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

Probabilistic Situation Assessment Framework for Multiple Interacting Traffic Participants in Generic Traffic Scenes

Stefan Klingelschmitt, Florian Damerow, Volker Willert, Julian Eggert

2016

Preprint:

This is an accepted article published in Intelligent Vehicles Symposium (IV). The final authenticated version is available online at: https://doi.org/[DOI not available]

2016 IEEE Intelligent Vehicles Symposium (IV)
Gothenburg, Sweden, June 19-22, 2016

Probabilistic Situation Assessment Framework for Multiple, Interacting Traffic Participants in Generic Traffic Scenes

Stefan Klingelschmitt[1], Florian Damerow[1], Volker Willert[1] and Julian Eggert[2]

***Abstract*— Situation recognition is a prerequisite for many advanced driver assistance systems as well as for partially and fully automated vehicles. Current situation recognition approaches focus mainly on estimating maneuvers of single scene entities. However, assessing multiple, possibly interacting, traffic participants simultaneously is crucial in complex traffic scenes and has hardly been investigated. Considering the variability and combinatorics of such scenarios, having specialized situation recognition systems covering each case directly is unrealistic. In this paper, we present a flexible framework for assessing generic traffic scenes with multiple interacting traffic participants. It is able to construct a fully interaction-respecting probabilistic situation assessment, while relying on reusable state-of-the-art single-entity-based maneuver predictions. The benefits and applicability are presented on a real-world data set. The evaluation indicates that the approach is not only able to reconstruct underlying interdependent probability distributions; it outperforms specially designed models, due to the reduced complexities of the single-entity-based recognition models.**

I. INTRODUCTION

An increased interest by academia, automotive manufacturers as well as the general public in Advanced Driver Assistance Systems (ADAS), partially automated vehicles up to fully autonomous vehicles can currently be observed. This is due to the reason that these techniques are expected to relieve the stress on the driver leading to more comfort while simultaneously reducing the number of accidents [1]. This is typically being realized by taking the current observed traffic scene, predicting it several seconds into the future and evaluate its criticality. Based on this assessment warnings or behavioral recommendations can be given up to a level where complete autonomous reactions are provided.

In order for such systems to account for the two goals (minimise stress, increase safety) they need to be able to deal with a large variety of possibly encountered scenarios with varying degrees of complexity. However, correctly assessing complex traffic scenarios is still a challenging task, as the prediction of traffic participants' future trajectories are associated with many uncertainties.

A common approach to address this problem is to partition the evolution of traffic scenes into discrete prototypical situation classes. Such situation classes combine a subset of traffic participants, specifying their typical associated spatio-temporal behavior patterns, e.g. like a car deceleration in order to give another car right of way before making a right turn at an intersection. Based on these prototypical situation classes, specific prediction models can be used to estimate future kinematic states of the traffic participants with higher precision. Such a decomposition requires a potent situation recognition system.

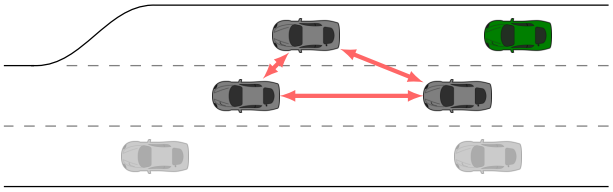

Fig. 1: Interaction dominated highway scenario.

A. Related Work

Systems for recognizing the current situation in inner city as well as highway scenarios have been widely discussed in literature. The authors of [2], [3], [4], [5] and [6] focus on predicting maneuvers at intersections. In [4] and [6] probabilistic maneuver predictions are performed using Gaussian Processes, while [5] are using Bayesian networks. A probabilistic approach capable of dealing with large amount of training data is presented in [3]. The authors of [2] use a Hidden Markov Model for their maneuver predictions at general intersections. The aforementioned approaches are reliable when the upcoming maneuver is solely influenced by the driver's intention. However, in the presence of other effecting traffic participants (e.g. yielding scenarios, lane changes in dense traffic) or a restricting context (e.g. traffic lights, occluded lanes) their performance will significantly suffer as the execution of a driver's intended maneuver will be strongly influenced by these factors.

Hence, approaches have been presented taking also the influence of possibly effecting traffic participants and context information into account. In [7] an approach is introduced which adds the influence of traffic light information and preceding vehicles to a solely kinematic feature based maneuver prediction leading to improved recognition rates. The approaches presented in [8] and [9] also incorporate the influence of preceding vehicles. The authors of [10] focus on determining the different characteristics of a right turn maneuver in the presence of an effecting vehicle with priority. In a similar setting, the authors of [11] propose a learning concept based on case-based reasoning for distinguishing between turn and stop maneuvers. The authors of [12] present an approach for determining various maneuvers

[1]S. Klingelschmitt, F. Damerow and V. Willert are with Control Methods and Robotics Lab, Technical University of Darmstadt, D-64283 Darmstadt, Germany stefan.klingelschmitt, fdamerow, vwillert@rmr.tu-darmstadt.de

[2]J. Eggert is with Honda Research Institute Europe GmbH, D-63073 Offenbach am Main, Germany jeggert@honda-ri.de

978-1-5090-1820-8/16/$31.00 ©2016 IEEE

in highways (e.g. lane change to the left or right). They incorporate relational features of nearby traffic participants to model the dependencies between the subject vehicle and possibly effecting vehicles.
All these approaches incorporate possibly effecting vehicles by making use of relational features, however they do not specifically include possible behavioral options of the considered effecting traffic participants. The authors of [13] do take account for this in their prediction model for highway entering scenarios, considering explicitly the maneuver predictions of vehicles on the neighboring lane.
Nevertheless, all of these approaches focus on predicting the behaviors of single scene entities. Very few approaches have been proposed for predicting behaviors of multiple, possibly interacting traffic participants, simultaneously. In [14], dependencies are neglected by assuming that the current behavior of a traffic participant is independent of other traffic participants' past behaviors. The situation recognition systems presented [15], [16] and [17] do also assume the stochastic independence of the regarded traffic participants behavior. However, [16] and [17] try to take the interactive character into account by subsequently adjusting the situation probabilities according to resulting, separately calculated collision probabilities.
The approach of [18] does incorporate such interdependencies indirectly via the estimated *expected* behavior. However, the calculation of the expected behavior depends on vehicles with priority. The question remains how such an expected behavior can be determined if the dependencies cannot be uniquely decomposed with priority rules (e.g. highway entering scenarios).

B. Contribution

In complex scenarios, the behavior of an intelligent vehicle might not just be influenced by one effecting object at a time but rather be influenced by several.[1] Thus, possible maneuvers of several effecting traffic participants need to be predicted. A simple example of a highway entering clarifies this circumstance, shown in Fig. 1. The behavior of the green vehicle strongly depends on how the three gray cars in front are going to behave as an interacting group. However, most of the current, well studied state-of-the-art maneuver recognition systems focus on estimating maneuvers of single traffic participants. Combining the separately obtained maneuver predictions into one common behavior prediction is not straightforwardly possible, as a more probability-theoretical discussion in Section III and IV implies. Approaches capable of dealing with multiple entities, either rely on unique stochastic independencies or neglect the dependencies completely.

Therefore, we provide an approach for dealing with traffic scenarios where several traffic participants need to be assessed simultaneously. The main contribution of our work is threefold. **First**, we propose a general framework for probabilistic situation assessment for generic traffic scenes with multiple interacting participants. It is based on more universally usable models determining *complete* conditional probability distributions. The framework is based on theoretically sound and widely discussed probabilistic sampling techniques. **Second**, we extend state-of-the-art single-entity-based recognition systems for obtaining these complete conditionals. **Third**, we reformulated the problem of reconstructing the desired joint probability distributions from the complete conditional probability distributions as a polynominal equation system, replacing costly sampling techniques.

The reminder of this paper is organized as follows. In Section II we discuss general problems arising in the context of probabilistic situation assessment in generic traffic scenes. Sections III - V introduce the developed framework for reconstructing joint probabilities. Experimental results on a real-world data set are presented in Section VI, showing the benefits of our approach.

II. PROBLEM STATEMENT

In this section, the challenges of assessing situations with multiple interacting traffic participants are discussed. At first the term *situation* is defined and the necessary probability theoretical definitions are stated.

A. Situation Hypotheses Definition

Each situation hypothesis h specifies a typical prototypical spatio-temporal behavior pattern of several traffic participants, starting from the current time t for a prediction horizon of Δt, as introduced [19]. Therefore, we define one discrete random variable B_i, for each currently present dynamic scene entity $i \in 1,\dots n$. Each random variable can take on different values (events), each of them specifying one prototypical spatio-temporal behavior for the scene entity it represents (e.g. right turn, going straight, right turn and yielding, lane change, ...). All possible prototypical behavior patterns (events) for one traffic participant i form its sample space $\Omega_i = \{b_i^1, b_i^2, \dots b_i^{|\Omega_i|}\}$. A single event is denoted by, e.g. b_i^1 where the subscript i indicates the random variable membership while the superscript distinguishes specific events. It should be noted that the sample space contains different events for each scene, as different road-layouts, traffic regulations and different present traffic participants lead to different possible events.
The discrete probability distribution for the random variable B_i is given by

$$p(\{b_i \in \Omega_i : B_i = b_i\}). \tag{1}$$

In the reminder of this paper the formal notation in (1) is abbreviated with the common lower case notation $p(b_i)$.

One way to define a situation hypothesis h is to take one possible event combination regarding all random variables. Hence, the set of possible situation hypotheses $\mathcal{H}$ is determined by

$$\mathcal{H} = \Omega_1 \times \Omega_2 \times \dots \times \Omega_n, \tag{2}$$

which leads to $m = \prod_{i=1}^{n} |\Omega_i|$ possible situation hypotheses $h \in \mathcal{H}$.

[1] The subject vehicle itself becomes an object of the situation assessment, if the own intended future behavior is not known (e.g. its driven by a human).

B. Problems Arising With Generic Traffic Scenes

For a probabilistic situation assessment we are interested in the probability distribution $p(h|\mathbf{x})$, which can be rewritten as $p(b_1, \ldots, b_n|\mathbf{x})$, given the relevant scene evidence $\mathbf{x}$. The scene evidence generally contains the relevant spatial layout and context information (e.g. traffic light states) as well as state information of present dynamic scene entities and their relation between each other.

Probabilistic situation recognition can be understood as a probabilistic classification problem, where each possible situation hypothesis corresponds to one class of the classification algorithm. Accordingly, the posterior class probabilities correspond to the desired probability distribution $p(b_1, \ldots, b_n|\mathbf{x})$. For scenes with multiple possibly interacting traffic participants, the number of possible situation hypotheses $|\mathcal{H}|$ increases rapidly and thus the dimensionality of the classifier's target space increases, which in turn increases the complexity of the classification algorithm drastically.
Furthermore, the variability of traffic scenes has to be taken into account. In general, each encountered traffic scene leads to a different set of possible situation hypotheses (e.g. different number of traffic participants with different possible behavior patterns). A classifier that is designed for one specific set of hypotheses (scenario) cannot be used for another set of situation hypotheses. Therefore, specialized systems for each possibly encountered scenario would be needed.

The addressed problems point out that a direct approximation of the joint probability distribution over multiple interacting traffic participants in generic traffic scenes is hardly feasible. Therefore, our approach decomposes the problem by making fully interaction-respecting maneuver predictions for each traffic participant separately. Subsequently these are used for reconstructing $p(b_1, \ldots, b_n|\mathbf{x})$. The decomposition is of great benefit, as the high target space dimensionality can be partitioned into several partial classifiers with reduced complexity. These partial classifiers can be more easily reused in other scenarios since each of them relies on fewer assumptions.

III. Interaction-respecting Single-entity Maneuver Predictions

The probabilistic single-entity-based approaches, discussed in Section I-A, are in fact estimating conditional probability distributions of the form

$$p(b_i|\mathbf{x}), \tag{3}$$

where the behavior b_i is conditioned on the currently observed scene evidence $\mathbf{x}$. This scene evidence can also include relational features, which is a direct way to reflect the influence of other traffic participants on participant i's behavior.
In general, the behavior of subject i will not only dependent on the current states of effecting traffic participants but also on their possible future behaviors. Therefore it is reasonable to add the behavioral alternatives of all effecting traffic participants explicitly to the conditional of (3), leading to

$$p(b_i|\mathbf{x}, b_1, \ldots, b_{i-1}, b_{i+1}, \ldots b_n). \tag{4}$$

The obtained fully interaction-respecting conditional probability distributions are called *complete* conditionals.

The concrete behavior information of all effecting traffic participants leads to a better discriminability of the behavioral options of traffic participant i. Furthermore, the target space dimensionality is reduced to $|\Omega_i|$, as $n-1$ random variables are shifted to the input space of the classification algorithm.

In the next step the question of how to construct a joint probability distribution from the proposed complete conditionals is addressed since a straightforward multiplication

$$\begin{aligned} p(b_1, b_2, ..., b_n|\mathbf{x}) \neq & p(b_1|\mathbf{x}, b_2, \ldots, b_n) \\ & p(b_2|\mathbf{x}, b_1, b_3, \ldots, b_n) \\ & \ldots \\ & p(b_n|\mathbf{x}, b_1, \ldots, b_{n-1}), \end{aligned} \tag{5}$$

does not provide a valid decomposition.

IV. Reconstructing Joint Probabilities using Complete Conditionals

An interaction-respecting decomposition of the joint probability distribution can be obtained by recursively applying the chain rule of probability, e.g.

$$\begin{aligned} p(b_1, b_2, \ldots, b_n|\mathbf{x}) = & p(b_1|\mathbf{x}, b_2, \ldots, b_n) p(b_2|\mathbf{x}, b_3, \ldots, b_n) \\ & \ldots p(b_{n-1}|\mathbf{x}, b_n) p(b_n|\mathbf{x}). \end{aligned} \tag{6}$$

It should be noted that that $n!$ different decompositions are possible. The obtained decomposition is characterized by the decreasing number of regarded dependent entities for the conditional probability distributions

$$p(b_2|\mathbf{x}, b_3, \ldots b_n), \ldots, p(b_{n-1}|\mathbf{x}, b_n), p(b_n|\mathbf{x}). \tag{7}$$

If the dependencies of multiple present scene entities can in fact be uniquely decomposed using the chain rule of probability (like e.g. being done in Bayesian networks), these probability distributions can be directly estimated using single-entity based behavior predictions.
Howsoever, in general a unique complete interaction-aware decomposition will not always be possible. Therefore, we refer to the distributions in (7) as marginal conditional distributions, since for the general interdependent case they correspond to marginals, e.g. pointed out by

$$p(b_2|\mathbf{x}, b_3, \ldots, b_n) = \sum_{b_1} p(b_2, b_1|\mathbf{x}, b_3, \ldots, b_n). \tag{8}$$

The question of how to determine such marginal conditional probabilities leads to the same problems as already discussed in II-B.

Therefore, we propose two systematic procedures for correctly inferring the desired joint probability distribution $p(b_1, b_2, ..., b_n|\mathbf{x})$ based solely on the complete conditionals.

A. Reconstructing $p(b_1, \ldots, b_n|\mathbf{x})$ using Gibbs Sampling

In many practical problems exactly inferring a probability distribution of interest is not directly possible or intractable. A common method for obtaining information about complicated distributions is to use sampling methods [20]. However, due to the complexity of many target distributions drawing samples from such distributions directly is, again, not possible. A popular sampling method, the *Gibbs sampling*, overcomes this problem by successively drawing samples from the complete conditional distributions. Gibbs sampling starts with a randomly initialized value for each random variable b_i. Starting from this initial state, a complete cycle of samples for each random variable is drawn for $\tau = 1, ..., T$ overall samples:

$$\begin{aligned} b_1^{(\tau+1)} &\sim p(b_1|\mathbf{x}, b_2^{(\tau)}, ..., b_n^{(\tau)}) \\ b_2^{(\tau+1)} &\sim p(b_2|\mathbf{x}, b_1^{(\tau+1)}, b_3^{(\tau)} ..., b_n^{(\tau)}) \\ &\vdots \\ b_{n-1}^{(\tau+1)} &\sim p(b_{n-1}|\mathbf{x}, b_1^{(\tau+1)}, b_2^{(\tau+1)} ..., b_{n-2}^{(\tau+1)}, b_n^{(\tau)}) \\ b_n^{(\tau+1)} &\sim p(b_n|\mathbf{x}, b_1^{(\tau+1)}, b_2^{(\tau+1)} ..., b_{n-1}^{(\tau+1)}), \end{aligned} \tag{9}$$

as stated in [20]. The resulting Markov chain can be used to calculate the sufficient statistics of the joint probability distribution $p(b_1, \ldots, b_n|\mathbf{x})$.

The theoretical background of Gibbs sampling is based on the *Hammersley-Clifford Theorem*, which states under which conditions a joint distribution is completely characterized by its conditional distributions [21]. For practical implementations these are satisfied as long as neither of the conditionals or the target joint distribution is anywhere zero. Furthermore, we assume that the complete conditional distributions, obtained by the single-entity-based maneuver classifiers, are in fact the correct complete conditionals of the desired joint distribution.

B. Direct Calculation of $p(b_1, \ldots, b_n|\mathbf{x})$

Convergence of sampling techniques, especially when considering large sample spaces, can be time consuming. Additionally, determining the degree of sufficient convergence towards the stationary distribution is a non trivial task. Therefore, we present a general approach for substituting the marginal conditional probabilities of the general joint probability decomposition in (6). The approach is based on alternating the chain rule of probability and *remarginalizing* over the resulting marginal cond. probabilities.

The approach is outlined on a simple example, consisting of two random variables b_1, b_2. First, the chain rule of probability is used to obtain

$$p(b_1, b_2|\mathbf{x}) = p(b_1|\mathbf{x}, b_2)p(b_2|\mathbf{x}), \tag{10}$$

already containing a complete cond. probability $p(b_1|\mathbf{x}, b_2)$. In order to substitute the marginal cond. probability $p(b_2|\mathbf{x})$ using the remaining complete cond. probability, we remarginalize $p(b_2|\mathbf{x})$ over b_1, resulting in

$$p(b_1, b_2|\mathbf{x}) = p(b_1|\mathbf{x}, b_2) \sum_{b_1'} p(b_1', b_2|\mathbf{x}). \tag{11}$$

Based on this, another possible decomposition of the joint probability distribution can be used leading to

$$p(b_1, b_2|\mathbf{x}) = p(b_1|\mathbf{x}, b_2) \sum_{b_1'} p(b_2|\mathbf{x}, b_1')p(b_1'|\mathbf{x}), \tag{12}$$

where the remaining complete cond. probability distribution $p(b_2|\mathbf{x}, b_1')$ is incorporated, while additionally including $p(b_1'|\mathbf{x})$. Remarginalizing over b_2 results in

$$p(b_1, b_2|\mathbf{x}) = p(b_1|\mathbf{x}, b_2) \sum_{b_1'} p(b_2|\mathbf{x}, b_1') \sum_{b_2'} p(b_1', b_2'|\mathbf{x}), \tag{13}$$

containing only cond. probabilities and the desired joint probability itself. Equating (10) with (13) reveals the substitution for the marginal conditional $p(b_2|\mathbf{x})$

$$p(b_2|\mathbf{x}) = \sum_{b_1'} p(b_2|\mathbf{x}, b_1') \sum_{b_2'} p(b_1', b_2'|\mathbf{x}). \tag{14}$$

In the same way a joint probability distribution consisting of three random variables can be decomposed, starting with

$$p(b_1, b_2, b_3|\mathbf{x}) = p(b_1|\mathbf{x}, b_2, b_3)p(b_2|\mathbf{x}, b_3)p(b_3|\mathbf{x}). \tag{15}$$

Applying the same scheme of alternating the chain rule of probability and remarginalizing, as applied in (10) - (13), leads to

$$\begin{aligned} p(b_1, b_2, b_3|\mathbf{x}) = p(b_1|\mathbf{x}, b_2, b_3) \\ \sum_{b_1'} p(b_2|\mathbf{x}, b_1', b_3) \frac{\sum_{b_2''} p(b_1', b_2'', b_3|\mathbf{x})}{\sum_{b_1'', b_2''} p(b_1'', b_2'', b_3|\mathbf{x})} \\ \sum_{b_1', b_2'} p(b_3|\mathbf{x}, b_1', b_2') \frac{\sum_{b_3''} p(b_1', b_2', b_3''|\mathbf{x})}{\sum_{b_1'', b_2'', b_3''} p(b_1'', b_2'', b_3''|\mathbf{x})}, \end{aligned} \tag{16}$$

where the cond. marginals of (15) are equal to

$$\begin{aligned} p(b_2|\mathbf{x}, b_3) &= \sum_{b_1'} p(b_2|\mathbf{x}, b_1', b_3) \frac{\sum_{b_2''} p(b_1', b_2'', b_3|\mathbf{x})}{\sum_{b_1'', b_2''} p(b_1'', b_2'', b_3|\mathbf{x})}, \\ p(b_3|\mathbf{x}) &= \sum_{b_1', b_2'} p(b_3|\mathbf{x}, b_1', b_2') \frac{\sum_{b_3''} p(b_1', b_2', b_3''|\mathbf{x})}{\sum_{b_1'', b_2'', b_3''} p(b_1'', b_2'', b_3''|\mathbf{x})}. \end{aligned}$$

The outlined procedure for two and three random variables can be generalized for an arbitrary number of n random variables. The general construction rule for substituting any marginal cond. probability reads as follows

$$\begin{aligned} p(b_i|\mathbf{x}, b_{i+1}, \ldots, b_n) = \\ \sum_{b_1', \ldots b_{i-1}'} p(b_i|\mathbf{x}, b_1', \ldots b_{i-1}', b_{i+1}, \ldots, b_n) \\ \frac{\sum_{b_i''} p(b_1', \ldots b_{i-1}', b_i'', b_{i+1}, \ldots, b_n|\mathbf{x})}{\sum_{b_1'', \ldots, b_i''} p(b_1'', \ldots b_i'', b_{i+1}, \ldots, b_n|\mathbf{x})}. \end{aligned} \tag{17}$$

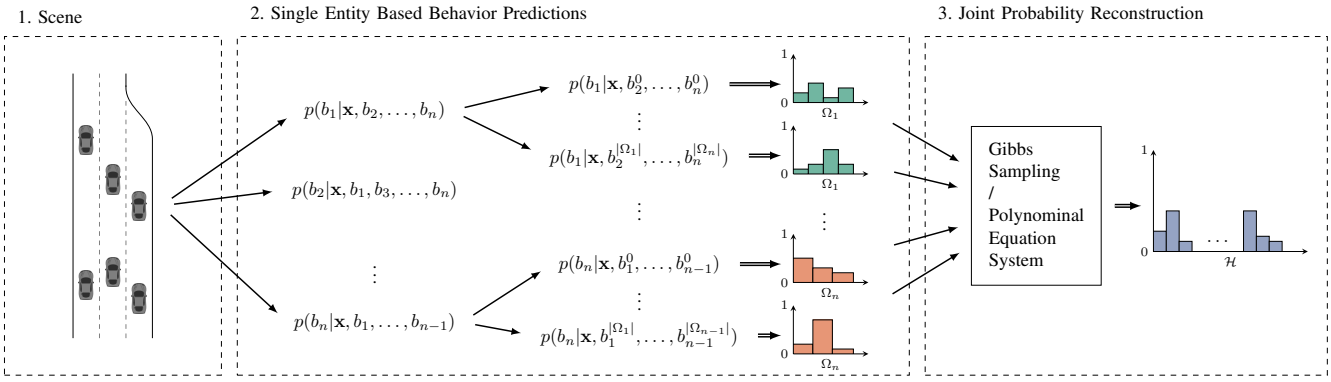

Fig. 2: Overview of the proposed framework, detailed in Section V.

Using this substitution, $p(b_1, \ldots, b_n|\mathbf{x})$ can be represented using only complete conditionals and itself. For each hypothesis in $\mathcal{H}$ we can instantiate one equation based on (6) and (17), where each single probability of the discrete joint distribution $p(b_1, \ldots, b_n|\mathbf{x})$ is an unknown. This leads to $|\mathcal{H}|$ polynominal equations with $|\mathcal{H}|$ unknown probabilities. The resulting polynominal equation system can be solved using numerical solvers, while regarding two additional constraints

$$p(h|\mathbf{x}) > 0, \forall h \in \mathcal{H}, \tag{18}$$

$$\sum_{h \in \mathcal{H}} p(h|\mathbf{x}) = 1. \tag{19}$$

The first constraint ensures the existence of the substitution rule in (17) and is equivalent to the prerequisite of Gibbs sampling stated in Section IV-A. The second constraint assures that the resulting numerical solution represents a probability distribution.

V. FRAMEWORK OVERVIEW

This section gives an overview and summary on how to perform a probabilistic situation assessment in generic traffic scenes using the proposed framework. The complete procedure is illustrated in Fig. 2. **First**, the set of possible situation hypotheses $\mathcal{H}$ is instantiated according to Section II-A. **Second**, each dynamic scene entity i is assessed separately. A single-entity-based prediction model is selected for obtaining the respective complete conditionals $p(b_i|\mathbf{x}, b_1, \ldots, b_{i-1}, b_{i+1}, \ldots b_n)$. One distribution for all possible event combinations $\prod_{j \in \{1,\ldots,n|j \neq i\}} \Omega_j$ in the conditional of the corresponding complete conditional distribution is determined. **Third**, based on the obtained complete conditionals the desired joint probability distribution can be reconstructed either by using the discussed sampling technique in Section IV-A or directly based on the polynominal equation system proposed in IV-B. The result is the fully interaction-aware joint probability distribution over all situation hypotheses.

It should be noted that if a random variable b_i is stochastically independent of one or more random variables, the complete conditional of b_i from (9) or (17), can be substituted with the conditional probability distribution specified by its Markov blanket.

VI. EXPERIMENTAL RESULTS

Lane changes are cooperative maneuvers that mainly occur on highways. This cooperative character gets reinforced on highway entering scenarios, where the vehicles are forced to perform a lane change before the on-ramp ends. Thus, we chose these scenarios to evaluate the performance of our fully interaction-respecting approach compared to a set of different independence assumptions.

The used data is extracted from the NGSIM real-world data set, recorded in 2005. It is publicly available at [22]. The Interstate-80 data is used for extracting entering scenarios in dense traffic. It contains 45 minutes of recorded vehicle trajectories.

A. Investigated Scenario

Entering scenarios involving three traffic participants are extracted: one vehicle on the on-ramp and two nearby vehicles on the most right lane, as illustrated in Fig. 3 by the three gray cars. These scenes pose an interesting scenario where three dynamic scene entities are forced to interact with each other. How these vehicles interact as a group might be interesting for several traffic participants behind or beside of this group as well as for assistance systems in vehicles which are part of the group.

According to the procedure introduced in Section II-A, we define possible situation hypotheses based on the combination of typically expected behavior patterns of each scene entity. For each vehicle on the rightmost lane (2, 3) one random variable b_2 respectively b_3 is defined. We define three expressive behavior patterns in relation to the expected longitudinal position obtained by a kinematic extrapolation

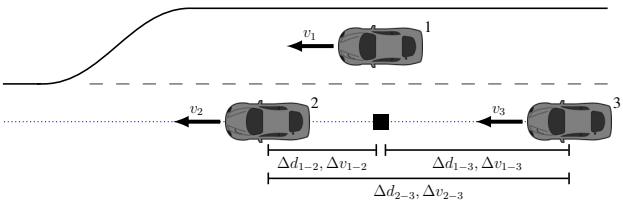

Fig. 3: Investigated scenario consistent of three scene entities. Additionally, the extracted relative features are shown.

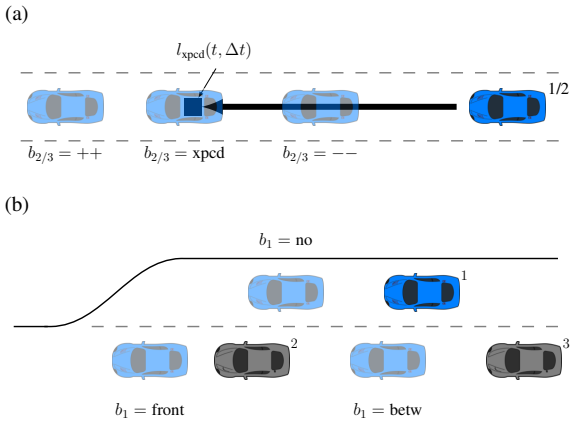

Fig. 4: (a) the three possible behavior patterns/events for vehicle 2 and 3. The kinematic extrapolated position $l_{\text{xpcd}}(t, \Delta t)$ is indicated by the square. Possible behavior patterns/events based on the actual observed $l(t + \Delta t)$ are indicated by the transparent blue vehicles. (b) the three possible events of vehicle 1. The transparent blue vehicles indicate each possibly occurring behavior pattern/event at at time $t + \Delta t$.

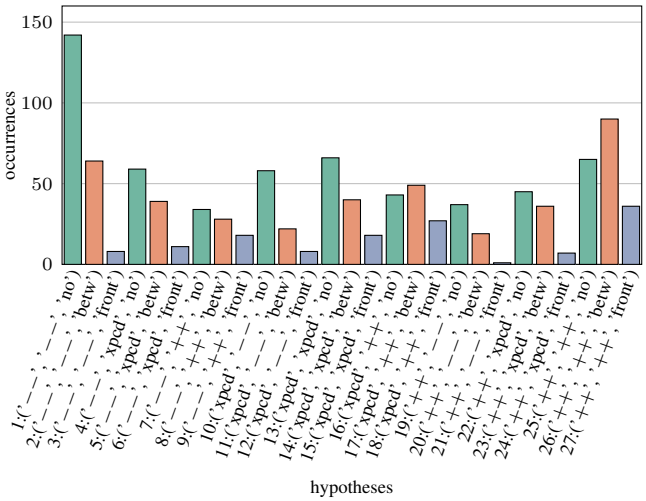

Fig. 5: Number of occurrences according to the 27 situation hypotheses. The hypotheses labels indicate one possible event combination over all three random variables in the form (b_3, b_2, b_1).

for a prediction horizon of Δt

$$l_{\text{xpcd}}(t, \Delta t) = \frac{1}{2} a(t) \Delta t^2 + v(t) \Delta t, \tag{20}$$

where $a(t)$ and $v(t)$ correspond to the currently measured acceleration and velocity at prediction time t. The typical behavior patterns are defined by comparing $l_{\text{xpcd}}(t, \Delta t)$ with the actual observed longitudinal position $l(t + \Delta t)$ at the end of the prediction horizon. If $l(t + \Delta t)$ is not within a typical car length ($\sim 4m$) around $l_{\text{xpcd}}(t, \Delta t)$ it is either a decelerated ($--$) or accelerated ($++$) behavior pattern compared to the otherwise expected behavior pattern (xpcd). The different events are illustrated in 4 (a). The sample spaces Ω_2 and Ω_3 are identically composed of the events

$$\Omega_2 = \Omega_3 = \{--, \text{xpcd}, ++\}. \tag{21}$$

Regarding vehicle 1, it is of interest in front of which vehicle (2 or 3) it is going to merge or if no merge can be observed during the prediction horizon Δt. The possible events are illustrated in Fig. 4 (b). Thus, the sample space Ω_1 is composed of

$$\Omega_1 = \{\text{betw}, \text{front}, \text{no}\}. \tag{22}$$

The set of possible hypotheses $\mathcal{H}$ is instantiated by all possible event combinations $\mathcal{H} = \Omega_1 \times \Omega_2 \times \Omega_3$ resulting in $|\mathcal{H}| = 27$ situation hypotheses.

Cases are extracted for vehicles on the on-ramp of the Interstate-80 data for a prediction horizon of $\Delta t = 5s$ before a lane change (t_{ch}) or no lane change ($t_{\text{no,ch}}$). The point in time where a vehicle 1 has completely entered the rightmost lane specifies t_{ch}. Additionally, examples for no lane change maneuvers are extracted. For these cases $t_{\text{no,ch}}$ is specified by a randomly chosen time assuring at least $t_{\text{ch}} - \Delta t > t_{\text{no,ch}}$. The states at $t = t_{\text{ch}} - \Delta t$ (respectively $t = t_{\text{no,ch}} - \Delta t$) are extracted for vehicle 1 as well as the two nearest vehicles on the rightmost lane at t. The extracted cases are categorized according to the hypotheses $\mathcal{H}$. The detailed distribution over the situation occurrences can be seen in Fig. 5.

B. Evaluated Models

We evaluate our approach for reconstructing the joint probability distribution $p(b_1, b_2, b_3|\mathbf{x})$ using the direct calculation presented in Section IV-B. A model based on three random variables has already been derived in (16) and is from now on referred to as $\mathcal{M}_{\text{recons}}$. The joint probability distribution can be obtained by solving the resulting $|\mathcal{H}| = 27$ polynominal equations. The needed complete conditional probability distributions $(p(b_1|\mathbf{x}, b_2, b_3), p(b_2|\mathbf{x}, b_1, b_3), p(b_3|\mathbf{x}, b_1, b_2))$ are learned from the extracted data and are discussed in VI-C.

In order to benchmark the proposed framework we train another model directly approximating $p(b_1, b_2, b_3|\mathbf{x})$

$$\mathcal{M}_{\text{direct}} : p(b_1, b_2, b_3|\mathbf{x}). \tag{23}$$

This model serves as reference. However, it should be noted that it is especially targeted at only the exact scenario that is investigated here. In general it suffers from the problems discussed in Section II-B.

Additionally, two models are created for reconstructing $p(b_1, b_2, b_3|\mathbf{x})$ from single-entity-based conditional distributions. They rely on specific independence assumptions regarding the three entities.
Firstly, we create a model using the chain rule of probability focusing on the most prominent dependencies

$$\mathcal{M}_{\text{decomp}} : p(b_1|\mathbf{x}, b_2, b_3) p(b_3|\mathbf{x}, b_2) p(b_2|\mathbf{x}), \tag{24}$$

assuming that vehicle 3 is independent of vehicle's 1 behavior ($b_3 \perp b_1$) since vehicle 1 is on the on-ramp. Additionally, it is assumed that vehicle 2 is independent of vehicles' 1 and 3 behaviors ($b_2 \perp b_1, b_3$) since they are located behind vehicle 2. However, theses assumptions will most likely violate the actual interaction-dominated character of the scenario.

Secondly, a model making the assumption that the scene entities behave completely independent ($b_1 \perp b_2 \perp b_3$) from

each other is evaluated

$$\mathcal{M}_{\text{indep}} : p(b_1|\mathbf{x})p(b_2|\mathbf{x})p(b_3|\mathbf{x}), \tag{25}$$

representing an even stronger assumption.

C. Learning (Conditional) Probability Distributions

For each of the conditional probability distributions present in models (16), (24) and (25) we train a probabilistic multiclass classifier, where each possible class corresponds to a possible event of the particular sample space Ω_i. As $|\Omega_i| = 3$ for $i = 1, \ldots, 3$, each classifier has 3 possible classes. The classifier for directly approximating $p(b_1, b_2, b_3|\mathbf{x})$, given by model (23), consists of $|\mathcal{H}| = 27$ classes.

We chose a logistic regression for each of the needed conditional probability distributions, as well as for the classifier directly approximating $p(b_1, b_2, b_3|\mathbf{x})$. The logistic regression is particular suitable since the obtained posterior probability distributions over the possible classes are in general well calibrated [23]. This is beneficial in terms of the actual interpretability of the probabilities as well as for the quality of the reconstruction. Besides this, it has other advantages like its training and run-time speed as well as the natural extensibility to multinomial classification problems.

The used feature vectors correspond to the information present in the conditional of the approximated probability distributions. They consist of a common part $\mathbf{x}$ representing the observable scene information. This information is identical for each trained logistic regression, ensuring the comparability of the evaluated models, regardless of the underlying independence assumptions. This means that e.g. the logistic regression for determining $p(b_2|\mathbf{x})$, assuming $b_2 \perp b_1, b_3$, nevertheless uses relative kinematic features dependent on scene entities 1 and 3. The common scene information is given by

$$\begin{aligned}\mathbf{x} = [&v_1, v_2, v_3, \Delta d_{1-2}, \Delta d_{1-3}, \Delta d_{2-3}, \\ &\Delta v_{1-2}, \Delta v_{1-3}, \Delta v_{2-3}]^T,\end{aligned} \tag{26}$$

where $\Delta d_{1-2}, \Delta d_{1-3}, \Delta d_{2-3}$ correspond to the relative longitudinal distance (i.e. the distance if they where in fact on the same lane). These can be obtained by a perpendicular projection of the vehicles' position on the rightmost lane. $\Delta v_{1-2}, \Delta v_{1-3}, \Delta v_{2-3}$ correspond to relative longitudinal velocities between scene entities. Fig. 3 illustrates the construction of the used features. [2] Acceleration information did not result in higher prediction accuracies for the investigated scenarios and was therefore not being used.

Additionally to the common present scene information $\mathbf{x}$, a categorical variable for each entity in the conditional (if present) is added representing the different behavioral options of that entity and expressing their interdependence.

In order to take the unevenly distributed situation observations from Fig. 5 into account, all models are trained and evaluated weighting the importance of each case with the inverse of its overall situation hypothesis prototype occurrence.

[2]The used features are not specific to the sensory setup of the NGSIM program. The same information could also be abstracted from more traditional setups, e. g. IMU and radar/laser detectors).

D. Evaluation

The evaluation is performed using a 4-fold stratified cross validation on the 1070 extracted observations. We chose ROC curves in order to compare our results, as these allow to take also the *false positives* into account. The *false positive rate* plays an important role for intelligent vehicle applications, as users typically get easily annoyed by false warnings. In addition, for this general evaluation we had no specific application in mind and hence no defined minimum reliability (*true positive rate*) or maximum false positive rate was specified. Thus, the ROC curve provides the most objective metric.

The resulting ROC curves for all four models, introduced in Section VI-B, can be seen in Fig. 6. The model based on our presented approach $\mathcal{M}_{\text{recons}}$ shows the best performance of all investigated models. Surprisingly, it shows notably better overall performance than $\mathcal{M}_{\text{direct}}$. This indicates that the direct model suffers from the high complexity of its target space (27 classes) compared to the models for obtaining the complete conditionals (3 classes each). Model $\mathcal{M}_{\text{decomp}}$ shows acceptable performance but nevertheless does not reach the overall performance of the proposed $\mathcal{M}_{\text{recons}}$. This indicates that for such scenarios a unique decomposition of the dependencies cannot be performed correctly. Consequently, the performance of $\mathcal{M}_{\text{indep}}$ deteriorates further, as this model relies on the most independence assumptions.

The ROC curves of Fig. 6 are based on all 27 hypotheses. However, not all hypotheses are in fact dominated by interactions. Therefore, we examine ROC curves that are only based on situation hypotheses prone to interaction. This leads to 6 hypotheses: $\mathcal{H}_{\not\perp} = \{2, 3, 6, 9, 11, 20\}$. Hypotheses $\{2, 11, 20\}$ are characterized by the vehicle on the on-ramp performing a lane change in front of vehicle 2, which in turn performs a deceleration maneuver. Accordingly, hypotheses $\{3, 6, 9\}$ are characterized by a merging maneuver in front of vehicle 3, which in turn performs a deceleration maneuver. Fig. 7 shows the results for this subset of hypotheses.

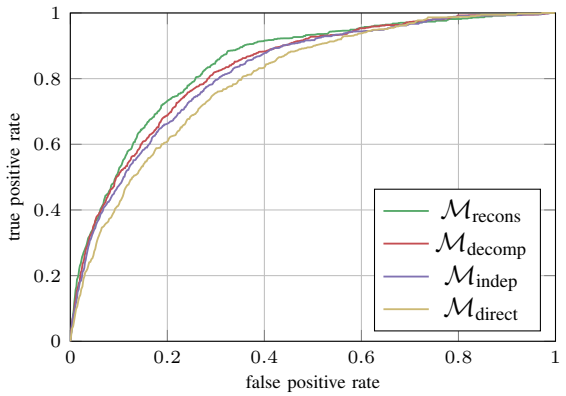

Fig. 6: ROC for all situation hypotheses $\mathcal{H}$.

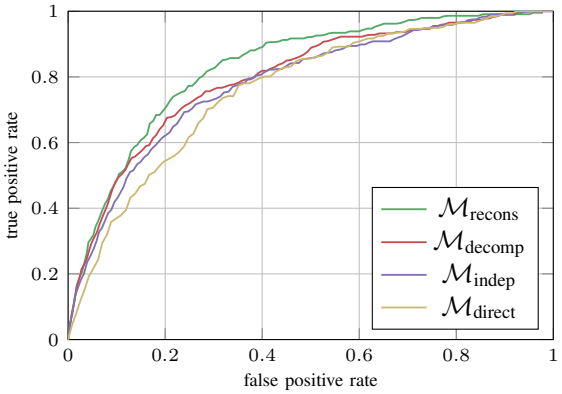

Fig. 7: ROC for situation hypotheses prominent for interaction, $\mathcal{H}_{\not\perp}$.

It can be seen that the general ordering is retained. However, the performance differences between the proposed approach and the other investigated models clearly increase for the hypotheses prone to interaction. A clear diversification of true positive rates in the false positive rate range between ~ 0.1 and ~ 0.4 can be observed.

A quantitative metric for comparing ROC curves can be obtained using the Area Under the Curve (AUC) metric. The increase in *AUC* from the proposed model $\mathcal{M}_{\text{recons}}$ in comparison to each of the other three evaluated models is shown in Tab. I.

	$\mathcal{H}$	$\mathcal{H}_{\not\perp}$
$\mathcal{M}_{\text{direct}}$	5.88%	9.07%
$\mathcal{M}_{\text{decomp}}$	1.68%	4.01%
$\mathcal{M}_{\text{indep}}$	2.92%	6.27%

TABLE I: Area Under the Curve increase by model $\mathcal{M}_{\text{recons}}$.

VII. CONCLUSION & FUTURE WORK

In this paper we presented a general framework for recognizing situations with multiple, interacting traffic participants. For this purpose we extended state-of-the-art single-entity-based maneuver recognition systems in order to obtain complete conditional distributions. On a real-world data set we showed that our system is able to reconstruct joint probability distributions, representing prototypical behavior patterns over multiple scene entities, using these complete conditionals. The benefits of the reduced model complexities of the entity based recognition system compared to learning a direct fully interaction-respecting situation recognition system could be shown.

As a next step we will test the presented framework on a larger variety of scenarios, including inner-city and highway scenarios, emphasizing its applicability for generic traffic scenes by making use of the reusable, general single-entity-based recognition models. Beyond that, we would like to investigate possibilities to utilize the proposed motion prediction of [24] within the presented framework to make interaction-aware motion predictions.

Furthermore, we think that the presented approach could also be applied successfully to cooperative scenarios where interactions dominate.

REFERENCES

[1] K. A. Brookhuis, D. de Waard, and W. H. Janssen, "Behavioural impacts of advanced driver assistance systems–an overview," *European Journal of Transport and Infrastructure Research*, vol. 1, no. 3, pp. 245–253, 2001.

[2] B. Tang, S. Khokhar, and R. Gupta, "Turn Prediction at Generalized Intersections," in *IEEE IV*, 2015, pp. 1399–1404.

[3] A. Bender, J. R. Ward, S. Worrall, and E. M. Nebot, "Predicting Driver Intent from Models of Naturalistic Driving," in *18th IEEE ITSC*, 2015.

[4] Q. Tran and J. Firl, "Online maneuver recognition and multimodal trajectory prediction for intersection assistance using non-parametric regression," in *IV*, 2014, pp. 918–923.

[5] S. Lefèvre, C. Laugier, and J. Ibañez-Guzmán, "Exploiting map information for driver intention estimation at road intersections," in *IEEE IV*, 2011, pp. 583–588.

[6] A. Armand, D. Filliat, and J. Ibañez-Guzmán, "Modelling stop intersection approaches using Gaussian processes," in *16th IEEE ITSC*, 2013, pp. 1650–1655.

[7] S. Klingelschmitt, M. Platho, H.-M. Gross, V. Willert, and J. Eggert, "Combining behavior and situation information for reliably estimating multiple intentions," in *IV*, 2014, pp. 388–393.

[8] M. Liebner, F. Klanner, M. Baumann, C. Ruhhammer, and C. Stiller, "Velocity-Based Driver Intent Inference at Urban Intersections in the Presence of Preceding Vehicles," *Intelligent Transportation Systems Magazine, IEEE*, vol. 5, no. 2, pp. 10–21, 2013.

[9] D. Petrich, T. Dang, G. Breuel, and C. Stiller, "Assessing map-based maneuver hypotheses using probabilistic methods and evidence theory," in *IEEE 17th ITSC*, 2014, pp. 995–1002.

[10] E. Ward and J. Folkesson, "Multi-classification of Driver Intentions in Yielding Scenarios," in *18th IEEE ITSC*, 2015.

[11] R. Graf, H. Deusch, F. Seeliger, M. Fritzsche, and K. Dietmayer, "A learning concept for behavior prediction at intersections," in *IV*, 2014, pp. 939–945.

[12] J. Schlechtriemen, A. Wedel, J. Hillenbrand, G. Breuel, and K.-D. Kuhnert, "A lane change detection approach using feature ranking with maximized predictive power," in *IEEE IV*, 2014, pp. 108–114.

[13] S. Bonnin, F. Kummert, and J. Schmudderich, "A Generic Concept of a System for Predicting Driving Behaviors," in *15th IEEE ITSC*, 2012, pp. 1803–1808.

[14] R. D. Tobias Gindele, Sebastian Brechtel, "A Probabilistic Model for Estimating Driver Behaviors and Vehicle Trajectories in Traffic Environments," in *13th IEEE ITSC*, 2010, pp. 1625–1631.

[15] D. Kasper, G. Weidl, T. Dang, G. Breuel, A. Tamke, A. Wedel, and W. Rosenstiel, "Object-Oriented Bayesian Networks for Detection of Lane Change Maneuvers," *Intelligent Transportation Systems Magazine, IEEE*, vol. 4, no. 3, pp. 19–31, 2012.

[16] A. Lawitzky, D. Althoff, C. F. Passenberg, G. Tanzmeister, D. Wollherr, and M. Buss, "Interactive scene prediction for automotive applications," in *IEEE IV*, 2013, pp. 1028–1033.

[17] E. Käfer, C. Hermes, C. Wöhler, H. Ritter, and F. Kummert, "Recognition of situation classes at road intersections," in *IEEE International Conference on Robotics and Automation*, 2010, pp. 3960–3965.

[18] S. Lefèvre, C. Laugier, and J. Ibañez-Guzmán, "Risk assessment at road intersections: Comparing intention and expectation," in *IEEE IV*, 2012, pp. 165–171.

[19] J. Eggert, S. Klingelschmitt, and F. Damerow, "The Foresighted Driver: Future ADAS Based on Generalized Predictive Risk Estimation," in *FAST-zero 2015 Symposium*, 2015.

[20] C. M. Bishop, *Pattern Recognition and Machine Learning*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.

[21] M. Johannes and N. Polson, "MCMC Methods for Continuous-Time Financial Econometrics," *Handbook of Financial Econometrics*, vol. 2, pp. 1–72, 2010.

[22] FHWA, "Home of the Next Generation SIMulation Community," 2015, accessed: 2015-11-18. [Online]. Available: http://ngsim-community.org/

[23] A. Niculescu-Mizil and R. Caruana, "Predicting Good Probabilities with Supervised Learning," in *ICML*, 2005, pp. 625–632.

[24] S. Klingelschmitt and J. Eggert, "Using Context Information and Probabilistic Classification for Making Extended Long-Term Trajectory Predictions," in *18th IEEE ITSC*, 2015.