Honda Research Institute EU
Honda Research Institute Europe GmbH
https://www.honda-ri.de/

Enabling Cooperative Pedestrian-Vehicle Interactions using an eHMI

Markus Amann, Malte Probst , Raphael Wenzel, Thomas H Weisswange

2024

Preprint:

This is an accepted article published in 35th IEEE Intelligent Vehicles Symposium (IV 2024). The final authenticated version is available online at: https://doi.org/10.1109/IV55156.2024.10588537 Copyright 2024 IEEE

Enabling Cooperative Pedestrian-Vehicle Interactions using an eHMI

Markus Amann*, Malte Probst*, Raphael Wenzel*, Thomas H. Weisswange*

***Abstract*— Interactions between humans and automated vehicles (AVs) will become increasingly common in future traffic. Cooperative behavior enables comfortable and efficient resolutions of such interactions through joint actions. Thus, AVs need the ability to communicate their intention and planned behavior to outside road users to facilitate cooperation among the interaction partners. Many recent studies suggest external human-machine-interfaces (eHMIs) as a feasible way to realize the communication between AVs and other road users. These studies often focus on the design of eHMIs and evaluate their usability based on subjective measures from the perspective of the outside road user. There is only little research on the objective benefits of explicit external communication on the resolution of interactions with AVs. The decision making behind the activation of the communication and the consideration of the actual interaction in the communication are still open research topics. In this work, we present a communication framework that facilitates cooperation between an AV and pedestrians by communicating the vehicle's yielding intention to give right of way. We demonstrate the proposed system's ability to improve the joint utility as well as the driving comfort by empirically evaluating simulated interactions between AVs and pedestrians. The results indicate that timely communication leads to more efficient and more predictable pedestrian-vehicle interactions.**

I. Introduction

Many everyday traffic situations require cooperation among traffic participants to enable comfortable, safe, and collision free interactions. The interaction between a vehicle and a pedestrian that intends to cross the road is one example for a common scenario that requires cooperation [1], [2]. In recent research, the interaction between automated vehicles (AVs) and vulnerable road users (VRUs), such as pedestrians, has received a lot of attention [1]–[4], since VRUs are exposed to higher risks compared to vehicles. While the crossing order of an interaction between an AV and a pedestrian is clear at signalized intersections or crosswalks, the relevance of cooperation becomes more significant in ambiguous traffic situations with unclear priorities. Fig. 1 shows an example situation of an interaction at an unsignalized crossing. The left-hand side depicts the interaction partners' continuous assessment of the situation and their uncertainty about the intention of the respective partner. Even if both are aiming for courteous behavior they might not be able to agree on a joint resolution. The successful resolution of a cooperative traffic situation requires knowledge about the intention and the planned behavior of the interaction partner. This information is transferred using communication. In the context of communication between traffic participants, it is common to distinguish between implicit and explicit communication. Implicit communication refers to information about intent and planned maneuvers that can be inferred implicitly from the dynamic state [5], [6]. However, implicit communication cues have been demonstrated to have limited effectiveness in traffic situations with low behavioral flexibility [1]. In traffic situations without clear priorities, explicit signals such as visual or auditory cues can help to strengthen the intended communication [1], [7] and allow for a more transparent and more efficient resolution. Studies [7]–[10] indicate that communication by the means of external human-machine-interfaces (eHMIs) can have a positive impact on the resolution of interactions between AVs and pedestrians, and that they can increase the acceptance of AVs significantly.

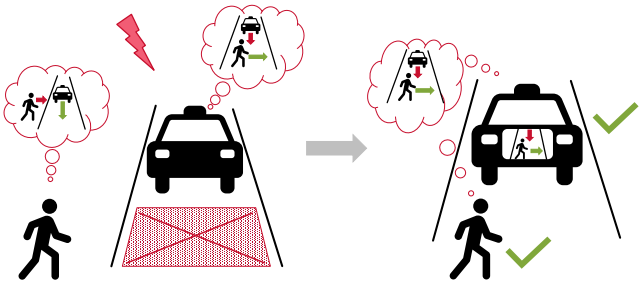


Fig. 1. Interaction of a vehicle and a pedestrian that has the intention to cross the road at an unsignalized crossing. The situation is resolved using communication by the means of an eHMI.

Several studies research the crossing behavior of pedestrians interacting with vehicles that are equipped with eHMIs to assess the overall effects of eHMIs on the comprehensibility of the communicated behavior [11]–[14]. The studies draw different conclusions, which indicates that there is no definite consensus about the usability and effectiveness of eHMIs. However, there is only little research on the timing of the communication and the effects of explicit communication on the driving comfort of the vehicle. Therefore, we propose a cooperative communication framework using explicit communication cues that facilitates cooperation between AVs and pedestrians and enables a more efficient and more comfortable resolution of crossing scenarios. The research questions of this work can be formulated as follows:

A. Can overall utility and driving comfort be increased by using explicit external communication?

B. What is the optimal timing for explicit communication in order to foster cooperative behavior?

This research was supported by the HEIDI project which has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101069538. Views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or CINEA. Neither the European Union nor the granting authority can be held responsible for them.

*Honda Research Institute Europe GmbH, Carl-Legien-Str. 30, 63073 Offenbach, Germany, Email: {firstname.lastname}@honda-ri.de

In the following, we discuss prior art on cooperative behavior planning and pedestrian-vehicle interactions focusing on eHMIs and the modeling of pedestrian crossing behavior. We present our approach for a communication strategy using a cooperative risk-based behavior planner and demonstrate its effectiveness in simulated pedestrian-vehicle interactions.

II. Related Work

A. Interaction-Aware Behavior Planning

Interaction-aware state-of-the-art behavior planners use different approaches to account for the bilateral negotiation between the interaction partners. In [15], the authors use the concept of a shared traffic space (STS), which is a limited resource that is shared by the interacting partners. They argue that traffic cooperation is about finding a passing order of the STS that leads to a high driving benefit for both interaction partners. They present a risk-based planning approach similar to [16] and make a bimodal behavior prediction to account for the different intentions of the interaction partner. This allows to consider the distinct ordering options.

In [17], a heuristic model that makes the vehicle's behavior decision based on the estimated pedestrian intention is presented. Similarly, [18] provides the advised vehicle speed to successfully resolve the negotiation for the right of way at intersections based on the estimated yielding intention of a pedestrian and the situation kinematics.

Other approaches, use game-theoretic models [19] or Partially Observable Markov Decision Processes [10] to account for the interaction. In [20], the authors present a game-theoretic approach using a leader-follower game. They formulate the planning task as an optimal control problem and enable cooperative behavior by considering joint costs of both interaction partners in the objective function.

Most approaches for interaction-aware planning are capable of anticipating the behavior of the interaction partner, thus, planning their own behavior ahead and reacting to behavioral changes. However, they fall short of the ability to explicitly influence the interactions partner's behavior by communicating their intention or joint behavior.

B. External Human-Machine-Interfaces

In [7], the authors conduct an extensive literature review on several types of HMIs and emphasize the high potential of eHMIs for interactions between AVs and VRUs to increase the level of objective and perceived safety, comfort, and efficiency. In [11], the authors evaluate the response of VRUs towards an AV that is equipped with two different types of visual eHMIs in a real-world user study. The study results do not show a significant difference in the reaction between the equipped vehicle and the baseline without an eHMI. Thus, they conclude that visual eHMIs are not necessarily required in an environment in which informal rules apply. In a second study [12], they compare the influence of the eHMIs on the performance of a pedestrian intention estimation framework. Similar to the first study, they do not see a significant difference comparing the prediction performance of an eHMI equipped vehicle to a baseline vehicle. However, in both studies they note that effects such as distraction through smartphone usage resulting in a safety maneuver or operating at a very low vehicle speed that is not perceived as dangerous might also have an effect on the results. In contrast to that, the studies [9], [13], [14] indicate that pedestrians do react differently when an approaching vehicle is equipped with an eHMI and that eHMIs have a positive impact on the perceived safety and crossing willingness. However, all studies highlight that the effect of eHMIs on the crossing behavior of pedestrians can only be measured when the approaching vehicle is actually yielding. The importance of yielding behavior for the crossing decision of a pedestrian is also underlined in [4] demonstrating a positive impact of smooth and early braking on the gap acceptance. However, the results indicate that about 25% of all participants would wait for the approaching vehicle to come to a full stop even at the earliest and smoothest braking trajectory. This finding reinforces the assumption of the aforementioned studies that eHMIs mainly work in addition to implicit yielding behavior. The different results of the presented studies outline the necessity for further research on eHMIs and motivate our research on explicit communication.

C. Pedestrian Behavior Modeling

Numerous studies have been conducted to identify the most relevant features in order to characterize and describe pedestrian crossing behavior interacting with both eHMIs equipped vehicles and ordinary vehicles. In [21], the decision making of a crossing pedestrian interacting with an ordinary vehicle in urban traffic is described by a continuous process flow. The model states that pedestrians make two main decisions during the interaction. First, a pedestrian makes a decision whether the time gap to the vehicle is large enough to cross (gap acceptance model). If this is not the case, the pedestrian has to decide whether to accept the yielding behavior, meaning the implicit communication of the vehicle (yielding acceptance model). They argue that the gap acceptance model can be described as a probabilistic binary decision of accepting or rejecting a gap to an oncoming vehicle. This modeling goes along with the findings of [4], which show that the first crossing decision is mainly based on the time gap independent of the vehicle's yielding behavior. If the pedestrian decides to reject a gap the decision making in the second stage is based on the vehicle kinematics. The average time gap at which pedestrians make their crossing decision in the first stage can be approximated to $4.3\,\mathrm{s}$.

In [22], the authors find larger values than [4] examining the accepted time gap between vehicles and pedestrians that are violating a red pedestrian traffic light signal. The average accepted time gap varies between $6.3\,\mathrm{s}$ for young people to $7.4\,\mathrm{s}$ for elderly showing that the gap acceptance varies depending on the age.

In [23], the authors do not only show the potential to generally improve interactions between pedestrians and vehicles by using a frontal braking light but they also show that such kind of eHMI can reduce the identification time of a braking maneuver from $1.5\,\mathrm{s}$ to less than $1\,\mathrm{s}$.

III. Method

A. Cooperative Behavior Planner

We use an established framework for risk-based behavior planning, which has previously been demonstrated in real-world scenarios [16]. The planning framework follows a scheme consisting of the modules *Understand*, *Predict*, *Plan & Select* and *Act*. Similar schemes are widely used in behavior planning for AVs [15], [19], [24]. Fig. 2 shows the overall structure of the complete system for cooperative communication. In this work, we focus on the *Communicate* module that makes the decision whether to communicate the AV's yielding to another road user to give right of way. The decision making and timing behind the communication are often neglected in studies that focus on the eHMI design.

The *Understand* module receives input about ego vehicle data and the surrounding environment. The module creates relational information between the ego vehicle and other road users. Similar to vehicles, we model pedestrians such that they move along paths, such as a kinematic extrapolation of the current orientation, since we assume that a pedestrian that is waiting at a road curb or walking towards it in order to cross the road would use a straight path that can be inferred from the sensed data. Moreover, we model the interaction between the ego vehicle and the pedestrian using the concept of STS introduced in [15]. Defining an encroachment or collision zone is a commonly used idea to model interactions between vehicles and pedestrians [2], [17]. As described in [15], each agent has two potential strategic options with respect to the passing order. They can either aim for passing the STS first, or passing second, and the two agents have to negotiate the order of passing during the interaction. We define the variables $d_{\text{STS},e}$ and $d_{\text{STS},o}$ which denote the distance between the STS and the ego vehicle or the other road user, respectively. This is illustrated in Fig. 3a.

The *Predict* module predicts a velocity profile for each of the two semantically different underlying intentions of the other agent. This allows to cover a wide range of possible resolutions without having to carry out an extensive prediction sampling which would lead to a significantly higher computational effort. Nevertheless, the proposed communication method also works for a different number of prediction options as long as all predictions refer to a STS such that they can be assigned to one of the two options of passing first or passing second. We denote the predicted options by the index $j = 0, ..., p$, with $p = 1$ in our case.

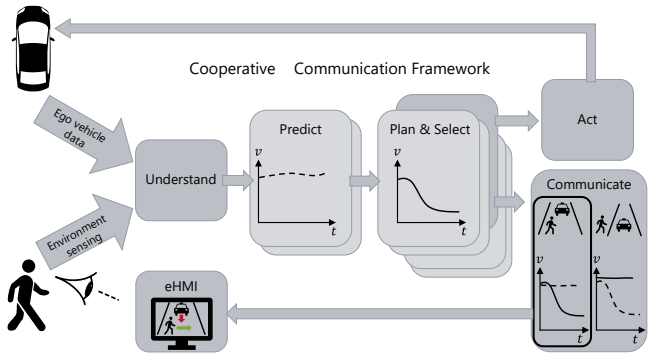


Fig. 2. Overview of the system structure. The presented *Communicate* module makes the decision whether to communicate to the pedestrian.

Based on the different predicted behavior options, the *Plan & Select* module plans the behavior of the ego vehicle. Both, the prediction and the planning are done for a time horizon of $T = 12\,\text{s}$ during each cycle of the whole behavior planner. To plan the ego behavior, a set of standardized velocity profiles, e.g., constant velocity, is generated as possible candidates. Additionally, two optimized velocity profiles are computed for each predicted behavior option. Given the predictions, the velocity profiles for the ego vehicle are optimized with respect to risk, utility, and comfort. For further details about the behavior optimization and risk modeling, see [16], [25]. For each prediction j, we denote the ego behavior options by $i = 0, ..., q$, with $q = 9$ including both, the two optimized velocity profiles and the standard candidates. Hence, there are $p \cdot q$ pairs of potential solutions. Each solution contains a prediction $v_{o,j}(t)$ and a corresponding plan $v^j_{e,i}(t)$. Note that a combination of an ego behavior $v^j_{e,i}(t)$ with a prediction $v_{o,j}(t)$ is only eligible if $v^j_{e,i}(t)$ is either a standard candidate or $v^j_{e,i}(t)$ has been optimized with respect to that particular prediction $v_{o,j}(t)$. This is denoted by the superscript j. We define the set $\mathbb{S}$ that contains all eligible behavior combinations $S_{ij} = \left(v^j_{e,i}(t), v_{o,j}(t)\right)$.[1] After the planning step, the *Plan & Select* module evaluates and selects the overall best behavior combination. The system chooses the best solution S^* which leads to the lowest costs regarding risk, utility, and comfort

$$S^* = \left(v_e^*, v_o^*\right) := \underset{S \in \mathbb{S}}{\arg\min}\, \varphi(S)\,. \tag{1}$$

The function $\varphi(S)$ denotes the joint cost evaluation for the combinations of predicted and planned behavior. The selected behavior is then actuated by the AV's controller.

In contrast to [15], which focuses on cooperation modeling and planning of cooperative behavior, we extend the planning framework by the *Communicate* module to enable communication between the AV and other road users.

B. Communication Framework

By selecting one of the combinations of predicted and planned behavior, the AV makes an explicit assumption about the pedestrian behavior. As long as the vehicle does not communicate, the pedestrian does not have any explicit information — neither about the vehicle's intention nor about the assumption that the vehicle makes about the pedestrian's behavior. In this work, we develop a communication framework that facilitates cooperation and reduces the vehicle's uncertainty about the pedestrian behavior. If an AV communicates its yielding intention and that it is safe to cross, expecting a pedestrian to cross before the vehicle, missed or misinterpreted communication resulting in a non-crossing pedestrian can cause inefficient behavior.

[1] For the sake of readability, we choose an easier notation and we omit the indices i and j when referring to a single, unspecified solution S.

However, it is unlikely that this will cause critical situations if the person does not show the expected crossing response. Communicating that it is not safe to cross and driving on under the expectation that the pedestrian will stop, can cause potentially dangerous situations. For this reason, we only want to communicate the AV's yielding intention and we do not want to demand a pedestrian to stop. This allows for reduced uncertainty about the pedestrian behavior without the risk of causing safety critical situations which require emergency maneuvers. Such situations are out of scope of cooperation from our perspective.

We formulate a heuristic decision making algorithm based on several semantic criteria which allows the *Communicate* module to convey the yielding intention to the pedestrian in the aforementioned communication relevant situations.

After S^* has been selected, the *Communicate* module uses the predicted behavior options for the pedestrian and the planned behavior of the vehicle to determine whether communication is necessary. For this purpose, we map all possible solutions $S \in \mathbb{S}$ to the strategic options of going first or letting the other agent pass first. During each planning cycle, we compute $s_e(t)$ and $s_o(t)$ as the distances that the agents will travel according to the predicted and planned behaviors by integrating over the respective velocity profiles. We define the subset $\mathcal{O} \subseteq \mathbb{S}$ which contains all solutions that lead to the pedestrian passing the STS first. We consider the moment at which an agent enters the STS, i.e. reaching d_{STS}, as the moment of passing because as soon as one agent enters the STS the order of passing is clear and the negotiation is finished. Thus, $\mathcal{O}$ is defined as

$$\mathcal{O} := \Big\{ S \in \mathbb{S} \,|\, \arg\min_{t \in \mathcal{T}} \big(s_e(t) \geq d_{\text{STS},e}\big) \geq \arg\min_{t \in \mathcal{T}} \big(s_o(t) \geq d_{\text{STS},o}\big) \Big\}. \quad (2)$$

The set $\mathcal{T}$ denotes the prediction and planning horizon. Similarly, the subset $\mathcal{E} \subseteq \mathbb{S}$ contains all solutions that lead to the vehicle passing the STS first such that

$$\mathcal{E} := \Big\{ S \in \mathbb{S} \,|\, \arg\min_{t \in \mathcal{T}} \big(s_e(t) \geq d_{\text{STS},e}\big) < \arg\min_{t \in \mathcal{T}} \big(s_o(t) \geq d_{\text{STS},o}\big) \Big\}. \quad (3)$$

To ensure completeness, if no agent reaches the STS within the prediction horizon a solution is classified as part of $\mathcal{O}$. Then, each possible solution is a unique member of one and only one of the two sets and the union of the two sets covers all possible options such that $\mathbb{S} = \mathcal{O} \cup \mathcal{E}$. Several semantic criteria need to be fulfilled to communicate that the vehicle is giving right of way. First, the set $\mathcal{O}$ that contains all solutions which lead to the pedestrian crossing before the vehicle must not be empty, meaning $\mathcal{O}$ has to fulfill

$$\mathcal{O} \neq \emptyset. \quad (4)$$

In addition, the optimal solution S^* which leads to the overall lowest joint costs must result in the pedestrian crossing before the vehicle which means

$$S^* \in \mathcal{O}. \quad (5)$$

This is necessary since the system is only meant to communicate if the pedestrian should cross first.

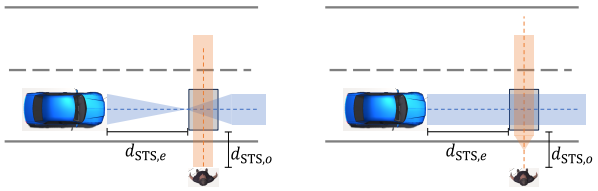
Likewise, the set $\mathcal{E}$ that contains all solutions in which the AV passes the STS before the pedestrian must not be empty

$$\mathcal{E} \neq \emptyset. \quad (6)$$

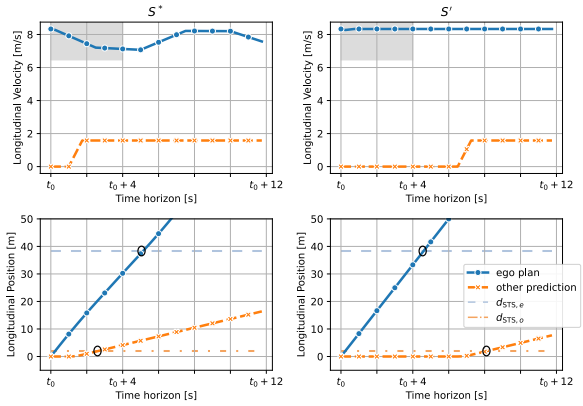
This condition can be explained by discussing the predicted and planned behavior from the respective perspective of the different interaction partners. If we look at the planned vehicle behavior from the viewpoint of the pedestrian, we assume that the different behavior plans represent possible prediction options from the perspective of the pedestrian. Only if there is at least one solution in which the vehicle crosses before the pedestrian, we assume that this outcome is within the pedestrian's scope of possible resolutions. Thus, communication indicating the AV's yielding intention is required to reduce the uncertainty about possible other behavior options. If $\mathcal{E}$ is empty and all solutions are part of $\mathcal{O}$, communicating that the pedestrian should cross first would be redundant as all options imply the same outcome.

If conditions (4), (5), and (6) are true, we determine $S' \in \mathcal{E}$ which describes the best combination of planned and predicted behavior that leads to the vehicle passing the STS before the pedestrian such that

$$S' = (v'_e, v'_o) := \arg\min_{S \in \mathcal{E}} \varphi(S). \quad (7)$$



(a) Examples depicting the two possible resolutions. The rectangular area in front of the interaction partners visualizes the STS. The width of the bars represents the velocity of the respective agent's behavior.



(b) Velocity and position profiles of the ego vehicle and the pedestrian for the selected solution S^* and the best alternative S' of one planning cycle. Starting at the current timestep t_0, the plots last until the end of the prediction horizon $T = 12\,\text{s}$. The black circles depict the time at which the agents reach the STS. The gray areas highlight the first $\Delta T = 4\,\text{s}$ which are used to determine if the AV has to decide for a distinct maneuver.

Fig. 3. Possible resolutions for a crossing interaction between a vehicle and a pedestrian based on the two different behavior intentions of the pedestrian.

This solution can be considered as the best alternative behavior which allows the AV to be first. We use S' to determine whether v_e^* and v_e' are sufficiently different from each other which implies that the AV has to explicitly decide for a certain maneuver. We make this distinction with respect to the longitudinal acceleration a_e^* and a_e' of the ego vehicle. We consider the first ΔT seconds of the planning horizon because changes in the near future have the largest impact on the evolution of the scenario. The average accelerations $\bar{a}_{e,\Delta T}^*$ and $\bar{a}_{e,\Delta T}'$ within ΔT are defined as

$$\bar{a}_{e,\Delta T}^{*\prime} = \frac{1}{\Delta T}\big(v_e^{*\prime}(t_0 + \Delta T) - v_e^{*\prime}(t_0)\big) \tag{8}$$

with t_0 denoting the current timestep. They must differ by at least ϵ in order to communicate such that

$$\left|\bar{a}_{e,\Delta T}^* - \bar{a}_{e,\Delta T}'\right| > \epsilon\,, \tag{9}$$

and we experimentally determine $\epsilon = 0.2\frac{\mathrm{m}}{\mathrm{s}^2}$ and $\Delta T = 4\,\mathrm{s}$. These parameters might need to be adjusted for real-world interactions. The condition is relevant because it prevents the system from extensive signaling and it only communicates if the communication has an actual impact on the uncertainty of the ego behavior since the AV has to make a maneuver decision according to the planned behavior options.

Fig. 3b shows the velocity and position profiles of the planned ego behavior and the predicted pedestrian behavior for a specific planning cycle. The left column relates to the solution S^* in which the ego vehicle decelerates to let the pedestrian cross whereas the right column depicts the alternative resolution S' in which the ego vehicle passes first. In the illustrated example, the criteria (4), (5), (6), and (9) are fulfilled such that the *Communicate* module would communicate that the AV is letting the pedestrian cross first.

IV. Experiments

To the best of our knowledge, there are no datasets available that allow evaluation of a communication framework, as presented in this paper, with real data. To address this problem, we use a microscopic traffic simulator with a pedestrian crossing model based on previous research results.

A. Simulation Setup

We simulate the interaction between a pedestrian and an approaching vehicle at an unsignalized, urban road as depicted in Fig. 3a. The vehicle starts driving with an initial velocity of $30\,\mathrm{km/h}$ which is at the same time its target velocity. We consider two types of scenarios. First, we analyze the interaction between a moving vehicle and a pedestrian that is *standing* at the road curb to get insights about the decision making of the AV and the timing of communication. In the second type of scenarios, we consider a pedestrian that is *walking* towards the road curb to understand the influence of the walking dynamics on the decision making.

B. Pedestrian Assumptions and Parameters

We assume that all pedestrians have the target to cross the road. However, they can have different intentions to achieve their goal, which are given by either crossing before or after the vehicle and which are considered in the two prediction modes. In the *standing* scenarios, the pedestrian stands at the side of the road with a distance of $d_{\mathrm{STS},o} = 2\,\mathrm{m}$ to the STS. We use a crossing model which is similar to the one presented in [21]. The pedestrian makes an initial crossing decision based on the time gap to the vehicle. For this purpose, we determine the time-to-arrival (TTA) of the vehicle based on the initial vehicle speed. In real-world interactions between the same pedestrian and different vehicles that person might have a tendency towards accepting generally smaller gaps or preferring larger gaps, however, the actual crossing time gap will always be different. As explained in II-C, studies come to different conclusions regarding the size of the accepted time gap. Therefore, we choose a distribution in between the results of the considered studies covering a wide spectrum of possible interactions. By sampling several time gaps from the normal distribution $\mathcal{N}(5.7, 2.0)$, we account for the uncertainty in the crossing decision. For simplicity reasons, the gap acceptance only depends on the vehicle's TTA and we don't take other modulators such as age, gender or group size into account. If the accepted time gap is smaller than the TTA the crossing decision is positive and the pedestrian will cross the road. If the crossing decision is negative, the pedestrian will keep standing until the vehicle has passed unless there is communication from the vehicle that it is willing to yield. Since there are no occlusions, we assume that pedestrian and vehicle are able to see each other at all times. Moreover, we assume that the pedestrians will always perceive, understand, and accept the communication, meaning that the pedestrian always crosses first if there is communication. We make this assumption to demonstrate the benefit of unequivocal communication with attentive and responsive pedestrians. We do not examine the case that pedestrians do not react properly to the signal since the behavior planner would switch the expected behavior and plan to pass first after some seconds if the pedestrian does not adhere to the communication.

If the pedestrian receives communication from the vehicle it will start crossing the road after a reaction time delay of $1\,\mathrm{s}$. In the *walking* scenarios, the simulation starts $2\,\mathrm{s}$ in advance with a pedestrian walking towards the road curb. We assume that, even though the pedestrian sees the oncoming vehicle, the crossing decision is made when the pedestrian reaches a distance of $d_{\mathrm{STS},o} = 2\,\mathrm{m}$. Thus, the pedestrian makes the crossing decision with respect to the vehicle's TTA after $2\,\mathrm{s}$. If the gap acceptance model leads to a negative decision the pedestrian decelerates such that it stops at $d_{\mathrm{STS},o} = 2\,\mathrm{m}$. Therefore, the *walking* scenarios evolve into the same geometric setting of a similar *standing* scenario, and the *standing* scenario can be considered as a special case of the *walking* scenario with zero velocity.

In [26], walking speeds range from $1.1 - 1.8\,\mathrm{m/s}$ for

ordinary walking. Zheng et al. [27] identify a similar average crossing speed of $1.58\,\mathrm{m/s}$ for jaywalkers. Thus, the simulated pedestrian walks with $1.5\,\mathrm{m/s}$ in our scenarios.

C. Metrics

We assess the performance of the proposed system with respect to the overall utility of both interaction partners and the driving comfort of the ego vehicle.

1) Utility Measure: Similar to [15], the joint utility measure $\mathcal{Q}_{\text{utility}}$ takes the resolution time of both agents into account and is defined as

$$\mathcal{Q}_{\text{utility}} = \frac{T_{\text{best},e} + T_{\text{best},o}}{T_{\text{actual},e} + T_{\text{actual},o}}\,. \tag{10}$$

The measure assumes values between 0 and 1. The numerator represents the best mutual resolution time that would be achieved if both agents could pass the STS at constant velocity with their respective target speed independent of potential collisions. The denominator represents the actual resolution time that both agents need to resolve the situation. Thus, $\mathcal{Q}_{\text{utility}}$ is equal to 1 for scenarios with the highest utility, in which both agents can pass without any interaction. We interpret $\mathcal{Q}_{\text{utility}}$ as a measure for cooperativity because small values mean less efficient interactions in which one or both agents need to reduce their velocity.

2) Driving Comfort: The main goal of the communication system is to reduce the behavior uncertainty allowing for more efficient traffic interactions. However, it is necessary to validate the system performance from the perspective of the ego vehicle. A system that constantly yields to pedestrians resulting in uncomfortable maneuvers will most likely not be accepted. A common measure to evaluate driving comfort is the longitudinal acceleration. We measure the driving comfort as the strongest deceleration $\mathcal{Q}_{\text{comfort}} = \min a_{\text{lon},e}$ within a scenario with 0 being the best value. Note: Since safety critical situations are out of scope, we don't assess the performance of the proposed system with respect to risk.

D. Empirical Results

We evaluate the *standing* and *walking* scenarios empirically. Therefore, we sample the gap acceptance model of 100 pedestrians from $\mathcal{N}(5.7, 2.0)$. We simulate the interaction of each pedestrian with a vehicle starting at 16 different time gaps for the *standing* scenarios and 28 time gaps for the *walking* scenarios between $2.5\,\mathrm{s}$ and $10\,\mathrm{s}$. We don't simulate interactions with an initial vehicle TTA less than $2.5\,\mathrm{s}$ as we consider them to be out of scope of cooperation. The gaps represent the time gap between the vehicle and the pedestrian at the moment when the pedestrian makes the crossing decision, meaning $d_{\text{STS},o} = 2\,\mathrm{m}$. We compare the behavior of the proposed *communicative* agent with the *cooperative* baseline agent that considers bimodal behavior planning but that is not able to communicate.

1) Standing: The empirical experiments with a *standing* pedestrian indicate that the *communicative* agent is able to improve the interaction both in terms of utility and driving comfort. Up to a starting TTA of $4.5\,\mathrm{s}$ the outcome of both the *cooperative* and *communicative* agents is identical. For

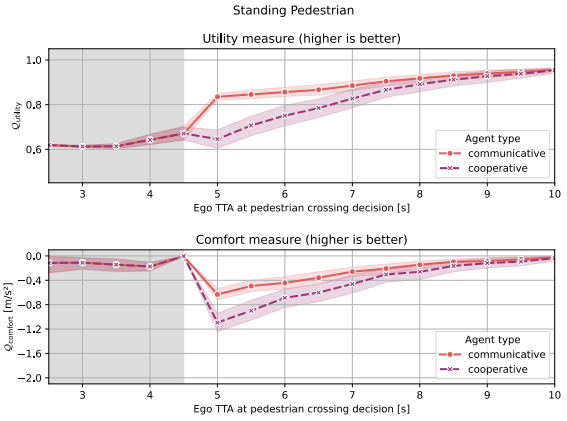


Fig. 4. Empirical results with a *standing* pedestrian. We compare a *cooperative* agent that has no ability to communicate with our *communicative* approach. The gray area represents the distances at which the communication framework decides not to communicate. In the white area the ego vehicle communicates that it gives right of way to the pedestrian.

larger TTAs, the results show a better performance of the *communicative* agent compared to the *cooperative* agent. Moreover, the *communicative* agent is able to reduce the variance of the resulting behavior. Fig. 4 depicts the results of interactions with a *standing* pedestrian. The two compared agents behave identically for TTA $\leq 4.5\,\mathrm{s}$, as illustrated by the gray area in the plots. On average, the scenarios are resolved with a low utility score $\mathcal{Q}_{\text{utility}}$ and a high comfort score $\mathcal{Q}_{\text{comfort}}$. Up to that point, the AV is not willing to accept a larger discomfort to let the pedestrian cross first. After the evaluation of the joint costs, the option of passing first is the overall better solution for the vehicle, unless the pedestrian decides to cross the road in front of the vehicle. If the pedestrian does so, the behavior of the pedestrian does not match the vehicle's prediction, thereby forcing the vehicle to decelerate unexpectedly. The gray area represents these distances at which the ego vehicle does not communicate to let the pedestrian cross. The results show a difference between the *cooperative* and *communicative* agents for TTA $> 4.5\,\mathrm{s}$. This TTA can be interpreted as the decision point at which the communication framework decides to communicate the vehicle's yielding decision. The decision point is an implicit result of the communication framework and it varies depending on the parametrization (proximity of pedestrian to STS, vehicle velocity, predicted walking speed, reaction time). The *communicative* agent is able to increase the scores of the quality measures by communicating its yielding intention. The *cooperative* agent can not achieve this improvement, even though the trade-off between utility and driving comfort results in a decision in favor of letting the pedestrian cross first. However, in case of interactions between the *cooperative* agent and a pedestrian that is not willing to accept the time gap the yielding decision leads to a mismatch between the actual pedestrian behavior and the best predicted option from the perspective of the vehicle. The *communicative* agent is able to convey the "correct" resolution to the pedestrian. This does not only increase the total utility but we also obtain an improvement in

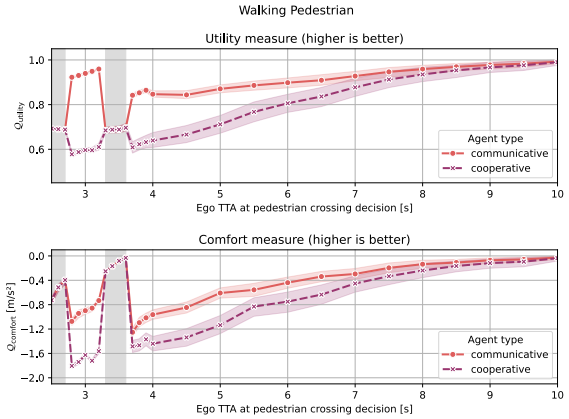


Fig. 5. Empirical results with a *walking* pedestrian. The plot shows the influence of the pedestrian's walking kinematics on the communication decision. $\mathcal{Q}_{\text{utility}}$ is a measure for the joint utility of both interaction partners and $\mathcal{Q}_{\text{comfort}}$ denotes the driving comfort of the ego vehicle.

driving comfort. The AV is able to change the initial decision of the pedestrian not to cross, thus, resolving the mismatch between actual and recommended behavior. This allows safe and early crossing of the pedestrian which facilitates more comfortable braking. As the initial time gap increases, more and more pedestrians are willing to cross the road according to their gap acceptance model. Thus, the difference between the two agents decreases and both quality measures converge towards their optimal score. This means a reduced impact of the communication for large time gaps.

2) Walking: Similar to the *standing* experiments, the *communicative* agent is able to improve the interaction with *walking* pedestrians. With a pedestrian walking towards the road curb we add a dynamic component to the decision making of the ego vehicle. Due to that dynamic component, the communication decision of the *communicative* agent changes compared to the *standing* scenarios. The *communicative* agent adjusts the communication decision to the walking behavior, thus, it signalizes its yielding intention at generally smaller time gaps. Furthermore, the walking behavior leads to a two-stage communication decision depending the dynamics of the pedestrian. Fig. 5 shows the outcome of the interactions with *walking* pedestrians. The decision to yield and to communicate to the pedestrian at generally smaller time gaps can be explained by two reasons. First, the prediction changes according to the constant walking behavior that is perceived at the very beginning of the scenario. Secondly, even though the pedestrian makes the crossing decision after $2\,\mathrm{s}$ when it reaches the road curb, the additional simulation time before the crossing decision gives the vehicle more time to anticipate the pedestrian behavior and make an earlier communication decision. Similar to the *standing* experiments, the results can be divided into sections where the *communicative* agent decides to communicate its yielding intention and sections where it does not communicate. Only now, there are four sections instead of two compared to the *standing* experiments. This can be explained by the influence of the walking dynamics on the communication decision. Between $2.5\,\mathrm{s}$ and $2.7\,\mathrm{s}$ the results of the *cooperative* and the *communicative* agents are identical because both agents would decide not to yield. We see a first difference between the two agents between $2.8\,\mathrm{s}$ and $3.2\,\mathrm{s}$. The yielding decision of both agents even for such small time gaps is caused by the walking behavior of the pedestrian. The vehicle infers that it is more beneficial if the pedestrian keeps walking and crosses first. However, when making the crossing decision, there are only a few pedestrians that would accept such small time gaps. In most cases, this results in a mismatch between the predicted behavior and the actual decision of the pedestrian. This mismatch causes a strong deceleration and a low utility for the *cooperative* agent. The decision point of the *communicative* agent, however, coincides with the walking movement of the pedestrian. Thus, the *communicative* agent signalizes its yielding intention which allows for a higher utility and smoother braking. A rejected yielding signal would lead to a similar result as for the *cooperative* agent. Between $3.3\,\mathrm{s}$ and $3.6\,\mathrm{s}$ this decision changes again and the vehicle does not decide to yield. This can be explained by the fact that now the AV perceives the deceleration of the pedestrian. The vehicle interprets this behavior such that the pedestrian is willing to stop and wait for the vehicle to pass, which matches the recommended behavior. After a TTA of $3.7\,\mathrm{s}$, the vehicle always decides to let the pedestrian cross independent of the pedestrian kinematics. Subsequently, the results of the *walking* scenarios evolve similar to the *standing* scenarios. As in the *standing* scenarios, the quality measures of both agents become more similar with increasing time gaps. The results indicate that the explicit communication does not only improve the average of the quality measures but it also reduces the variance of the resulting behavior with attentive pedestrians that will accept the communication. This allows a more predictable, thus, more robust interaction.

V. Conclusion

In this work, we present a communication framework that builds upon a multimodal interaction-aware behavior planner and enables cooperation between an AV and pedestrians that want to cross the road. We model the interaction by determining the passing order of a STS. We use a pedestrian crossing model based on related studies from the literature, both on interactions with eHMIs equipped vehicles and ordinary vehicles. We develop a heuristic communication module that communicates the AV's yielding decision to the pedestrian to give right of way. We use a simple, yet generic scenario in which a pedestrian intends to cross an unsignalized road as this scenario implies unclear priority. However, the proposed system can also be applied to traffic situations with clear priority, such as interactions at crosswalks, by considering the priority in the cost function of the planned behavior.

Relating to research question *A*, whether explicit external communication helps improving the overall utility and driving comfort, the experiments show that our *communicative* approach can reduce the uncertainty about the pedestrian's behavior compared to the *cooperative* approach without

communication. Thus, the overall utility is increased while at the same time improving the driving comfort by reducing the AV's yielding deceleration. Regarding research question *B*, what the optimal timing for communication would be, we obtain a decision point at which the system decides to yield to pedestrians and let them cross. The decision point varies depending on the parametrization and the setting of the situation as well as the perceived pedestrian movement. This makes the proposed system adaptable to various scenarios with different geometrical and dynamic properties.

The presented approach is limited to communicating that the vehicle is giving right of way, which means that situations with a different optimal resolution cannot be improved by the system. Moreover, the benefit of the communication relies on the assumption that a waiting pedestrian would perceive and accept the communicated signal by the AV. Related real-world studies show that the reaction to eHMIs can be diverse, depending on subjective properties such as personal crossing preferences or attentiveness. Incorrectly inferred crossing intentions can also lead to inefficient behavior.

Extending the communication framework by a tracking module that accounts for mismatches of the behavior with the recommendation can help to reasonably adjust behavior recommendations if necessary. To make the presented approach more generalizable, other modulators such as pedestrian age, attentiveness, and comprehensibility of the communication can be modeled. Furthermore, the performance of the proposed system will be validated in a small-scale user study.

REFERENCES

[1] A. Habibovic, V. Fabricius, J. Andersson, M. Klingegård, T. Lagström, A. Sirkka, J. Fagerlönn, C. Edgren, R. Fredriksson, K. Stas, D. Saluäär, and P. Larsson, "Communicating intent of automated vehicles to pedestrians," *Frontiers in Psychology*, vol. 9, p. 1336, 08 2018.

[2] A. Dietrich, M. Tondera, and K. Bengler, "Automated vehicles in urban traffic: The effect of kinematics and ehmis on pedestrian crossing behavior," *Advances in transportation studies*, vol. 2020, pp. 73–84, 2020.

[3] A. Schieben, M. Wilbrink, C. Kettwich, R. Madigan, T. Louw, and N. Merat, "Designing the interaction of automated vehicles with other traffic participants: design considerations based on human needs and expectations," *Cognition, Technology & Work*, vol. 21, pp. 69–85, 02 2019.

[4] K. Tian, A. Tzigieras, C. Wei, Y. M. Lee, C. Holmes, M. Leonetti, N. Merat, R. Romano, and G. Markkula, "Deceleration parameters as implicit communication signals for pedestrians' crossing decisions and estimations of automated vehicle behaviour," *Accident Analysis & Prevention*, vol. 190, p. 107173, 2023.

[5] D. Dey and J. Terken, "Pedestrian interaction with vehicles: Roles of explicit and implicit communication," in *9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 09 2017, pp. 109–113.

[6] P. B. Anirudh Sripada and J. de Winter, "Automated vehicles that communicate implicitly: examining the use of lateral position within the lane," *Ergonomics*, vol. 64, no. 11, pp. 1416–1428, 2021.

[7] K. Bengler, M. Rettenmaier, N. Fritz, and A. Feierle, "From hmi to hmis: Towards an hmi framework for automated driving," *Information*, vol. 11, p. 61, 01 2020.

[8] Y. E. Song, C. Lehsing, T. Fuest, and K. Bengler, "External hmis and their effect on the interaction between pedestrians and automated vehicles," in *1st International Conference on Intelligent Human Systems Integration (IHSI 2018)*, 01 2018, pp. 13–18.

[9] G. de Clercq, A. Dietrich, J. P. Nuñez Velasco, J. de Winter, and R. Happee, "External human-machine interfaces on automated vehicles: Effects on pedestrian crossing decisions," *Human Factors The Journal of the Human Factors and Ergonomics Society*, vol. 61, pp. 1351–1370, 12 2019.

[10] M. Matthews, G. Chowdhary, and E. Kieson, "Intent communication between autonomous vehicles and pedestrians," *arXiv preprint arXiv:1708.07123*, 08 2017.

[11] W. M. Alvarez, M. A. De Miguel, F. García, and C. Olaverri-Monreal, "Response of vulnerable road users to visual information from autonomous vehicles in shared spaces," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*. IEEE, 2019, pp. 3714–3719.

[12] W. M. Alvarez, F. M. Moreno, O. Sipele, N. Smirnov, and C. Olaverri-Monreal, "Autonomous driving: Framework for pedestrian intention estimation in a real world scenario," in *2020 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2020, pp. 39–44.

[13] R. Izquierdo, S. Martín, J. Alonso, I. Parra, D. Fernández-Llorca, and M. A. Sotelo, "Human-vehicle interaction for autonomous vehicles in crosswalk scenarios: Field experiments with pedestrians and passengers," in *26th IEEE International Conference on Intelligent Transportation Systems ITSC 2023*. IEEE, 2023.

[14] D. Eisele, C. B. Bubeck, and T. Petzoldt, "Equipping vehicles with novel ehmis potentially changes how pedestrians interact with vehicles without ehmis," in *26th IEEE International Conference on Intelligent Transportation Systems ITSC 2023*. IEEE, 2023.

[15] R. Wenzel, M. Probst, T. Puphal, T. H. Weisswange, and J. Eggert, "Asymmetry-based behavior planning for cooperation at shared traffic spaces," in *2021 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2021, pp. 1008–1015.

[16] M. Probst, R. Wenzel, T. Puphal, M. Komuro, T. H. Weisswange, N. Steinhardt, B. Bolder, B. Flade, Y. Sakamoto, Y. Yasui, and J. Eggert, "Automated driving in complex real-world scenarios using a scalable risk-based behavior generation framework," in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*. IEEE, 2021, pp. 629–636.

[17] B. Varga, D. Yang, and S. Hohmann, "Intention-aware decision-making for mixed intersection scenarios," in *2023 IEEE 17th International Symposium on Applied Computational Intelligence and Informatics (SACI)*. IEEE, 2023, pp. 369–374.

[18] S. Gupta, M. Vasardani, and S. Winter, "Negotiation between vehicles and pedestrians for the right of way at intersections," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 3, pp. 888–899, 2018.

[19] K. Li, Y. Chen, M. Shan, J. Li, S. Worrall, and E. Nebot, "Game theory-based simultaneous prediction and planning for autonomous vehicle navigation in crowded environments," in *26th IEEE International Conference on Intelligent Transportation Systems ITSC 2023*. IEEE, 2023.

[20] C. Burger, J. Fischer, F. Bieder, Ö. Ş. Taş, and C. Stiller, "Interaction-aware game-theoretic motion planning for automated vehicles using bi-level optimization," in *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2022, pp. 3978–3985.

[21] B. Schroeder, N. Rouphail, K. Salamati, E. Hunter, B. Phillips, L. Elefteriadou, T. Chase, Y. Zheng, V. P. Sisiopiku, and S. Mamidipalli, "Empirically-based performance assessment & simulation of pedestrian behavior at unsignalized crossings," Southeastern Transportation Research, Innovation, Development and Education Center (STRIDE), Tech. Rep., 2014.

[22] T. Giuffrè, T. Campisi, and G. Tesoriere, "Implications of adaptive traffic light operations on pedestrian safety," *IOSR Journal of Mechanical and Civil Engineering (IOSR-JMCE)*, vol. 13, pp. 58–63, 11 2016.

[23] T. Petzoldt, K. Schleinitz, and R. Banse, "Potential safety effects of a frontal brake light for motor vehicles," *IET Intelligent Transport Systems*, vol. 12, no. 6, pp. 449–453, 2018.

[24] J. F. Medina-Lee, V. Trentin, J. L. Hortelano, A. Artuñedo, J. Godoy, and J. Villagra, "$IA(MP)^2$: Framework for online motion planning using interaction-aware motion predictions in complex driving situations," *IEEE Transactions on Intelligent Vehicles*, pp. 1–16, 2023.

[25] T. Puphal, M. Probst, and J. Eggert, "Probabilistic uncertainty-aware risk spot detector for naturalistic driving," *IEEE Transactions on Intelligent Vehicles*, vol. 4, no. 3, pp. 406–415, 2019.

[26] J. Zębala, P. Ciępka, and A. Reza, "Pedestrian acceleration and speeds," *Problems of Forensic Sciences*, vol. 91, pp. 227–234, 2012.

[27] Y. Zheng, T. Chase, L. Elefteriadou, B. Schroeder, and V. P. Sisiopiku, "Modeling vehicle–pedestrian interactions outside of crosswalks," *Simulation Modelling Practice and Theory*, vol. 59, pp. 89–101, 2015.