

Article

Data-Driven Test Scenario Generation for Cooperative Maneuver Planning on Highways

Christian Knies * and Frank Diermeyer Institute of Automotive Technology, Technical University of Munich, 85748 Garching, Germany;
diermeyer@ftm.mw.tum.de

* Correspondence: knies@ftm.mw.tum.de

Received: 19 October 2020; Accepted: 15 November 2020; Published: 18 November 2020



Abstract: Future automated vehicles will have to meet the challenge of anticipating the intentions of other road users in order to plan their own behavior without compromising safety and efficiency of the surrounding road traffic. Therefore, the research area of cooperative driving deals with maneuver-planning algorithms that enable vehicles to behave cooperatively in interactive traffic scenarios. To prove the functionality of these algorithms, single test scenarios are used in the current body of literature. The use of a single, exemplary scenario bears the risk that the presented approach only works in the presented scenario and thus no general statement can be made about the performance of the algorithm. Furthermore, there is a risk that fictitious traffic scenarios may be solved which do not occur in reality. Therefore, we present a procedure for generating test scenarios based on real-world traffic datasets that require cooperation of at least one of the involved vehicles and thus are challenging from the perspective of cooperation. This procedure is applied to a large highway traffic dataset, resulting in a test scenario catalog that allows a comprehensive performance evaluation. The extracted scenarios are clustered according to the cooperative actions used to solve the respective scenario, which enables a more detailed understanding of the underlying cooperative mechanisms. In order to serve as a basis for making comparisons between different behavior planners and thus contribute to the development of future maneuver planning algorithms, a tool to extract the test scenarios from the used traffic dataset is made publicly available.

Keywords: cooperative driving; automated vehicles; connected vehicles; CAV; test scenarios

1. Introduction

Human drivers behave cooperatively in road traffic by anticipating the intentions of other vehicles and adapting their own behavior accordingly. In this way, they support the surrounding vehicles in achieving their own tactical goals. This cooperative behavior enhances traffic flow and prevents critical situations. If automated vehicles (AV) are not able to perform such cooperative maneuvers, safety and traffic flow in the future will decrease as the number of AVs increase. Therefore, researchers in the field of cooperative driving are working on maneuver planning algorithms that are able to interact with the surrounding traffic in a cooperative way. The published work in this field proves the functionality of their approaches based on a small number of scenarios, e.g., [1–4]. However, the main question is not whether a maneuver planner works in a single scenario, but how well it does in a wide range of realistic scenarios. Only such a comprehensive performance evaluation enables the comparison of different approaches of maneuver planners, which is necessary for the research field of cooperative driving to make progress.

Generating significant test scenarios is the subject of current research in the field of AV. However, most publications (e.g., [5–8]) focus on critical scenarios needed to validate AV safety. In addition, more recent publications also consider the behavior of AV alongside safety criticality, testing the

completion of driving tasks [9–11]. However, these only consider the performance of the AV under test and not the effects on other road users in terms of cooperative behavior. With the CommonRoad framework, Althoff et al. [12] introduced a platform for composing and sharing motion planning problems that consist of a scenario, a vehicle model, vehicle parameters, and cost functions. For each scenario, a ranking is provided, where users can upload and compare the results of their maneuver planners. Since CommonRoad only hosts the uploaded scenarios, it is dependent on the content generated by its community.

Generating scenarios for testing AV cooperative behavior has been a less addressed topic. Initial works in this field propose artificially creating test scenarios. Whereas Lizenberg et al. [13] use a Simulation of Urban Mobility (SUMO) to detect interesting situations from the perspective of cooperation, Hallerbach et al. [14] create specific scenarios from a predefined logical scenario by randomly choosing the respective parameters from limited parameter spaces. After simulating the generated scenarios, a combined metric uses a threshold to identify critical situations.

The main aim of this paper is to present a test scenario catalog for cooperative maneuver planning that is based on real traffic data. The advantage of the data-driven concept lies in the automated and realistic generation of scenarios. No manual effort is required to create the scenarios, and they are not constrained by the behavior model of traffic simulation programs or the imagination of the user generating the scenarios either. A further contribution to the state of the art is achieved by introducing a scenario selection approach that is able to decide if a certain scenario requires cooperation, which would thereby ensure that only relevant scenarios are included in the test scenario catalog. The presented scenario generation pipeline is applied to a state-of-the-art highway dataset and a tool to extract the computed scenarios is made publicly available on www.github.com/TUMFTM/test_scenarios_cooperation.

2. Materials and Methods

After explaining the concept and process of scenario extraction, this section presents the applied metrics for cooperation, the simulation environment with its driver models as well as the dataset on which the analyses were performed.

2.1. Concept

The goal in creating the scenario dataset is to provide the possibility of testing different maneuver planners against each other. In terms of methodology, it is therefore advisable to base the scenario generation on the quality criteria of test theory. These comprise objectivity, reliability and validity as main criteria ([15], S. 179) and standardization, economy, and practicability as subsidiary criteria ([16], p. 485 f.). The following section explains how these quality criteria are considered in the concept of test scenario generation.

Objectivity: If a test achieves the same results regardless of the person performing it, the test is objective. Since the test in our case represents the execution of a simulation, it is objective if the conditions of the simulation are clearly defined. Therefore, the description of the roadway characteristic, starting position (longitudinal and lane) and initial speed, as well as the duration of the simulation, are determined for each scenario.

Reliability: A test that produces the same results in repeated measures is reliable. Since a simulation doesn't undergo external influences, as e.g., field tests do, the simulative test of a single scenario is inherently reliable. Because different maneuver planners are better at different scenarios, a reliable comparison must include a large number of diverse situations and varying roadway characteristics, such as differing lane numbers and acceleration lanes. This was taken into account by using a large traffic dataset that was recorded in several locations for the scenario extraction.

Validity: In contrast to the model validation domain, where validity refers to the consistency between model and reality, validity in terms of test theory means that a test measures exactly what it is supposed to measure. In the case of cooperative behavior planning, this comprises the metrics used to

measure cooperation, the test mode in which the maneuver planners are tested as well as the selection of test scenarios. The applied metrics are the result of a previous work [17] and are presented briefly in Section 2.3. The test mode must ensure that the tested maneuver planners not only behave intelligently in situations where they need the cooperation of other vehicles but also when they are in the position to support others. Therefore, the test mode intends that the maneuver planner being tested is applied to all vehicles of a scenario within the same simulation. For this reason, no trajectory is specified for any of the vehicles, but rather the starting situation. From this starting situation, the maneuver planners are applied until the determined duration of the scenario is reached. In order to be valid for testing cooperative behavior, the test scenarios must benefit from cooperation. In the context of cooperative driving, Düring and Pascheka [18] state that cooperative behavior affects the utility of the cooperatively acting agent and the utility of at least one other agent in a way that the total utility increases with respect to a reference utility. Furthermore, the cooperative action must be performed intentionally. This concept implies that every scenario needs to be simulated twice, by a reference driver model and by a completely cooperative driver model. We then regard a scenario as valid if at least one vehicle accepts a lower individual utility to increase the total utility in the cooperative solution compared to the solution of the reference model. The individual drawback of at least one vehicle is not a requirement according to the definition of Düring and Pascheka [18], but in order to imply the intention to increase the other's utility and not only the ego utility, this is a necessary assumption. In the sense of Düring and Pascheka [18], cooperative actions with an individual drawback of the acting agent are referred to as altruistic-cooperative.

Standardization: A standardized test delivers a reference for its test results. Within the generation of the test scenarios, a cooperative global planning algorithm is used that computes the best possible solution of the scenarios within a given discretization. The results of this fully cooperative planner represent the lower limit of achievable costs, whereas the reference driver model without behavior planning ability can be regarded as the upper cost limit. Both algorithms are described in Section 2.4.

Economy: From the perspective of economy, the resources required for the execution of a test should not be unnecessarily high. Therefore, we use a scenario-based approach instead of a large-scale traffic simulation that allows only scenarios that are relevant for testing cooperative behavior to be computed.

Practicability: A test is practicable if the test method is appropriate for the research purpose. Since the goal of the test scenarios is to assess how a cooperative behavior planner would perform in road traffic, the scenarios contained in the test data set should be as natural and realistic as possible. The easiest way to achieve this is to extract the scenarios from a real traffic data set. This approach ensures that the test scenario generation is not restricted by manual inputs such as defined parameter spaces or a certain behavior model. A further aspect of practicability is that not every scenario that could be solved cooperatively in theory would be solved cooperatively in practice because of the acceptance of the driver of the AV. For example, it cannot be assumed that drivers are willing to accept a high individual disadvantage for a small overall advantage. From which ratio of ego-vehicle-costs to surrounding-vehicle benefit a cooperative action is accepted depends on the degree of cooperativity of the AV. According to [1,17], this ratio lies between 0 (uncooperative because no extra costs are accepted) and 1 (fully cooperative because every extra cost is accepted for an overall benefit). Since there is no established value regarding to what extent AV behave cooperatively, we assume an intermediate value of 0.5. This means for our scenarios that the drawback of the cooperating vehicle has to pay off twice-once to compensate the extra costs of the cooperating vehicle to an overall cost difference of 0 and the second time to reach the overall benefit in size of the drawback. Scenarios in which the overall cost benefit is smaller than the drawback of the cooperating vehicle are assumed to have no practical relevance and are therefore discarded.

2.2. Process

The process of scenario generation is divided into the following phases: scenario extraction, pre-selection, strategic goal generation, and evaluation of cooperation. This process is repeated several times for every roadway characteristic of the dataset.

In traffic data sets, vehicles are recorded as a continuous flow. In order to generate scenarios from this, consecutive sequences of vehicles must be extracted. Therefore, we place a reference line in the beginning of the recorded area (dashed line in Figure 1) and select the n next vehicles with n being the desired number of vehicles in the scenario. When one vehicle is replaced by another among the n next vehicles, a new scenario is created. In this way, the entire data set is processed into scenarios without duplicates and the maximum number of vehicles per scenario is set to 4. This upper limit of vehicles is restricted by the cooperative planner's exponential growth of computing time, which is described in detail in Section 2.4.2. If there are fewer than four vehicles in a recording frame, the scenario is extracted with the respective number of vehicles.

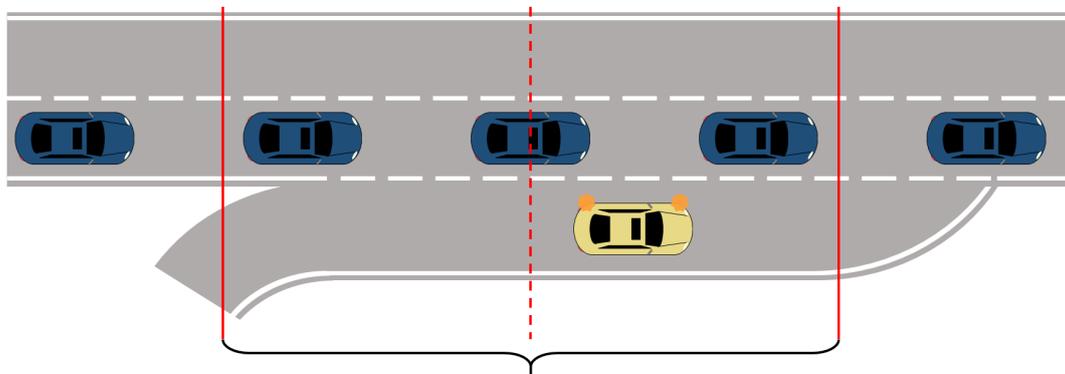


Figure 1. Scenario extraction with four vehicles.

In order not to simulate all extracted scenarios and thereby save computing time, a pre-selection filters out obviously irrelevant scenarios. This includes situations where the lowest time to collision [19] is greater than the duration of the simulation because no interaction is supposed to take place in the considered time period. In addition, scenarios extracted from bound or congested traffic states, which can be determined by an average speed of less than 60 km h^{-1} ([20], p. 88), are discarded because congestions can only be modeled with a large-scale traffic simulation and not with a small number of vehicles in a scenario-based approach.

The remaining scenarios consist of a roadway model and the movement states of the vehicles in their starting position. In addition to the initial state, the vehicles must be given a strategic goal that the behavior models are to pursue in the subsequent simulations. Since the presented work only considers motorway-like roads without intersections or similar, the strategic goal is to drive at a given speed, also referred to as desired velocity. This speed is determined individually for each vehicle based on its respective state. If a vehicle has a leader, the desired velocity might be higher than the current speed due to the obstruction of the front vehicle. In this case, the desired velocity is set to the vehicle's maximum velocity throughout the entire recording. If a vehicle has no leader car, it is supposed to already drive at its desired velocity, which in this case is therefore set to the current speed.

Starting from the initial state, the scenarios are simulated with the reference and the cooperative behavior planner. According to the concept described in Section 2.1, the maximum individual drawback as well as the overall cost benefit are determined by the results of the simulations. If a vehicle in the scenario performs a cooperative action and thereby accepts an individual drawback that leads to an overall cost benefit that is equal or higher than the cost drawback of the cooperating vehicle, the scenario is considered valid and relevant for testing cooperative behavior and thus added to the scenario catalog.

The action that leads to the individual drawback of the cooperating vehicle and thereby causes the benefit for the surrounding vehicles is considered the main cooperative action. The extracted scenarios are clustered by their main cooperative action in order to generate a deeper understanding of the underlying cooperative mechanisms in highway road traffic.

2.3. Metrics

A core aspect of the scenario generation process is the evaluation of cooperation in the performed simulations. In our previous work [17], we presented such a metric for measuring cooperation. The basic ideas of this paper are presented in this section briefly, and a detailed description can be found in the original publication [17].

The cost function is structured in the sub-fields of safety, energy efficiency, time efficiency, and lateral maneuvers. The safety metric considers the distance headway of each vehicle and calculates the reaction time required to avoid a collision in case of an emergency braking of the lead vehicle. Situations with required reaction times of less than 0.5 s are classified as unsafe and result in a negative rating.

At the maneuver level, energy efficiency is mainly influenced by braking. Therefore, the energy that is converted into heat by the brakes during deceleration is estimated and used as a measure of energy efficiency.

The evaluation of time efficiency is based on the concept of vehicle individual desired velocity. By deviating from this desired velocity, the vehicle loses time, which causes an increasingly negative evaluation with growing time loss.

Lane changes are regarded as the only lateral maneuvers on highways and are charged with costs for each time step when they occupy more than one lane. The single cost terms are combined by weighting factors to a metric for cooperation. Determining the weight factors is also described in detail in the original publication [17].

2.4. Simulation Environment

As explained in Section 2.2, the scenarios extracted from the data set are simulated with a reference and a completely cooperative maneuver planner to filter out the relevant ones for the test scenario catalog. These simulations are performed in a self-developed environment that meets the basic requirements for modeling the maneuver planning level of highway-like road traffic. The implemented roadway model consists of a variable number of straight lanes and highway entrances that are 250 m long. The longitudinal dynamics of the vehicles is modeled by a basic kinematic approach:

$$\begin{pmatrix} s'_i \\ v'_i \end{pmatrix} = \begin{pmatrix} 1 & \Delta t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} s_i \\ v_i \end{pmatrix} + \begin{pmatrix} \frac{1}{2}\Delta t^2 \\ \Delta t \end{pmatrix} a_i \quad (1)$$

where the control input of vehicle i is represented by the acceleration a_i and the velocity v_i and position s_i in the next time step (denoted by ') change accordingly. The simulation time step is set to $\Delta t = 0.1$ s.

The lateral direction is represented by discrete lanes. Therefore, lane changes comprise a fixed duration in which the changing vehicle occupies both involved lanes. Lane changes have a mean duration of 4.8 s for cars [21,22] and 7.7 s to 8 s for trucks [22,23]. Since, for the simulation, the duration of a lane change comprises only the time the vehicles block both lanes and not the entire lane change, the lane change durations are modeled at 4 s for cars and 6 s for trucks.

To fulfill the criterion of objectivity as stated in Section 2.1, the duration of the performed simulations must be specified, and the simulation time should be long enough that the scenarios can be solved completely. Theoretically, we regard a scenario as solved when the vehicles reach a state in which they can drive at their desired velocity without needing to brake or change lanes, or, in other words, when the cost terms presented in Section 2.3 become zero. From this point on, the vehicles follow each other without changing their order. Practically, a vehicle can end up in a situation in which it follows a slightly slower leader, but the costs for a lane change would be higher than the costs for the

time loss when staying behind. Therefore, the time costs will not become zero in all scenarios but will asymptotically approximate a constant value near zero. For the scenario generation process, we used a simulation time of 20 s and validated it using a plot that shows the cost terms over simulation time averaged for all scenarios (Appendix A). The value of 20 s proved to fulfill the demands of zero costs for energy and lane changes as well as asymptotic approximation for time costs.

2.4.1. Reference Behavior Planner

The reference planner should provide a basic and reactive driving behavior without tactical maneuver planning. Such behavior can be generated by standard driver models from the field of traffic simulation. Hamdar [24] compares the difference of these models in terms of their macroscopic behavior in the form of fundamental diagrams and their microscopic behavior on the basis of their trajectories. Since the “Intelligent Driver Model” (IDM) as a predecessor of the Improved Intelligent Driver Model (IIDM) achieves realistic results in both areas and at the same time uses only a few parameters, the IIDM with its associated lane change model, “Minimizing Overall Braking Induced by Lane change” (MOBIL), is used as the reference planner.

The IIDM behaves similar to an Adaptive Cruise Control which regulates to the set speed v_0 during free flow and maintains a specified time gap T to the front vehicle during car-following. The maximum acceleration and the approaching behavior are considered by the parameters a and b . A detailed description of the model can be found in ([25], p. 187 ff.). Table 1 shows the parameters applied to the simulation.

The MOBIL model [20] makes decisions about lane changes based on two criteria, one for safety and one for incentives. The safety criterion considers the necessary braking of the approaching vehicle in the target lane. If its hypothetical acceleration falls below the permissible threshold value b_{safe} , the lane change is not performed. The incentive criterion weighs the hypothetical accelerations of all participating vehicles with and without lane changes against each other. The needs of the other drivers are weighted with a politeness factor $0 \leq p \leq 1$, where $p = 0$ represents egoistic and $p = 1$ altruistic behavior. If the accumulated advantage resulting from the lane change exceeds threshold Δa , the lane change is executed. According to the European traffic rules, passing in the right lane is forbidden in non-congested traffic states, as determined by a velocity over 60 km h^{-1} . The obligation to drive in the rightmost possible lane is not considered (parameter $a_{bias} = 0$) because the data show that it is hardly observed in real traffic. The parameters used for simulation are listed in Table 1.

Whereas most of the parameters of the IIDM model were set according to the recommendations of model’s authors [25,26], a different parameterization was chosen for the time gap T . The standard values of the time gap led to strong decelerations at the beginning of the simulations because the distances of the drivers kept in reality were smaller than the default value of 1.0 s ([25], p. 190). Instead, a value of $T = 0.5 \text{ s}$ resulted in the expected behavior.

Table 1. IIDM and MOBIL parameters.

Parameter	Car	Truck
Acceleration coefficient a	1.4 m s^{-2} [26]	0.7 m s^{-2} [26]
Deceleration coefficient b	2.0 m s^{-2} [26]	2.0 m s^{-2} [26]
Desired velocity v_0	individual max. velocity	individual max. velocity
Time gap T	0.5 s	0.5 s
Minimum distance s_0	2.0 m [26]	4.0 m [26]
Acceleration exponent δ	4 [26]	4 [26]
Safe deceleration b_{safe}	-2.0 m s^{-2} ([25], p. 244)	-2.0 m s^{-2} ([25], p. 244)
Threshold Δa	0.1 m s^{-2} ([25], p. 244)	0.1 m s^{-2} ([25], p. 244)
Right lane bias a_{bias}	0 m s^{-2} ([25], p. 244)	0 m s^{-2} ([25], p. 244)
Politeness factor p	0.2 ([25], p. 245)	0.2 ([25], p. 245)

2.4.2. Cooperative Behavior Planner

The scenario generation approach presented in Section 2.1 needs a cooperative solution for each traffic scenario in order to evaluate if it benefits from cooperative behavior and can therefore be regarded as a valid test scenario. However, it is not easy to prove that a maneuver planner works sufficiently well for this purpose. One option is to use a procedure that centrally optimizes the global utility for all vehicles and inherently finds the best solution within the selected discretization level for every planning problem.

In order to model the maneuver planning level of a vehicle, a tree structure is used that contains all possible states of a scenario. Starting from the beginning of the scenario, the root node, all vehicles perform actions over a discrete time step, which leads to the next node in the next time step. Each node contains the movement state of all vehicles. In this way, the tree expands through time until the end of the planning horizon t_{end} . Figure 2 shows this approach using an example scenario with one vehicle.

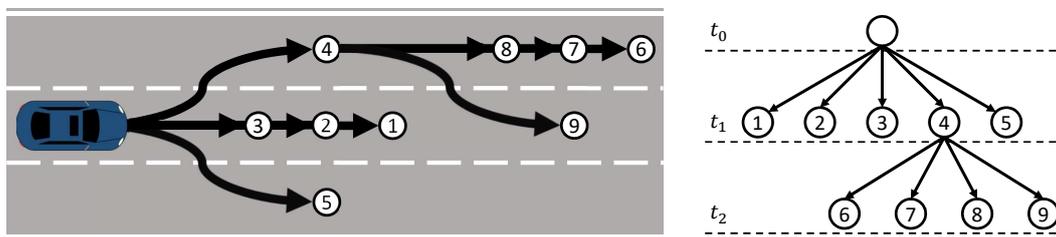


Figure 2. Structure of the search tree.

The optimal solution of the scenario within the selected discretization in time and actions is the sequence of nodes that causes the lowest costs according to the cost function presented in Section 2.3. Due to the exponential growth of tree states:

$$n_{states} = A^{mt_{end}/\Delta t} \tag{2}$$

with number of actions A , number of vehicles m and number of time steps $t_{end}/\Delta t$ the search tree becomes too large to calculate each state. Therefore, a tree search method is used to find the best solution for the scenario with the least possible computation effort. In a previous work [27], we solved the scenarios by applying the Monte Carlo Tree Search Method (MCTS). However, since there is no guarantee for the optimality of the solution in the MCTS procedure, an A* algorithm is used instead.

The A* algorithm aims to find the best sequence of nodes from the start node to a goal node. Therefore, it expands iteratively the nodes of the tree, as depicted in Figure 2, by performing one possible action for each vehicle in the scenario, which again leads to a new node. Each node n is associated with costs $g(n)$ according to the metric described in Section 2.3 applied to all previous actions up to the start node. Furthermore, a node is evaluated by a heuristic $h(n)$ which estimates the minimum costs to reach a goal node from the respective node n . The decision of which node to expand is based on the function $f(n) = g(n) + h(n)$. The algorithm puts each new node in a queue sorted by the minimum value of $f(n)$. At each iteration, the node with the lowest value is taken from the queue and being expanded. The new successor nodes are sorted into the queue and again the node with the lowest value of $f(n)$ is chosen. This procedure continues until a goal node has the lowest f value. If $h(n)$ is guaranteed to underestimate the real costs to reach a goal node, the path to the chosen goal node represents the least-cost sequence of nodes and therefore the optimal solution. Since the algorithms purpose is to operate the vehicles with smallest possible costs over the length of a scenario, all nodes that lie on the planning horizon t_{end} are valid goal nodes. The definition of the vehicles' possible actions as well as the derivation of $h(n)$ are described in the following. For a more detailed description of the A* search algorithm, please refer to ([28], p. 93 ff.).

As described in Section 2.1, the results of the cooperative behavior planning are compared to those of the reference planner. To ensure that the differences between the two methods result from intelligent planning and not from an unequal action space, the actions must be constrained by the same limits. In the IIDM model, the parameter a represents the upper limit for positive accelerations. Negative accelerations can reach any value up to full braking, but the range of a comfortable deceleration is given by parameter b . The action set must therefore reflect the normal driving range between a and b and still be able to provide stronger braking if necessary. To achieve this with a limited number of actions, the situation-dependent IIDM acceleration is also an element of the action set in the cooperative planner. The limits of normal driving behavior are represented by constant accelerations with a and b . Within these limits, the vehicles can maintain their speed, coast with slight deceleration, or follow their leading vehicles according to IIDM acceleration. Lane changes to the left or right form the actions in the lateral direction. In summary, each vehicle can choose from the following options at each state transition:

- Accelerate: constant acceleration of value a
- Maintain speed: constant acceleration of value 0
- Coast: deceleration according to driving resistance forces, at least -0.5 m s^{-2}
- Decelerate: constant deceleration of value b
- Car-following: acceleration according to IIDM model
- Lane change to left lane
- Lane change to right lane

When implementing the A* method, the heuristic used is essential. It estimates the minimum expected future costs for each node. As long as these future costs are not overestimated, the A* method is guaranteed to find the best possible solution within the applied discretization. The speed of the vehicles in their current state can serve as the basis for this best-case estimation. Since the maximum acceleration a_{max} is limited upward by parameter a and downward by full braking, the desired speed cannot be reached immediately. The minimum costs $h(n)$ incurred by deviating from the desired speed v_{des} (compare Section 2.3) can therefore be calculated as:

$$h(n) = \int_0^{t_{max}} \left| \frac{v_{des} - v}{v_{des}} \right| dt = \begin{cases} (1 - \frac{v_{start}}{v_{des}})t_{max} - \frac{a_{max}}{2v_{des}}t_{max}^2 & \text{if } v_{start} \leq v_{des} \\ (\frac{v_{start}}{v_{des}} - 1)t_{max} + \frac{a_{max}}{2v_{des}}t_{max}^2 & \text{else} \end{cases} \quad (3)$$

where $v = v_{start} + a_{max}t$ describes the future velocity v obtained by constant acceleration with a_{max} from current speed v_{start} and $t_{max} = \frac{v_{des} - v_{start}}{a_{max}}$ represents the time until v_{des} is reached.

Depending on the size of a scenario in terms of duration, number of vehicles and discretization in time, it may not be solved with the available amount of working memory or within reasonable computing time. In this case, a timeout fixes the parent node of the best-rated node of the deepest tree level. The previous states of this node are not changed anymore and the search continues from the fixed node. In such cases, the computed solution is not guaranteed to be optimal.

2.5. Dataset

The scenario selection process presented in Section 2.2 requires naturalistic traffic data as the input source. For the scope of this paper, the highD dataset [29] that was generated by a video drone capturing vehicle trajectories within a 420 m long segment of highways was used. It contains recordings from both driving directions of six different locations in Germany with an overall recording time of 147 h, of which the upper driving direction was used for the scenario extraction process. The videos were captured between 8:00 AM and 5:00 PM during sunny and windless weather indicating dry road conditions and unrestricted visibility. The included roadway models cover two- and three-lane motorways as well as three-lane sections with entrance lanes.

3. Results

The results chapter is organized in three subsections. First, a case study illustrates the scenario selection process, followed by a description of the extracted scenario catalog. In order to provide a more detailed understanding of the cooperative mechanisms in highway road traffic, the last section presents a classification of the extracted scenarios according to the cooperative actions performed.

3.1. Case Study

In order to identify the scenarios that are to be part of the scenario catalog, a naturalistic driving dataset is split into scenarios, which are then evaluated using the reference and the cooperative planner, as described in Section 2.2. This assessment process is illustrated using an exemplary scenario. The solution computed by the planners is depicted by means of two plots each. The first plot (Figure 3), referred to as the lateral plot, shows the scenario from the bird's eye perspective in each relevant time step. Besides the start and end of the scenario, relevant time steps are the first time steps of and after a lane change. The second plot (Figure 4) shows the velocity of all involved vehicles in solid lines as well as their desired velocities in horizontal dashed lines. The vertical dashed lines indicate the time steps where a lane change starts or ends. These time steps are the relevant time steps depicted in the bird's eye perspective plot and are assigned by unique numbers.

The start situation is the same for both simulations. The lead vehicle 3 (red) drives in lane 2 at its desired velocity of 26 m s^{-1} , followed by vehicle 1 (orange) at a lead distance of 26.4 m and a speed of 24.3 m s^{-1} . Both vehicles are approached by the small blue truck (vehicle 0) approaching at a considerably higher velocity of 34 m s^{-1} . Due to its greater desired speed of 34.6 m s^{-1} , it might want to overtake vehicles 1 and 2 during the scenario. In lane 1, vehicle 2 (green) catches up at 35.4 m s^{-1} and might block the blue truck from making a future lane change. The leftmost lane 0 is not occupied.

Figures 3 and 4 show the solution of the reference IIDM planner. The blue truck changes to the left in time step $t = 0.2 \text{ s}$ in order to overtake vehicles 1 and 3. To ensure a safe distance to vehicle 1 during the lane change, it decelerates subject to the IIDM until the end of the lane change at $t = 6.1 \text{ s}$. Since vehicle 2 now has a slower object in front, it also starts a lane change to the free lane 0 at $t = 4.3 \text{ s}$. Until the end of its lane change at $t = 8.2 \text{ s}$, it must consider the slower blue truck as its lead vehicle and therefore decelerate. After their lane changes, vehicles 0 and 2 accelerate toward their respective desired velocities subject to the IIDM. The decelerations, the associated time loss, and the lane changes result in scenario costs of 7.6 according to the evaluation function presented in Section 2.3.

The cooperative planner solves the scenario in a different way as depicted in Figures 5 and 6. In the computed solution, vehicle 1 accelerates over its desired velocity (dashed lines in Figure 6), making way for vehicle 0. Thus, the blue truck has more room for its lane change, allowing it to maintain its initial speed. Therefore, vehicle 2 is not affected by the blue truck until the end of the scenario. After the lane change, vehicles 0 and 1 approach their respective desired velocities. The scenario costs for the solution computed by the cooperative planner amount to 0.9 according to the cost function introduced in Section 2.3. In comparison to the reference planner solution, vehicle 1 accepts a drawback in terms of a deviation from its desired velocity in order to reduce the overall costs for this scenario. Therefore, it can be regarded as a scenario that benefits from cooperation.

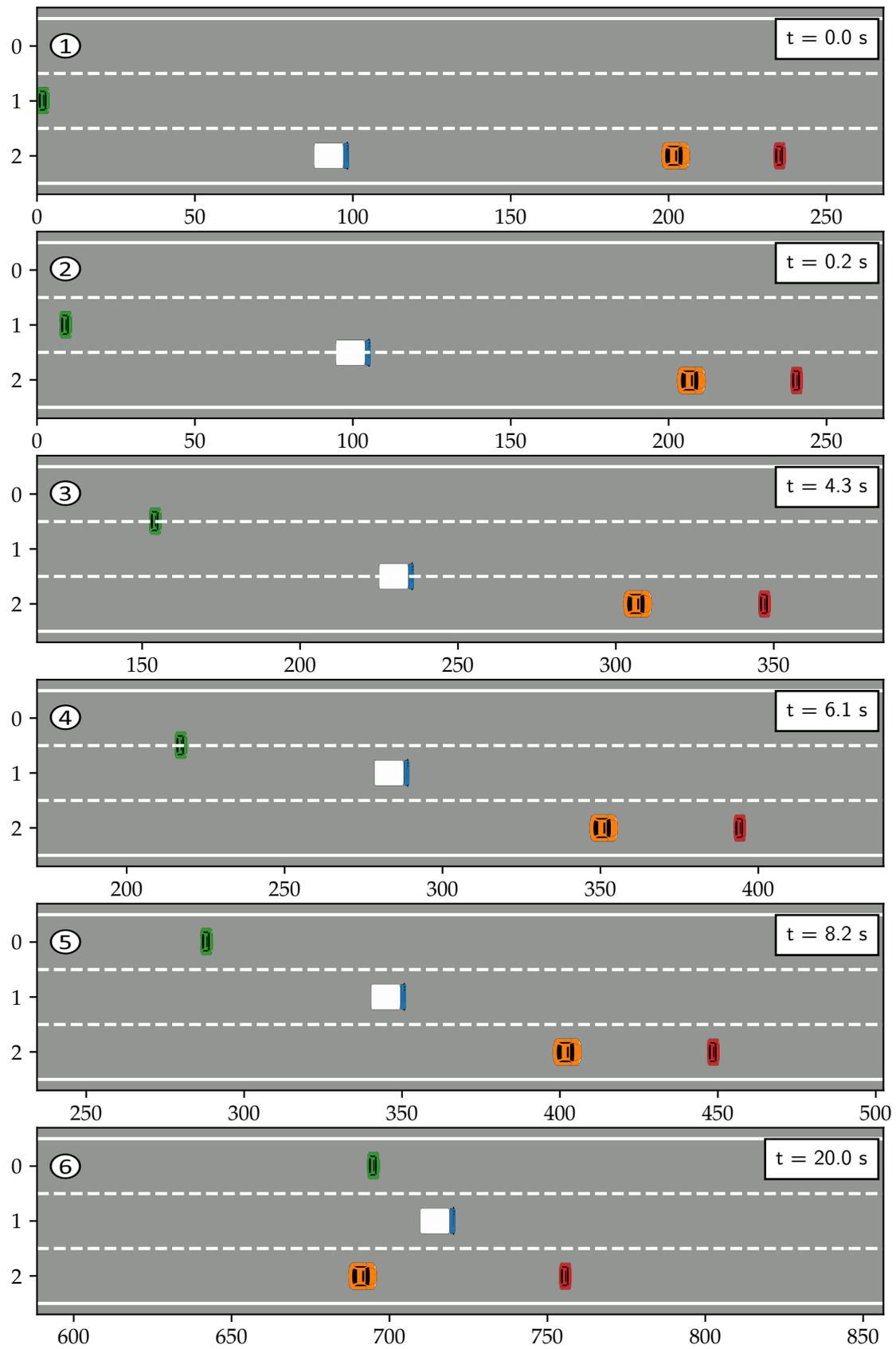


Figure 3. Lateral plot of the simple planner.

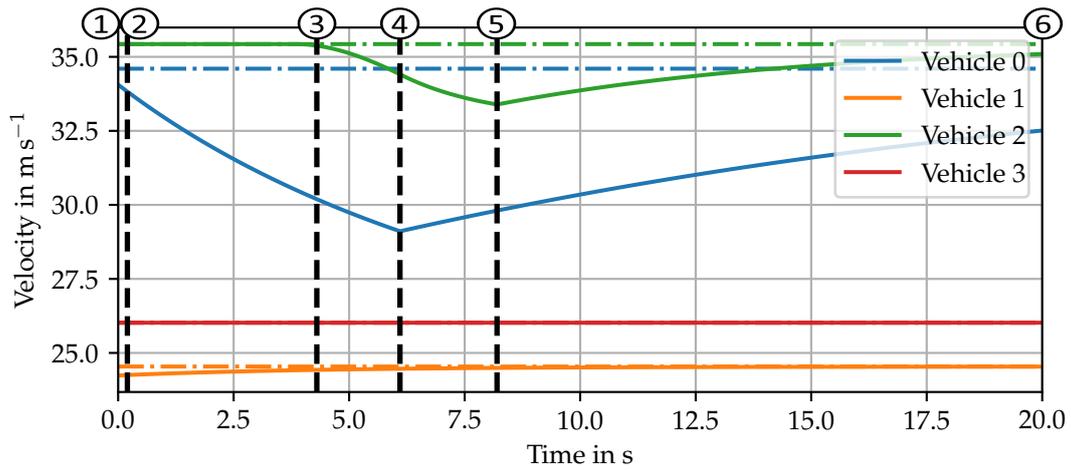


Figure 4. Longitudinal plot of the simple planner.

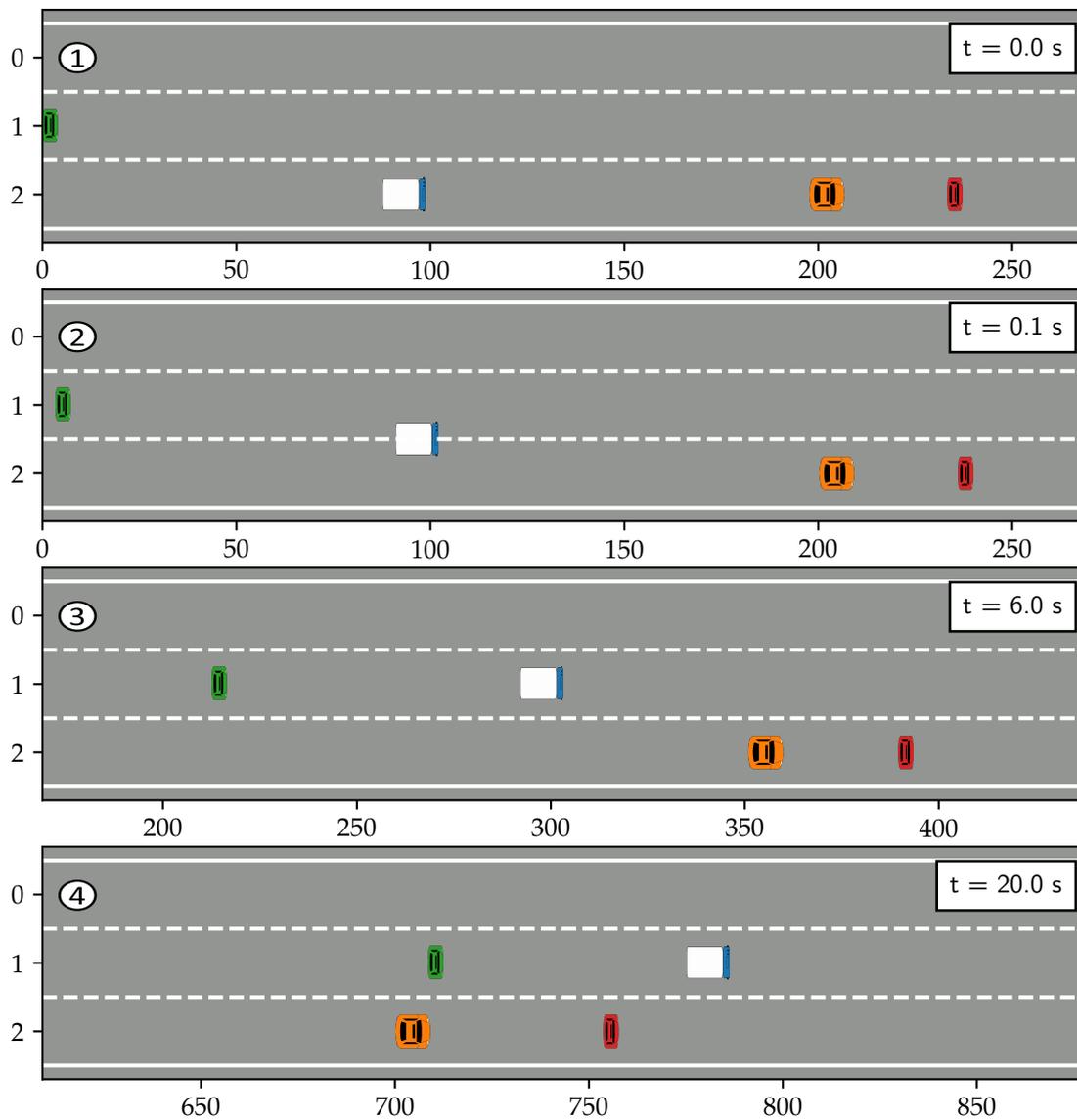


Figure 5. Lateral plot of the global planner.

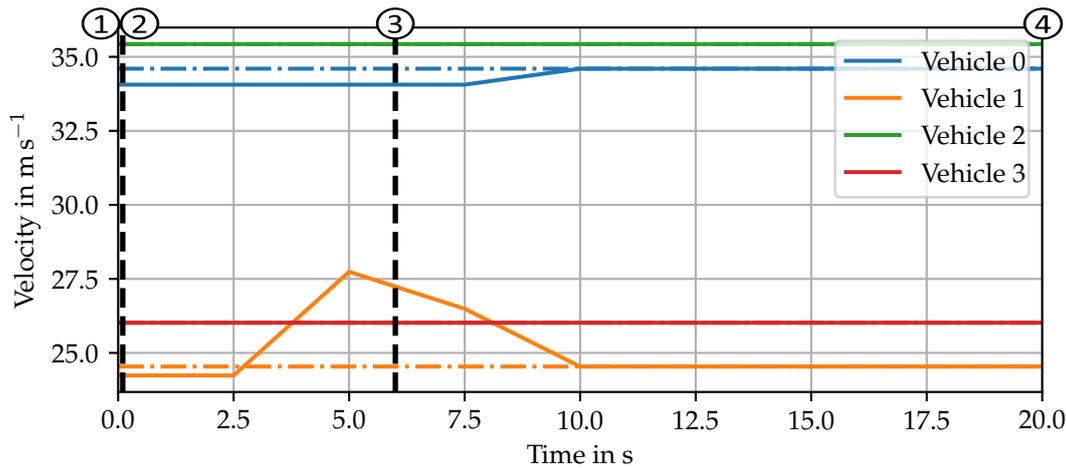


Figure 6. Longitudinal plot of the global planner.

3.2. Scenario Catalog

The described simulations are performed with all scenarios extracted from the data set. The result is a list of scenarios, each rated by the maximum drawback of all involved vehicles and the overall costs of the solutions of the reference and the cooperative planner. The decision as to which scenario should be part of the test catalog is made according to the quality criteria of validity and practicability described in Section 2.1. From the perspective of validity, a scenario is cooperative if at least one vehicle accepts a drawback in the cooperative solution in comparison to the standard behavior of the reference planner with the goal of reducing the overall scenario costs. In order not to classify negligible drawbacks as a cooperative action, a threshold is introduced. The value of this threshold is set equal to the costs of a lane change (0.1), which is considered the smallest clearly cooperative action. In addition to the requirement of an individual drawback, the cooperative solution of a valid test scenario must cause lower costs than the solution with the reference behavior. However, according to the concept of practicability described in Section 2.1, we consider a scenario only as relevant if the overall cost benefit of the cooperative solution is at least equal to the individual drawback of the cooperating vehicle. All scenarios that fulfill these requirements are regarded as valid and practically relevant and are therefore included in the test catalog.

Table 2 illustrates the number of scenarios, the evaluated recording time, and the average cooperation costs of both behavior models for each roadway characteristic. Due to the varying recorded time, the number of extracted scenarios varies between 16 for the three-lane merge scenario and 1902 for the three-lane scenario. The average cooperation costs also differ in terms of roadway characteristics. With the reduction from 3 to 2 lanes, the average costs for both behavior models increase by 75%. For the merge scenario, the average costs raise by a factor of 7 compared to the three-lane scenario without a merge lane. In terms of all roadway characteristics, the reference behavior model shows considerably higher costs (factors 2–3) compared with the fully cooperative behavior model. It is expected that behavior models that use cooperative maneuver planning, but are not capable of controlling all vehicles in a scenario, will fall between the performance of the fully cooperative and the reference planner.

Table 2. Structure of the test scenario catalog.

Roadway	Number of Scenarios	Recording Time	Score Ref. Model	Score Coop. Model
Two-lane	337	192 min	1.80	0.84
Three-lane	1902	759 min	1.04	0.48
Three- and acc. lane	16	30 min	7.52	2.73
Overall	2255	981 min	1.20	0.55

The selected scenarios cover the full speed range of uncongested highway traffic (Table 3). Differences between the solutions of the behavior models are particularly apparent in the minimum velocity, which drops to 0 m s^{-1} in the reference model, but is at least 12.8 m s^{-1} in the cooperative model. The range of acceleration shows that the cooperative model manages to solve the selected scenarios with common longitudinal driving behavior, whereas the reference model needs to perform hard braking maneuvers. Both minimum velocity and acceleration occur on ending lanes after unsuccessful merging maneuvers.

Table 3. Range of velocity and acceleration.

Behavior Model	Range of Velocity	Range of Acceleration
Reference model	0 m s^{-1} to 55.9 m s^{-1}	-9 m s^{-2} to 1.4 m s^{-2}
Cooperative model	12.8 m s^{-1} to 58.4 m s^{-1}	-2.6 m s^{-2} to 1.4 m s^{-2}

3.3. Scenario Clusters

For a more detailed understanding of the cooperative mechanisms in the dataset, a clustering of the main cooperative actions is performed. The main cooperative action is the action that the cooperating vehicle (determined by the highest drawback) performs in order to reduce overall costs of the scenario. According to the cooperation metric's cost terms, the drawback can be dominated by higher safety, time, energy or lateral costs. Since braking induces lower velocity, which affects the time metric, the energy costs are not considered separately. The safety metric is also irrelevant because the cooperative planner does not produce unsafe solutions, whereby the safety costs cannot be higher than in the reference solution. The remaining terms are the metrics for time efficiency and lateral maneuvers. Within the time metric, increased costs for the cooperative solution can be caused by higher ($V+$) or lower ($V-$) speed compared to the desired velocity. Additional lateral costs can be induced by a lane change to the left (LCL) or to the right (LCR). The analysis of these scenario clusters delivers the following cooperative mechanisms:

LCR: The cooperating vehicle performs a lane change to the right. According to the passing-on-the-right prohibition, it enables its follower to pass in the cleared lane.

LCL: The lane change to the left clears the initial lane of the cooperating vehicle. Because of the passing-on-the-right prohibition, the cleared lane cannot be used to pass the cooperating vehicle. Instead, the cleared lane becomes the target lane of another vehicle's lane change from the right. Reasons for the enabled lane change of the benefiting vehicle are slower lead vehicles or that the lane ends.

V+: The scenario group of cooperation through increased velocity is split into two subgroups. The first one, hereinafter referred to as $V+$, is characterized by the cooperating vehicle staying in its lane. In this group, the acceleration opens a gap behind the vehicle, which allows a faster following vehicle to have more time for a lane change, enabling it to maintain greater speed (see example scenario). In the second sub group ($V+$ LCR), the acceleration is followed by a lane change to the right. According to the pure LCR scenarios, the goal of the cooperating vehicle's action is to let a faster vehicle pass. Due to a vehicle on the right lane blocking an immediate lane change, the cooperating vehicle accelerates to overtake the blocking vehicle and to clear the lane as fast as possible to prevent its follower from losing speed. The main difference between the $V+$ LCR and the LCR group lies in the definition of the cooperative action. Since in the LCR group the cooperative behavior planner performs the LCR in contrast to the reference model, the LCR is the main cooperative action. In the $V+$ LCR group, both behavior models perform the LCR, and the cooperative action lies in the acceleration to clear the lane earlier. The combination of $V+$ with a LCL is irrelevant in the computed cooperative solutions.

V-: The $V-$ scenarios are also split into two subgroups. Similar to the $V+$ group, the distinction can be made based on future lane changes of the cooperating vehicle. In the first subgroup ($V-$), the cooperating vehicle decelerates without changing lanes afterward. The deceleration opens a gap in

front of the vehicle that is used by other vehicles to merge in from the left or the right lane. In the second subgroup (V− LCL), the cooperating vehicle performs a lane change to the left after decelerating or maintaining its speed below its desired velocity. The reason why the LCL is not performed immediately is that approaching vehicles in the left lane that would have to slow down if the cooperating vehicle were to cut in front of it. By decelerating or keeping its speed low, the cooperating vehicle maintains a safe distance to its leading vehicle while letting the faster car in the left lane pass before the lane change. The combination of V− with a LCR is irrelevant in the computed cooperative solutions.

Figure 7 shows the relative frequency of the cooperative actions for straight and merge scenarios. The LCL is the most prevalently performed action in merge scenarios (56.2%), followed by the V− LCL (18.8%), the V− (12.5%) and the LCR (12.5%). The V+ and V+ LCR actions do not occur in merge scenarios. The scenarios with straight lanes are dominated by the LCR action (85.8%), whereas V− (4.7%), V+ (3.4%), V+ LCR (3.2%), LCL (2.4%) and V− LCL (0.5%) follow at a great distance. The extraction tool provided on www.github.com/TUMFTM/test_scenarios_cooperation includes a tag with the main cooperative action for each scenario in order to analyze the performance of cooperative behavior planners with respect to the underlying cooperative mechanisms.

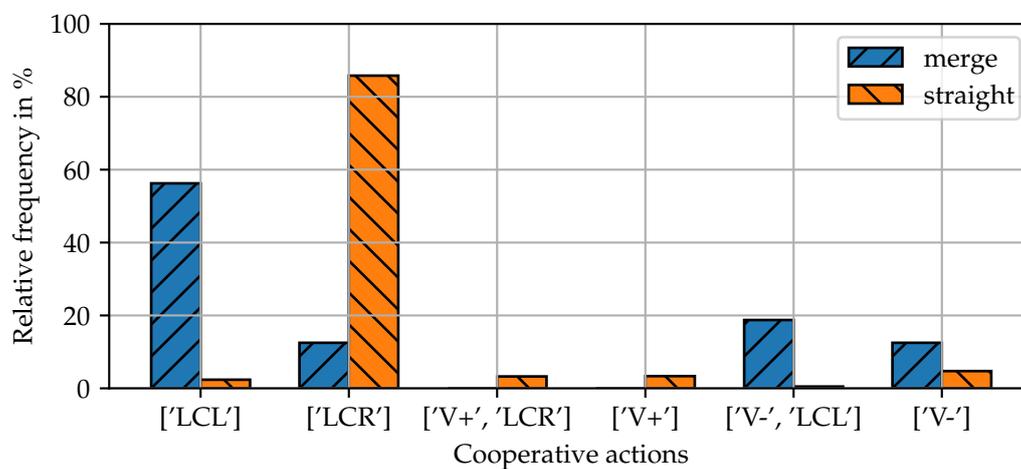


Figure 7. Relative frequency of cooperative actions.

4. Discussion

The intention of the presented work is to contribute to the development of future AV that are not only safe, but can also move cooperatively in road traffic. Therefore, we presented a method for extracting test scenarios from real-world traffic data that are challenging from the perspective of cooperative behavior. By applying this method to a highway traffic dataset, we created a catalog of test scenarios that allows comparison and benchmarking of different cooperative maneuver planning algorithms. Since the used highD dataset [29] is only accessible after permission, the scenario catalog cannot be shared directly. Instead, we provide a Python tool that extracts the scenarios of the test catalog by their recording- and vehicle IDs, provided there is access to the root dataset.

In contrast to the state-of-the-art of test scenario generation for AV, which so far addressed the topics of safety evaluation and driving task completion of single vehicles, the presented work aims to test an AV's ability of cooperatively interacting with other road vehicles. This requires test scenarios that demand cooperative behavior of at least one of the involved vehicles. The presented scenario extraction approach can select scenarios with this characteristic, which represents the main contribution to the state-of-the-art of AV test scenario generation.

Furthermore, the presented results extend the state-of-the-art for test scenarios for cooperative maneuver planning, which has so far consisted of artificially generated scenarios, by a data driven approach. It was shown that, due to the presented selection process of scenarios, a condensed

representation of cooperative situations in road traffic is possible. Since the data set created with this method contains a large number of scenarios that are difficult to think up, but which occur in real traffic, it can also be used in the development process of new behavior planners.

The extracted scenarios are clustered according to the most cooperative action that is performed by one of the vehicles in the scenario. Analyzing these scenario clusters shows that the exact sequence of the scenarios varies, but that the cooperative mechanisms are the same within each cluster. Therefore, the performance evaluation of a tested cooperative behavior planner allows conclusions to be drawn about the situations in which the planner works well and in which there is room for improvement.

One limitation of the test scenario catalog is that the amount of data in the root dataset is limited, which results in a small number of two-lane and especially three-lane merge scenarios. Therefore, the result of these roadway types could be less representative compared to the three-lane scenarios. This can only be overcome by a larger amount of data. Since the presented method is independent from the root dataset, it could be applied to a larger dataset released in the future. In case of a much larger dataset, it might not be applicable to use all extracted scenarios for testing. Therefore, the selection of representative test scenarios as well as the exclusion of similar ones will be a topic of future work.

Another point of discussion is whether a foresightful planning automated vehicle ended up in difficult situations like the ones in the dataset or if it would be able to avoid them beforehand. In the second case, the test scenarios, which are derived from human-operated traffic, would not be suitable for AV. Since a traffic situation arises as a result of the actions of all involved vehicles, and the surrounding vehicles of an automated vehicle will be mainly driven by humans in the near future, a behavior planner for AV must also be able to manage today's difficult traffic situations.

The solution of some scenarios of the global planner show non-humanlike behavior, such as creating gaps behind a vehicle, e.g., when vehicle 1 performs in the exemplary scenario in Section 3. These types of actions show that an automated vehicle could be capable of cooperating better than a human driver due to rational maneuver planning and comprehensive sensor information. However, human subject research must show the extent to which AV driving behavior may deviate from human driving behavior in order to still be accepted by human passengers.

The test mode of the methodology presented in this paper intends that there is no ego perspective in a scenario. Instead, the behavior model being tested is applied to all vehicles of the scenario independently. This approach ensures that the behavior model not only behaves intelligently when it is in a situation where it needs the cooperation of others, but also recognizes when other vehicles need help from the ego vehicle. However, one disadvantage to this approach is that the behavior model is only confronted with driving behavior similar to its own. To test the robustness in the interaction with uncooperative, non-planning agents, individual vehicles could be replaced by a reactive driver model in a future work.

A further limitation of the generated scenarios is the maximum number of vehicles, which is set to four. With an increasing number of vehicles, the tree of possible states grows exponentially (see Equation (2)), which leads to strongly increasing computing times. Therefore, it would not be reasonable to compute a dataset with more involved vehicles. In order to test a behavior model with more surrounding vehicles, a large-scale traffic simulation would be more appropriate. However, a large-scale traffic simulation does not offer a fully cooperative solution as a comparison and involves the computation of a large number of situations that do not require cooperative behavior, which makes testing more inefficient. On the other hand, a large-scale traffic simulation would offer the possibility of analyzing the effect of cooperatively behaving AV throughout the entire traffic flow and can therefore be considered as a complementary approach to scenario-based tests in future work.

Author Contributions: Conceptualization, C.K.; methodology, C.K.; software, C.K.; writing—original draft preparation, C.K.; writing—review and editing, C.K. and F.D.; supervision, F.D.; project administration, F.D.; funding acquisition, F.D. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the German Federal Ministry for Economic Affairs and Energy within the research project IMAGinE (Intelligent Maneuver Automation—cooperative hazard avoidance in real time).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

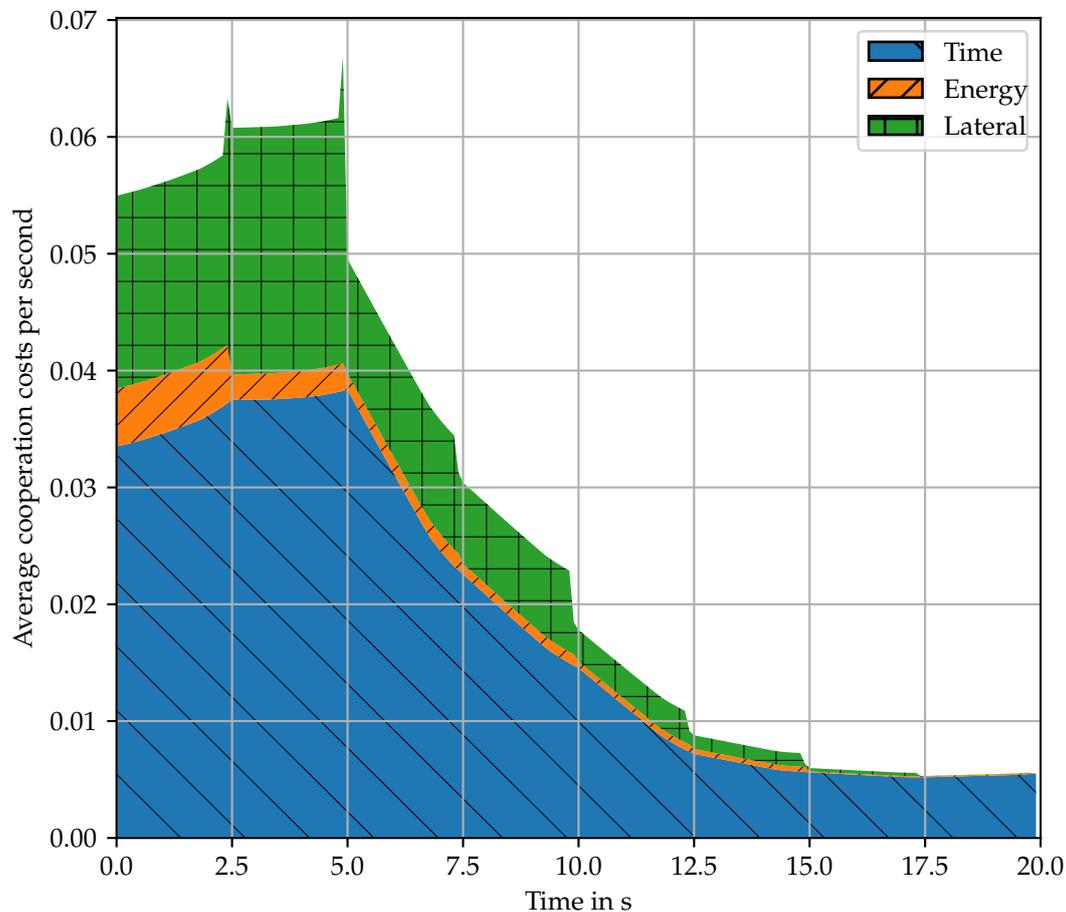


Figure A1. Costs over time for global planner.

References

1. Lenz, D.; Kessler, T.; Knoll, A. Tactical Cooperative Planning for Autonomous Highway Driving using Monte-Carlo Tree Search. In Proceedings of the IEEE Intelligent Vehicles Symposium, Gothenburg, Sweden, 19–22 June 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 447–453.
2. Fisac, J.F.; Bronstein, E.; Stefansson, E.; Sadigh, D.; Sastry, S.S.; Dragan, A.D. Hierarchical Game-Theoretic Planning for Autonomous Vehicles. In Proceedings of the International Conference on Robotics and Automation (ICRA), Montreal, QC, Canada, 20–24 May 2019; pp. 9590–9596.
3. Kurzer, K.; Engelhorn, F.; Zöllner, J.M. Decentralized Cooperative Planning for Automated Vehicles with Continuous Monte Carlo Tree Search. In Proceedings of the IEEE Intelligent Transportation Systems Conference, Maui, HI, USA, 4–7 November 2018; IEEE: Piscataway, NJ, USA, 2018.
4. Kurzer, K.; Zhou, C.; Zöllner, J.M. Decentralized Cooperative Planning for Automated Vehicles with Hierarchical Monte Carlo Tree Search. In Proceedings of the IEEE Intelligent Vehicles Symposium, Changshu, China, 26–30 June 2018; IEEE: Piscataway, NJ, USA, 2018.
5. Jenkins, I.R.; Gee, L.O.; Knauss, A.; Yin, H.; Schroeder, J. Accident Scenario Generation with Recurrent Neural Networks. In Proceedings of the International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 4–7 November 2018; pp. 3340–3345. [[CrossRef](#)]
6. Klueck, F.; Li, Y.; Nica, M.; Tao, J.; Wotawa, F. Using Ontologies for Test Suites Generation for Automated and Autonomous Driving Functions. In Proceedings of the IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW), Memphis, TN, USA, 15–18 October 2018; pp. 118–123.

7. Mullins, G.E.; Stankiewicz, P.G.; Gupta, S.K. Automated generation of diverse and challenging scenarios for test and evaluation of autonomous vehicles. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), Singapore, 29 May–3 June 2017; pp. 1443–1450. [[CrossRef](#)]
8. Rocklage, E.; Kraft, H.; Karatas, A.; Seewig, J. Automated scenario generation for regression testing of autonomous vehicles. In Proceedings of the IEEE International Conference on Intelligent Transportation Systems (ITSC), Yokohama, Japan, 16–19 October 2017; pp. 476–483. [[CrossRef](#)]
9. Feng, S.; Feng, Y.; Yu, C.; Zhang, Y.; Liu, H.X. Testing Scenario Library Generation for Connected and Automated Vehicles, Part I: Methodology. *IEEE Trans. Intell. Transp. Syst.* **2020**, 1–10. [[CrossRef](#)]
10. Feng, S.; Feng, Y.; Sun, H.; Bao, S.; Zhang, Y.; Liu, H.X. Testing Scenario Library Generation for Connected and Automated Vehicles, Part II: Case Studies. *IEEE Trans. Intell. Transp. Syst.* **2020**, 1–13. [[CrossRef](#)]
11. Feng, S.; Feng, Y.; Sun, H.; Zhang, Y.; Liu, H.X. Testing Scenario Library Generation for Connected and Automated Vehicles: An Adaptive Framework. *IEEE Trans. Intell. Transp. Syst.* **2020**, 1–10. [[CrossRef](#)]
12. Althoff, M.; Koschi, M.; Manzinger, S. CommonRoad: Composible benchmarks for motion planning on roads. In Proceedings of the IEEE Intelligent Vehicles Symposium, Los Angeles, CA, USA, 11–14 June 2017; pp. 719–726. [[CrossRef](#)]
13. Lizenberg, V.; Knapp, S.; Mannale, R.; Wendel, V.; Köster, F. Simulationsbasierte Bewertungs- und Vergleichsmethodik für Abstimmungsverfahren in kooperativen Fahrfunktionen. *AAET Autom. Vernetztes Fahr.* **2019**, *20*, 49–65.
14. Hallerbach, S.; Xia, Y.; Eberle, U.; Koester, F. Simulation-Based Identification of Critical Scenarios for Cooperative and Automated Vehicles. *SAE Int. J. Connect. Autom. Veh.* **2018**, *1*, 93–106. [[CrossRef](#)]
15. Bailey, C.A. *A Guide to Qualitative Field Research*; Pine Forge Press: London, UK; Thousand Oaks, CA, USA; New Delhi, India, 2007.
16. Himme, A. Gütekriterien der Messung: Reliabilität, Validität und Generalisierbarkeit. In *Methodik der Empirischen Forschung*; Albers, S., Klapper, D., Konradt, U., Walter, A., Wolf, J., Eds.; Gabler: Wiesbaden, Germany, 2009; pp. 485–500.
17. Knies, C.; Fank, J.; Diermeyer, F. How to Measure Cooperation? Cost Functions for Cooperative Maneuver Planning on Highways. In Proceedings of the IEEE Intelligent Transportation Systems Conference (ITSC), Auckland, New Zealand, 27–30 October 2019; pp. 2483–2489. [[CrossRef](#)]
18. Düring, M.; Pascheka, P. Cooperative decentralized decision-making for conflict resolution among autonomous agents. In Proceedings of the IEEE International Symposium on Innovations in Intelligent Systems and Applications, Alberobello, Italy, 23–25 June 2014; IEEE: Piscataway, NJ, USA, 2014.
19. Hayward, J.C. Near Misses as a Measure of Safety at Urban Intersections, Master’s Thesis, Pennsylvania State University, University Park, PA, USA, 1971.
20. Kesting, A.; Treiber, M.; Helbing, D. General Lane-Changing Model MOBIL for Car-Following Models. *Transp. Res. Rec. J. Transp. Res. Board* **2007**, *1999*, 86–94. [[CrossRef](#)]
21. Yang, M.; Wang, X. Modeling Lane Change Gap Acceptance and Duration Using Shanghai Naturalistic Driving Data. In Proceedings of the CICTP, Beijing, China, 5–8 July 2018; American Society of Civil Engineers: Reston, VA, USA, 2018.
22. Moridpour, S.; Sarvi, M.; Rose, G. Modeling the Lane-Changing Execution of Multiclass Vehicles under Heavy Traffic Conditions. *Transp. Res. Rec. J. Transp. Res. Board* **2010**, *2161*, 11–19. [[CrossRef](#)]
23. Merala, R.; White, K. *Tractor Semitrailer Left Turns and Lane Changes*; SAE Technical Paper Series; SAE International: Warrendale, PA, USA, 2010. [[CrossRef](#)]
24. Hamdar, S. Driver Behavior Modeling. In *Handbook of Intelligent Vehicles*; Springer Reference; Eskandarian, A., Ed.; Springer: London, UK, 2012; pp. 537–558.
25. Treiber, M.; Kesting, A. Car-Following Models Based on Driving Strategies. In *Traffic Flow Dynamics*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 181–204.
26. Kesting, A.; Treiber, M.; Helbing, D. Enhanced Intelligent Driver Model to Access the Impact of Driving Strategies on Traffic Capacity. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **2010**, *368*, 4585–4605. [[CrossRef](#)]
27. Knies, C.; Hermansdorfer, L.; Diermeyer, F. Cooperative Maneuver Planning for Highway Traffic Scenarios based on Monte-Carlo Tree Search. In *AAET—Automatisiertes & Vernetztes Fahren*; ITS Mobility e.V.: Brunswick, Germany, 2019; Volume 20, pp. 10–25.

28. Russell, S.; Norvig, P. *Artificial Intelligence: A Modern Approach, Global Edition*; Pearson Education Limited: Harlow, UK, 2016.
29. Krajewski, R.; Bock, J.; Kloeker, L.; Eckstein, L. The highD Dataset: A Drone Dataset of Naturalistic Vehicle Trajectories on German Highways for Validation of Highly Automated Driving Systems. In Proceedings of the International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 4–7 November 2018; pp. 2118–2125. [[CrossRef](#)]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).