Ethical V2X: Cooperative Driving as the Only Ethical Path to Multi-Vehicle Safety

Sidorenko, Galina; Thunberg, Johan; Vinel, Alexey

Published in:
2023 IEEE 98th Vehicular Technology Conference (VTC2023-Fall)

DOI:
10.1109/VTC2023-Fall60731.2023.10333432

2023

Document Version:
Peer reviewed version (aka post-print)

Ethical V2X: Cooperative Driving as the Only Ethical Path to Multi-Vehicle Safety

Galina Sidorenko, Johan Thunberg and Alexey Vinel

***Abstract*—We argue that an information exchange between vehicles via the vehicular communications is the foundation for ethical driving. In other words – autonomous vehicles must be cooperative to be able to resolve ethical dilemmas in a multi-vehicle scenario. We show this by exploring the minimal setting of a longitudinal driving in a formation of three vehicles.**

***Index Terms*—Cooperative vehicles, automated driving, vehicular safety, ethical dilemmas, V2X communications.**

I. Introduction

Ethical dilemmas in autonomous driving are widely discussed [1]–[3]. More specifically, the severity of a potential collision, which is characterized via a notion of *harm*, can be taken into account while making driving decisions [4]. Recently, ethical trajectory planning algorithm for autonomous vehicles has been proposed in [5].

We argue that information exchange between the vehicles via vehicular communications (V2X) is the foundation for ethical driving. In other words – autonomous vehicles *must be cooperative* to be able to resolve ethical dilemmas in a multi-vehicle scenario. We show this by exploring the minimal setting of a longitudinal driving in a formation of three vehicles. We elaborate on the coordinated emergency braking concept [6] and apply the safety assessment approach similar to [7] to analyze the scenario of interest and present the following contributions:

- We formulate an ethical dilemma in longitudinal autonomous driving and conceptualize a V2X protocol to communicate information needed to perform ethical decision-making during the emergency braking.
- We develop a mathematical framework for the evaluation of the ethical dilemma in a scenario of a three-vehicles formation in a cooperative autonomous driving setting and provide guidelines for the design of an ethical emergency braking strategy.

The paper is structured as follows. Section II introduces the scenario and the ethical dilemma. Section III presents the mathematical framework for calculating the harm function in the considered scenario. Numerical examples illustrating the achievement of different ethical objectives are provided in Section IV. The role of the V2X for ethical autonomous driving and the outline of respective future work is discussed in Section V.

Galina Sidorenko, Johan Thunberg and Alexey Vinel are with Halmstad University, Halmstad, Sweden. Alexey Vinel is also with the Karlsruhe Institute of Technology (KIT), Germany. Corresponding author is Galina Sidorenko: galina.sidorenko@hh.se. We acknowledge the support of the Knowledge Foundation Project "SafeSmart", ELLIIT Strategic Research Network and Helmholtz Program "Engineering Digital Futures".

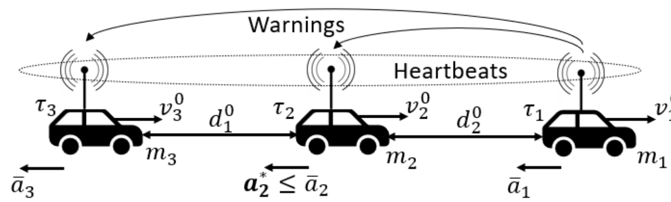

Fig. 1: Considered scenario.

II. Ethical Dilemma

Let us consider a formation of three autonomous vehicles (Fig. 1) with maximum braking capacities $\bar{a}_i$. At point of time τ_1 the vehicle 1 starts emergency braking due to, for example, an obstacle appeared on the road. As the response, vehicles 2 and 3 start braking at some moments τ_2 and τ_3, respectively. Let us assume vehicles 1 and 3 apply their maximum deceleration $\bar{a}_1$ and $\bar{a}_3$ and we focus on the behaviour of vehicle 2 which does not necessary perform maximum braking with $\bar{a}_2$, but can select a reduced deceleration $a_2^* \leq \bar{a}_2$ instead. The choice of an appropriate deceleration a_2^* depends on the given goal or interest of the vehicles involved. This is where ethical considerations come into play, encompassing a set of principles, values and norms that guide decision making. If the choice of deceleration a_2^* allows for avoidance of any collisions among the considered three vehicles, the resulting harm would be zero. However, in the case of any collisions between the vehicles, the harm should quantify the severity of the accident. Thus, the possible objectives could be, for instance, to choose a_2^* to reduce the individual harms H_i of the vehicles or to reduce the total harm H for the entire formation.

III. Solution Based on Harm

A. Main notations

The main notations that are used in the paper are summarized in Table I.

B. Scenario-related assumptions

Here, we list all the assumptions for the considered scenario:

- we abstract from the specific cooperative adaptive cruise controllers and assume that vehicle 2 maintains a constant velocity v_2^0 during interval $[\tau_1, \tau_2]$, and the vehicle 3 maintains a constant velocity v_3^0 during interval $[\tau_1, \tau_3]$;
- we consider central collinear collisions meaning that vehicles after a collision keep moving along the same central line without any angular changes or rotational motion;

Notation	Definition
d_i^0	distance between vehicles i-th and $(i+1)$-th at the moment when emergency braking starts, i.e., at $t = \tau_1$
v_i^0	velocity of the i-th vehicle at the moment when emergency braking starts, i.e., at $t = \tau_1$
τ_i	time when the i-th vehicle enters emergency braking by applying a_i^*. For vehicles 1 and 3, $a_i^* = \bar{a}_i$
$\bar{a}_i$	maximum deceleration capability of the i-th vehicle
a_i^*	chosen deceleration of the i-th vehicle during emergency braking, $0 \leq a_i^* \leq \bar{a}_i$
m_i	mass of the i-th vehicle
$t_{i-1,i}^*$	time of collision between $(i-1)$-th and i-th vehicles
$v_i(t^*)$	velocity of the i-th vehicle directly before the impact occured at t^*
$u_i(t^*)$	velocity of the i-th vehicle directly after the impact occured at t^*
$\Delta v_i(t)$	relative velocity $v_{i+1}(t) - v_i(t)$ at the moment t
e_i	coefficient of restitution for the impact between vehicles $(i+1)$-th and i-th
H_i	harm for vehicle i
H	summed harm for all 3 vehicles

TABLE I: Main notations.

- if a collision occurs, two collided vehicles obtain new velocities defined by conservation of momentum and coefficient of restitution e_i;
- after collision, vehicles move along the same central line and keep decelerating with new decelerations $\alpha_i a_i$ where a_i is deceleration applied by the vehicle i before the collision, and α_i is some coefficient;
- we consider only one pairwise collision, i.e., maximum one collision for every pair of neighbouring vehicles. This implies that, for example, after the collision between vehicles 1 and 2, vehicle 1 no longer has any impact on the movement of vehicle 2. In other words, vehicle 2 after collision obtain a new velocity and keeps decelerating while vehicle 1 does not affect this motion.

C. If collisions are avoidable

In the case of three vehicles, where 1-st and 3-rd vehicles apply their maximum deceleration, $\bar{a}_1$ and $\bar{a}_3$ respectively, the deceleration of the 2-nd vehicle can be chosen. It can be the case when the deceleration of the middle vehicle can be chosen in a such way that no collisions occur between any pair of vehicles.

Previously, in [8], formulas for calculating safe distance d_{i-1}^0 between two vehicles, $(i-1)$-th and i-th, were given. Keeping such a distance guarantees no collisions between the considered vehicles in the case of the emergency braking. By expanding those formulas [8] to take into account two time delays τ_{i-1} and τ_i and then inverting them, we can calculate corresponding a_i. Such a value of a_i is "safe" meaning that no collisions happen for the given parameters d_{i-1}^0, a_{i-1} and τ_{i-1}, τ_i if the i-th vehicle applies at least a_i during braking scenario. If the calculated a_i is non-feasible, then the collision between $(i-1)$-th and i-th vehicles is unavoidable. If it is feasible, we can choose any deceleration in the interval $[a_i, \bar{a}_i]$ and avoid collision with the preceding vehicle, $(i-1)$-th.

For the given deceleration a_{i-1} of the previous vehicle, the safe deceleration a_i that allows to avoid collisions between vehicles $(i-1)$-th and i-th is determined by inequalities:

$$a_i \geq \begin{cases} \frac{(v_{i-1}^0 - v_i^0 - a_{i-1}\tau)^2}{2(d_{i-1}^0 - a_{i-1}\frac{\tau^2}{2} + (v_{i-1}^0 - v_i^0)\tau_i)} + a_{i-1} \\ \text{if } d_{i-1}^0 \leq \frac{v_i^0(v_{i-1}^0 + a_{i-1}(\tau_{i-1}+\tau_i)) - v_{i-1}^0(v_{i-1}^0 + 2a_{i-1}\tau_{i-1})}{2a_{i-1}} \\ a_{i-1}\frac{(v_i^0)^2}{2d_{i-1}^0 a_{i-1} + (v_{i-1}^0)^2 + 2a_{i-1}(v_{i-1}^0\tau_{i-1} - v_i^0\tau_i)} \quad \text{otherwise} \end{cases} \tag{1}$$

Analogically, we can find "safe" deceleration a_i having fixed deceleration a_{i+1} of the following vehicle. Such safe a_i that allows avoiding a collision between vehicles i-th and $(i+1)$-th is determined by inequalities:

$$a_i \leq \begin{cases} \frac{2a_{i+1}(d_i^0 + (v_i^0 - v_{i+1}^0)\tau_{i+1}) - (v_i^0 - v_{i+1}^0)^2}{2d_i^0 + 2(v_i^0 - v_{i+1}^0)\tau_i + a_{i+1}\tau^2} \\ \text{if } d_i^0 \leq \frac{v_{i+1}^0(v_{i+1}^0 + 2a_{i+1}\tau_{i+1}) - v_i^0(v_{i+1}^0 + a_{i+1}(\tau_i + \tau_{i+1}))}{2a_{i+1}} \\ a_{i+1}\frac{(v_i^0)^2}{-2d_i^0 a_{i+1} + (v_{i+1}^0)^2 + 2a_{i+1}(v_{i+1}^0\tau_{i+1} - v_i^0\tau_i)} \quad \text{otherwise} \end{cases} \tag{2}$$

Formulas (1) and (2) together provide a safe interval A_i (if exists) of a_i such that any deceleration from this region allows to avoid collisions between three vehicles, i.e., $(i-1)$, i, and $(i+1)$.

D. If collisions are unavoidable

If collision among the considered three vehicles is unavoidable, the harm metric should reflect the severity of the occurred crash. On the path of presenting the final algorithm for calculating harm for the considered scenario with three vehicles, several essential aspects are discussed. In more details, the current section proceeds with presenting different harm functions, following by formulas for calculating vehicles' velocities after a collision. Further, formulas for calculating time of collision and harm function are presented for a two vehicle case that serve as a basis for the final algorithm.

Harm metrics: Different harm metrics can be employed to assess severity of the occurred collision between two vehicles. Without a loss of generality, in the current subsection, we assess harm for the vehicle 1 involved into a collision with vehicle 2.

According to accident research [9], there is a strong correlation between the probability of being injured and the collision induced speed change $\delta v_1 = u_1 - v_1$, i.e., the difference between ego vehicle's velocity after u_1 and before v_1 a collision. In other words, δv_1 can serve as an indicator of the severity of the crash.

In [5], the harm for the vehicle 1 involved in the collision with the vehicle 2 is modelled as a logistic function of relative velocity $\Delta v_1 = v_2 - v_1$ before the collision and vehicles masses. In other words, the harm monotonically increases with:

$$H_1 = \frac{m_2}{m_1 + m_2}\Delta v_1, \tag{3}$$

where m_1 and m_2 are masses of the corresponding vehicles. Note that if in metric (3) one considers square root of velocity change, i.e., $\frac{m_2}{m_1+m_2}(\Delta v_1)^2$, then such a new metric will reflect the kinetic energy absorbed by the vehicle 1 during the collision. Furthermore, using definition of coefficient of restitution that relates velocities of two involved vehicles before and after the impact:

$$e_1 = \frac{u_1 - u_2}{v_2 - v_1}, \quad 0 \le e_1 \le 1. \tag{4}$$

one can derive that

$$\delta v_1 = \frac{m_2}{m_1 + m_2}(e_1 + 1)\Delta v_1. \tag{5}$$

Thus, all those metrics are similar to each other and can be considered as different ways of quantifying the same underlying concept. In the paper, we will consider metric (3) as defining harm in a collision.

Velocities after impact: In accidents, the speed change of a vehicle in a collision can be calculated by the equation for conservation of momentum:

$$m_1 v_1 + m_2 v_2 = m_1 u_1 + m_2 u_2. \tag{6}$$

and definition of coefficient of restitution (4). Having these two equations and considering central collinear collisions, vehicles velocities after collision can be calculated as:

$$u_1 = v_1 + \frac{m_2}{m_1 + m_2}(1 + e_1)\Delta v_1 \tag{7}$$

$$u_2 = v_2 - \frac{m_1}{m_1 + m_2}(1 + e_1)\Delta v_1 \tag{8}$$

Two vehicle case: Below in the current subsection, we refer to the two vehicles in the considered emergency braking scenario as considered system. Without loss of generality we denote those vehicles as $(i-1)$ and i. By parameters of the system, we mean d^0_{i-1}, v^0_{i-1}, v^0_i, a_{i-1}, a_i, τ_{i-1}, τ_i. To find the moment of collision and corresponding harm for two vehicles, we can split all possible cases into two groups. The first group A contains a more interesting case when both vehicles have started decelerating before collision, meaning that $t^*_{i-1,i} \ge \max(\tau_{i-1}, \tau_i)$. The second group B covers all the other cases, i.e., when at least one vehicle has not initiated emergency braking and keeps moving with its constant speed, i.e., $t^*_{i-1,i} \le \max(\tau_{i-1}, \tau_i)$.

To obtain solution for both cases, let us consider two vehicles moving along the road with initial inter-vehicle distance d^0_{i-1}. We assume that the i-th vehicle moves with a constant velocity v^0_i until entering emergency braking at time τ_i. The following second order equations describe the considered scenario:

$$\begin{aligned}
&\dot{x}_{i-1} = v_{i-1},\\
&\dot{v}_{i-1} = \begin{cases} 0 & \text{for } 0 \le t < \tau_{i-1},\\ -a^*_{i-1} & \text{for } \tau_{i-1} \le t < T_{i-1}, ,\\ 0 & \text{for } t \ge T_{i-1}, \end{cases}\\
&\dot{x}_i = v_i,\\
&\dot{v}_i = \begin{cases} 0 & \text{for } 0 \le t < \tau_i,\\ -a^*_i & \text{for } \tau_i \le t \le T_i, \end{cases},
\end{aligned} \tag{9}$$

with initial conditions: $x_{i-1}(0) = d^0_{i-1}$, $x_i(0) = 0$, $v_{i-1}(0) = v^0_{i-1}$, $v_i(0) = v^0_i$. Here, T_i is the time when the i-th vehicle comes to a full stop.

The time of collision can be defined as the moment $t^* = t^*_{i-1,i}$ when $x_{i-1}(t^*) = x_i(t^*)$ and $v_{i-1}(t^*) < v_i(t^*)$. The last condition on velocities excludes cases when the vehicle i touches the preceding vehicle without a collision, i.e., when directly after t^*, the inter-vehicle distance either stays zero or increases again. This can happen if $a_{i-1} < a_i$ [8]. Note that equations (9) also consider possible collisions when the $(i-1)$-th vehicle has already stopped, i.e., when $T_{i-1} \le t^* \le T_i$. Here, the notation t^* was introduced to save space for the representation of formulas in the current subsection.

Below, we provide the time of collision and the corresponding harm for the cases in group A followed by the same metrics derived for group B. By considering solution of differential equations (9) when $t \ge \max(\tau_{i-1}, \tau_i)$ (group A), the time of collision can be derived as:

$$t^*_{i-1,i} = \tau_i + \begin{cases} \frac{-(v^0_{i-1} - v^0_i) - a_{i-1}\tau - \sqrt{D_I}}{a_i - a_{i-1}} & if \;\; (AI)\\ \frac{v^0_i - \sqrt{D_{II}}}{a_i} & if \;\; (AII). \end{cases} \tag{10}$$

Here,

$$\begin{aligned} D_I = {} & (v^0_{i-1} - a_{i-1}\tau - v^0_i)^2 - \\ & - 2\left(d^0_{i-1} + (v^0_{i-1} - v^0_i)\tau_i - \frac{a_{i-1}\tau^2}{2}\right)(a_i - a_{i-1}), \end{aligned} \tag{11}$$

$$\begin{aligned} D_{II} = {} & (v^0_i)^2 - \frac{a_i}{a_{i-1}}(v^0_{i-1} - a_{i-1}\tau)^2 - \\ & - 2a_i\left(d^0_{i-1} + (v^0_{i-1} - v^0_i)\tau_i - \frac{a_{i-1}\tau^2}{2}\right), \end{aligned} \tag{12}$$

and conditions defining cases are:

$$\textit{(AI)}: \begin{cases} D_I \ge 0;\\ \frac{-(v^0_{i-1} - v^0_i - a_{i-1}\tau) - \sqrt{D_I}}{a_i - a_{i-1}} \le \frac{v^0_{i-1}}{a_{i-1}} - \tau \end{cases}$$

$$\textit{(AII)}: \begin{cases} D_{II} \ge 0;\\ \frac{v^0_i - \sqrt{D_{II}}}{a_i} \ge \frac{v^0_{i-1}}{a_{i-1}} - \tau \end{cases}$$

Here, $\tau = \tau_i - \tau_{i-1}$. Case *(AI)* corresponds to the collision when both vehicles are moving, and case *(AII)* to a collision of vehicle i with the stand-still vehicle $(i-1)$. Note that collision happens, i.e. feasible t^* exists, only if parameters of the system are such that either the square root expression D_I is not negative, and the resulting time t^* happens earlier than the first vehicle comes to a full stop, or the square root expression D_{II} is not negative, and the resulting time t^* happens after the first vehicle comes to a full stop. If parameters of the system are such that none of the cases *(AI)* and *(AII)* are valid, then no collision occurs in the considered scenario after both vehicle started emergency braking (for group A).

The corresponding harm at the moment of the collision is proportional to the relative speed $\Delta v_{i-1} = v_i - v_{i-1}$ at the moment t^* which is defined as:

$$\Delta v_{i-1}(t^*) = \begin{cases} \sqrt{D_I} & if \;\; (AI)\\ \sqrt{D_{II}} & if \;\; (AII). \end{cases} \tag{13}$$

It can be shown that the time of collision t^* is a monotonically increasing function of deceleration a_i whereas relative speed $\Delta v_{i-1}(t^*)$ monotonically decreases with a_i. When considering the dependence on a_{i-1}, the opposite relationship holds.

For the cases in the second group B, the moment of collision t^* and corresponding harm can be found by considering the same equations (9) and assuming that the time of collision either $t^* \leq \min(\tau_{i-1}, \tau_i)$ or $\min(\tau_{i-1}, \tau_i) \leq t^* \leq \max(\tau_{i-1}, \tau_i)$. This leads to formulas below:

$$t^* = \begin{cases} \frac{d^0_{i-1}}{(v^0_i - v^0_{i-1})} & if \;\; (BIII) \\ \tau_{i-1} + \frac{(v^0_{i-1} - v^0_i) + \sqrt{D_{IV}}}{a_{i-1}} & if \;\; (BIV) \\ \frac{d^0_{i-1} + v^0_{i-1}\tau_{i-1} + \frac{(v^0_{i-1})^2}{2a_{i-1}}}{v^0_i} & if \;\; (BV) \\ \tau_i + \frac{-(v^0_{i-1} - v^0_i) - \sqrt{D_{VI}}}{a_i} & if \;\; (BVI) \end{cases} \tag{14}$$

Here,

$$D_{IV} = (v^0_{i-1} - v^0_i)^2 + 2a_{i-1}(d^0_{i-1} + (v^0_{i-1} - v^0_i)\tau_{i-1}), \tag{15}$$

$$D_{VI} = (v^0_{i-1} - v^0_i)^2 - 2a_i(d^0_{i-1} + (v^0_{i-1} - v^0_i)\tau_i), \tag{16}$$

and conditions defining cases are:

$$\textit{(BIII)} : \begin{cases} \frac{d^0_{i-1}}{(v^0_i - v^0_{i-1})} \leq \min(\tau_{i-1}, \tau_i) \\ v^0_i - v^0_{i-1} > 0 \end{cases}$$

$$\textit{(BIV)} : \begin{cases} D_{IV} \geq 0; \\ 0 \leq \frac{(v^0_{i-1} - v^0_i) + \sqrt{D_{IV}}}{a_{i-1}} \leq \min\left(\tau, \frac{v^0_{i-1}}{a_{i-1}}\right) \end{cases}$$

$$\textit{(BV)} : \begin{cases} \frac{v^0_{i-1}}{a_{i-1}} \leq \frac{d^0_{i-1} + (v^0_{i-1} - v^0_i)\tau_{i-1} + \frac{(v^0_{i-1})^2}{2a_{i-1}}}{v^0_i} \leq \tau \end{cases}$$

$$\textit{(BVI)} : \begin{cases} D_{VI} \geq 0; \\ \tau \leq \frac{(v^0_{i-1} - v^0_i) + \sqrt{D_{IV}}}{a_{i-1}} \leq 0 \end{cases}$$

The relative velocity at the moment of collision is defined as:

$$\Delta v_{i-1}(t^*) = \begin{cases} v^0_i - v^0_{i-1} & if \;\; (BIII) \\ \sqrt{D_{IV}} & if \;\; (BIV) \\ v^0_i & if \;\; (BV) \\ \sqrt{D_{VI}} & if \;\; (BVI). \end{cases} \tag{17}$$

In the case *(BIII)*, the collision occurs when both vehicles have not yet started to decelerate. Although this case is unlikely to happen in practice due to the reaction of an adaptive cruise controller in place, it is still included here for the sake of comprehension. Cases *(BIV)* and *(BV)* occur when $\tau_{i-1} \leq \tau_i$ and the $(i-1)$-th vehicle is decelerating while the i-th continues moving without deceleration. In the case *(BIV)*, both vehicles are moving at the time of collision while in the case *(BV)*, the collision happens with the $(i-1)$-th vehicle being standstill. Finally, the case *(BVI)* covers scenario when $\tau_{i-1} \geq \tau_i$ and the $(i-1)$-th vehicle has started decelerating, but the i-th has not. As was mentioned above, those cases are of lesser interest for consideration since they are less likely to occur with respect to the cases in the group A.

Summarizing, the formulas (13), (17) constitute the full final model for calculating relative velocity at the moment of collision between two vehicles. In order to calculate harm as a metric (3), the relative velocity should be multiplied by the corresponding relation of the masses.

Algorithm of defining the harm for a three vehicle case: Recall that for the considered scenario with three vehicles, we assume that all parameters are fixed except for the deceleration of the middle vehicle. To find "the best" deceleration a^*_2, the Algorithm 1 can be applied. The output of this algorithm is the harm function $H = \sum_i W_i H_i$ calculated as a weighted sum of harms for individual vehicles: H_1, H_2 and H_3. Coefficients W_i are non-negative and in general, characterize "the importance" of the involved vehicles (the elaboration on the ethical principles behind deciding on the values of these coefficients is out of our scope). With two zero coefficients H turns into an individual harm, whereas all W_i equal to 1 constructs the total harm with an equal importance of all the vehicles: $H = \sum_i H_i$.

Algorithm 1 Constructing harm function

for every $a_2 \in [0, \bar{a}_2]$ **do**
    find time of the first collision $t^{c1} = \min\left(t^c_{1,2}, t^c_{2,3}\right)$ where $t^c_{1,2}$ and $t^c_{2,3}$ are calculated with (10), (14)
    calculate respective harm at the moment t^{c1} with use of formulas (13), (17)
    calculate new velocities after the collision ((7),(8))
    assign new current time $t = t^{c1}$, new initial velocities and relative distances equal to those at $t = t^{c1}$
    find time of the second collision t^{c2} with (10), (14)
    calculate respective harm at the moment t^{c2} with the use of the formulas (13), (17)
end for
return H

To find "the best" a^*_2, one should first check using formulas (1)-(2) whether the safe deceleration interval A_2, where all individual harms are zero due to no-collisions, exists. If such a safe A_2 exists and feasible, i.e., $A_2 \cap [0; \bar{a}_2]$, the deceleration a^*_2 should be chosen from this interval, i.e., $a^*_2 \in A_2$. If this is not true, the Algorithm 1 should be executed to get an understanding of the ways to decrease the desired harm objectives. Such an order of actions allows for reduced calculations in the case when A_2 is feasible and is summarized by the Algorithm 2 below. We leave the design of the exact procedure for selecting a robust solution a^*_2 based on the constructed H for the future work.

Algorithm 2 Checking for ¨zero harm¨

Calculate the interval A_2 of a safe a_2 such that no collisions occur between pairs of vehicles using (1)-(2)
if safe A_2 is feasible **then**
    return A_2
else
    construct the harm function ▷ Algorithm 1
end if

IV. Numerical examples

We demonstrate how the presented approach can serve as a base for the strategy of choosing braking deceleration a_2^*. Furthermore, we discuss how the vehicle 2 can gain understanding regarding the order of collisions within the formation, which is preferred to reduce the harm objective.

Figures 2 and 3 correspond to the case when there exists a feasible safe interval A_2 of the deceleration a_2. One can see that any deceleration $4.46 \leq a_2^* \leq 5.34$ applied by the second vehicle in the emergency scenario allows avoiding any collisions. Thus, for such a_2^*, the harm is zero. Fig. 3 depicts the individual and total harm constructed by the choice of respective coefficients W_i. For any $a_2^* < 4.46$, the collision between vehicles 1 and 2 is unavoidable. For any $a_2^* > 5.34$, the collision between vehicles 2 and 3 is unavoidable. However, between 4.46 and 5.34 vehicles stop without any collisions. As can be seen, this safe interval A_2 does not include the maximum braking capacity of the vehicle 2. Therefore, the ethical cooperation between vehicles allows for smarter and safer braking approach compared to just a straightforward strategy, when the maximum braking is applied by all vehicles.

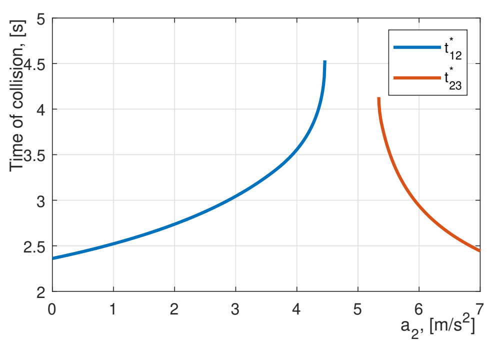

Fig. 2: Time of collision versus deceleration a_2 for a 3-vehicle scenario. Here, $\bar{a}_1 = 6$, $\bar{a}_2 = 7$, $\bar{a}_3 = 6\ m/s^2$; $v_1^0 = 20$, $v_2^0 = 18$, $v_3^0 = 20\ m/s$; $\tau_1 = 0$, $\tau_2 = 0.5$, $\tau_3 = 0.8\ s$; $d_1^0 = 12$, $d_2^0 = 10\ m$.

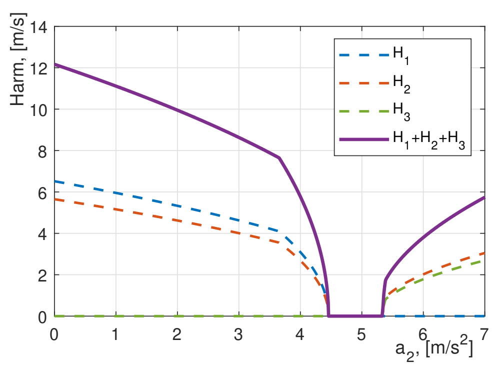

Fig. 3: Harm function H versus deceleration a_2. Here a_2 can be chosen such that no collisions occur.

Now, let us consider a situation when there exists no safe deceleration a_2 allowing to avoid all the collisions. This means that regardless of the chosen deceleration rate, there will always occur at least one collision among three considered vehicles. The proposed Algorithm 1 for constructing the harm function is executed for the three vehicles placed at the time $\tau_1 = 0$ at the distances $d_1^0 = 5\ m$, $d_2^0 = 7\ m$. All other parameters are chosen the same as for the previous example, and coefficients of restitution e_1 and e_2 are chosen as 0.3. In general, the magnitude of the coefficient of restitution depends on several parameters including the materials of the bumper/body, the surface geometry, and the impact velocity. According to crash reports, the coefficient of restitution is observed to be inversely dependent on the relative speed Δv_1 prior to impact, and typically ranges between 0 and 0.3 for rear-end collisions [10], [11], but can go up to 0.6 for some low-speed collisions [12].

Fig. 4 shows time of the collision where red and blue lines correspond to the first collision in the considered triple whereas purple and orange – to the second collision. One can see that for deceleration $a_2 < 5.01$, the vehicle 2 experiences a collision with the vehicle 1 followed additionally by the second collision with the vehicle 3 if $a_2 \geq 4.14$. For deceleration $a_2 > 5.01$, the vehicle 2 first experiences a collision with the vehicle 3 followed additionally by the second collision with the vehicle 1 if $a_2 \leq 6.62$. When $a_2 = 5.01\ m/s^2$, the order of collisions among three considered vehicles is switched. Since masses m_1, m_3 as well as velocities v_1, v_3 are independent, the harm function experiences a jump.

The corresponding individual harm for each vehicle separately as well as the total harm are presented on Fig. 5. To save the vehicle 1, i.e., to reach zero level of H_1, the deceleration a_2 should be chosen less than $6.62\ m/s^2$. In order to save the vehicle 3, i.e., to ensure $H_3 = 0$, the deceleration a_2 should be chosen less than $4.14\ m/s^2$. However, for the vehicle 2, there exists no a_2 such that $H_2 = 0$.

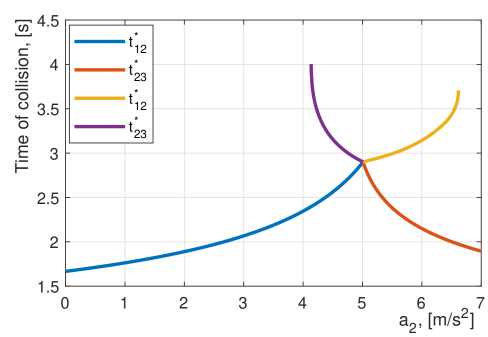

Fig. 4: Time of collision versus deceleration a_2 for a 3-vehicle scenario. Parameters for a simulation were chosen the same as for Fig. 2 except initial distances.

V. Ethical V2X

We now elaborate on the necessity of the V2X communication in the considered setting. The longitudinal autonomous driving can be maintained thanks to a cooperative adaptive cruise control enabled by a continuous exchange of broadcast *heartbeat* messages (aka cooperative awareness messages). These messages normally include current position and velocity of a sending vehicle. Moreover, the automatic emergency braking can be made cooperative via *warning* messages (aka decentralized environmental notification messages) repeatedly emitted by the vehicle 1 for the vehicles 2 and 3. However, it

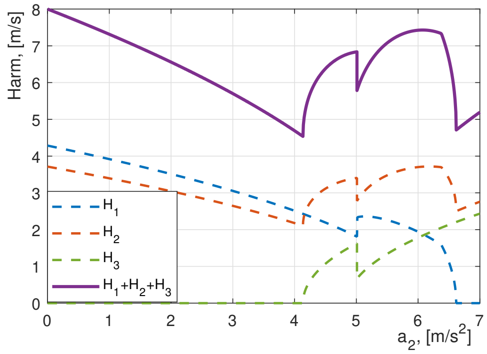

Fig. 5: Harm function H versus deceleration a_2. Here, for any choice of a_2, at least one collision happens.

is possible to design the system based on a non-cooperative paradigm, for example, by relying on frontal radar and camera sensors [7]. Vehicles can benefit from the cooperation [13], but they are not necessarily supposed to be cooperative to perform autonomously in a studied scenario. Importantly, however, vehicles must be cooperative for ethical driving.

Firstly, in the considered formulation of the ethical dilemma, the vehicle 2 requires the parameters of the vehicles 1 and 3 such as their "importance" W_i, their masses m_i, their braking capabilities $\bar{a}_i$ as well as other information needed for the estimation of the coefficients of restitution (for example, their types). These information elements are rather static and can be encapsulated into a low frequency container of the heartbeats. Secondly, the vehicle 2 requires the knowledge of relative velocities at the moment of impact. In our assumptions their calculation are based on the velocities v_i^0 of the vehicles 1 and 3 and respective inter-vehicular distances d_i^0 at the moment of braking. Those can be retrieved from the high frequency container of the heartbeats. Thirdly, the vehicle 2 needs to know the moments of braking of the vehicle 1 and the vehicle 3. The former τ_1 can be directly included into the warnings emitted by the vehicle 1 to inform about the braking start. The latter can be calculated by the vehicle 3 as $\tau_3 = \tau_1 + \delta_3$, where δ_3 can be decided in advance and provided to the vehicle 2 based on the current estimated radio link quality between the vehicle 1 and the vehicle 3 so that the probability of not receiving at least one warning from the vehicle 1 is made small. Similar reasoning can be applied by the vehicle 2 for the decision on the choice of $\tau_2 = \tau_1 + \delta_2$ based on the quality of its link from the vehicle 1.

In an event that either warning from the vehicle 1 does not come either to the vehicle 2 within δ_2 or to the vehicle 3 within δ_3, or required heartbeats are not received respectively, the ethical braking plan will not be possible to fulfill. We denote the probability of this event as ϵ. Let us assume the IEEE 802.11p V2X technology is used and T be heartbeats/warnings broadcasting period. The carrier-sense multiple access priorities are adopted so that the warnings and the heartbeats do not interfere. The warnings from the vehicle 1 might not be received by the vehicle 2 or the vehicle 3 due to the channel errors. Let p be the propagation-induced loss probability increment per vehicle [7]. The heartbeats from the vehicles 1 and 3 might interfere with each other with probability q. This leads to $\delta_2 = T \log_{1-(1-p)(1-q)} \epsilon$ and $\delta_3 = T \max(\log_{2p} \epsilon, \log_{1-(1-p)(1-q)} \epsilon)$. For $T = 100$ ms, $p = 0.03$, q computed for IEEE 802.11p parameters as explained in [14] and $\epsilon = 0.01$, the ethical braking requires $\delta_2 \approx 140$ ms and $\delta_3 \approx 160$ ms.

We believe that there is a huge potential in designing ethical V2X protocols which would support resolving ethical dilemmas. As a first step, a thorough design and performance evaluation of the protocol sketched above is needed with a special attention to the corner cases, when due to the communications unreliability, the ethical braking plan cannot be accomplished. Furthermore, alternative formulations of the ethical dilemmas are of interest, for example, cooperative maneuvering at the intersection is worth looking into from the ethical perspective. Finally, further work is needed to incorporate more realism in modelling of the collisions.

References

[1] G. Li, Y. Li, Z. Gao, and F. Chen, "Ethical thinking of driverless cars," in *Proceedings of the 2019 International Conference on Modern Educational Technology*, ser. ICMET 2019. New York, NY, USA: Association for Computing Machinery, 2019, p. 91–95. [Online]. Available: https://doi.org/10.1145/3341042.3341055

[2] V. Behzadan, J. Minton, and A. Munir, "Trolleymod v1.0: An open-source simulation and data-collection platform for ethical decision making in autonomous vehicles," in *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, ser. AIES '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 391–395. [Online]. Available: https://doi.org/10.1145/3306618.3314239

[3] Y. Wang, X. Hu, L. Yang, and Z. Huang, "Ethics dilemmas and autonomous vehicles: Ethics preference modeling and implementation of personal ethics setting for autonomous vehicles in dilemmas," *IEEE Intelligent Transportation Systems Magazine*, vol. 15, no. 2, pp. 177–189, 2023.

[4] N. De Moura, R. Chatila, K. Evans, S. Chauvier, and E. Dogan, "Ethical decision making for autonomous vehicles," in *2020 IEEE Intelligent Vehicles Symposium (IV)*, 2020, pp. 2006–2013.

[5] M. Geisslinger, F. Poszler, and M. Lienkamp, "An ethical trajectory planning algorithm for autonomous vehicles," *Nature Machine Intelligence*, vol. 5, no. 2, pp. 137–144, 2023.

[6] C. Bergenhem, K. Meinke, and F. Ström, "Quantitative safety analysis of a coordinated emergency brake protocol for vehicle platoons," in *Leveraging Applications of Formal Methods, Verification and Validation. Distributed Systems*, T. Margaria and B. Steffen, Eds. Cham: Springer International Publishing, 2018, pp. 386–404.

[7] G. Sidorenko, J. Thunberg, K. Sjöberg, A. Fedorov, and A. Vinel, "Safety of automatic emergency braking in platooning," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 3, pp. 2319–2332, 2022.

[8] G. Sidorenko, A. Fedorov, J. Thunberg, and A. Vinel, "Towards a complete safety framework for longitudinal driving," *IEEE Transactions on Intelligent Vehicles*, vol. 7, no. 4, pp. 809–814, 2022.

[9] D. J. Gabauer and H. C. Gabler, "Comparison of delta-v and occupant impact velocity crash severity metrics using event data recorders," in *Annual proceedings/association for the advancement of automotive medicine*, vol. 50. Association for the Advancement of Automotive Medicine, 2006, p. 57.

[10] A. F. Tencer, "Low speed rear end automobile collisions and whiplash injury, the biomechanical approach," *Medicine, Law & Society*, vol. 12, no. 2, pp. 1–20, 2019.

[11] J. Zhou, H. Peng, and J. Lu, "Collision model for vehicle motion prediction after light impacts," *Vehicle System Dynamics*, vol. 46, no. S1, pp. 3–15, 2008.

[12] R. M. Brach, "Modelling of low-speed, front-to-rear vehicle impacts," 2003.

[13] G. Sidorenko, D. Plöger, J. Thunberg, and A. Vinel, "Emergency braking with acc: How much does v2v communication help?" *IEEE Networking Letters*, vol. 4, no. 3, pp. 157–161, 2022.

[14] A. Baiocchi, I. Turcanu, N. Lyamin, K. Sjöberg, and A. Vinel, "Age of information in ieee 802.11p," in *2021 IFIP/IEEE International Symposium on Integrated Network Management (IM)*, 2021, pp. 1024–1031.