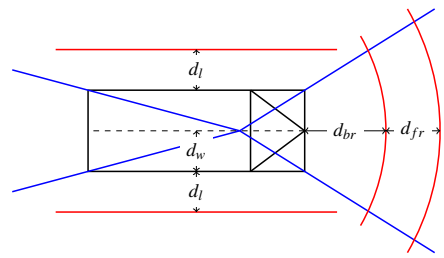

Fig. 2: AV with all the proximity limits

If a road user is predicted to be inside one of these zones, then a cost, calculated by equation 6, is added to the reward. Lateral frontiers are placed at a distance equal to d_l from the AV's body (d_w). The frontal limit is given by d_{fr} plus d_{br} (breaking distance). Both d_{fr} and d_l take into account the displacement of other road users during t_{pol}.

$$s_{\text{prox}} = c_{\text{st}} + w_v \cdot \left[\Delta v_{t+1}^{\text{proj}} - \Delta v_t^{\text{proj}}\right] \tag{6}$$

In equation 6, c_{st} is the static cost related to the road user's presence inside the AV's limits. Both Δv measure the difference between the road user's and AV's projected velocity onto the distance vector between them; the projections can be seen in Figure 3. Such quantity shows if the AV is converging on (Δv negative) or diverging from (Δv positive) the road user. Therefore the weight w_v increases or decreases how fast the AV changes direction due to possible collisions.

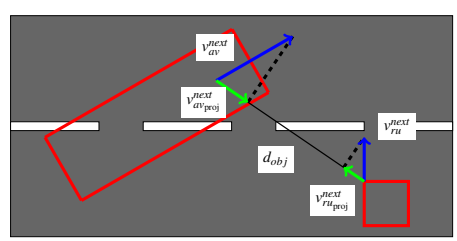

Fig. 3: Calculation of velocity projections

IV. EVALUATING DILEMMA SITUATIONS

Until now the AV has been controlled without any ethical considerations. A dilemma situation is characterized by the

AV's inability to avoid an accident. It can be detected when all $a_i \in A$ cause an accident. In such a case, the reward function as proposed by equation 3 does not account for the moral implications arising from the chosen course of action. These actions necessarily have a moral component, since they impose a risk distribution in the environment which has consequences for all road users [2].

The approach presented from this point onwards uses a quantitative variable to measure the consequences of an action for each road user, and a deliberation process to decide which action is more appropriate given a moral theory. It intervenes when, and only when, a dilemma situation is detected (since in non-dilemma situations the AV, as it was defined, should be able to cruise only with analytical parameters). Three different deliberation processes are proposed, one based on rawlsian contractualism [26], one based on utilitarianism [27] and another based on egalitarianism [28].

A. Defining Harm

To estimate an accident's severity, the most popularly employed measurement is the difference of velocity (Δv) between road users prior to a collision [29], [30]. But Δv does not consider the accident context (road user's mass and collision mechanics, for example). One can start considering that $h \propto \vec{J}$, c_{vul}, h representing the road user's harm, J for the impulse applied during collision and c_{vul} for a measure of physical vulnerability. The expression then becomes $h \propto \Delta\vec{p}$, c_{vul}, resulting in the equation 7, for a road user k:

$$^{k}h(s_t, a_t, s'_t) = {}^{k}c_{vul} \cdot \left(\|\vec{v}_f - {}^{k}\vec{v}_i\| \right) \tag{7}$$

To calculate the velocity after collision, v_f, both road users (k and l) are considered as punctual masses and all dissipation forces are neglected in equation 8.

$$m_k {}^{k}\vec{v}_i + m_l {}^{l}\vec{v}_i = (m_k + m_l)\vec{v}_f \tag{8}$$

For collisions with pedestrians, it is considered that the pedestrian's $\vec{v}_f$ is equal to the AV's $\vec{v}_i$. Thus, the AV's velocity does not change, and its harm is equal to zero.

A plethora of studies exist in accidentology addressing collisions for pedestrians or vehicles. The role of c_{vul} is to represent the inherent physical vulnerability of each road user. It is proportional to injury probability given a Δv between two road users involved in a collision. This is an indirect measurement, since it is velocity dependent, but it should give an estimate of the physical differences between pedestrian and vehicles.

The accuracy level of this evaluation depends on which information is used to plot the risk vs Δv curves. Ideally one would need data that considers angle of collision, compatibility between vehicles, vehicle masses and different types of road users but, regrettably, such detailed data is hard to find. These curves are more abundant for fatality risk and frontal collisions. For the simulation in section V, they will be adopted as the source for the c_{vul} determination ([30] and [31]).

B. Action deliberation

Using the harm value, defined by equation 7 for each road user, an action could be chosen in a dilemma situation. But the transition probability should also be taken into account, and is covered by equation 9 which defines expected harm.

$$^{k}h_{exp}(s_t, a_t) = \sum_{s_{t+1} \in S} p(s_{t+1}|s_t, a_t)\, {}^{k}h(s_t, a_t, s'_t) \tag{9}$$

In dilemma scenarios, some form of ethical reasoning should be used to deliberate about an action. Since most ethical theories are created for personal forms of practical reason, or states and societies, some degree of conversion from this context to the dilemma situations an AV might face is required. Three different ethical theories were chosen:

1) **Contractarian**: Two principles of justice are proposed by [26], the first requires that all individuals have access to an equal and maximal degree of liberty and rights, which is translated in our context as a road user's right to not be targeted or directly disadvantaged in the AV's deliberation . The second principle requires that if there is some injustice, characterized by a difference of rights or treatment, it should be beneficial to the worst off. Thus, starting from the most equal state, when the standard deviation is minimal for an action a (equation 10; $h_{exp}(a)$ represents the set ${}^{i}h_{exp}(a)$, $0 \leq i \leq n$, n being the number of road users in the scene), the largest expected harm is minimized, as long as the harm for all others does not increase (equation 11).

$$a_{int} = \underset{a \in A}{\operatorname{argmin}}\, \sigma_{h_{exp}}(a) \tag{10}$$

$$\underset{a}{Min} \left\{ {}^{i}h_{exp}(a) \right\}, \text{ subject to } {}^{j}h_{exp}(a) \leq {}^{j}h_{exp}(a_{int}) \tag{11}$$

2) **Utilitarian**: This approach is rooted in the Greatest Happiness Principle, which states that the righteousness of an action is proportional to how much happiness it promotes, and consequently its wrongfulness is proportional to how much pain it produces [27]. As such, utilitarianism defends the maximization of the total amount of good (or happiness) as a method to find the most rightful action. Using the second formulation of the same problem, one can minimize the total amount of expected harm to find the "least wrongful action", as equation 12 proposes.

$$a_{eth} = \underset{a}{\operatorname{argmin}} \sum_{i=0}^{n} {}^{i}h_{exp}(a) \tag{12}$$

3) **Egalitarian**: Beyond the idea to minimize the overall expected harm, egalitarianism also proposes to divide the expected harm avoiding large discrepancies between each road user [28]. Therefore the variance of each road user's expected harm at each action is used as a weight for the sum of expected harms. Equation 13 represents such idea.

$$a_{eth} = \underset{a}{\operatorname{argmin}} \sum_{i=0}^{n} var_{h_{exp}}(i,a) \cdot {}^{i}h_{exp}(a) \tag{13}$$

As it can be seen, a road user is represented by his expected harm. The number of passengers inside the vehicle, or the pedestrians in close proximity, are not considered into the decision making.

C. Solving the MDP

As previously stated in section III, the MDP to be solved consists of a discrete state space and a discrete action set, with a discount criterion and four transitions. An iteration process uses equation 14, known as the Bellman's equation, to maximize the value for all states.

$$V_{t+1}(s_t) = \max_{a \in A} \left[R(s_t, a, s_{t+1}, e) + \gamma \cdot \sum_{s_{t+1}} P(s_{t+1}|s_t, a) V_t(s_{t+1}) \right] \tag{14}$$

When the convergence is achieved, arriving at a mean squared error smaller that ε, the policy is extracted using the equation 15.

$$\pi^*(s_i) = \underset{a \in A}{\operatorname{argmax}} \sum_{s_{t+1} \in S} P(s_{t+1}|s_t, a) V_t(s_{t+1}) \tag{15}$$

Some states may have actions that cause collisions and actions that do not. Those that cause accidents are discarded during the iteration of 14. If a dilemma situation arises, then one of the three deliberation methods of subsection IV-B is used to select $\pi^*(s_i)$ action.

But one should also consider the future consequences of these actions that do not provoke an accident immediately. This is done projecting future collisions into previous states. Figure 4 shows how this mechanism works. Red states show that a collision occurred during the transition, and in such a case the AV breaks and stops. State $s_{2,2}$ is critical, since all its actions will cause an accident. When the deliberation is done for $s_{1,0}$, which is not critical *per se*, it realizes that $s_{2,2}$ is. Thus $s_{1,0}$ is considered to be critical and the information used in the ethical deliberation of $s_{2,2}$ is transfered to $s_{1,0}$ to be used in its ethical deliberation. The same procedure is done for $s_{0,0}$.

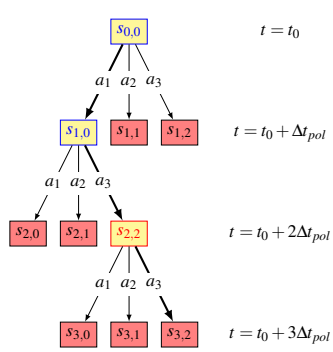

Fig. 4: Projection of future collision into previous states

V. SIMULATION RESULTS

Using the WeBots simulator, the aforementioned algorithms were tested in the scenario shown by figure 5. While the AV is traveling in its lane, two pedestrians, P1 and P2 suddenly cross into the street (start moving at $t_{simul} = 4s$). The AV cannot stop, so it decides to avoid them using the opposite lane. At this moment, an incoming vehicle, VEH, is detected, right in its immediate front, creating a critical situation: the AV can collide with the vehicle, the pedestrians or with the physical barriers along the sidewalk's external borders. Each moral profile will be used to determine the best course of action.

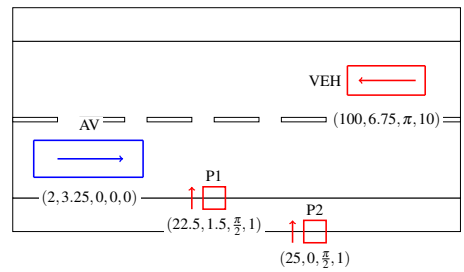

Fig. 5: Initial simulation setup (other road users' configuration is represented by (x, y, θ, v), not in scale)

All MDP parameters and the physical information of road users' are shown in table I.

TABLE I: Road users' physical properties

	AV	**Pedestrian**	**Vehicle**
Mass (kg)	$2105+80$	80	$1065+80$
Width, Height (m)	$4.853, 1.65$	$0.625, 0.625$	$3.475, 1.310$
Wheelbase (m)	2.933	–	2.55

A. Decision making without dilemma situations

The AV should not produce accidents or critical situations on its own. Varying the initial position of the vehicle (x_{veh}) and the velocity direction of one of the pedestrians (θ_{ped}) relative to figure 5, different AV behaviors can be observed. Each policy predicts two seconds ahead, with one second used for control ($(\dot{v}, \dot{\phi})$ are directly fed into the AV). The AV needs at least two transitions to verify if the next action will not cause an accident, and that a state has actions which will not necessarily lead to a dilemma situation. The weights of equations 4 and 6 were set heuristically to obtain a result representative of the MDP defined in section III.

The first of the AV's behavior plots shown in figure 6, for $x_{veh} = 100$ is acceptable, avoiding the pedestrian and returning to the other lane afterwards. All road user states are considered observable and known, therefore at the beginning of a policy calculation the AV can assert correctly all its states and predict changes throughout the decision horizon. But since the AV only has the capacity to predict two seconds ahead of time, it starts the maneuver without knowing how it will end. This can be seen more clearly in the second

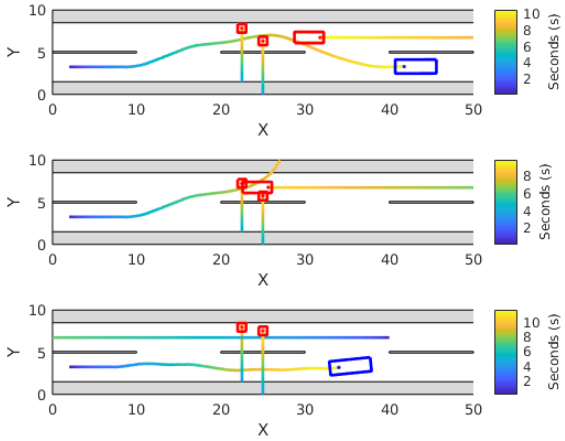

Fig. 6: AV's behavior for $x_{veh} = 100$, 80, 40m

plot, for $x_{veh} = 80$ when after the beginning of the maneuver, it reaches a critical situation and simply collides with the wall beyond the sidewalk. Adding one mode transition to the policy produces a viable trajectory, seen in figure 7 first plot, but increases the total calculation time from $t = 60.1$s to $t = 802.4$s. In figure 8, we can see that there is a difference between velocities and directions in $t \in [3, 7]$. The capacity to predict whether the other vehicle will end up too close to the AV causes the AV to adjust its direction to return to the correct lane as fast as possible.

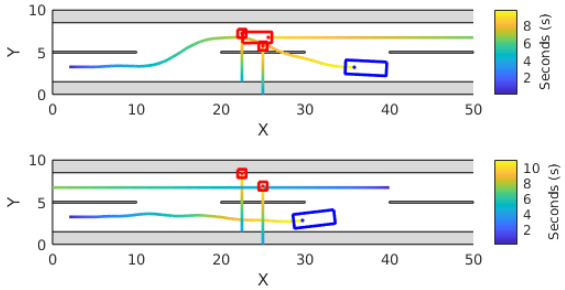

Fig. 7: AV's behavior for $x_{veh} = 80, 40$m with 5 transitions

Starting the other vehicle at $x_{veh} = 40$m (third plot in figure 6) creates a situation in which the AV cannot execute the avoidance trajectory. It is perhaps intuitive for a human driver to stop and wait for the pedestrians, but here the AV only reduces its velocity at a certain point and continues to move until the path is clear. Even with five transitions (second plot in figure 7) the behavior is similar, meaning that it needs more transitions to receive the delayed reward from the future.

Some conclusions can be drawn from these three different situations: it is essential for the decision-making of the autonomous vehicle that the decision horizon be increased. However, the exponential nature of the state space discovery process (subsection II-A) must be avoided. One way to do this is to sample the most probable next states given a current state and an action, focusing only on these discovered next states (as some POMDP solvers do). Another problem

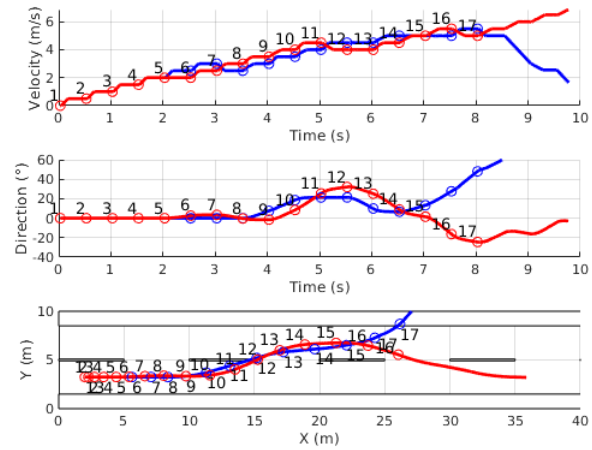

Fig. 8: AV's velocity, direction and trajectory for $x_{veh} = 80$m

occurs when the AV's velocity increases too much, causing a single turn left action to create a dilemma situation, since all actions and the transition time (Δt_{pol}) are fixed. The AV cannot compensate the increase of velocity into each action consequence, therefore it chooses to turn right due to the negative influence the cost has into each neighboring state. To avoid these results, the actions and/or policy time should be dynamic and dependent on the AV's velocity. Risk evaluation is also necessary to stop the AV from starting a maneuver that cannot be completed in one policy, as done in figure 6. Defining the state space without a fixed number of transitions and using the end of a maneuver as a sign of policy termination would always allow the AV to start and end in low-risk states.

Changing the direction of the first pedestrian from $\frac{\pi}{2}$ rad to $\frac{3\pi}{4}$ rad creates a different situation, but the results obtained are similar. Figure 9 shows the trajectory at $x_{veh} = 100$ m and 40 m and figure 10 at $x_{veh} = 40$ m for five transitions.

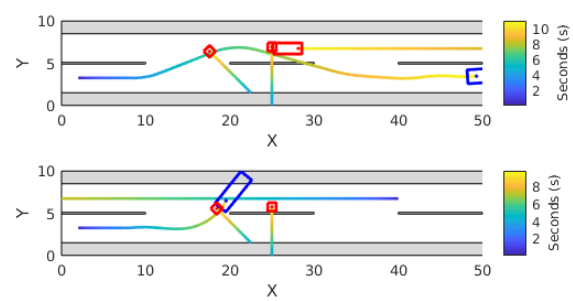

Fig. 9: AV's behavior for $x_{veh} = 100$, 40m and $\theta_{ped} = \frac{3\pi}{4}$rad

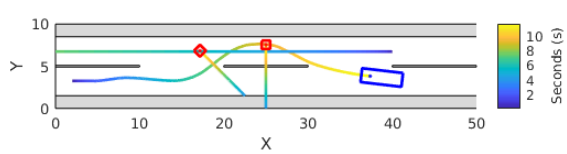

Fig. 10: Traj. for x_{veh}=40m, θ_{ped}=$\frac{3\pi}{4}$rad for 5 transitions

In comparison with figure 6, figure 9 produces similar behavior, where an additional transition allows the AV to avoid an accident. For $x_{veh} = 40$ and 5 transitions, it makes the AV choose to accelerate to avoid an accident, which in turn causes it to invade the sidewalk. With only four transitions the collision cost was not detected, only the sidewalk invasion is avoided, which causes the accident.

Another critical point in the implementation is the weights used for the reward function. Such weights were defined heuristically for the specific use case considered. Even so, they do not reproduce in the AV's behavior an ideal trajectory, which can be seen in figure 6 between $[5, 15]$ at the X axis. The small but perceptible change in AV's direction is not illegal but pointless.

B. Decision making in dilemma situations

As mentioned in the previous subsection, the ethical dilemma scenario is created moving VEH abruptly in just one time step (also, the x coordinate of P1 and P2 are changed to 20 and 22.5, respectively). At this point (where the AV is in the same state for all policies tested) the AV must deliberate about an action, using one of the deliberation processes proposed in subsection IV-B. The velocity difference between each road user (Δv) is not significant enough to be identified in the fatality probability versus Δv graphic given by [30] and [31], thus the constant c_{vul} used for the collisions will be equal to one.

As it can be seen in figure 11, the result for all cases is a collision with the pedestrian. But each policy chooses a different action, which imposes different collision consequences. Take the contractarian policy, table II: instead of decreasing the velocity right before the collision (actions 1 through 5), it chooses to maintain it ($v_{\text{AV}} = 4m/s$) to escape a collision with the other vehicle (action a_5). The AV ends up colliding with P1 (in the next calculated policy) with ($v_{\text{AV}} = 4m/s$).

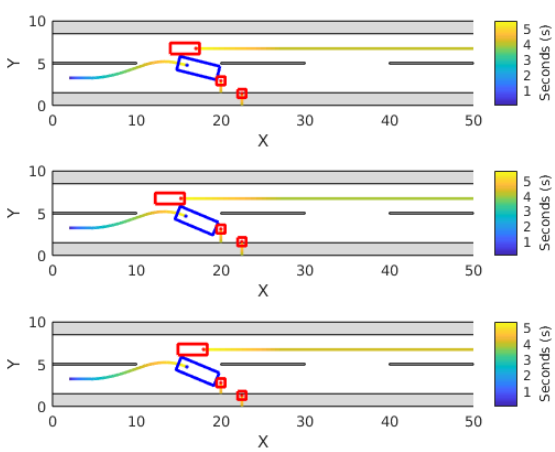

Fig. 11: Trajectories for contractarian, utilitarian and egalitarian policy, respectively

For the utilitarian policy, (second plot of figure 11), action a_0 is chosen because it minimizes the total amount of harm,

TABLE II: Expected harms and $\sigma_{h_{exp}}$ for contractarian policy

Action $(m/s^2, \circ/s)$		AV	P1	VEH	P2	$\sigma_{h_{exp}}$
a_0	$(-1, -40)$	3.64	0	6.94	0	2.89
a_1	$(-1, -20)$	3.64	0	6.94	0	2.89
a_2	$(-1, 0)$	3.61	0	6.90	0	2.87
a_3	$(-1, 20)$	3.54	0	6.77	0	2.82
a_4	$(-1, 40)$	3.49	0	6.76	0	2.81
a_5	$(0, -40)$	0.383	4.38	0.730	4.00	1.83

as can be seen in table III. It decreases the velocity to $v_{\text{AV}} = 3.5m/s$ and then increases to $v_{\text{AV}} = 4m/s$ as to escape the other vehicle. The final collision velocity, in the next calculated policy is $v_{\text{AV}} = 3.1m/s$.

TABLE III: Expected harms and Σh_{exp} for utilitarian policy

Action $(m/s^2, \circ/s)$		AV	P1	VEH	P2	Σh_{exp}
a_0	$(-1, -40)$	0.364	3.90	0.694	0	4.96
a_1	$(-1, -20)$	3.28	0.433	6.25	0	9.96
a_2	$(-1, 0)$	3.63	0	6.93	0	10.57
a_3	$(-1, 20)$	3.58	0	6.83	0	10.40
a_4	$(-1, 40)$	3.55	0	6.81	0	10.36

The egalitarian policy, trying to achieve a compromise between total expected harm minimization and dispersion minimization actually increases the velocity right after the ethical deliberation, and ends up colliding with the pedestrian with $v_{\text{AV}} = 4.5m/s$. Action a_{10} predicts that both pedestrians will be struck by the AV. This policy, in this example, produces more negative consequences than the other two, but maybe with another formulation (for example not using variance as weight, but another measure) it can represent a compromise between a contractarian and a utilitarian policy.

TABLE IV: Expected harms and total cost for egalitarian policy

Action $(m/s^2, \circ/s)$		AV	P1	VEH	P2	Total
a_0	$(-1, -40)$	3.77	3.34	7.20	0	95.06
a_1	$(-1, -20)$	3.76	3.05	7.18	0	98.07
a_5	$(0, -40)$	0.385	4.69	0.735	0	50.00
a_6	$(0, -20)$	3.46	2.73	6.61	0	77.50
a_{10}	$(1, -40)$	0.431	4.91	0.824	4.60	46.05

VI. CONCLUSION AND NEXT STEPS

The deployment of autonomous vehicles presents an opportunity to decrease the number of traffic accidents, but one also must consider the moral issues that can arise if this autonomous system provokes accidents. To explore the consequences of different ethical approaches when a collision is inevitable, a framework was proposed, consisting of a measure of harm for collisions and different interpretations of ethical theories that could be used as an action deliberation method in dilemma scenarios. However, we do not advocate a specific ethical approach.

In terms of future work, the MDP algorithm must be adapted to consider maneuvers through its policy calculation, passing through the increase in prediction capacity. A more efficient state space sampling, together with a dynamic method for transition probability estimation, is essential to accomplish such an objective. It would likewise be interesting to consider higher velocities to measure the trade-offs between c_{vul} and velocity variation during collisions. Additionally, the harm calculation would be more precise if detailed Δv diagrams were available and more informative if probability of serious injury was used as a risk parameter.

For the ethical deliberation, it is also necessary to consider deontological rules, like [20] and [21], since a quantification of actions is a soft constraint, while morally there are actions which can be considered unjustifiable, even from a harm-minimization standpoint.

Acknowledgment

This research has been conducted as part of the AVEthics project funded by the French National Agency for Research, grant agreement number ANR-16-C22-008.

References

[1] J. Fleetwood, "Public health, ethics, and autonomous vehicles," *American journal of public health*, vol. 107, no. 4, pp. 532–537, 2017.

[2] N. J. Goodall, "Ethical decision making during automated vehicle crashes," *Transportation Research Record*, vol. 2424, no. 1, pp. 58–65, 2014.

[3] J. A. Michon, "A critical view of driver behavior models: what do we know, what should we do?" in *Human behavior and traffic safety*. Springer, 1985, pp. 485–524.

[4] S. Ulbrich and M. Maurer, "Probabilistic online pomdp decision making for lane changes in fully automated driving," in *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*. IEEE, 2013, pp. 2063–2067.

[5] C. Urmson, J. Anhalt, D. Bagnell, C. Baker, R. Bittner, M. Clark, J. Dolan, D. Duggins, T. Galatali, C. Geyer *et al.*, "Autonomous driving in urban environments: Boss and the urban challenge," *Journal of Field Robotics*, vol. 25, no. 8, pp. 425–466, 2008.

[6] O. Sigaud and O. Buffet, *Markov decision processes in artificial intelligence*. John Wiley & Sons, 2013.

[7] S. Brechtel, T. Gindele, and R. Dillmann, "Probabilistic mdp-behavior planning for cars," in *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2011, pp. 1537–1542.

[8] J. Wei, J. M. Dolan, J. M. Snider, and B. Litkouhi, "A point-based mdp for robust single-lane autonomous driving behavior under uncertainties," in *2011 IEEE International Conference on Robotics and Automation*. IEEE, 2011, pp. 2586–2592.

[9] H. Bai, D. Hsu, and W. S. Lee, "Integrated perception and planning in the continuous space: A pomdp approach," *The International Journal of Robotics Research*, vol. 33, no. 9, pp. 1288–1302, 2014.

[10] S. Brechtel, T. Gindele, and R. Dillmann, "Probabilistic decision-making under uncertainty for autonomous driving using continuous pomdps," in *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2014, pp. 392–399.

[11] C. Hubmann, J. Schulz, M. Becker, D. Althoff, and C. Stiller, "Automated driving in uncertain environments: Planning with interaction and uncertain maneuver prediction," *IEEE Transactions on Intelligent Vehicles*, vol. 3, no. 1, pp. 5–17, 2018.

[12] E. Galceran, A. G. Cunningham, R. M. Eustice, and E. Olson, "Multi-policy decision-making for autonomous driving via changepoint-based behavior prediction: Theory and experiment," *Autonomous Robots*, vol. 41, no. 6, pp. 1367–1382, 2017.

[13] Z. Qiao, K. Muelling, J. Dolan, P. Palanisamy, and P. Mudalige, "Pomdp and hierarchical options mdp with continuous actions for autonomous driving at intersections," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2018, pp. 2377–2382.

[14] B. Mirchevska, C. Pek, M. Werling, M. Althoff, and J. Boedecker, "High-level decision making for safe and reasonable autonomous lane changing using reinforcement learning," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2018, pp. 2156–2162.

[15] A. Rizaldi, J. Keinholz, M. Huber, J. Feldle, F. Immler, M. Althoff, E. Hilgendorf, and T. Nipkow, "Formalising and monitoring traffic rules for autonomous vehicles in isabelle/hol," in *International Conference on Integrated Formal Methods*. Springer, 2017, pp. 50–66.

[16] F. Pusse and M. Klusch, "Hybrid online pomdp planning and deep reinforcement learning for safer self-driving cars," in *2019 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2019, pp. 1013–1020.

[17] C. Xu, W. Zhao, and C. Wang, "An integrated threat assessment algorithm for decision-making of autonomous driving vehicles," *IEEE Transactions on Intelligent Transportation Systems*, 2019.

[18] B. M. McLaren, "Computational models of ethical reasoning: Challenges, initial steps, and future directions," *IEEE intelligent systems*, vol. 21, no. 4, pp. 29–37, 2006.

[19] R. C. Arkin, P. D. Ulam, and B. Duncan, "An ethical governor for constraining lethal action in an autonomous system," Georgia Institute of Technology, Tech. Rep., 2009.

[20] V. Bonnemains, C. Saurel, and C. Tessier, "Embedded ethics: some technical and ethical challenges," *Ethics and Information Technology*, vol. 20, no. 1, pp. 41–58, 2018.

[21] F. Berreby, G. Bourgne, and J.-G. Ganascia, "Modelling moral reasoning and ethical responsibility with logic programming," in *Logic for programming, artificial intelligence, and reasoning*. Springer, 2015, pp. 532–548.

[22] V. Estivill-Castro, "Game theory formulation for ethical decision making," in *Robotics and Well-Being*. Springer, 2019, pp. 25–38.

[23] D. Leben, "A rawlsian algorithm for autonomous vehicles," *Ethics and Information Technology*, vol. 19, no. 2, pp. 107–115, 2017.

[24] J. E. Pickering, M. Podsiadly, and K. J. Burnham, "A model-to-decision approach for the autonomous vehicle (av) ethical dilemma: Av collision with a barrier/pedestrian (s)," *IFAC-PapersOnLine*, vol. 52, no. 8, pp. 257–264, 2019.

[25] S. M. Thornton, S. Pan, S. M. Erlien, and J. C. Gerdes, "Incorporating ethical considerations into automated vehicle control," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 6, pp. 1429–1439, 2016.

[26] J. Rawls, *A Theory of Justice. Rawls*. The Belknap, 1971.

[27] J. S. Mill, "Utilitarianism (1863)," *Utilitarianism, Liberty, Representative Government*, pp. 7–9, 1859.

[28] B. John, *Equality*. John Wiley Sons, Ltd, 2017, ch. 9, pp. 174–201. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119451266.ch9

[29] L. Evans, "Driver injury and fatality risk in two-car crashes versus mass ratio inferred using newtonian mechanics," *Accident Analysis & Prevention*, vol. 26, no. 5, pp. 609–616, 1994.

[30] Q. Hussain, H. Feng, R. Grzebieta, T. Brijs, and J. Olivier, "The relationship between impact speed and the probability of pedestrian fatality during a vehicle-pedestrian crash: a systematic review and meta-analysis," *Accident Analysis & Prevention*, vol. 129, pp. 241–249, 2019.

[31] D. Richards, "Relationship between speed and risk of fatal injury: pedestrians and car occupants," Tech. Rep. 16, 2010. [Online]. Available: https://www.roadsafetyobservatory.com/Evidence/Details/11281