Probabilistic, Variable and Interaction-aware Situation Recognition

Dem Fachbereich
Elektrotechnik und Informationstechnik
der Technischen Universität Darmstadt
zur Erlangung des akademischen Grades
eines Doktor-Ingenieurs (Dr.-Ing.)
vorgelegte Dissertation

von

Stefan Klingelschmitt, M. Sc.

Referent:	Prof. Dr.-Ing. J. Adamy
Korreferent:	Prof. Dr. rer. nat. B. Sendhoff

Darmstadt 2018

Klingelschmitt, Stefan:
Probabilistic, Variable and Interaction-aware Situation Recognition
Darmstadt, Technische Universität Darmstadt,
Jahr der Veröffentlichung der Dissertation auf TUprints: 2018
URN: urn:nbn:de:tuda-tuprints-80662
Tag der mündlichen Prüfung: 10.09.2018

Erklärung laut §9 PromO

Ich versichere hiermit, dass ich die vorliegende Dissertation allein und nur unter Verwendung der angegebenen Literatur verfasst habe. Die Arbeit hat bisher noch nicht zu Prüfungszwecken gedient.

Acknowledgments

The submission of this thesis concludes a PhD project that started in 2013 together with the Control Methods and Robotics Lab at Technische Universität Darmstadt and the Honda Research Institute GmbH in Offenbach. During this time I received support from many people for which I am very thankful for. First of all, I would like to thank Professor Jürgen Adamy for having me as a research associate at the Control Methods and Robotics Lab. He has created a fun and stimulating environment to work. Moreover, the opportunity to supervise dozens of student projects and being involved in teaching itself has equipped me with valuable skills for tasks to come. A special thanks goes to Dr. Volker Willert whom I could always ask for help when I didn't know a way forward and for his inspiring theoretical suggestions as well as to Birgit Heid for helping me with everything not related to my research.

I would like to thank Professor Bernhard Sendhoff for the chance to work on this PhD project in cooperation with the Honda Research Institute. Also, I would like to thank Professor Edgar Körner for giving me the chance for doing my Master Thesis at the Honda Research Institute which finally lead to this PhD Project. A special thanks goes to Dr. Julian Eggert for supervising this thesis. He gave me the freedom to explore my personal interest in machine learning as well as data science in order to apply it to the problem statements of this thesis.

Furthermore, I would like to say thanks to my colleagues at the Honda Research Institute and the Control Methods and Robotics Lab with whom I had interesting conversations and drank hundreds of coffees with. In particular Andrea Schnall, Martin Buczko, Moritz Schneider and Saman Khodaverdian who probably drank thousands of coffees with me. Sebastian Bernhard and myself have known each other, worked and learned together since the first semester of our studies, for which I am very thankful for. Additionally, I would like to thank Dr. Jörg Deigmöller for the nice conversations we had during our commute, Dr. Sven Rebhan for his additional support during my PhD project and Dr. Matthias Platho for giving me a head start into the project.

A special thanks goes to my parents who always supported me during my

studies and beyond so that I could focus on my professional development. Also, many thanks to Martin Buzcko and Andreas Balz for proof reading this thesis.

Last but not least I would like to thank Gesine who started this project as my girlfriend, became my wife and ended this project with me by giving birth to our daughter. Thanks for the everlasting support, for always cheering me up at times where nothing seemed to work and reminding me that there are more important things in live than this thesis. I could not have finished this without you.

"...it ain't about how hard you hit. It's about how hard you can get hit and keep moving forward."

Rocky Balboa, Dir. Sylvester Stallone, 20th Century Fox, 2006, Film.

Contents

Abbreviations and Symbols

Abbrevations

ADAS	Advanced Driver Assistance Systems
AUC	Area Under the Curve
BWD	Backward
CRF	Conditional Random Field
FN	False Negative
FP	False Positive
FPR	False Positive Rate
FWD	Forward
GPS	Global Positioning System
HMM	Hidden Markov Model
IVSS	Intelligent Vehicle Safety Systems
LDA	Linear Discriminant Analysis
NGSIM	Next Generation SIMulation
OvA	One vs. All
OvO	One vs. One
PC	Personal Computer
pdf	probability density function
pmf	probability mass function
ROC	Receiver Operating Characteristic
SVM	Support Vector Machine
TN	True Negative
TTI	Time To Intersection
TP	True Positive
TPR	True Positive Rate
WGS84	World Geodetic System 1984
e.g.	exempli gratia
i.e.	id est
r.v.	random variable
–	decelerated behavior pattern
++	accelerated behavior pattern
xpcd	expected behavior pattern
betw	merge between two vehicles
front	merge in front of vehicles
no	no lane change maneuver

Symbols

Notation

x	Scalar
$\hat{x}$	Estimated quantity
x_t	Quantity at time t
X	Random Variable
$\mathbf{x}$	Column vector
$\mathbf{x}^T$	Row vector
$\lvert\mathcal{X}\rvert$	Cardinality of set $\mathcal{X}$
$p(\cdot)$	Probability distribution or value
$\mathrm{KL}(\cdot \,\|\|\, \cdot)$	Kullback-Leibler divergence
$\lVert\cdot\rVert_1$	L1-norm
$\perp$	Stochastic independence between two r.v.
$\not\perp$	Stochastic dependence between two r.v.
$\mathcal{M}$	Models for estimating prob. distributions
$\mathbf{1}, \mathbf{2}, \ldots$	multi-index for events of multivariate r.v.

Latin Uppercase Letters

$\mathbf{A}$	Stochastic reconstruction matrix
$\mathbf{A}^{\mathbf{s}}_r$	reconstruction matrix for distribution $p(b_{r:n}\|b^{\mathbf{s}}_{1:r-1})$
B	r.v. for the discrete behavior choices of a vehicle
B_i	r.v. for the discrete behavior choices of vehicle i
$\mathbf{C}$	Matrix of quotients $c^{k,l}$
H	Hidden state in a CRF
$\mathbf{I}$	Identify matrix
$\mathcal{I}_i$	Dummy event set of vehicle i
$\mathcal{I}_{i:j}$	Dummy event combination set for vehicles i to j
$\mathcal{M}_{\text{spec}}$	Specialized model for the scene
$\mathcal{M}_{\text{4-way}}$	Coupling model trained on 4-way intersections
$\mathcal{M}_{\text{3/4-way}}$	Coupling model trained on 4-way and 3-way intersections
$\mathcal{M}_{\text{coup}}$	Coupling model
$\mathcal{M}_{\text{norm}}$	General model that uses normalization
$\mathcal{M}_{\text{gen}}$	Generative model
$\mathcal{M}_{1,\perp}$	Model for vehicle 1
$\mathcal{M}_{1,\not\perp}$	Model for vehicle 1 dependent on other vehicles
$\mathcal{M}_{1,\perp}$	Model for vehicle 2
$\mathcal{M}_{1,\not\perp}$	Model for vehicle 2 dependent on other vehicles

$\mathcal{M}_{\text{LGS}}$	Reconstruction model based on solving linear equation systems
$\mathcal{M}_{\text{BWD, LGS}}$	Reconstruction model based on solving linear equation systems in recursive manner
$\mathcal{M}_{\text{marg}}$	Reconstruction model that directly calculates missing marginal distributions
$\mathcal{M}_{\text{decomp}}$	Model using decomposition of joint probabilities
$\mathcal{M}_{\text{recons}}$	Model based on reconstructing joint probabilities
$\mathcal{M}_{\text{indep}}$	Model based on independence assumptions
$\mathcal{M}_{\text{direct}}$	Model directly estimating joint probabilities
$\mathcal{M}_{\text{pair}}$	Model based on reconstructing joint probabilities using pairwise probability coupling
$\mathcal{N}$	Gaussian distribution
$\mathbf{Q}$	Stochastic matrix for pairwise probability coupling
$\tilde{\mathbf{Q}}$	$\mathbf{Q}$ with an arbitrary row replaced by $[1, 1, \ldots, 1]^T$
$\mathbb{R}^{\delta}$	δ-dimensional real vector space
$\mathcal{T}$	Trajectory
$\mathcal{T}_g$	Trajectory g
Y	Exemplary discrete r.v.
Z	Exemplary continuous r.v.

Latin Lowercase Letters

a_t	Acceleration at time t
$a^{o,u}$	Elements of $\mathbf{A}$, 2 r.v. case
$(\mathbf{a}^{f,h})^T$	Row of stochastic matrix $\mathbf{A}$, 2 r.v.
$(\mathbf{a}^{\mathbf{k}})^T$	Generalization of $\mathbf{a}^{f,h}$ for n r.v.
$a^{\mathbf{k},\mathbf{l}}$	Elements of $\mathbf{A}$ in n r.v. case
b	Event of r.v. B
$\mathbf{b}_{i:j}$	Multivariate r.v. comprising vehicles i to j
$b^{(0)}$	Randomly chosen starting event
$b^{(\tau)}$	Event drawn at sample τ
$c^{k,l}$	Quotient of complete conditionals of two vehicle's events k, l
d	Distance to intersection
$\hat{d}$	Interpolated distance to intersection
$\Delta d_{i,j}$	Relative longitudinal distance between vehicle i and j
e_i	Expected behavior of vehicle i

f	Index for indexing events $(1, \ldots, \lvert\Omega_i\rvert)$ of r.v. i
g	Index for different trajectories $\mathcal{T}_g$
h	Index for indexing events $(1, \ldots, \lvert\Omega_i\rvert)$ of r.v. i
$\mathbf{h}$	Helper nodes r.v. in a Bayesian network
i	Index for a vehicle
j	Index for a vehicle
k	Index for indexing events $(1, \ldots, \lvert\Omega_i\rvert)$ of r.v. i
$\mathbf{k}$	Multi-index for $\mathbf{1}, \ldots, \lvert\mathbf{\Omega_{1:n}}\rvert$
$\mathbf{l}$	Multi-index for $\mathbf{1}, \ldots, \lvert\mathbf{\Omega_{1:n}}\rvert$
l	Index for indexing events $(1, \ldots, \lvert\Omega_i\rvert)$ of r.v. i
l_{xpcd}	Expected longitudinal travelled distance
m	Number of events in Ω
n	Number of all r.v./vehicles
n^+	Number of positive cases
n^-	Number of negative cases
$\tilde{n}$	Number of all cases
$n^{+,-}$	Number of all cases converted to binary problem
n_{spec}	Number of needed specialized models
n_{bin}	Number of needed binary classifiers
n_τ	Number of drawn, complete overall samples
$n_\mathcal{T}$	Number of trajectories
n_g	Number of samples of trajectory g
o_i	Observations of vehicle i
o	Index for coefficients $(1, \ldots, \lvert\Omega_{i:j}\rvert)$ of reconstruction matrix $\mathbf{A}$ for vehicle i to j
u	Index for coefficients $(1, \ldots, \lvert\Omega_{i:j}\rvert)$ of reconstruction matrix $\mathbf{A}$ for vehicle i to j
$\mathbf{p}$	Vector containing all probability values
$\mathbf{p}_r^{\mathbf{s}}$	probability vector for distribution $p(b_{r:n} \lvert b_{1:r-1}^{\mathbf{s}})$
$q^{k,l}$	Element of $\mathbf{Q}$
$\mathbf{q}^T$	Row of $\mathbf{Q}$
r	Recursion depth
s	Index for indexing events $(1, \ldots, \lvert\Omega_i\rvert)$ of r.v. i
$\mathbf{s}$	Multi-index for sample space $\mathbf{1}, \ldots, \lvert\mathbf{\Omega_{1:r-1}}\rvert$
t	Time
Δt	Time difference
t_{ch}	Time at which a lane change occurs
$t_{\text{no,ch}}$	Randomly chosen time at which no lane change occurs

v	Velocity
$\mathbf{v}$	Eigenvector
$\hat{v}_{1,\hat{d}}$	Interpolated velocity of veh. 1 at $\hat{d}$
$\hat{v}_{1,\hat{d},g}$	Interpolated velocity of veh. 1 at $\hat{d}$ of trajectory g
v_t	Velocity at time t
$\Delta v_{i,j}$	Relative longitudinal velocity difference between vehicle i and j
w	Categorical variable indicating priority
$\mathbf{x}_i$	Scene evidence of vehicle i
$\mathbf{x}$	Multivariate r.v. consisting of the scene evidence
$\mathbf{x}_\subset$	Multivariate r.v., subset of scene evidence $\mathbf{x}$
$\mathbf{x}_{\setminus\subset}$	Multivariate r.v., remaining part of scene evidence $\mathbf{x}$
$\hat{\mathbf{x}}_\subset$	Prediction of $\mathbf{x}_\subset$
y	Event of r.v. Y
z	Event of r.v. Z

Greek Letters

α	exponent for matrix multiplication
β	Multiplicator for chance level pooling
ϵ	Error term
λ	Eigenvalue
$\mu_{\hat{d}}$	Mean velocity at interpolated distance $\hat{d}$
σ	Standard deviation
$\sigma^2_{\hat{d}}$	Variance of velocity at interpolated distance $\hat{d}$
τ	Specific sample/iteration
$\tilde{\Omega}_{\setminus k}$	Complete sample space without event b^k
${}^{\mathrm{I}}\Omega$, ${}^{\mathrm{I}}\Omega$, ${}^{\mathrm{III}}\Omega$, ${}^{\mathrm{IV}}\Omega$	Sample spaces of scene I, II, III, IV
$\bar{\Omega}_i$	Reduced sample space of vehicle i
$\bar{\Omega}_{i:j}$	Reduced sample space of vehicles i to j
Ω	Sample space
Ω_i	Sample space of the behavior of vehicle i
$\Omega_{i:j}$	Combined sample space of multivariate r.v. $\mathbf{b}_{i:j}$
$\Omega_\subset$	Incompatible events of the scene
$\tilde{\Omega}$	Sample space for all ever possible maneuvers
$\Omega^{k,l}$	Binary sample space for events b^k and b^l

Abstract

Future advanced driver assistance systems (ADAS) as well as autonomous driving functions will extend their applicability to more complex highway scenarios and inner-city traffic. For these systems it is a prerequisite to know how an encountered traffic scene is most likely going to evolve. Situation recognition aims to predict the high level behavior patterns traffic participants pursue. Thus, it provides valuable information that helps to predict the next few seconds of a traffic scene. The extension of ADAS and autonomous driving functions to more complex scenarios poses a problem to state-of-the-art situation recognition systems due to the variability of the encountered scene layouts, the presence of multiple interacting traffic participants and the concomitant large number of possible situation classes. This thesis proposes and discusses approaches that tackle these challenges.

A novel discriminative maneuver estimation framework provides the possibility to assess traffic scenes with varying layout. It is based on reusable, partial classifiers that are combined online using a technique called pairwise probability coupling. The real-world evaluations indicate that the assembled probabilistic maneuver estimation is able to provide superior classification results.

A novel interaction-aware situation recognition framework constructs a probabilistic situation assessment over multiple traffic participants without relying on independence assumptions. It allows to assess each traffic participant individually by using maneuver estimation systems that determine complete conditional distributions. A real-world evaluation outlines its applicability and shows its benefits.

The challenges associated with the increasing number of possible situation classes are addressed in two ways. Both frameworks allow to reuse classifiers in different contexts. This reduces the number of models required to cope with a large variety of traffic scenes. Moreover, a situation hypotheses selection scheme provides an efficient way for reducing the number of situation hypotheses. This lowers the computational demands and eases the load on subsequent systems.

Kurzfassung

Zukünftige intelligente Fahrerassistenzsysteme (FAS) und autonome Fahrfunktionen werden darauf abzielen, auch mit komplexen Autobahnszenarien und innerstädtischen Verkehr umgehen zu können. Für solche Systeme ist es eine Grundvoraussetzung einschätzen zu können, wie sich eine vorgefundene Szene wahrscheinlich entwickeln wird. Die Situationserkennung beschäftigt sich damit abstrakte Verhaltensmuster der beteiligten Verkehrsteilnehmer vorherzusagen und ist damit ein entscheidender Teil der Szenenvorhersage. Die Ausweitung des Anwendungsbereiches zukünftiger FAS und autonomer Fahrfunktionen stellt ein Problem aufgrund der Variabilität der vorgefundenen Szenen, der Anwesenheit von mehreren, interagierenden Verkehrsteilnehmern und die damit einhergehende große Anzahl von möglichen Situationen für aktuell verfügbare Situationserkennungstechniken dar. Die vorgelegte Arbeit befasst sich mit Ansätzen, diese Probleme zu lösen.

Ein neuartiger, diskriminativer Manövererkennungsansatz bietet die Möglichkeit, auf unterschiedliche Verkehrsszenen angewendet zu werden. Die Idee des Ansatzes basiert darauf, wiederverwendbare Teile des Klassifikationssystems online zu kombinieren. Evaluationen auf realen Verkehrsszenarien zeigen die Vorteile des vorgeschlagenen Ansatzes.

Ein neuartiger, Situationserkennungsansatz bietet die Möglichkeit, für eine Szene mit mehreren, interagierenden Verkehrsteilnehmern eine probabilistische Situationsschätzung vorzunehmen. Dabei ist der Ansatz auf keinerlei Unabhängigkeitsannahme angewiesen und erlaubt es trotzdem, mögliche Manöver jedes Verkehrsteilnehmers einzeln zu schätzen. Die Idee dieses Ansatzes basiert darauf, die kompletten bedingten Wahrscheinlichkeitsverteilungen über die Manöver jedes Verkehrsteilnehmers zu schätzen. Evaluationen auf realen Verkehrsszenarien bestätigen die Anwendbarkeit des vorgestellten Ansatzes.

Die Probleme, die mit der hohen Anzahl von möglichen Situationen einhergehen, werden auf zwei Arten in dieser Arbeit angegangen. Beide vorgestellten Ansätze erlauben es bereits parametrisierte Situationsklassifikatoren wiederzuverwenden und in neuen Kontexten zu kombinieren. Das reduziert die Anzahl an benötigten Klassifkationsmodellen, da nicht

jede vorgefundene Szene ein spezialisiertes Klassifikationsmodell benötigt. Darüber hinaus bietet eine Situationshypothesenselektion die Möglichkeit, effizient den aufgespannten Situationsraum zu verkleinern. Diese Maßnahme reduziert den rechnerischen Aufwand, und die verringerte Anzahl von Situationshypothesen entlastet nachfolgende Systeme.

1 Introduction

In the last decade the capabilities of technical systems have been continuously improved. They constantly expand further into humans' everyday life and their development progressively heads into the direction of independence and autonomy. These developments do not exclude the automotive industry.
An increased interest by academia, automotive manufacturers as well as the general public in Advanced Driver Assistance Systems (ADAS), partially automated vehicles and fully autonomous vehicles can currently be observed. Representative for these observations are already available - like assistance systems for lane keeping, parking, velocity adaptation or front collision avoidance [118] as well as the media-effective presentation of several self-driving car projects from universities, *IT*-companies and car manufacturers. Fig. 1.1 illustrates the accompanied interest of the general public using Google Trends. It reveals an almost exponentially increase in search requests like "self-driving car", "driverless car" or "autonomous ve-

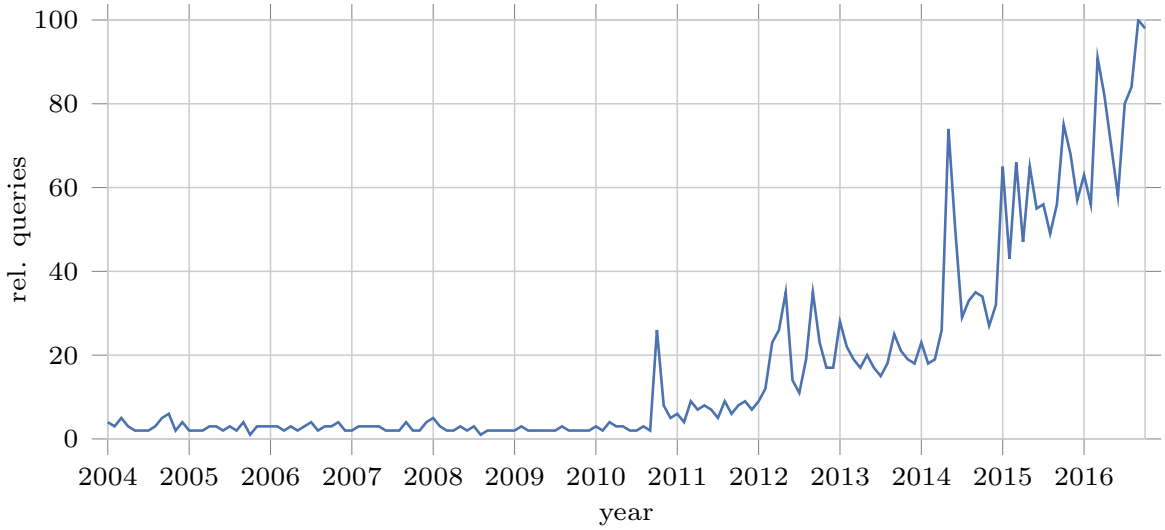


Figure 1.1: Google Trends for "self-driving car" [2]. The figure depicts relative numbers in relation to the maximum number of received search request per month.

hicle" in the last seven years. The visible peaks correlate to media-effective presentations:

- October 2010: Professor Alberto Broggi and his team at the University of Parma complete a 13000 kilometer long test run from Parma to the Shanghai Expo in China with four autonomous vehicles [15],
- May 2012: Google officially introduces its self-driving car project by administering a driving test by the Nevada motor vehicle examiners [47],
- May 2014: Google announces 100 self-designed autonomous test vehicles [112],
- March 2015: Tesla announces that its *Autopilot* will be available in mid 2015 per software update in their cars equipped with the needed hardware setup [60].

There are two main reasons for the great interest in such systems:
First, future ADAS and automated driving promise to increase the **safety** of general traffic. The vast majority of accidents can be traced back to human error. E.g. in 2012, 86% of the traffic accidents registered in Germany can be assigned to human error [101]. Therefore, technical systems assisting the driver or completely overtaking the driving task will be crucial for reducing traffic accidents. As an example, comprehensively available lane changing assistants could reduce lane change-associated accidents by up to 25% according to studies of the *Deutscher Verkehrsicherheitsrat* [32].
Second, future ADAS and automated driving promise to increase the **comfort** for passengers. Technical systems that take over more and more responsibility of the driving task relieve the stress on the driver. Taking into consideration the time that a significant part of the population spends commuting every day emphasizes the need for such systems. E.g. 22% of employed persons in Germany need to commute between 30 and 60 kilometers to work on daily basis [103].

1.1 Problem Formulation

Assistance systems that are available today are mainly reactive assistance systems. An example poses the currently available adaptive cruise controls. Such a system starts reacting to a sudden lane entering event of a vehicle in front after it passes the lane marking, leaving the ACC little to

no time to adapt its velocity. Such abrupt velocity adaptations lead to discomfort for the driver and constitute a hazardous maneuver for following vehicles.
Consequently, assistance systems and autonomous functionalities that allow for taking responsibility and stress of the driver need to be proactive. This means that they are able to anticipate the future development of a traffic scene and provide adequate behavioral reactions or warnings. For this reason more and more systems are introduced that take the possible evolution of traffic scenes into account.

Such proactive systems typically share a common structure: Fig. 1.2 illustrates that these systems consist of an environment perception, traffic scene prediction and a module for realizing the intended target functionality. The environment perception addresses problems like detecting the drivable area [69], self localization [22] as well as classification and localization of static and dynamic objects [106].
Based on this information the next few seconds of the traffic scene are predicted. Predicting the evolution of a traffic scene is a demanding task as traffic participants' future trajectories involve many uncertainties. Because of this, it is common to decompose the prediction task into two steps:
First, predicting the discrete prototypical situation classes which group a subset of traffic participants and specify their typical associated spatio-temporal behavior patterns (maneuvers), e.g. a car decelerating to give another traffic participant right of way before making a turn at an intersection. Second, based on the prototypical situation class specific motion prediction models estimate future kinematic states.
With the information on likely future developments of the scene the system can provide safe and comfortable reactions.

The thesis at hand focuses on the recognition of discrete situation

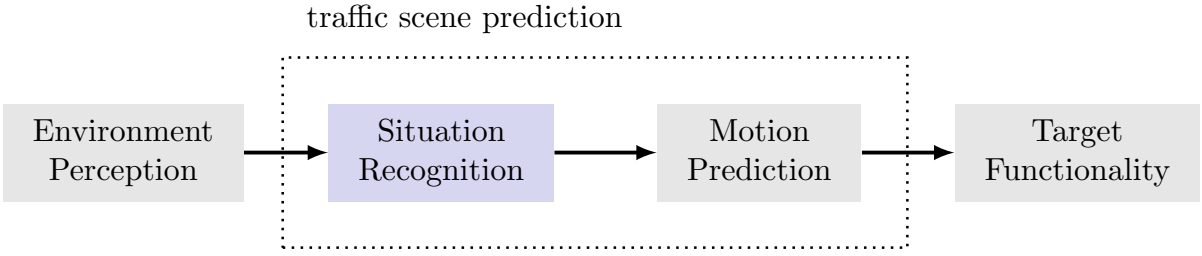


Figure 1.2: Typical structure of ADAS and automated driving functions.

classes. Current situation recognition approaches are designed for fixed scenes with a limited number of traffic participants. However, especially complex scenes like inner-city intersections with high variability or highway enterings result in the most discomfort and accidents [28, 102]. Hence, future situation recognition systems need to be able to tackle complex scenes. This leads to three main research questions that are tackled in this thesis.

1. How can situation recognition systems be adapted to the currently encountered scene?

General traffic undergoes a high variability and situation recognition systems need to deal with this variability. For example, inner-city traffic scenes differ strongly in available turning options or the number of possibly affecting traffic participants like pedestrians, bicyclists or other vehicles. The presence of other acting traffic participants leads to the second research question.

2. How can situation recognition systems be designed so that they are able to deal with multiple interacting traffic participants?

In traffic situations with more than one traffic participant interactions can play an important role and influence the future development of traffic scenes. Thus, situation recognition systems need to deal with interacting traffic participants. Moreover, the presence of several traffic participants has another major impact on situation recognition systems which leads to the third research question.

3. How can situation recognition systems deal with a large number of possible situation classes?

Considering several traffic participants simoultaneously leads to an combinatorial increase of possible situation hypotheses. This circumstance induces computational challenges to the situation recognition system itself as well as to subsequent systems in the processing chain responsible for motion prediction or criticality assessment.

1.2 Contribution

The main contributions of this thesis tackle the above stated research questions and summarize as follows:

- Novel discriminative as well as generative approaches for estimating the maneuver of single traffic participants [31, 35, 66, 67].
- A novel framework for adapting state-of-the-art discriminative maneuver estimation approaches to generic traffic scenes by decomposing the classification task into binary subproblems [67].
- A novel interaction-aware situation recognition framework that decomposes the situation recognition with multiple traffic participants into single-entity maneuver predictions without the need for independence assumptions [64].
- Two novel analytic methods for solving discrete sampling problems where the probability distribution are preexisting in their functional form, e.g. the interaction-aware situation recognition framework [64].
- A novel situation hypotheses reduction technique that tackles the combinatorial increasing number of situation hypotheses associated with multiple traffic participants.

The contributions have been tested and evaluated on real-world data sets ranging from simple turn predictions to complex intersection scenes with several traffic participants as well as challenging highway enterings.

1.3 Thesis Outline

The flowchart in Fig. 1.3 depicts the structure of the remainder of this thesis. Chapters containing novel contributions, developed during the PhD project, are highlighted in blue.
Chapter 2 explains the terms *maneuver* and *situation*. It also gives a short description of the data sets used for the conducted evaluations. Additionally, it introduces the required probability theoretical concepts and their notation.
Chapter 3 discusses related works in the field of maneuver estimation and outlines their problems associated with variable traffic scenes. Subsequently, it introduces the variable discriminative maneuver estimation

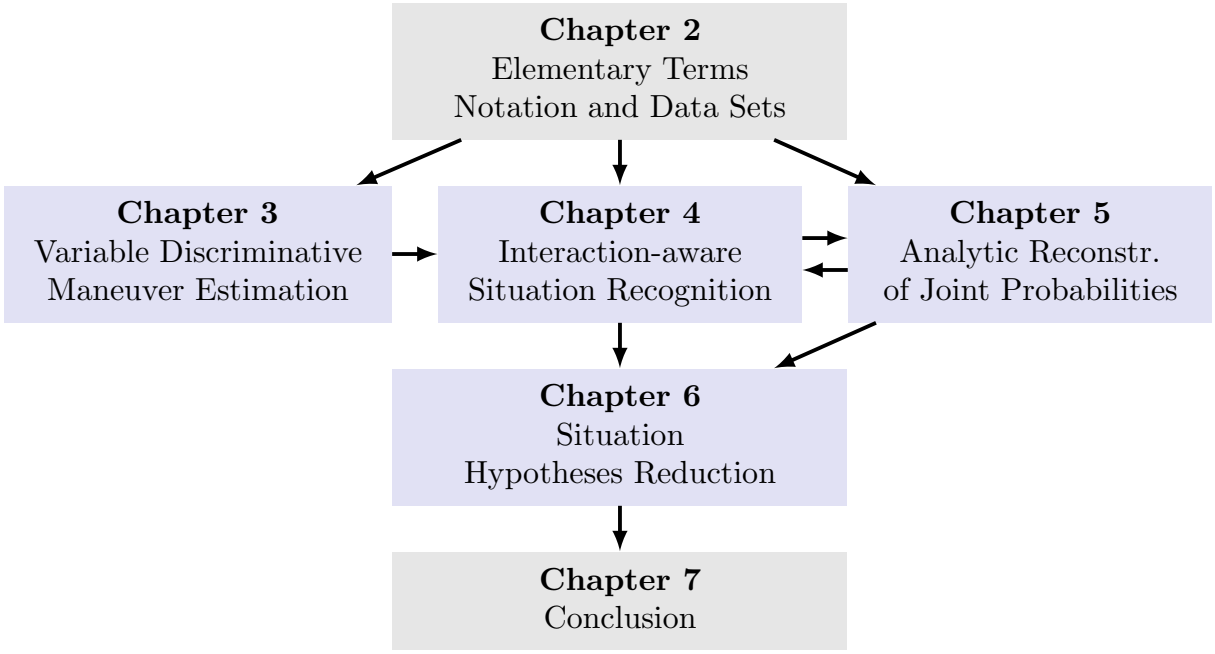


Figure 1.3: Thesis structure. Chapters with contributions are highlighted in blue.

framework and outlines its applicability with various tests on real-world traffic. Maneuver estimation systems form the basis of situation recognition.
Chapter 4 starts with discussing related work of situation recognition that considers multiple traffic participants, simoultaneously. Moreover, it states the need for a general interaction-aware situation recognition approach that does not rely on independence assumptions. For this purpose, an interaction-aware situation recognition framework is proposed that decomposes the problem into assessing each traffic participant individually while taking the influence of other vehicles' maneuvers explicitly into account. Either probabilistic sampling techniques or one of the novel analytic reconstruction schemes of Chapter 5 restore the holistic situation recognition overall traffic participants. Tests on real-world traffic scenarios outline the applicability of the approach. Chapter 6 introduces an approach for reducing the number of situation hypotheses that are investigated in detail. These reduction techniques allow for better scalability of situation recognition systems targeted at multiple traffic participants.
Finally, Chapter 7 concludes the thesis and outlines possible future research directions.

2 Elementary Terms, Notation and Data Sets

This chapter formally recapitulates elementary terms, introduces their notation, and describes the data sets, as well as evaluation metrics used for quantifying the performance of the approaches proposed throughout this thesis.

2.1 Probability Notation

Probability is a measure describing how likely an event is going to happen. Events are either continuous, for example a velocity value, or discrete, for example the state of a traffic light. A probability distribution comprises the probability values of all possible events in functional or tabular form. Fig. 2.1 shows two exemplary probability distributions for (a) a continuous case and (b) a discrete case.

There are many misconceptions regarding the notation and usage of terms related to probability theory. Therefore, this section provides a formal introduction for terms regularly occurring throughout this thesis.

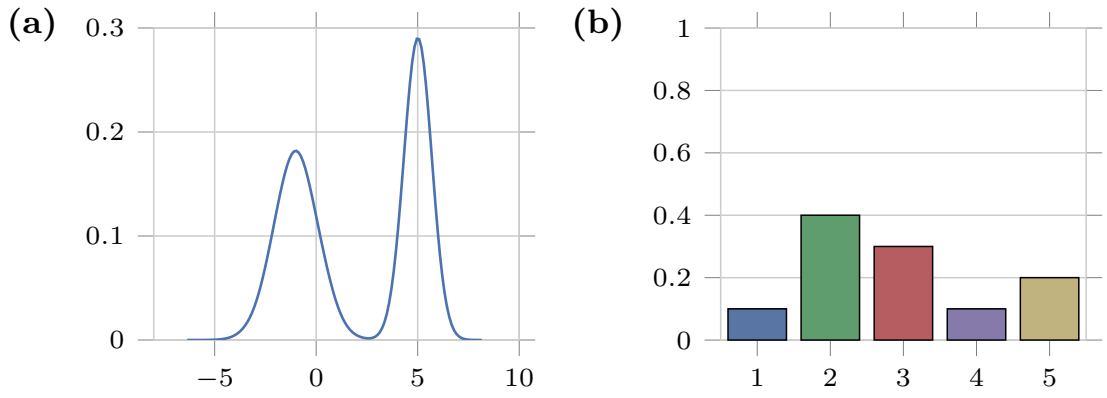


Figure 2.1: Probability distribution examples.

2.1.1 Probability Mass and Density Functions

An upper case notation, e.g. Y, denotes a discrete random variable. The lower case notation y marks possible events of random variable Y. All possible events form a set $\Omega = \{y^1, y^2, \ldots, y^{|\Omega|}\}$ called the sample space. A single event is denoted by, e.g. y^1 where the superscript distinguishes specific events. Furthermore, the notation $|\Omega|$ corresponds to the cardinality of the sample space depicting the number of events in Ω.
The *probability mass function* (*pmf*) for random variable Y is given by

$$p(\{y \in \Omega : Y = y\}), \tag{2.1}$$

containing probability values for each event $y \in \Omega$.
A concrete probability value is denoted by

$$p(Y = y^k). \tag{2.2}$$

In the further course of this thesis, the formal notations of (2.1) and (2.2) are abbreviated with the common lower case notations $p(y)$ and $p(y^k)$, respectively. Similarly to these simplified notations,

$$\sum_{y' \in \Omega} p(Y = y') = \sum_{y' \in \Omega} p(y')$$

denotes the summation over the events of a random variable.

For the continuous case, a random variable Z's events z take on values from an infinite set. Thus, the sample space corresponds to a range of continuous values, typically $\Omega = \mathbb{R}$. The *probability density function* (*pdf*) for random variable Z is given by

$$p(\{z \in \Omega, Z = z\}) \tag{2.3}$$

and is the equivalent of a *pmf* for the continuous case. Concrete density values indicated by $p(Z = z)$ are not to be confused with concrete event probabilities like defined in (2.2) for the discrete case. The probability of a continuous random variable exactly taking on a specific value z is necessarily 0.
The same lower case abbreviation as for discrete random variables is used, thus the notation of the *pdf* in (2.3) simplifies to $p(z)$.

The general term *distribution* describes either a probability mass functions or probability density functions in theoretical considerations regarding both discrete or continuous variables, as well as in cases where discrete and continuous variables are mixed within one distribution.

2.1.2 Multiple Random Variables

In many cases probability values over n random variables are of interest. Different random variables are denoted in two ways. The first option is to use different letters for different random variables, e.g. Y or Z. This option is chosen if the random variables represent different types of random variables, e.g. discrete behavior options of a traffic participant and continuous measurements. In addition to that, a subscript index i distinguishes between different random variables of the same type, e.g. $Y_1, \ldots, Y_n$.
Stochastic independency between random variables is constituted by $Y_1 \perp Y_2$, whereas dependency between variables is depicted by $Y_1 \not\perp Y_2$.

Joint Distribution

A *joint pmf* or *pdf* represents the probability , respectively density value, over several random variables taking on specific events at the same time. It is denoted by the lower case notation of (2.1) for a *pmf* or (2.3) for a *pdf* where random variables are separated by commas, e.g.

$$p(y, z).$$

Conditional Distribution

A *conditional pmf* or *pdf* represents the distribution of a random variable, given the values of another random variable as evidence. It is denoted by separating the target random variable y from its conditional z by a vertical bar, e.g.

$$p(y|z).$$

It is also possible that a conditional distribution's target space or conditional consists of several random variables. An example is given by

$$p(y_1, y_2|y_3, \ldots, y_n).$$

Combined Random Variables

With an increasing number n of random variables the notation of probability distributions suffers from decreasing clarity. Therefore, this section introduces a notation for combining random variables to improve readability.

The bold lower case notation for the set $\mathbf{y}_{i:j}$ summarizes successive random variables i to j, e.g.

$$p(y_1, \ldots, y_n) = p(\mathbf{y}_{1:n}),$$

or

$$p(y_2|y_1, y_3 \ldots, y_n) = p(y_2|y_1, \mathbf{y}_{3:n}).$$

Moreover, the sample spaces of such successive random variables are combined by applying the Cartesian product

$$\Omega_{i:j} = \Omega_i \times \ldots \times \Omega_j,$$

leading to the sample space $\Omega_{i:j}$.

It should be noted that the bold lower case notation for the combination of random variables in a set $\mathbf{y}_{i:j}$ can be interpreted as a new multivariate random variable on the sample space $\Omega_{i:j}$.

Multi-Index Notation

Furthermore, this section introduces a convenient way for dealing with indexing several events simultaneously. This notation is restricted to discrete random variables since an indexation of continuous events is not reasonable.

The combined sample space of all discrete random variables $i = 1, \ldots, n$ contains all $|\Omega_{1:n}|$ event combinations. Thus, $\Omega_{1:n}$ of $\mathbf{y}_{1:n}$ is composed of the events

$$\begin{aligned}
\mathbf{y}^{\mathbf{1}}_{1:n} &\rightarrow \{y^1_1, y^1_2, \ldots, y^1_n\} \\
\mathbf{y}^{\mathbf{2}}_{1:n} &\rightarrow \{y^1_1, y^1_2, \ldots, y^2_n\} \\
\mathbf{y}^{\mathbf{3}}_{1:n} &\rightarrow \{y^1_1, y^1_2, \ldots, y^3_n\} \\
&\vdots \\
\mathbf{y}^{|\mathbf{\Omega_{1:n}}|}_{1:n} &\rightarrow \{y_1^{|\Omega_1|}, y_2^{|\Omega_2|}, \ldots, y_n^{|\Omega_n|}\},
\end{aligned}$$

where the bold multi-index $\mathbf{k} = \mathbf{1}, \ldots, |\mathbf{\Omega_{1:n}}|$ indexes concrete event combinations in lexicographical order. Hence, it addresses a distribution's concrete event combinations on its combined sample space, e.g.

$$p(\mathbf{y}^{\mathbf{2}}_{1:n}) = p(y^1_1, y^1_2, \ldots, y^2_n).$$

Furthermore, it is used in cases where the joined notation of the corresponding random variables is not possible, e.g. for conditional distributions. Hence, also subsets $\mathbf{y}_{i:j} \subset \mathbf{y}_{1:n}$ or single random variables $y_i \in \mathbf{y}_{1:n}$ are indexed with the bold multi-index. By doing so, the subsets or single random variables take on the events that are specified by the multi-index (of the complete sample space $\Omega_{1:n}$) for the respective random variables, e.g.

$$p(y_1^{\mathbf{2}}|\mathbf{y}_{2:n}^{\mathbf{2}}) = p(y_1^1|y_2^1, y_3^1, \ldots, y_n^2).$$

2.2 Scene, Scenario, Maneuver and Situation

This section describes frequently recurring terms associated with situation assessment. Related literature often uses these terms in varying context and with different meanings. Thus, the following description provides definitions and clarifies their usage throughout this thesis.

2.2.1 Scene & Scenario

A commonly occurring term in conjunction with advanced driver assistance systems and automated driving is *scene*. A scene is a snapshot based on the current environment perception at a concrete point in time [111]. The origin of this environment perception might be the result of sensory setups of one or more test vehicles or from an objective observer, e.g. rooftop mounted sensors or simulation environments. It contains information about present dynamic entities (e.g. vehicles, pedestrians), their kinematic information and relations to each other, as well as static entities, such as the road layout (e.g. crossing type, lane options) or the currently applying traffic rules (e.g. yield signs). Digital map data, like *OpenStreetMap*, typically forms the basis for the road layout information. A scene is an objective description of the currently present circumstances. Objective means that it is free of interpretation of its future development or intentions of its acting entities. Its description, however, is a subjective result in terms of non-comprehensive perception and imperfect sensory measurements.

A *scenario* describes a sequence of scenes over a certain time interval [111]. It contains information on the temporal evolution of dynamic scene entities' states as well as information of entering and leaving dynamic scene entities. A *trajectory* comprises the evolution of traffic participants'

kinematic states (e.g. position, velocity, heading). Thus, trajectories form a large portion of the information in a scenario .

2.2.2 Maneuver

Future advanced driver assistance systems and automated driving functions rely on future trajectories of surrounding traffic participants. The evaluation of all future state evolutions is hardly feasible for complex traffic scenarios due to their continuous target space and different characteristics. For this reason it is common to categorize similar trajectories into prototypical spatio-temporal behavior patterns, i.e. *maneuvers*.
The spatial road layout strongly constrains the motion of a vehicle lateral to the driving direction. Fig. 2.2 (a) shows an example of a 3-way crossing. The intersection paths determine the typical spatial behavior patterns, in the case of Fig. 2.2 (a) either a left turn or a straight intersection crossing. The temporal aspect addresses the circumstance that different maneuvers are not only characterized by different path alternatives but also by their temporal execution. Traffic rules, driving comfort and the presence of other traffic participants typically influence the temporal and longitudinal driving behavior. Taking a closer look at a variety of recorded trajectories reveals that they can usually be clustered into typical temporal behavior patterns. Fig. 2.2 (b) shows the same recorded trajectories as in (a) with additional temporal information. The left turning trajectories clearly divide into two groups, where one group of shows significantly slowed down trajectories due to an oncoming vehicle with priority (red). The other group performed an undisturbed left turn (blue). Thus, a maneuver describes typical spatial as well as temporal behavior patterns.

Predicting a traffic participant's (human or machine) future maneuver is a challenging task since its intention is typically a hidden state. Thus, prediction systems rely on the observable information for predicting a traffic participant's maneuver. Ambiguous behavior patterns and imperfect sensor measurements lead to additional uncertainties during the prediction process. Hence, making discrete predictions probably leads to many false maneuver estimations. This problem is reinforced in conjunction with longer prediction horizons and increasing numbers of possible maneuver alternatives. Therefore, a probabilistic treatment, quantifying the involved uncertainties, is reasonable.
A discrete random variable B models the future spatio-temporal behavior

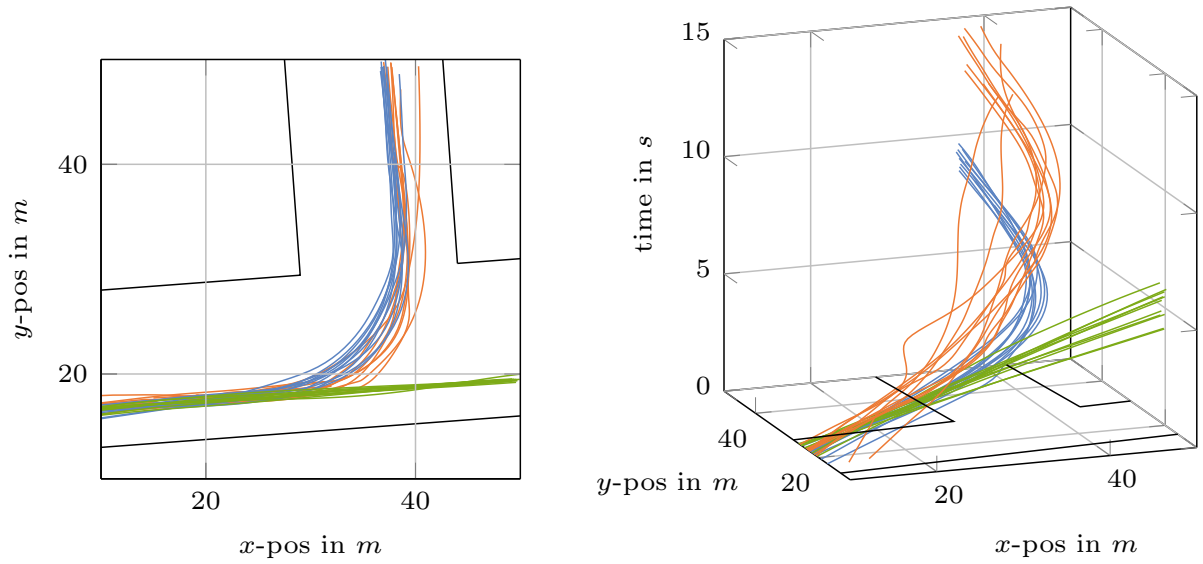


Figure 2.2: (a) Illustration of spatial clustering for recorded trajectories on a 3-way intersection crossing. Blue and red trajectories correspond to a left turn, the green trajectories to a straight intersection crossing. Part (b) highlights the temporal aspect and clusters of these trajectories, where the red trajectories are effected by oncoming traffic with priority.

pattern, i.e. maneuver, of a traffic participant. All possible maneuvers (events) form the sample space $\Omega = \{b^1, b^2, \ldots, b^n\}$. The distribution

$$p(\{b \in \Omega : B = b\}|\mathbf{x}),$$

quantifies the uncertainty regarding the maneuver estimation conditioned on the currently observable scene evidence $\mathbf{x}$. The scene evidence $\mathbf{x}$ is a multivariate random variable, see Section 2.1.2, comprising information of the scene in the form of continuous (e.g. velocities, relative distances) and discrete (e.g. traffic light state, right of way information) evidence.

2.2.3 Situation

A *situation* is the generalization over maneuvers of several traffic participants. Hence, different situation hypotheses correspond to different maneuver combinations of relevant traffic participants. A discrete random variable B_i represents each traffic participant's possible maneuvers. The sample space Ω_i comprises the maneuver alternatives for one of n regarded

traffic participants. The Cartesian product of all n sample spaces

$$\Omega_{1:n} = \Omega_1 \times \Omega_2 \times \ldots \times \Omega_n,$$

determines the set of possible situation hypotheses. The distribution

$$p(b_1, b_2, \ldots, b_n|\mathbf{x}), \tag{2.4}$$

quantifies the uncertainty regarding the situation recognition. It is conditioned on the currently observable scene evidence $\mathbf{x}$. The combined random variables notation of Section 2.1.2 offers the alternative notation $p(\mathbf{b}_{1:n}|\mathbf{x})$ for (2.4).

2.3 Data Sets

This section introduces the data sets that are used throughout this thesis. The data sets play a key role for developing the approaches revolving around statistical learning methods as well as for evaluating their performance. To emphasize the applicability of the proposed approaches the evaluations in this thesis exclusively use real-world data.

2.3.1 NGSIM, Interstate-80

The NGSIM data sets have been recorded with the goal of supporting the development of next generation driver behavior algorithms. They are publicly available at [40] and consist of recordings of the Interstate-80, the US Highway 101 and the Lankershim Boulevard. The evaluations focus on the Interstate-80 data set, which has been recorded in 2005. The 7 digital video cameras, mounted on a 30-story high building, recorded about 500 meters of the Interstate-80 near Emeryville, California. In total, 45 minutes of dense traffic, split into 3 streams, are available:

- 4:00pm to 4:15pm,
- 5:00pm to 5:15pm,
- 5:15pm to 5:30pm.

The data consists of tracked vehicle trajectories with a discrete sampling rate. Besides the available kinematic information, like detailed position, velocities and accelerations, additional information reaching from lane assignments to vehicle dimension are available for each tracked vehicle. The sampling rate of the vehicle trajectories is $10Hz$ Fig. 2.3 shows an image extracted from the raw video streams.

Figure 2.3: Extracted image from the NGSIM video stream recorded by camera 2 from the 4:00 pm to 4:15 pm stream.

2.3.2 IVSS, Sävenäs

The data set originates from the *IVSS* projects, [76]. These projects were instantiated by the Swedish Road Administration to drive the development for future road safety. A part from the gathered data has been used in [25] and consists of recordings of an intersection at Sävänas near Gothenburg (WGS84; Lat: 57.7327, Long: 12.0543). It is a non-signalized 3-way intersection in a part rural, part industrial area with a speed limit of $50km/h$. The vehicles coming from the northern street have to yield to traffic on the main road due to a yield sign. Fig. 2.4 depicts the layout of the intersection. The intersection is technically a 4-way intersection, however, the southern road immediately terminates in a private parking lot.
The data set has been recorded in 2007 by two cameras operating at $20Hz$ mounted on a building at the southwest corner of the intersection. Vehicle trajectories from approximately 18 hours of video recordings were extracted with a Kalman filter approach in [56]. The available data consists of absolute vehicle positions, velocities, acceleration and orientation.

2.3.3 Intersection Variability

During the PhD project a data set consisting of a large variety of different, unsignalized intersection approaches in a residential area was created. In total the data set consists of 112 intersection approaches at 85 different intersections. In contrast to the previous two data sets, recorded from an observer's point of view, this data set is composed of recordings from the test drive's ego perspective. The data is composed of two different recording setups. The first part of the data set has been recorded at the Honda Research Institute Europe GmbH in Offenbach am Main. The

(a) **(b)**

Figure 2.4: (a) Satellite image from the Sävenäs crossing. (b) Point of view from the private parking lot. Source: [1]

second part is extracted from the raw data provided by KITTI Vision Benchmark [42].

Recordings Offenbach

The test car at the Honda Research Institute is equipped with a forward-facing laser scanner and a stereo-camera. A consumer-grade GPS is used to determine the ego-vehicle's global position and the vehicle's velocity is obtained by tapping the CAN-Bus. A smoothing moving average filter is applied on the derived longitudinal acceleration information. The resulting ego-vehicle's trajectory information is resampled at $10Hz$. Furthermore, it is enriched by context information indicating whether the vehicle has right of way or possibly needs to yield to another vehicle. The annotation of the context information is the result of a subsequent inspection of the video recording footage.

KITTI

The KITTI Vision Benchmark [42] was originally developed to provide benchmarks on the field of odometry/SLAM and 3D object detection. The test car is equipped with a velodyne, as well as a localization unit that combines GPS, GLONASS and an IMU. Additionally, front facing stereo cameras are installed. This test car setup allows the same data to be extracted as the recordings conducted at the Honda Research Institute. Thus, the available raw data streams from the residential category are used to extract unsignalized intersection approaches to expand the data set. Fig. 2.5 shows an extracted unsignalized intersection approach.

Figure 2.5: Exemplary intersection approach from the KITTI Vision Benchmark [42].

2.4 Evaluation Metrics

The task of predicting the development of a scene in terms of either maneuver or situation estimation corresponds to a classification problem. The performance of such prediction systems depends on many factors such as the chosen classification algorithm and/or feature design. Thus, a variety of metrics exist to objectively compare the impact of different design decisions. The calculation of these metrics commonly assume a two-class classification problem (labels: *positive* and *negative*) and are based upon the categorization of predicted class labels into the four groups of Tab. 2.1. These categorizations form the basis for calculating the majority of evaluation metrics.
In case of probabilistic classification, the decision whether a newly predicted case corresponds to a *positive* or *negative* is based on the predicted class membership exceeding a threshold θ, i.e.

$$\begin{aligned} positive &\rightarrow p(b = positive) \geq \theta, \\ negative &\rightarrow p(b = positive) < \theta. \end{aligned} \tag{2.5}$$

true condition → prediction ↓	positive	negative
positive	*true positive* (TP)	*false positive* (FP)
negative	*false negative* (FN)	*true negative* (TN)

Table 2.1: Confusion matrix of a two-class classification problem.

Receiver Operating Characteristic (ROC)

In automobile context, maneuver and situation estimation systems form the foundation of subsequent warning or automated driving functions. Thus, it is important to correctly classify the future scenario to provide appropriate warnings or reactions. However, too many false warnings or reactions to events that do not occur easily annoy occupants. These two requirements are quantified by the *true positive rate* (TPR) and the *false positive rate* (FPR). The TPR relates the correctly predicted *positive* outcomes (TP) and the overall number of *positive* test examples n^+ by

$$TPR = \frac{TP}{TP + FN} = \frac{TP}{n^+},$$

whereas the FPR is setting the falsely predicted *positives* (FP) into relation to the overall number of *negative* test examples n^- using

$$FPR = \frac{FP}{FP + TN} = \frac{FP}{n^-}.$$

Hence, the trade-off between correctly classifying positive instances and unnecessary classified instances is expressed by tuples (FPR, TPR), i.e. the receiver operating characteristic (ROC).
This characteristic depends on the chosen value for the threshold θ in (2.5). Varying this threshold leads to unique tuples of FPR and TPR. To obtain a classifier's detailed characteristics, the threshold θ is increasingly varied in the interval $[0, 1]$. Fig. 2.6 illustrates an exemplary resulting ROC curve for 5 different thresholds θ in the interval $[0, 1]$.

Besides the capabilities to vividly illustrate the trade-off between FPR and TPR they have the advantage of being insensitive to uneven class distributions in the test set ($n^+ > n^-$ or $n^+ < n^-$).

A meta-metric to compare the ROC curves of several classifiers with a single number is given by the area under the curve (AUC). The AUC quantifies the area between the ROC curve of a classifier and its x-axis. The light blue area in Fig. 2.6 shows an example for the AUC of the displayed ROC curve. The AUC takes on values between 0 and 1, where values close to 1 are desirable and a value of 0.5 reflects the performance of a classifier operating at chance level.

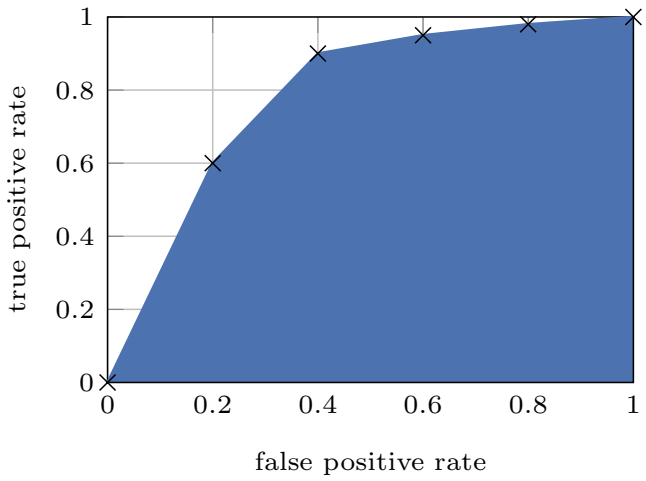


Figure 2.6: Exemplary *ROC* for five thresholds θ.

Multiclass Problems

The previous explanations addressed the evaluation of two class classification problems. However, the sample space of a maneuver and especially situation prediction problem is concerned with discriminating a large number of different outcomes of a scene. In general, classification problems with more than two classes are referred to as *multiclass* problems. In this case the 2-by-2 table of Tab. 2.1 extends to $|\Omega|_{1:n}$-by-$|\Omega|_{1:n}$ making its categorizations into *TP* and *FP* ambiguous [38].
For this reason, defining a *TP* for the case that the probability of an event $\mathbf{b^k}$ is greater or equal than the current θ and actually corresponds to the observed event $\tilde{\mathbf{b}}$

$$TP : (p(\mathbf{b}^{\mathbf{k}}_{1:n}|\mathbf{x}) > \theta) \wedge (\mathbf{b}^{\mathbf{k}}_{1:n} = \tilde{\mathbf{b}}),$$

and *FP*s as all events which probabilities equal or exceed θ and do not occur

$$FP : (p(\mathbf{b}^{\mathbf{k}}_{1:n}|\mathbf{x}) > \theta) \wedge (\mathbf{b}^{\mathbf{k}}_{1:n} \neq \tilde{\mathbf{b}}),$$

reduces the multiclass problem into a binary evaluation problem. This definition makes it possible that each of the $\tilde{n}$ tested cases possibly has several *FP*s (maximum $|\Omega_{1:n}| - 1$).
With these definitions the *TPR* is calculated as

$$TPR = \frac{TP}{n^+} = \frac{TP}{\tilde{n}},$$

and the *FPR* by

$$FPR = \frac{FP}{n^-} = \frac{FP}{N(|\Omega_{1:n}| - 1)},$$

for the multiclass case. This conversion leads to

$$n^{+,-} = \tilde{n} + \tilde{n}(|\Omega_{1:n}| - 1)$$

evaluated predictions.

3 Variable Discriminative Maneuver Estimation

Future ADAS and autonomous driving functions aim to reduce the number of accidents and relief the stress on the driver. To ensure a safe and comfortable system behavior they need to predict the next few seconds of a traffic scene. This includes estimating a traffic participant's future executed maneuver. Of particular importance are traffic scenes that form a high cognitive load on the drivers, since these scenes lead to an increased number of human errors. An example for traffic scenarios with a high variability is inner-city traffic. E.g. in the year of 2014, 69,3% of all registered accidents in Germany took place in urban environments [102]. Therefore, there is a strong demand for flexible, adaptable maneuver estimation systems that can deal with scenarios affected by a large variability like generic inner-city traffic. However, the applicability of maneuver estimation systems in variable environments has hardly been investigated.
Hence, this chapter introduces a novel and fully adaptable maneuver estimation framework that relies on discriminative classification techniques. The approach focuses on dividing the multiclass maneuver estimation problem into several binary classification tasks. To cover the possible maneuver alternatives of a currently encountered scene these binary classifiers are combined online. Thereby, the resulting maneuver estimation system is adaptable like generative approaches while providing the performance of discriminative classification techniques.

The remainder of this chapter is organized as follows. Section 3.1 discusses general problems arising in the context of probabilistic maneuver estimation for generic traffic scenes, reviews the current state of the art, and outlines the contribution of this chapter. Section 3.2 introduces the theoretical background of the proposed approach. Its applicability to the problem of adaptable maneuver estimation is shown in Section 3.3. The chapter concludes by presenting experimental results on two real-world data sets, showing the benefits of the proposed approach in Section 3.4.

The main results of this chapter have been published in [67].

3.1 Problem Statement

The difficulty regarding maneuver estimation for generic traffic scenes is that the sample space Ω generally contains different events for each encountered scene. This is due to the reason that different possibly present traffic participants and road-layouts lead to different possible events, i.e. maneuvers. The examples in Fig. 3.1 illustrate this circumstance. For each of the presented four scenes, the blue subject vehicle has different maneuver alternatives, hence varying sample spaces Ω. In Fig. 3.1 (a) Ω is mainly composed of the three path alternatives (left, right or straight) for an undisturbed, free drive, since no other traffic participant is present that could affect the execution of the subject vehicle's maneuver. In b), the left turn can be considerably influenced by the green car, e.g. it slows down to yield, stops completely or it makes a fast left turn in front of the green vehicle. Likewise, a slower preceding vehicle, as in c), can influence every possible maneuver significantly. A maneuver can even be completely missing, e.g. the right turn option in d). Many more possible maneuvers can occur when considering additional traffic participants (e.g. pedestrians or cyclists), varying infrastructure (e.g. traffic lights or yield sign) or different road-layouts (e.g. roundabouts). This problem also affects highway and rural road maneuver estimation systems and is not just limited to inner-city scenarios.

This section starts by reviewing related work regarding the topic of maneuver estimation. Afterwards, it addresses the probability theoretical problems that maneuver estimation systems are confronted with when applied to general traffic scenes and it states the main contribution.

3.1.1 Related Work

Recently, a variety of approaches addressing the problem of maneuver estimations have been presented. The following literature review focuses on approaches that estimate the maneuver of one traffic participant at a time in a probabilistic manner. Besides, a variety of approaches have been proposed that either use rule/case-based reasoning (e.g. [86], [45]) or acquire different types of logic(e.g. [5], [51]). Furthermore, approaches assessing several traffic participants simultaneously and potentially

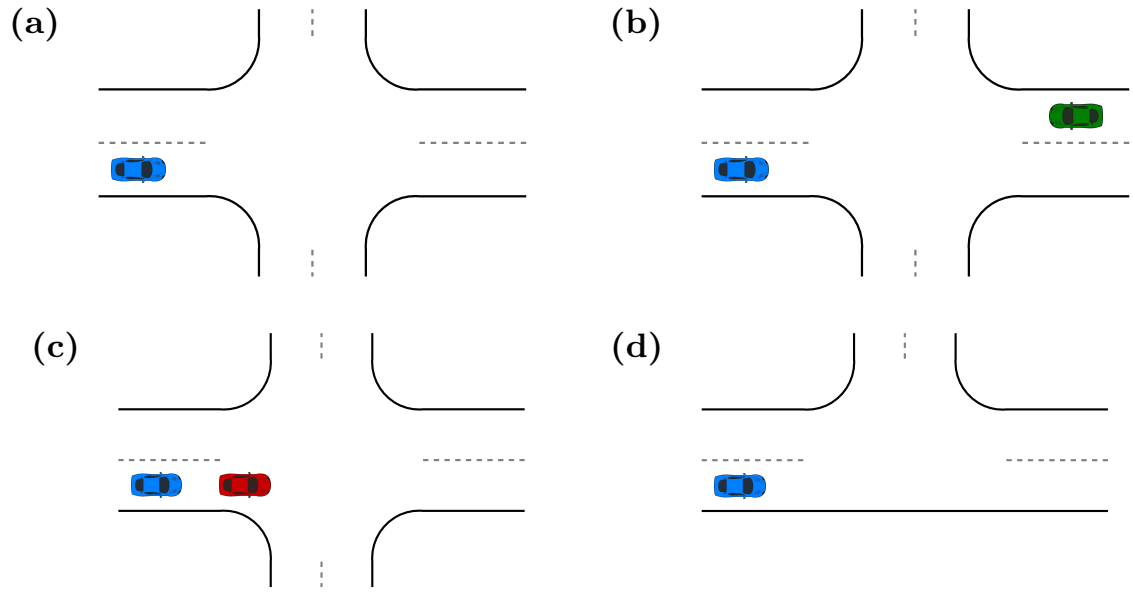


Figure 3.1: Four different scenarios, leading to four different sample spaces Ω.

incorporating interactions between them are discussed in Chapter 4. Very few of the presented approaches directly discuss the implications of a differing number of maneuver alternatives to their proposed approaches. In order to review the capabilities of the related approaches in terms of their flexibility they are divided into two categories. This categorization follows two well-known concepts from pattern recognition which are utilized to solve the task of probabilistic maneuver estimation. These differences in tackling the maneuver estimation task are crucial for their applicability to generic traffic scenes. A detailed discussion concerning these differences is provided in Section 3.1.2. For a more general discussion of maneuver estimation systems the reader is referred to the exhaustive survey collated in [75].

In the field of pattern recognition it is typically differentiated between two types of models for solving such classification tasks: *discriminative* models and *generative* models ([11], [16], [54]). Discriminative approaches are optimized to model the desired posterior probability distribution $p(b|\mathbf{x})$ directly. The generative approaches model the joint probability distribution over the classes and the observed data, i.e. $p(b, \mathbf{x})$. The desired posterior conditional distribution can be obtained using probability theoretical calculations and/or inference algorithms. In the simplest case generative approaches model the joint distribution $p(b, \mathbf{x})$ directly. Thus, the desired

conditional distribution can be obtained using

$$p(b|\mathbf{x}) = \frac{p(b, \mathbf{x})}{p(\mathbf{x})}, \tag{3.1}$$

where $p(\mathbf{x})$ represents the prior distribution over the observations $\mathbf{x}$ and can be obtained by marginalization, i.e. $p(\mathbf{x}) = \sum_b p(\mathbf{x}, b)$.
However, directly modeling high dimensional joint probabilities is affected by the "curse of dimensionality", especially in combination with discrete sample spaces. This is due to the combinatorial "explosion" of possible states of such high dimensional joint distributions. For this reason, the majority of generative approaches model their joint distribution by a factorization into partial conditional distributions while not infrequently incorporating helper and/or unobservable latent variables. For instance this is done by Bayesian networks or hidden Markov models. The downside of these factorizations is that it typically is a challenging task to infer conditional distribution, like $p(b|\mathbf{x})$. Therefore, specialized inference algorithms like *belief propagation* [83] or the *forward algorithm* [94] have been introduced. Nevertheless, a popular and easily inferable type of generative models arises from applying Bayes' rule to (3.1). The inference task is reduced to the calculation of

$$p(b|\mathbf{x}) = \frac{p(\mathbf{x}|b)p(b)}{p(\mathbf{x})}, \tag{3.2}$$

where the core of such models is the generative distribution $p(\mathbf{x}|b)$ that directly models the relationship between a certain class b (maneuver) and the observed data $\mathbf{x}$. $p(b)$ represents the prior probability distribution over the investigated classes (maneuvers).

These differences in tackling probabilistic classification problems and therefore the maneuver estimation problem are the reasons why this section divides the related work into discriminative and generative approaches. The probability theoretical differences are an important factor when considering the flexible applicability of maneuver estimation approaches in generic traffic scenes, and thus form the basis for the following literature discussion.
We contributed tow approaches, one targeted at discriminative and on targeted at generative maneuver estimation, that are discussed in this literature review as related work. However, their detailed description and discussion are beyond the scope of this thesis due to differing thematic focuses.

Discriminative Maneuver Estimation Approaches

The approaches categorized as discriminative adapted models are based on building one model in a discriminative manner in order to distinguish between a fixed and predefined set of maneuver hypotheses. This model can be used to directly asses the current maneuver based on its inherent discrimination function. Fig 3.2 shows the general structure of these models, where the conditional maneuver estimates are obtained by a not further specified classification model $\mathcal{M}$ that takes the current scene information $\mathbf{x}$ as input. Discriminative models are typically created using statistical machine learning methods in order to determine the free parameters of the underlying discrimination function.

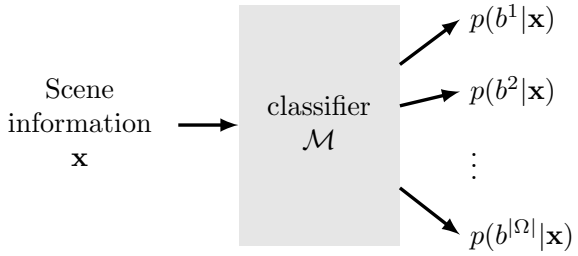


Figure 3.2: General structure of discriminative maneuver estimation approaches. One model directly discriminating between m maneuver hypotheses.

The approaches presented in this section differ from each other mainly in their underlying classification algorithms, target scenario and incorporated scene information.
Discriminative maneuver estimation systems for highway scenarios have been proposed in several works. An early work investigating the applicability of discriminative models for detecting upcoming lane change events is presented in [79]. They used a support vector machine (SVM) and came to the conclusion that they are well suited for the task and offer superior detection classification[1]. The authors of [82] presented a real-time capable approach based on a Bayesian extension of the SVM, the relevance vector machine (RVM). They additionally discern between lane changes to the right and left. The incorporation of head position was identified as the most expressive feature in order to make reliable predictions up to 3 seconds before the maneuver was actually performed. However, driver

[1] Though support vector machines do not naturally output class probabilities they can be easily extended to do so, e.g. using Platt scaling [93].

monitoring is rarely existent and even if available, the features cannot be used to predict other traffic participants maneuvers.

Hence, approaches incorporating more and more context features in order to boost classification performance have been presented. The authors of [97] present a combined approach for predicting a vehicle's future lateral and longitudinal motion based on a prior probabilistic maneuver estimation. They discriminate between three different maneuvers: lane change to the left, lane change to the right, and lane following. In addition to the subject vehicle's kinematic states they take up to eight surrounding vehicles into account to reliably estimate the maneuver alternatives. Random decision forests are used to assign probabilities to the individual maneuver alternatives.

Another popular area of application for maneuver estimation systems is inner-city traffic. Inner-city scenes' maneuvers are mainly composed of different turning options and thereby occurring disturbances from road infrastructure (e.g. traffic lights) and/or other affecting traffic participants (e.g. yielding). The authors of [41] used multilayer perceptrons (MLP) in order to discern between different reoccurring maneuver primitives in inner-city traffic. Their trained multilayer perceptron is able to reliably differentiate between the investigated maneuver hypotheses *stopped*, *braking* and *other* with a prediction horizon of 3 seconds. Although it is important to recognize deceleration patterns reliably, inner-city traffic is also dominated by turning maneuvers. Hence, the authors of [108] used a combination of log-linear model and conditional random fields (CRF) to classify turning maneuvers at intersections. Fig 3.3 illustrates the developed CRF. Their discriminative approach differentiates between a right or respectively left turn, and a straight going maneuver at a time without considering the presence of other traffic participants. However, other dynamic traffic participants can drastically influence the dynamic execution of maneuvers, e.g. following and yielding scenarios and are therefore crucial for maneuver estimation systems in inner cities.

Hence, the author of this thesis and his colleagues have presented an approach in [66] where they used a combination of Bayesian networks and a discriminatively learned logistic regression in order to reliably estimate *stopping at traffic light*, *right turn*, *going straight* and *car following* maneuvers. Fig. 3.4 shows the structure of the Bayesian network. A slow driving vehicle in front of the subject vehicle can drastically influence the temporal execution of maneuvers, e.g. a vehicle in front that is currently preparing a right turn maneuver significantly slows down the preceding

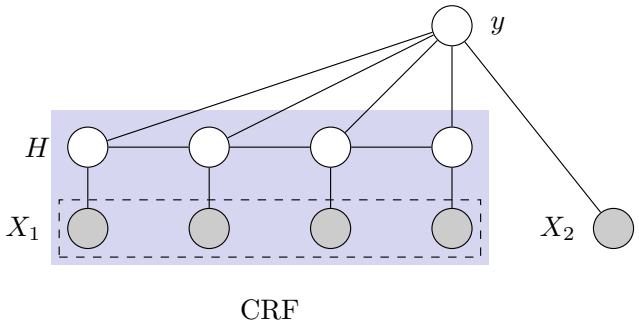


Figure 3.3: Overview of the system the author proposed in [108]. The blue rectangle indicates die conditional random field for modeling the temporal dependencies. It consists of the hidden states H and the input sequence X_1. The hidden states from the CRF and the features from the current time step X_2 form the input of the log-linear model y.

vehicle. Hence, the kinematic states of the preceding vehicle have to be incorporated explicitly. This approach is able to reliably estimate the currently pursued maneuver several seconds before entering the intersection. Another approach directly incorporating other traffic participants kinematic states has been presented in [114]. They investigated the classification performance of several discriminative classification algorithms in yielding scenarios and differentiated between 4 frequently occurring temporal behavior patterns (*go*, *creep*, *stop* and *no action*) while performing a right turn. The possibly affecting traffic participants with priority have been explicitly incorporated into the feature vector of the classification algorithms.
The presented approaches offer reliable estimations in the scenarios they were designed for due to their discriminative adopted models. However, they are in general highly specialized for a given scenario and cannot straightforwardly be reused in different contexts, e.g. missing or additional possible maneuver alternatives. This is a problem since designing specialized models for all possible scenes is unfeasible.

A hierarchical approach addressing the mentioned problems is proposed in the successive publications of [20], [18], and [19]. The hierarchical structure is based on the idea of having different models with varying degree of scenario-specific specialization. Based on the current context these specialized models can either be turned on or off. Either way the general nodes make maneuver predictions that have to eventually compete with the predictions obtained by the specialized nodes. One example they discussed and evaluated investigates lane change maneuvers. They differentiate be-

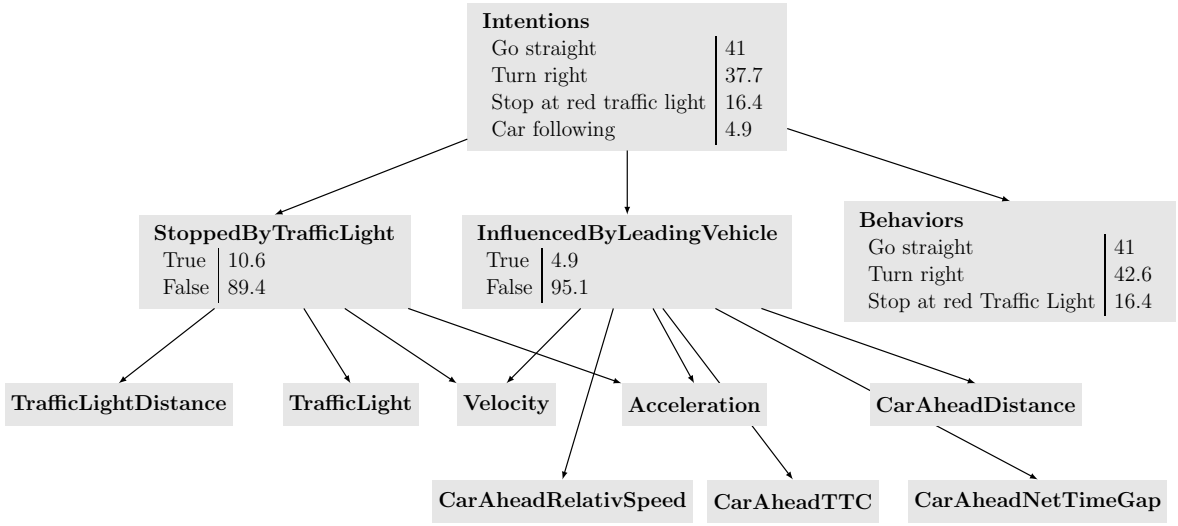


Figure 3.4: Bayesian network architecture used in [66] in order to discern between 4 different maneuvers. It is based on a discriminatively learned *Behavior* node and so called configurations which model the influence of a preceding vehicle and a present traffic light.

tween lane following and lane change to the left. The node-specific maneuver predictions are based on single-layer perceptrons. Figure 3.5 illustrates the general idea of the proposed framework. While this approach is the only discriminative maneuver estimation system considering changing environments, it suffers from the number of specialized systems required to deal with a wide variety of scenarios, since the specialized systems cannot cope with variability.

Generative Maneuver Estimation Approaches

Generative maneuver estimation approaches focus on modeling the joint distribution $p(b, \mathbf{x})$ either directly or by one arbitrary possible factorization of the joint distribution. The target conditional maneuver probabilities $p(b|\mathbf{x})$ can be obtained by inference algorithms specific to the underlying model. Fig. 3.6 shows the general procedure for using this type of generative models for probabilistic maneuver estimation.

Bayesian networks are popular for modeling complicated joint distributions. They use a possible factorization to tackle the complexity of the underlying joint distribution. Hence, Bayesian networks are used in several maneuver estimation systems, e.g. [59], [73], [99], [100], and [115]. In [73], a comprehensible Bayesian network structure is presented for inferring on which lane a vehicle will exit an intersection. Bayesian networks offer the

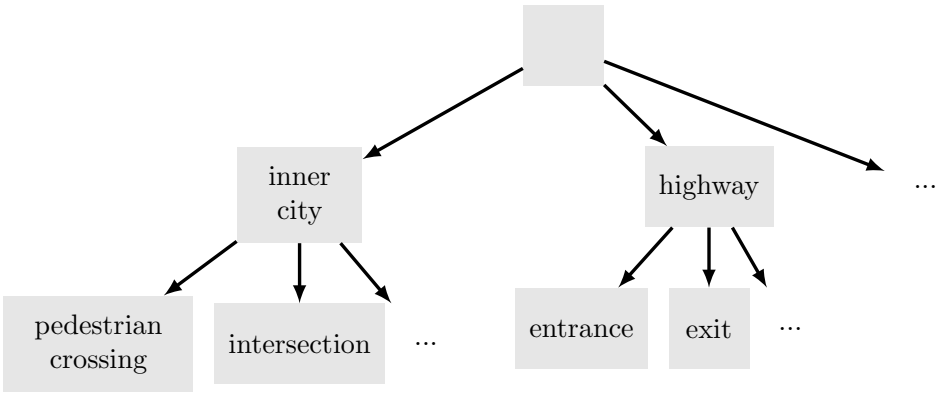


Figure 3.5: Schematic drawing of the hierarchical framework the authors proposed in [20]. Each node, with exception of the root node, represents a scenario-specific maneuver classifier. Digital map data is used to determine the active nodes on each level.

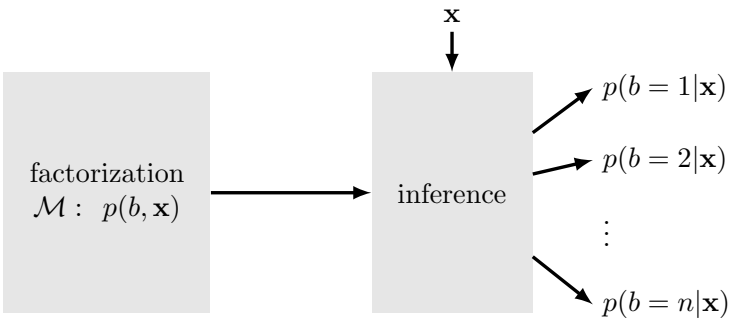


Figure 3.6: General structure of generative maneuver estimation systems approaches. An inference algorithm is needed to determine the target conditional distribution $p(b|\mathbf{x})$ based on the model for the joint distribution $p(b, \mathbf{x})$ and the observed scene evidence $\mathbf{x}$.

possibility to deal with uncertain evidence. This feature is used to incorporate uncertain information regarding the intersection entrance of the subject vehicle. However, in context of a greater variety of maneuvers and additional scene evidences Bayesian networks suffer from their increasing complexity and can therefore not be designed in the same straightforward manner as proposed in [73].

For this reason, the authors of [99] and [100] used a hand-designed Bayesian network with a variety of helper nodes, from now on referred to as $\mathbf{h}$, to parametrize the complicated target distribution. As a result the Bayesian networks describing the joint distribution are dependent on conditional distributions of the form

$$p(b, \mathbf{x}) \propto p(\mathbf{h}|\mathbf{x}),$$

or

$$p(b, \mathbf{x}) \propto p(\mathbf{x}|\mathbf{h}),$$

or a combination for both.

Another option for dealing with the complex structure sophisticated Bayesian networks suffer from is presented by the authors of [59]. They introduced the "object-oriented Bayesian networks", which allow for a hierarchical modeling of subproblems within the maneuver estimation, e.g. detecting a lane marking crossing before actually modeling the lane change detection. In a subsequent publication ([115]) the authors address the real-time applicability of such systems.

The authors of [117] present an alternative method for modeling the joint distribution $p(b, \mathbf{x})$. It is based on their previously published trajectory prediction framework [116]. The target joint distribution is modeled using a Bernoulli-Gaussian mixture model covering the discrete maneuver alternatives b and the feature vector $\mathbf{x}$ consisting of the current kinematic states of the subject vehicle, a possibly present leading vehicle and the traffic light state.

The discussed approaches modeling the joint distributions using Bayesian networks or Gaussian mixture-models suffer from their complicated structure and their costly inference algorithms. Furthermore, due to their arbitrary factorizations and the usage of helper nodes it cannot be expected that these systems can be straightforwardly adapted to additional or missing maneuver alternatives. This is due to the fact that the conditional distributions, accomplishing the complexity reduction, are highly dependent on a predefined maneuver set. Section 3.1.2 discusses the probability theoretical problems occurring when using such models

for variable traffic scenes.

A popular sub-category of generative models use a maneuver-specific factorization of the joint probability distribution based on conditional likelihoods $p(\mathbf{x}|b)$ to determine $p(b|\mathbf{x})$ according to (3.2). Fig. 3.7 shows the common structure of this type of approaches. The maneuver-specific models investigate to which degree the currently observed scene evidence fits the scene evidence that can be expected by a specific maneuver hypotheses. The presented approaches differ mainly in their determination/specification which $\mathbf{x}$ can be expected by certain maneuver hypotheses and the utilized similarity measure.

One possibility to tackle this problem is to use models from the field of pattern recognition that directly output $p(\mathbf{x}|b)$. The authors of [57], [104], and [105] build separate hidden Markov models for each maneuver in order to determine $p(\mathbf{x}|b)$. A similar approach is proposed in [13], where a linear discriminant analysis is used to model $p(\mathbf{x}|b)$.

Another possibility to obtain an estimate of the conditional likelihood is to perform a detailed prediction for a subset of the kinematic states $\hat{\mathbf{x}}_{\subset}$ and to compare it with the actual measured scene evidence $\mathbf{x}_{\subset}$. This

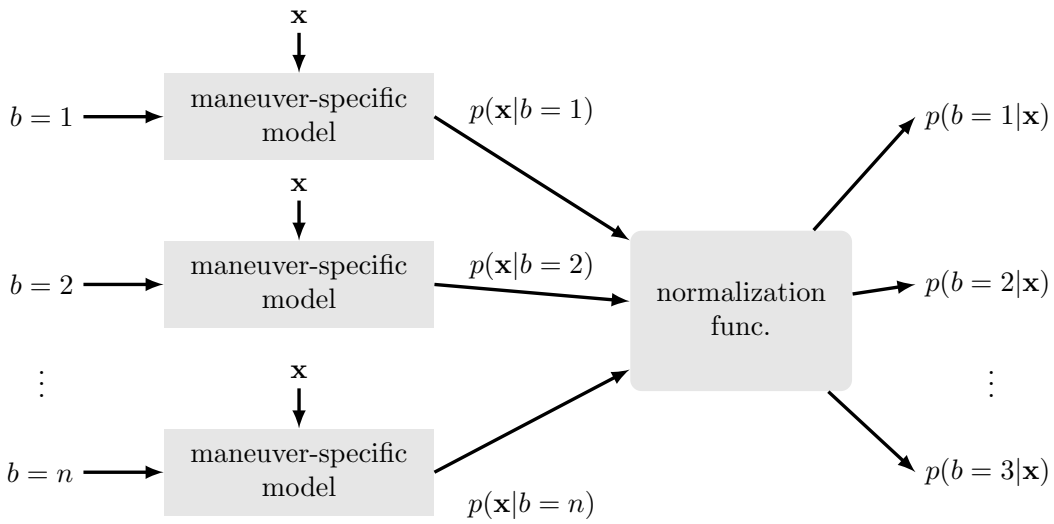


Figure 3.7: Typical system structure of generative models relying on the determination of the conditional likelihood $p(\mathbf{x}|b)$. Each maneuver is investigated individually, either by direct probabilistic models (e.g. hidden Markov models), maneuver-specific motion predictions (e.g. Gaussian processes or driver models), indicator functions or trajectory similarity measures. The subsequent normalization yields the target distribution $p(b|\mathbf{x})$.

results in a distribution of the form $p(\mathbf{x}_{\subset}|\mathbf{x}_{\setminus\subset}, b)$ for which

$$p(\mathbf{x}|b) \propto p(\mathbf{x}_{\subset}|\mathbf{x}_{\setminus\subset}, b)$$

is assumed. Such an approach is pursued by the authors in [109], [110], and [4]. They use maneuver-specific Gaussian process regression models for predicting the mean and variance of the expected velocity for that maneuver hypotheses. The resulting normal distribution determines the conditional likelihood $p(\mathbf{x}|b)$ for each maneuver hypotheses. Gaussian processes are powerful nonlinear non-parametric regression methods, however, they suffer from high computational costs and their assumption of a homoscedastic variance. The authors of [4] specifically address the homoscedasticity and propose to use a heteroscedastic extension for Gaussian processes, which in turn further facilitate the high computational demands. The same idea is pursued by the authors of [77]. Instead of using nonlinear regression models for predicting maneuver-specific velocities they use a parametric driver model for predicting maneuver-specific velocities. They use an extension of the intelligent driver model that incorporates the influence of road curvature, and thus, is able to discriminate between turning maneuvers. The velocity estimates for different maneuver hypotheses form the mean of a normal distribution with fixed variance for inferring $p(\mathbf{x}|b)$. In a subsequent publication [78] they incorporate more elaborate features like gaze direction. However, such features are rarely available, especially in the case where maneuvers of other traffic participants are estimated. Another approach based on their work is presented by [89]. The authors coupled the velocity estimates with an extended Kalman Filter.
These approaches rely on splitting the available scene evidence into two parts. The first part is used to predict the states of the second part (e.g. velocity) for the subsequent comparison. As a consequence, potential is being lost since only a part of the available scene evidence is used to determine the conditional likelihood $p(\mathbf{x}|b)$.

Nevertheless, instead of dynamically predicting expected kinematic states based on a subset of the available scene evidence, they can be directly compared with maneuver-specific static indicator functions. These indicator functions are designed in a way that they return a value between 0 and 1, where a 1 stands for perfect compliance between the observed evidence and the respective indicator function[2]. The authors of [26] and [27] discriminate between turn alternatives at intersections using indicator

[2]It should be noted that this is not an inherent way of modeling probabilities but can be interpreted as such.

functions including, among other things, the steering wheel angle, gaze direction and velocities. A similar approach is presented in [98] , where the authors propose the usage of indicator functions for detecting lane change maneuvers. The indicator functions' parameters are optimized using evolutionary algorithms.

A remaining option for determining $p(\mathbf{x}|b)$ is to use trajectory comparison techniques. The idea behind these type of approaches is to compare previously encountered and/or predicted trajectories of a certain type of maneuver with the currently observed trajectory. The authors of [58] propose a motion matching technique based on the *longest common subsequence*. Additionally, they choose a quaternion-based trajectory representation so that their approach becomes rotational invariant. Trajectory comparison techniques, such as the *longest common subsequence* typically suffer from high computational costs. Moreover, two circumstances strongly influence the shape of a trajectory in traffic scenes that need to be taken into account. The highly structured environment and the holonomic constraints of vehicles typically allow less lateral than longitudinal movements. However, a slight lateral offset that is rather uncritical for trajectory similarity, e.g. not driving at the exact center of a lane, accumulates to a large overall difference. Therefore, the author of this thesis and his colleagues proposed a novel trajectory similarity measure in [31, 63]. It explicitly differentiates between lateral and longitudinal errors for the comparison. Hence, the approach properly asses the similarity of trajectories that have a slightly lateral offset. Instead of relying on a motion database for acquiring maneuver-specific prototypical trajectories, the approach uses a general driver model developed in [36]: the *foresighted driver model*.

Comparison and Discussion

The presented discriminative approaches promise to offer reliable maneuver estimations in the scenarios they were designed for. Additionally, as discriminative models model the target distribution $p(b|\mathbf{x})$ directly, the calculation during runtime is typically short, e.g. logistic regression, neural networks or support vector machines. This is an important factor when considering the application for automotive purposes due to their comparatively low computational requirements. However, these discriminative approaches are typically designed for one scenario and cannot be straightforwardly adapted to changing environments.

Generative approaches on the other side, model the target conditional distribution indirectly and solve a more general problem, which might be unnecessary for the task at hand [113]. Nevertheless, this extra knowledge makes them generally more flexibly applicable to changes of the underlying probability distribution, e.g. missing or additional maneuver alternatives. However, building such general models comes at a price since they are not adapted to optimally solve the maneuver estimation task at hand and therefore often do not reach the same amount of generalization capability and performance compared to models adapted in a discriminative manner ([17], [84]). This hypothesis is strengthened by the domain-specific work of [3], [53], [79] and [107]. The authors of [107] compare different generative and discriminative maneuver estimation systems for turning scenarios. Their evaluation showed that the system based on maneuver specific hidden Markov models is outperformed by the systems based on one discriminative model. In [3] a comparison of support vector machines and hidden Markov models is performed for classifying the driver behavior as violating or compliant with similar results. Another work showing the benefits of using discriminative approaches has been presented in [53], where a LSTM recurrent neural network based maneuver estimation outperforms several variants of hidden Markov models while discerning between "left lane change", "right lane change", "left turn", "right turn" and "driving straight" maneuvers.

At this point it is important to mention that contrary results have been presented in [52], where support vector machines and random forest classifier based maneuver estimation systems have been outperformed by hidden Markov models. However, a closer look reveals that this can primarily be explained by the inferior temporal modeling used for the discriminative algorithms compared to the generative approaches. Time-series data of 5 seconds is concatenated in a brute force manner into one feature vector resulting in an input space dimensionality of 3840 for the SVM approach. As stated in [33], such excessive input space dimensionality needs a tremendous amount of training data in order to achieve acceptable generalization capabilities, however, the authors use a training data set consisting of a few hundred examples.

The addressed general performance differences between generative and discriminative approaches are a simplified depiction. Generative approaches can in fact reach the same performance levels when adapted with the objective of optimizing the classification task of determining $p(b|\mathbf{x})$. This means that the distributions required in order to obtain $p(b|\mathbf{x})$ from

general models (e.g. $p(\mathbf{x}|b)$ or $p(b, \mathbf{x})$) are optimized in a discriminative manner to yield the best possible $p(b|\mathbf{x})$ ([17], [21]). A variety of discriminative adaptation of generative models have been proposed, e.g. naive Bayes [50], hidden Markov model (HMM) [29], linear discriminant analysis (LDA) [23] and Gaussian mixture models [62]. As an example a Bayesian network [46] can be adapted in a discriminative manner, i.e. the conditional probability tables are optimized in order to yield the best performance for target distribution $p(b|\mathbf{x})$ on the training data instead of focusing on modeling the joint distribution $p(\mathbf{x}, b)$ or $p(\mathbf{x}|b)$ as accurate as possible. As a result, these distribution do not capture the underlying probability distribution and thus cannot be used in the same generative manner and suffer from the same problems discriminative approaches are confronted with in context of differing maneuver alternatives, as discussed in the following.

3.1.2 Probability Theoretical Considerations

The literature review reveals that the related works hardly take the variability of traffic scenes into account. Therefore, the following reviews the general capabilities of presented discriminative and generative maneuver estimation systems to be adapted to a currently encountered scene and the concomitant scene-specific sample space Ω.

The set $\tilde{\Omega}$ is the union of all maneuvers of each possibly encountered scene. This union comprises maneuvers from inner-city scenes as well as from highway and rural road scenes. Consequently, the sample space for a current scene is given by

$$\Omega = \tilde{\Omega} \setminus \Omega_{\subset},$$

with $\Omega_{\subset}$ containing all events which are incompatible with that scene (e.g. right turn for Fig. 3.1 d)). As a result the sample space Ω is a unique combination of discrete maneuver alternatives for each encountered scene, and thus models for obtaining the discrete target distribution $p(b \in \Omega|\mathbf{x})$, $b \in \Omega$ need to be adapted for each scenario. If this is not the case, constant models do either assign probabilities to maneuvers that are not possible given the current context or do not assess context-specific additional maneuvers at all.

The discriminative approaches discussed in Section 3.1.1 directly estimate the desired conditional probability $p(b|\mathbf{x})$ using one discrimination function. While these approaches promise to deliver superior classification

results, they can hardly be adapted to generic traffic scenes since their discrimination function is bound to its predefined maneuver set. Nevertheless, two naive approaches come to mind for using discriminate approaches for generic traffic.

Normalization

One model can be adapted, discriminating between all theoretically possible maneuvers $\tilde{\Omega}$. This approach suffers from two major drawbacks: First, the complexity of the required classification algorithm rapidly increases with increasing $|\tilde{\Omega}|$. Second, models trained for estimating $p(b|\mathbf{x}), b \in \tilde{\Omega}$ make also probability estimates for not possible maneuvers $\bar{b} \in \Omega_{\subset}$. These additional and unnecessary probability values can be removed from the resulting posterior conditional distribution by normalization

$$p_{\text{norm}}(b|\mathbf{x}) = \frac{p(b|\mathbf{x})}{\sum_{b' \in \Omega} p(b'|\mathbf{x})}. \tag{3.3}$$

However, it is not guaranteed that the internal discrimination function of a model estimating $p(b|\mathbf{x}), b \in \tilde{\Omega}$ also yields valid predictions for the reduced sample space Ω. The example in Fig. 3.8 clarifies this circumstance. This example assumes that the comprehensive sample space of all ever possible maneuvers is given by $\tilde{\Omega} = \{\text{right}, \text{straight}, \text{left}\}$. A discriminative model discerning between these $|\tilde{\Omega}| = 3$ maneuvers is applied to a scenario with no possible right turn option, due to the road-layout. However, the model outputs probability values for each maneuver alternatives, regardless of

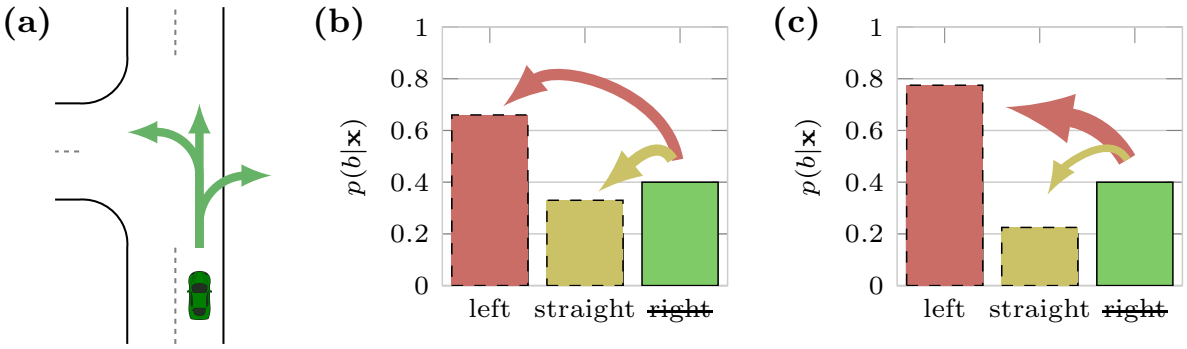


Figure 3.8: Illustration of the problems associated with normalization. (a) general maneuver estimations system is applied to a special scenario. (b) shows the probability mass functions obtained for a straightforward normalization. (c) indicates a probability mass function that is more likely to occur given a slow intersection approach and the information that only one turn option is possible.

the context. Tab. 3.1 shows the exemplary maneuver probabilities for a typical intersection approach with moderate velocity.

(a)

b^j	$p(b = b^j \mid \mathbf{x})$
left	0.4
straight	0.2
right	0.4

(b)

b^j	$p(b = b^j \mid \mathbf{x})$
left	2/3
straight	1/3
~~right~~	0

Table 3.1: Exemplary conditional probability table for a vehicle slowly approaching a T-Crossing. (a) shows the full conditional probability table for a general systems assuming that right and left turns as well as a straight going maneuver is possible. (b) shows the normalized distribution, excluding the probability value for the right turn maneuver.

Such a posterior conditional probability table suggests that the system is uncertain whether the vehicle is conducting a right or left turn. Taking a closer look at the conditional probability of Tab. 3.1 reveals that the maneuver estimation system is rather confident that a turning maneuver is going to be performed in contrast to a straight going maneuver. Ignoring the right turn maneuver and normalizing the remaining maneuver hypotheses according to (3.3) retains the original ratio between the right turn and straight going maneuver ($\frac{0.4}{0.2} = 2 = \frac{2/3}{1/3}$). Despite, a straightforward exclusion and normalization does not take advantage of the original assessment that a turn maneuver is most likely to happen. A specialized system, targeted at discerning between left turns and straight going maneuvers, and thus indirectly taking advantage of the knowledge that another turn option is not possible, would drastically increase the probability of a left turn occurring compared to the general system making maneuver predictions also for a right turn.

Additionally, special care has to be taken with regard to the evidence vector $\mathbf{x}$ as not all values might be existing in the current scene (e.g. the curve radii of a non-existing right turn or a relative velocity to a not existing preceding vehicle). Missing values pose problems for the majority of classification algorithms, and thus need to be treated carefully.

Specialization

For each possibly encountered scenario a specialized model can be created. Such an approach suffers from two major drawback: First, in order to have

an appropriate model for each theoretically occurring scene

$$n_{\text{spec}} = 2^{|\tilde{\Omega}|}, \tag{3.4}$$

models would be needed for the maneuver estimation of a single traffic participant[3]. With an increasing number of possibly ever considered maneuvers such an approach becomes quickly infeasible due to the exponentially growing number of needed models.
Second, such specialized models are targeted at one concrete scene and are therefore trained with examples of that exact scenario. However, it is questionable whether these models reach the same degree of generalization capabilities when considering training data sets of finite size. This means that such models cannot benefit from regularities present in slightly different scenes to improve a model's generalization capabilities. For example, a specialized model targeted at discerning right turns from straight going maneuvers cannot utilize regularities from other observed right turns observed at crossings with additional possible left turn option.

Generative Models and Generic Traffic

The generative approaches, discussed in Section 3.1.1 create a general model for the joint probability directly or explicitly use a factorization of $p(b, \mathbf{x})$ conditioned on the likelihood $p(\mathbf{x}|b)$. This design decision does affect the applicability of generative models to generic traffic scenes. The crucial point for their applicability is whether their factorizations include distributions of the form

$$p(b, \mathbf{x}) \propto p(b|\mathbf{h}),$$

where $\mathbf{h}$ represents a set of helper nodes, i.e. additional random variables that are added to model the underlying joint distribution or

$$p(b, \mathbf{x}) \propto p(b|\mathbf{x}_{\subset}),$$

where $\mathbf{x}_{\subset}$ is a subset of the evidence feature. The generative maneuver estimation systems presented in [100], [115] rely on these type of factorizations. In both cases neglecting an event/maneuver of b is not possible and leads to the same problems discussed while addressing the normalization problem of discriminative approaches. This is due to the reason that

[3]The number of specialized systems reduces to $n_{\text{spec}} = 2^{|\tilde{\Omega}|} - |\tilde{\Omega}| - 1$, assuming that at least two maneuvers are possible at at time.

those missing maneuvers are potentially in the target space of the helper factorizations like $p(b|\mathbf{x}_{\subset})$ or $p(b|\mathbf{h})$.

The group of generative approaches based on estimating the likelihood functions $p(\mathbf{x}|b)$ investigates which maneuver b explains the observed evidence $\mathbf{x}$ best. Probabilistic maneuver estimates can then be obtained using the Bayes' Rule

$$p(b|\mathbf{x}) = \frac{p(\mathbf{x}|b)p(b)}{p(\mathbf{x})} \propto p(\mathbf{x}|b)p(b). \tag{3.5}$$

The prior distribution $p(\mathbf{x})$ is most often used to ensure that the conditional probability estimates obtained from likelihood functions guarantee $\sum_{b\in\Omega} p(b|\mathbf{x}) = 1$ and is thus rarely modeled explicitly. The distribution $p(b)$ represents the prior probability of a certain maneuver but is usually assumed to be uniform to not bias the maneuver estimation. Hence, (3.5) simplifies to

$$p(b|\mathbf{x}) \propto p(\mathbf{x}|b).$$

Under these circumstances, not making likelihood predictions for $\bar{b} \in \Omega_{\subset}$ achieves the adaptation to the currently possible sample space. Using a generative approach for the example in Fig. 3.8 (a) illustrates the adaptation possibilities compared to the discriminative classifier that discerns between right, left and straight turn maneuvers. The scene does not allow for a right turn maneuver and thus solely the models that estimate the likelihood $p(\mathbf{x}|b = \text{left})$ and $p(\mathbf{x}|b = \text{straight})$ are selected. The exclusion of $b = \text{right}$ does not have an influence on these probability values since b is on the right side of the conditional and the models are not conditioned on a predefined set of maneuvers. This is not the case for generative models that are learned in a discriminative manner.

In cases where approaches do not assume a uniform prior, like [13], such a naive approach is not possible due the fact that the prior $p(b)$ is specified for a fixed maneuver set.

3.1.3 Contribution

Universally applicable future advanced driver assistance systems and automated driving functions cope with highly variable and complex traffic scenes. Hence, there is a need for flexible maneuver estimation systems that can be adapted to the currently encountered scene online. However, flexible maneuver estimation systems have been hardly investigated so far.

The probability theoretical discussion in Section 3.1.2 outlined that generative maneuver estimation approaches, especially those modeling $p(\mathbf{x}|b)$ can be adapted to different possible maneuver hypotheses, i.e. different sample spaces $\Omega \neq \tilde{\Omega}$. However, the literature reveals that generative models do often not achieve the same recognition performance compared to discriminative models.

Therefore, this chapter addresses the question of how a flexible maneuver estimation system can be created while simultaneously maintaining the performance levels of systems specially targeted at restricted traffic scenes (i.e. discriminative adapted models). The proposed framework is based on learning discriminative models while maintaining the possibility to be easily adapted to generic traffic scenes online. It is based on learning a pool of discriminative, probabilistic binary classifiers. Given the currently entered scene, the corresponding binary classifiers are selected capturing only the currently possible maneuver alternatives. The obtained probabilistic, binary predictions are combined online using a technique called *pairwise probability coupling*. This approach enables the maneuver prediction to benefit from discriminative models while it simultaneously offers the ability to be adaptable to generic traffic scenes. The advantages of the proposed approach are evaluated on a real-world data set.

3.2 Pairwise Probability Coupling

The idea of decomposing the overall classification task into smaller subproblems that are combined online forms the basis for the proposed adaptable discriminative maneuver estimation system.
Assembling complicated multiclass classification problems from partial, simpler classifiers has been widely discussed in pattern recognition literature. The need for such techniques arises since many classification algorithms are restricted to binary problems. Thus, the question of how these can be used for multiclass problems arises. Typically, binary classification results, covering all possible class combinations, are used to determine the most likely class by a voting scheme. However, probability estimates of class memberships pose a problem since straightforward voting schemes can no longer be applied. The problem formulation of assembling probability distributions over more than two classes with binary classifiers is identical to the challenges encountered when decomposing complicated, variable maneuver predictions into smaller, reusable sub-problems.
Therefore, this section starts with introducing the probability theoretical

formalism describing binary one vs. one (OvO) predictions and discussing the possibility to obtain multiclass probability estimates from binary predictions by formulating the underlying optimization problem and subsequently addressing its general solvability.

3.2.1 Binary One vs. One Predictions

This section starts by stating additional probability theoretical formalisms to describe the probability distributions obtained from binary classifiers. In order to make binary, OvO predictions for multi-class classification problems, all $\binom{|\Omega|}{2}$ binary subsets $\Omega^{k,l} \subset \Omega$ are extracted. For each of the $\Omega^{k,l}$ subsets a binary classifier, trained to discriminate between events b^k and b^l, determines the binomial distribution

$$p(\{b \in \Omega^{k,l} : B^{k,l} = b\}|\mathbf{x}) = p^{k,l}(b|\mathbf{x}), \tag{3.6}$$

with $B^{k,l}$ being a discrete random variable that can take on the respective events $B^{k,l} \in \Omega^{k,l} = \{b^k, b^l\}$. Note that the set $\Omega^{k,l}$ and $\Omega^{l,k}$ are identical. This means that the distributions $p^{k,l}(b|\mathbf{x})$ and $p^{l,k}(b|\mathbf{x})$ are identical, too. As a consequence, both distributions originate from the same binary classifier.

The concrete probabilities of

$$p^{k,l}(B^{k,l} = b^k|\mathbf{x}) = p^{k,l}(b^k|\mathbf{x}),, \tag{3.7}$$

$$p^{k,l}(B^{k,l} = b^l|\mathbf{x}) = p^{k,l}(b^l|\mathbf{x}), \tag{3.8}$$

denote the chance of the events in $\Omega^{k,l}$ happening. Additionally, in (3.6), (3.7), and (3.8) the same lower case abbreviation, as introduced in Section 2.1 is applied. These estimates form the basis for the subsequent probability coupling.

3.2.2 Optimization Problem

An approach for obtaining probabilistic multiclass estimates has been presented in [119]. The idea of their proposed pairwise probability coupling is based on optimizing the approximated conditional probability distribution $\hat{p}(b|\mathbf{x}) \approx p(b|\mathbf{x})$ of the desired target distribution with the objective

$$\frac{p^{k,l}(b^k|\mathbf{x})}{p^{k,l}(b^k|\mathbf{x}) + p^{k,l}(b^l|\mathbf{x})} = \frac{\hat{p}(b^k|\mathbf{x})}{\hat{p}(b^k|\mathbf{x}) + \hat{p}(b^l|\mathbf{x})}. \tag{3.9}$$

This objective is based on the idea that the probabilities obtained from the binary distributions should keep the same ratio as the estimated target probabilities $\hat{p}(b|\mathbf{x})$. A closer look at (3.9) reveals that the denominator $p^{k,l}(b^k|\mathbf{x})+p^{k,l}(b^l|\mathbf{x})$ sums up to 1, since it is equal to the complete output of the binary classifiers.
Thus, (3.9) simplifies to

$$p^{k,l}(b^k|\mathbf{x}) = \frac{\hat{p}(b^k|\mathbf{x})}{\hat{p}(b^k|\mathbf{x}) + \hat{p}(b^l|\mathbf{x})}, \tag{3.10}$$

where another interpretation of the optimization objective shows, namely that the probabilities $p^{k,l}(b^k|\mathbf{x})$ should take on the same value as $\hat{p}(b^k|\mathbf{x})$ normalized to the binary subset $\Omega_{k,l}$. Rearranging (3.10) leads to

$$\hat{p}(b^k|\mathbf{x}) = p^{k,l}(b^k|\mathbf{x})(\hat{p}(b^k|\mathbf{x}) + \hat{p}(b^l|\mathbf{x})).$$

This equation can be instantiated for each event $b^l \in \Omega$, where $l \neq k$, resulting in $m-1$ equations that all describe $\hat{p}(b^k|\mathbf{x})$ with its different partner events b^l. Adding up these $m-1$ equations leads to

$$(m-1)\hat{p}(b^k|\mathbf{x}) = \sum_{l=1,l\neq k}^{|\Omega|} (\hat{p}(b^k|\mathbf{x}) + \hat{p}(b^l|\mathbf{x}))p^{k,l}(b^k|\mathbf{x}). \tag{3.11}$$

A factorization of (3.11), with regard to the unknown probability $\hat{p}(b^k|\mathbf{x})$ yields

$$\hat{p}(b^k|\mathbf{x}) = \hat{p}(b^k|\mathbf{x})\Big(\frac{\sum_{l=1,l\neq k}^{|\Omega|} p^{k,l}(b^k|\mathbf{x})}{m-1}\Big) + \sum_{l=1,l\neq k}^{|\Omega|} \hat{p}(b^l|\mathbf{x})\frac{p^{k,l}(b^k|\mathbf{x})}{m-1}.$$

This equation is instantiated for each $k \in 1,\ldots,m$. The problem of finding appropriate probability values for $\hat{p}(b|\mathbf{x})$ based on all k equations can be reformulated as a linear equation system of the form

$$\mathbf{p} = \mathbf{Q}\mathbf{p}, \tag{3.12}$$

with

$$q^{k,l} = \begin{cases} \sum_{s=1,s\neq k}^{|\Omega|} p^{k,s}(b^k|\mathbf{x})/(m-1) & \text{if } k = l, \\ p^{k,l}(b^k|\mathbf{x})/(m-1) & \text{if } k \neq l\ . \end{cases} \tag{3.13}$$

The $(m \times 1)$ vector $\mathbf{p}$ contains the m probability values of $\hat{p}(b|\mathbf{x})$.

Equation (3.13) contains $|\Omega|$ equations with $|\Omega|$ unknown probabilities $\mathbf{p}$ that can be rewritten as a homogeneous linear equation system

$$\mathbf{0} = (\mathbf{Q} - \mathbf{I})\mathbf{p}. \tag{3.14}$$

As a consequence, the desired probability distribution $\hat{p}(b|\mathbf{x})$ can be determined directly. However, since the coefficient matrix $\mathbf{Q}$ differs in its dimension and values due to the varying cardinality of the sample space and the binomial distributions from the *OvO* classifiers this is not immediately obvious. Thus, the following provides a proof that it is in general possible to determine $\hat{p}(b|\mathbf{x})$ using the coefficient matrix $\mathbf{Q}$.

3.2.3 Analogy to Discrete Markov Chains

The linear matrix equation of (3.12) has the same functional form as discrete Markov chains defined on a finite state space. A Markov chain describes a specific kind of stochastic model, i.e. a sequence of possible events. The special characteristic of a Markov chain is that the probability of an event occurring only depends on the state of the model in the previous step [16].

The analogy of (3.12) to discrete Markov chains allows for a profound discussion regarding its general solvability and approaches for determining $\mathbf{p}$. Hence, the following formally examines the prerequisites for the existence of a unique stationary distribution as well as providing three alternatives for determining the stationary distribution.

Finite State Space and Stochastic Matrices

In order to guarantee that (3.12) in fact describes a discrete Markov chain the states' spaces need to be fixed and finite as well as $\mathbf{Q}$ needs to be a stochastic matrix. The state space of that discrete Markov chain corresponds to the possible maneuver alternatives Ω. The set of possible maneuvers is fixed and finite for a given scene.

Two properties determine whether a matrix is stochastic or not. The **first** property states that the elements of $\mathbf{Q}$ need to be in the range

$$0 \leq q^{k,l} \leq 1, \ \forall k, \forall l \in 1, \ldots, m. \tag{3.15}$$

Equation (3.13) specifies two possible construction rules for the elements $q^{k,l}$. The elements for $k \neq l$ are necessarily in the range $0 \leq q^{k,l} \leq 1$ since the probabilities $p^{k,l}(b^k|\mathbf{x})$ are in the range $[0, 1]$ by definition and

the division by a positive integer $(m-1)$ does not alter this circumstance. For the elements where $l = k$ applies the nominator consists of a sum of probability values and therefore guarantees that $0 \leq q^{k,l}$. Moreover, the characteristic of the nominator being a sum of probability values assures

$$\sum_{s=1, s\neq k}^{|\Omega|} p^{k,s}(b^k|\mathbf{x}) \leq m-1. \tag{3.16}$$

Thus, $0 \leq q^{k,l} \leq 1$ is also guaranteed since the cardinality of the index set $s \in 1, \ldots, |\Omega| : s \neq k$ is equal to $m-1$ and the maximum values of each probability in (3.16) are limited to 1.

The **second** property of a stochastic matrix requires that $\mathbf{Q}$ is either row stochastic

$$\sum_{k=1}^{|\Omega|} q^{k,l} = 1,$$

or column stochastic

$$\sum_{l=1}^{|\Omega|} q^{k,l} = 1,$$

or both.

The sum of an arbitrary column l of $\mathbf{Q}$ consists of

$$\sum_{s=1, s\neq l}^{|\Omega|} \frac{p^{l,s}(b^l|\mathbf{x})}{m-1} + \sum_{k=1, k\neq l}^{|\Omega|} \frac{p^{k,l}(b^k|\mathbf{x})}{m-1}, \tag{3.17}$$

based on (3.13) where the first addend represents the case $k = l$. The two addends are independent from each other due to their differing summation indices. Nevertheless, the set for each index of summation is identical and the summation indices are unified. Thus,

$$\sum_{s=1, s\neq l}^{|\Omega|} \frac{p^{l,s}(b^l|\mathbf{x}) + p^{s,l}(b^s|\mathbf{x})}{m-1}, \tag{3.18}$$

is an equivalent representation of (3.17). The two binary distributions $p^{l,s}(b|\mathbf{x})$ and $p^{s,l}(b|\mathbf{x})$ are identical since they originate from the same OvO classifier, determining the class probabilities of the binary subset $\Omega^{l,s}$. For

the investigated column, l takes on a fixed value and the summation index s is never identical to l. Hence, for the nominator of (3.18)

$$p^{l,s}(b^l|\mathbf{x}) + p^{s,l}(b^s|\mathbf{x}) = 1,$$

applies.

Combining this observation with (3.18) results in

$$\sum_{s=1,s\neq l}^{|\Omega|} \frac{1}{m-1}, \tag{3.19}$$

for a column sum of $\mathbf{Q}$. The set $s \in 1,\ldots,|\Omega| : s \neq l$ has a cardinality of $m-1$. Thus, (3.19) is identical to 1 for arbitrary columns l.

Existence of the Stationary Distribution

After proving that $\mathbf{Q}$ is a stochastic matrix, the following investigates the property indicating whether $\mathbf{p}$ converges to a stationary value, i.e. that the stationary distribution of (3.12) exists. For this purpose $\mathbf{Q}$ needs to be irreducible and aperiodic. The term irreducibility implies that all states can be reached by a Markov chain, independently of its starting point $\mathbf{p}^{*(0)}$ ([80]). An aperiodic Markov chain does not oscillate in a regular periodic movement between different subsets of the state space [95].

This is guaranteed if

$$0 < q^{k,l} < 1,\ \forall k, \forall l \in 1,\ldots,m.$$

This is a stricter form of the proved property in (3.15). However, this property is also met if the range of the binary probability values is limited to the interval $(0,1)$. For practical implementations this restriction does not pose a problem since many classification algorithms, e.g. logistic regressions, are naturally bound to the interval $(0,1)$. If the binary probability distributions are obtained from classification algorithms where this is not the case, e.g. random forest classifiers, slightly altering the distributions without changing their implication can be applied. Table 3.2 shows an example of such a modification, where the extreme values 0 and 1 are removed without notably changing the underlying distribution.

(a)

b	$p^{k,l}(b\|\mathbf{x})$
b^k	1
b^l	0

(b)

b	$p^{k,l}(b\|\mathbf{x})$
b^k	.999
b^l	.001

Table 3.2: Example for adapting conditional probability tables in order to guarantee a stationary solution to (3.12) without changing the implication of the probability distributions.

Determining the Stationary Distribution

After proving the existence of a unique stationary distribution for the discrete Markov chain, the following illustrates ways to determine the stationary distribution. In general there are three ways for obtaining the stationary distribution $\mathbf{p}$, based on the stochastic matrix $\mathbf{Q}$.

First, the most obvious way to determine $\mathbf{p}$ is to actually calculate the discrete Markov chain, starting from an arbitrary initial distribution $\mathbf{p}^{*(0)}$. For a sufficient large number of iterations

$$\mathbf{p} = \lim_{\alpha \to \infty} \mathbf{Q}^{\alpha} \mathbf{p}^{*(0)}, \tag{3.20}$$

can be obtained, where α corresponds to the exponent of $\mathbf{Q}$.

Second, the stationary distribution can be obtained by calculating the corresponding eigenvector $\mathbf{v}$ of the eigenvalue $\lambda = 1$. This circumstance is a consequence of the definition of eigenvectors and eigenvalues λ given by

$$\mathbf{Q}\mathbf{v} = \lambda \mathbf{v}, \tag{3.21}$$

as (3.21) is equivalent to (3.12) for the eigenvalue $\lambda = 1$. Note that each irreducible, stochastic matrix has the eigenvalue 1 with eigenspace dimension of 1. Thus the stationary distribution can be calculated based on the eigenvector $\mathbf{v}$. It is identical to the normalized eigenvector

$$\mathbf{p} = \frac{\mathbf{v}}{||\mathbf{v}||_1}.$$

Third, the linear equation system, formulated in (3.12), can be used to determine $\mathbf{p}$. As stated in [39], the rank of an irreducible, aperiodic, stochastic matrix is exactly $m - 1$. Thus, in order to find a unique solution of the linear equation system (3.12), an arbitrary row of $\mathbf{Q}$ can be discarded due to its redundancy. However, adding the information that $\mathbf{p}$

corresponds to a valid probability distribution with $||\mathbf{p}||_1$, e.g.

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} (\mathbf{q}^1)^T \\ \vdots \\ (\mathbf{q}^{|\Omega|-1})^T \\ \underbrace{1\,1\,1\,\ldots\,1}_{|\Omega|-\text{times}} \end{pmatrix} \mathbf{p} = \tilde{\mathbf{Q}}\mathbf{p}, \tag{3.22}$$

where $(\mathbf{q}^1)^T, \ldots, (\mathbf{q}^{|\Omega|-1})^T$ correspond to the $|\Omega| - 1$ first rows of $\tilde{\mathbf{Q}}$, results in an inhomogeneous linear equation system with full rank. Note that discarding the $|\Omega|$-th row of $\tilde{\mathbf{Q}}$ is just one example. In general each of the $|\Omega|$ rows could have been discarded.

All of the three discussed solution procedures yield the unique stationary distribution. However, no general recommendations can be given, which of these procedures should be used, as the time required for determining $\mathbf{p}$ is strongly dependent of the efficiency of the underlying implemented algorithm as well as the size $|\Omega|$ of $\mathbf{p}$.

3.2.4 Distinction to One vs. All Strategies

The discussed pairwise probability coupling is based on *OvO* predictions. Alternative techniques for obtaining multiclass predictions from binary classifiers are based on one vs. all predictions. One vs. all (*OvA*) strategies have the advantage of needing fewer binary classifiers compared to *OvO* strategies [81]. Nevertheless, this strategy is not further pursued due to the following reason. The objective of the proposed approach is to make flexible maneuver prediction that can be adapted to the currently encountered scene. This means that the sample space Ω only contains maneuvers that are actually possible in the current scene. However, a *OvA* strategy implicitly incorporates impossible maneuver alternatives $\Omega_{\subset}$, since these maneuvers are used as "all" examples during the training process. Hence, the binomial probability distribution of each binary *OvA* classifier k actually represents that b either corresponds to the associated behavior pattern b^k or that it corresponds to a behavior pattern of the set $\tilde{\Omega}_{\setminus k}$ without k. Thus, an *OvA* approach considers all maneuver hypotheses, regardless of the current scene, and is therefore not further investigated. Fig. 3.9 illustrates the differences in using the data for training a classifier right vs. straight (a) in comparison to right vs. all (straight and left) (b). In cases where no left turn option is possible, a classifier trained on the data of Fig.

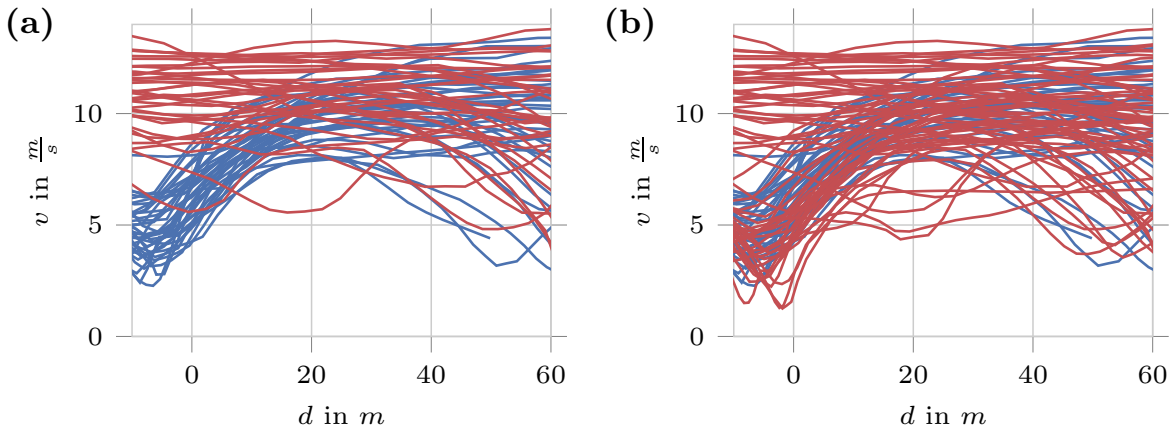


Figure 3.9: Shows the velocity profiles extracted from the data set consisting of right and left turns as well as straight going maneuvers, described in Section 2.3.3. The left part (a) shows training velocity profiles for right turn (blue) and straight going maneuvers (red). Such a data set can be used to train a *OvO* classifier for differentiating between those two maneuvers. The right part (b) shows training velocities profiles for right turn (blue) and left turn/straight going maneuvers (red). Such a data set can be used to train a *OvA* classifier, determining the likelihood of a right turn maneuver.

3.9 (b) would not capture the underlying probability distribution regarding right turns and straight going maneuvers, since it is heavily influenced by the additional training examples that do not pose valid maneuvers for the encountered scene. Additionally, the recorded left turn data affects the discrimination capabilities of such a classifier, since a right and left turn options are hard to discriminate solely based on kinematic data, [14]. For these reasons the proposed approach focuses on an *OvO* approach.

3.3 Application to Maneuver Estimation

This section illustrates the complete approach of discriminative maneuver estimation in generic traffic scenes. The idea is to assemble the conditional probability distribution $p(b|\mathbf{x})$ for the sample space Ω of a currently encountered scene from partial, reusable classifiers online. For this purpose a binary, probabilistic maneuver estimation for each combination of two events in Ω is performed. The separately estimated class probabilities for each combination are then combined online using the pairwise probability coupling presented in Section 3.2, resulting in the desired posterior conditional distribution $p(b|\mathbf{x})$.
The underlying binary classifiers of the proposed approach only discrimi-

nate between two maneuvers b^k, b^l at a time. Therefore, they are reusable in different scenes, where the overall number and types of maneuvers might vary, but the two maneuvers b^k and b^l are still possible events.

A prerequisite for such an approach is a pool of trained binary classifiers covering all possible pairs of two in $\tilde{\Omega}$, leading to

$$n_{\text{bin}} = \binom{|\tilde{\Omega}|}{2},$$

overall prototypical classifiers. This number scales much more favorably compared to that of specialized recognition systems, in (3.4).

Fig. 3.10 illustrates the complete process of probabilistic maneuver estimates using binary classifiers. A step-by-step explanation is given in the following:
First, the information on the currently encountered scene needs to be collected. This includes identifying dynamic objects, abstract context information and gathering information on the road layout using digital maps. Second, based on the acquired scene information the $|\Omega|$ possible prototypical maneuvers are identified and selected. These form the sample space Ω with $b^j, j = 1, \ldots, |\Omega|$. This rule-based selection process is based on structural information like digital map data, e.g. drivable path alternatives and different temporal execution patterns provoked by other scene

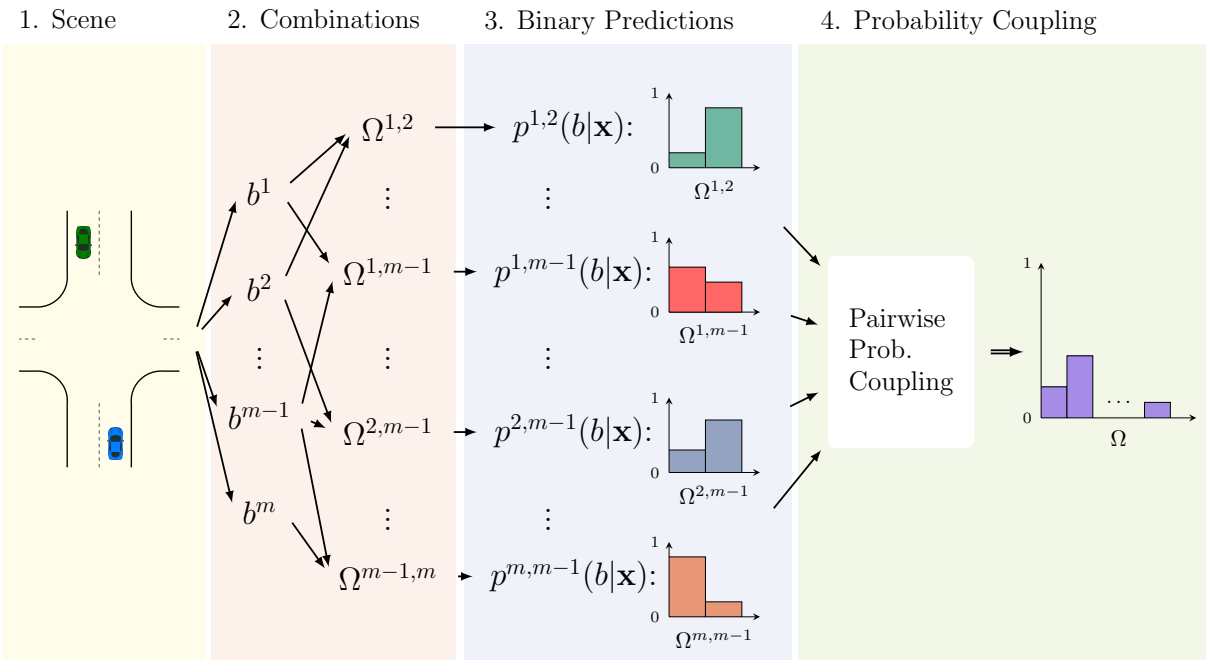


Figure 3.10: Detailed system overview, see Section 3.3 for more details.

entities. Subsequently, all $\binom{|\Omega|}{2}$ binary combinations $\Omega^{k,l}$ are instantiated. Third, for each combination, the corresponding binary classifier is selected. The relevant scene information, extracted in the first step, forms the feature vector $\mathbf{x}$. The class probabilities $p^{k,l}(b|\mathbf{x})$ for each combination are estimated using their respective classifier, separately.
Fourth, the individual binomial distributions are combined using the presented pairwise probability coupling. This results in the entire probability distribution $p(b|\mathbf{x})$ over the currently possible maneuvers $b \in \Omega$.

3.4 Experimental Results

Due to differing numbers of present traffic participants and road layouts intersection scenes are particularly variable. Thus, intersection scenes are used to evaluate the proposed approach. Two different evaluations for two different data sets are performed. The first evaluation focuses on investigating the reusability of the partial, binary classifiers in changing environments. The second evaluation focuses on investigating the performance on a complex intersection with several dynamic traffic participants resulting in a larger number of possible maneuver alternatives. First, the used evaluation metrics are explained.

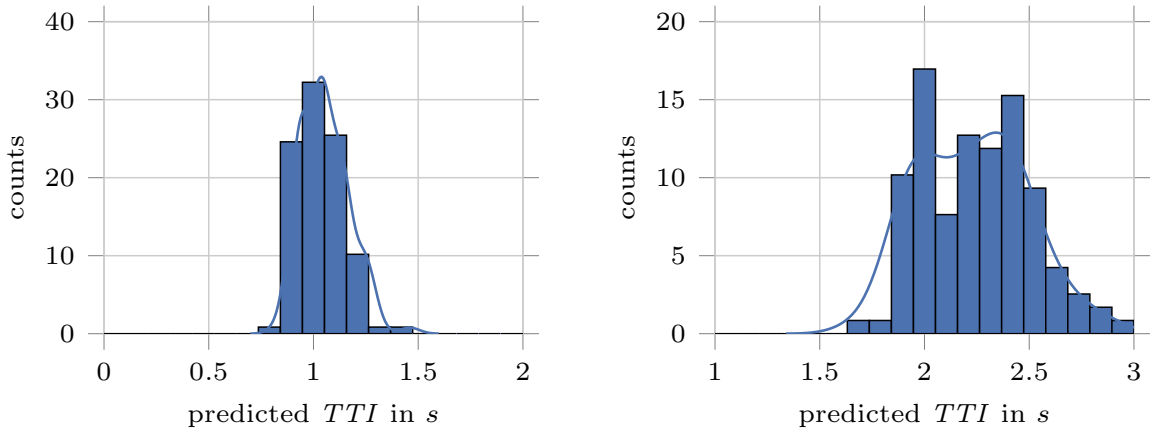
Evaluation Metrics

Maneuver estimation is affected by many uncertainties. Hence, it is also reasonable to evaluate the probabilistic output of the evaluated models. This can be done using ROC curves introduced in Section 2.4. These provide an objective metric for comparing the performances over the complete spectrum of FPR and TPR.

The AUC values are determined for different estimated time to intersections (TTI). This means that the maneuver estimation for a $TTI = 2s$ is triggered when the estimated TTI first fell below 2 seconds. In [66], we showed that assuming a constant velocity for the TTI calculation leads to conservative estimations for intersection approaching scenes. Fig. 3.11 confirms this finding for the data used herein and is thus used for the following evaluations.

3.4.1 Reusability of Binary Classifiers

In order to validate the reusability of the underlying binary classifiers, intersection approaches from the data set described in Section 2.3.3 are



(a) $TTI = 1s$, measured mean $= 1.05s$ **(b)** $TTI = 2s$, measured mean $= 2.23s$

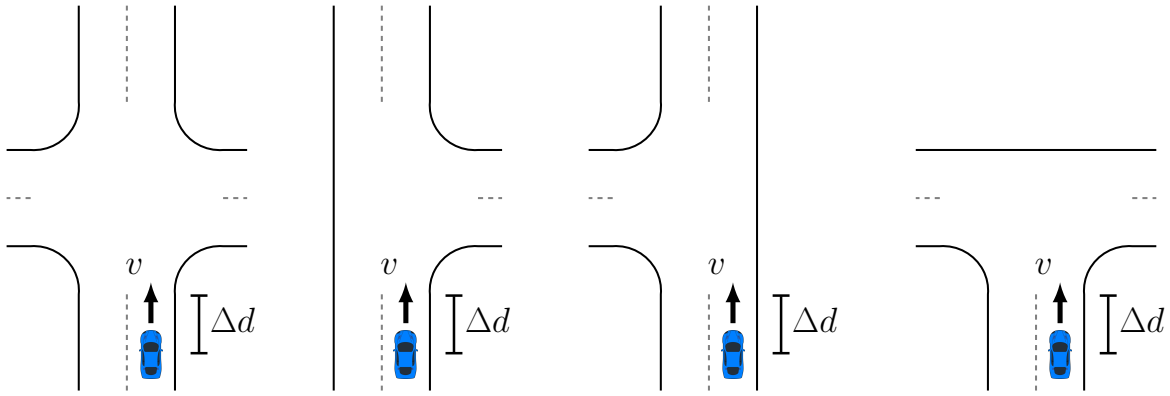
Figure 3.11: Histograms of the actual measured TTIs at a specific predicted TTI, respectively.

used. The data set consists of 112 3-way (three arms) and 4-way (4 arms) intersection approaches in an urban environment. 3 No other dynamic scene entities which could possibly interfere with the intended maneuver were present. The four different types of possible intersection approaches are illustrated in Fig. 3.12.

Evaluated Models

The binary classifiers are trained on 4-way intersection approaches for their reusability on 3-way intersections. The proposed pairwise probability coupling approach needs $\binom{|^{\mathrm{I}}\Omega|}{2} = 3$ binary classifiers for handling 4-way intersections. For the three possible maneuver alternatives of ${}^{\mathrm{I}}\Omega$ (right turn, left turn and straight crossing), the resulting binary subsets are identical to the sample spaces of the 3-way intersection crossings with sample spaces ${}^{\mathrm{II}}\Omega$, ${}^{\mathrm{III}}\Omega$ and ${}^{\mathrm{IV}}\Omega$. Tab. 3.3 clarifies this circumstance by showing the needed binary classifiers for each the four intersections in Fig. 3.12. Hence, each 3-way intersection approach can be assessed using the respective binary classifier, originally created for 4-way intersection approaches.

The subsequent evaluation compares three models. The first model $\mathcal{M}_{\mathrm{spec}}$ serves as reference. It is based on specialized classifiers for each possible 3-way intersection (b), (c), (d) in Fig. 3.12. These are trained using a 4-fold cross validation on the 3-way intersection approaches of the extracted data.



(a) ${}^{\mathrm{I}}\Omega = \{r, l, s\}$ **(b)** ${}^{\mathrm{II}}\Omega = \{r, s\}$ **(c)** ${}^{\mathrm{III}}\Omega = \{l, s\}$ **(d)** ${}^{\mathrm{IV}}\Omega = \{r, l\}$

Figure 3.12: 4- and 3-way intersection types present in the data set of [65] with the according possible maneuver alternatives (r = right turn, l = left turn, s = straight crossing). In total 112 cases (a): → 36 cases, (b): → 20 cases, (c): → 29 and (d) → 27 cases are examined.

${}^{\mathrm{I}}\Omega$	${}^{\mathrm{II}}\Omega$	${}^{\mathrm{III}}\Omega$	${}^{\mathrm{IV}}\Omega$
straight vs. right	straight vs. right		
straight vs. left		straight vs. left	
right vs. left			right vs.left

Table 3.3: Needed binary classifiers for the particular sample spaces of the four intersection types in Fig. 3.12 depicts.

The second model $\mathcal{M}_{\text{4-way}}$ is based on the proposed pairwise probability coupling approach. It is trained only on the 4-way intersection approaches. Hence, the underlying binary classifiers of this model have not been confronted with any 3-way intersection.
The third model $\mathcal{M}_{\text{4/3-way}}$ is, again, based on the proposed pairwise probability coupling approach. Additionally, training data from the 4-fold cross validation of the 3-way intersection data is used. As a result, model $\mathcal{M}_{\text{4/3-way}}$ is trained on data from 3-way and 4-way crossing, which not only adapts the binary classifiers to 3-way intersection but additionally increases the size of the training data set.

For such turning scenarios in free drive, logistic regression is successfully applied in [66] and [63]. Therefore, it is also the basis for all classifiers needed to determine the conditional distribution $p(b|\mathbf{x})$ in this evaluation.

Expressive information for distinguishing turning maneuvers are mainly composed of the distance to the intersection Δd and the velocity v of the

subject vehicle, both indicated in Fig. 3.12. Additionally, whether the subject vehicle has right of way strongly influences the temporal behavior pattern of the intended maneuver. The priority information is encoded in a binary variable w. The scene evidence contains the following features

$$\mathbf{x} = [v, \Delta d, w]^T.$$

Results

Fig. 3.13 shows the resulting AUC curves for a TTI range from $0s$ to $2.5s$. Model $\mathcal{M}_{\text{4-way}}$, never confronted with a 3-way intersection approach during training, almost reaches the performance of the specialized model $\mathcal{M}_{\text{spec}}$. This result emphasizes the possibility of reusing the same binary, partial classifiers in different scenarios, even if they never encountered that exact scenario during training. Furthermore, the results for $\mathcal{M}_{\text{4-way}}$ indicate that the generalization capabilities of reusable partial classifiers can exceed the performance of specialized systems. This is due to the fact that more training data from different scenarios can be acquired, which in turn strengthens the generalization capabilities of the underlying classifiers. E.g. in this evaluation the binary classifiers can be trained with data from 3-and 4-way intersection approaches. Tab. 3.4 shows the detailed number of training examples for each evaluated model and its underlying binary respectively specialized classifier. The table depicts that the number of training examples for the individual binary classifiers of $\mathcal{M}_{\text{4/3-way}}$

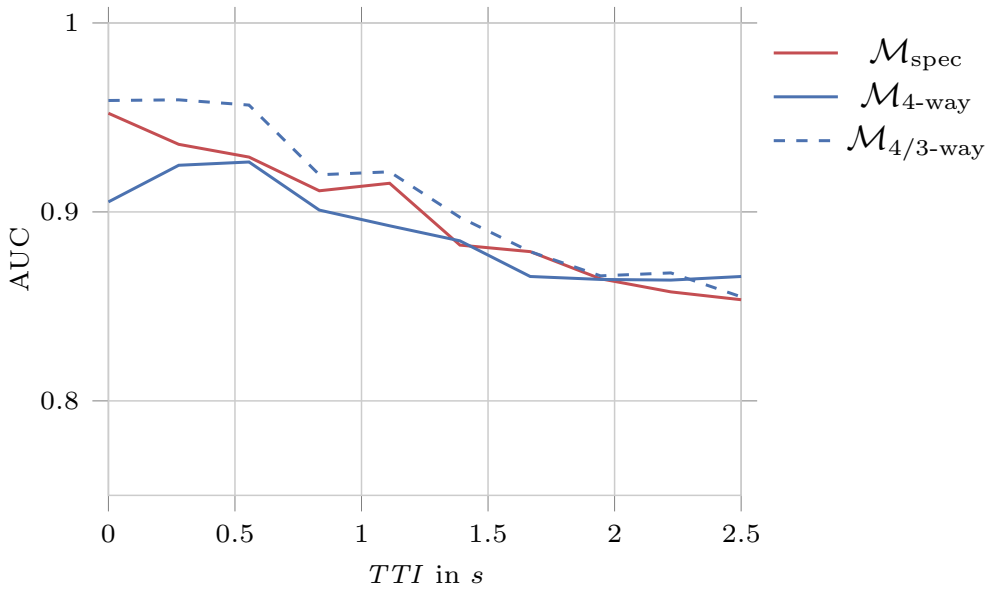


Figure 3.13: AUC for different TTIs.

exceeds those of the competing models. This explains the superior performance of model $\mathcal{M}_{4/3\text{-way}}$ due to its increased generalization capabilities.

Ω	$\mathcal{M}_{\text{spec}}$	$\mathcal{M}_{4\text{-way}}$	$\mathcal{M}_{4/3\text{-way}}$
s vs. r	15	26	41
s vs. l	21	28	49
r vs. l	20	18	38

Table 3.4: $\mathcal{M}_{\text{spec}}$ has 75% of the actual present data (see Fig. 3.12) available for training due to 4-fold cross validation. The data of the 36 available 4-way intersection approaches split into 18 straight going, 8 right and 10 left turning maneuvers and lead to the number of training examples depicted in column $\mathcal{M}_{4-way}$. This additional training data leads to an increased number of training examples for $\mathcal{M}_{3/4-way}$.

3.4.2 Intersection with Varying Number of Affecting Traffic Participants

The second evaluation focuses on comparing the pairwise probability coupling approach on a more complicated scene with the discussed alternatives in Section 3.1.2 for making maneuver estimations in generic traffic scenes.

The used data set was recorded by rooftop mounted cameras, filming a three-way-crossing near Gothenburg, Sweden. Section 2.3.2 gives a detailed introduction to the data set.

Scenarios are extracted for a subject vehicle approaching the intersection, as illustrated by the blue vehicle in Fig. 3.14 (b). The blue vehicle's spatial path alternatives, a left turn and a straight crossing path form the basis for the possible maneuvers $\tilde{\Omega}$. However, the temporal execution might be influenced by a preceding vehicle, represented by the vehicle in red, or an oncoming vehicle with priority, represented by the vehicle in green. Hence, the scene layout results in several distinct maneuver alternatives per path alternative, based on the affecting vehicles.

Tab. 3.5 shows the $|\tilde{\Omega}| = 7$ investigated prototypical spatio-temporal behavior patterns (maneuvers). The "left - free drive" maneuver characterizes intersection crossings where the blue vehicle is not influenced by the preceding. This means that its velocity profile resembles the velocity profile of left turns where no preceding vehicle was present at all. In contrast, the maneuver "left - influenced by preceding vehicle"

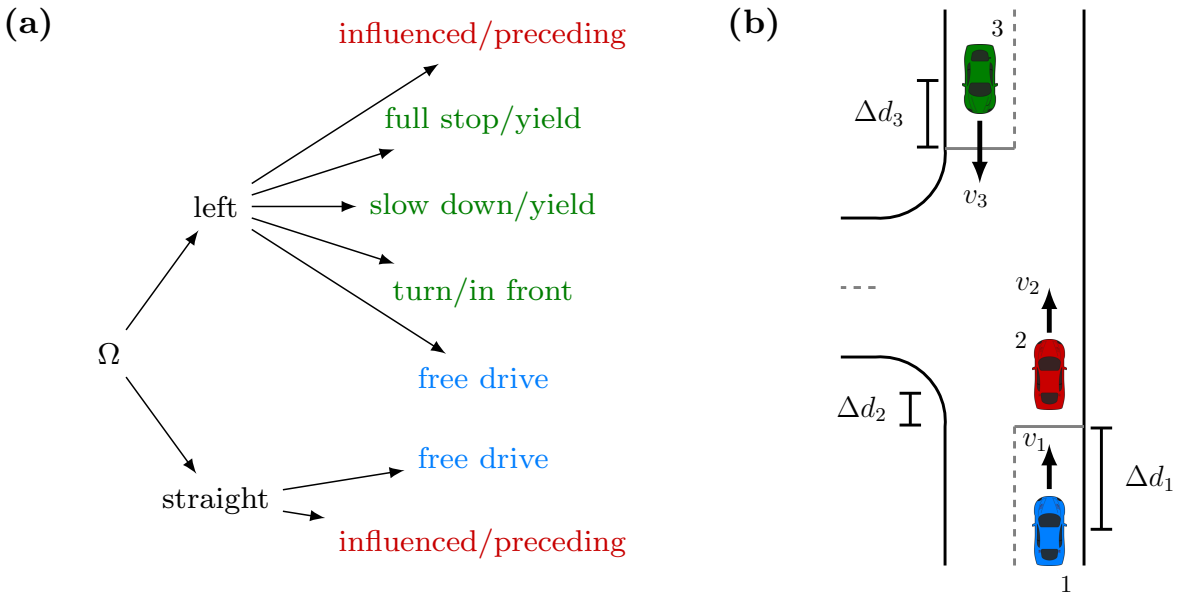


Figure 3.14: (a) The unfolding of the maneuver alternatives, starting from the spatial paths to the concrete prototypical temporal execution. (b) The spatial-layout and the abstract position of the three possibly present traffic participants as well as the used features.

$\tilde{\Omega}$	maneuver
b^1	left - free drive
b^2	left - influenced by preceding vehicle
b^3	left - approach slowly and yield to an oncoming vehicle
b^4	left - full stop and yield to an oncoming vehicle
b^5	left - turn shortly in front of an oncoming vehicle
b^6	straight - free drive
b^7	straight - influenced by preceding vehicle

Table 3.5: Different possible events/maneuvers for the subject vehicle.

characterizes a crossing with a significant deviation from the free drive velocity profile. The calculation of a prototypical free drive velocity helps to heuristically distinguish between these two maneuver alternatives during the creation the training and test data set. Fig. 3.15 (a) shows the prototypical velocity profile for left turns in free drive. A prototypical velocity profile consists of mean and variance of the velocity at each distance to intersection. Due to this probabilistic modeling, an unlabeled velocity profile's probability of belonging to a free drive maneuver can be calculated. Appendix A.1.1 explains the detailed calculation. For example

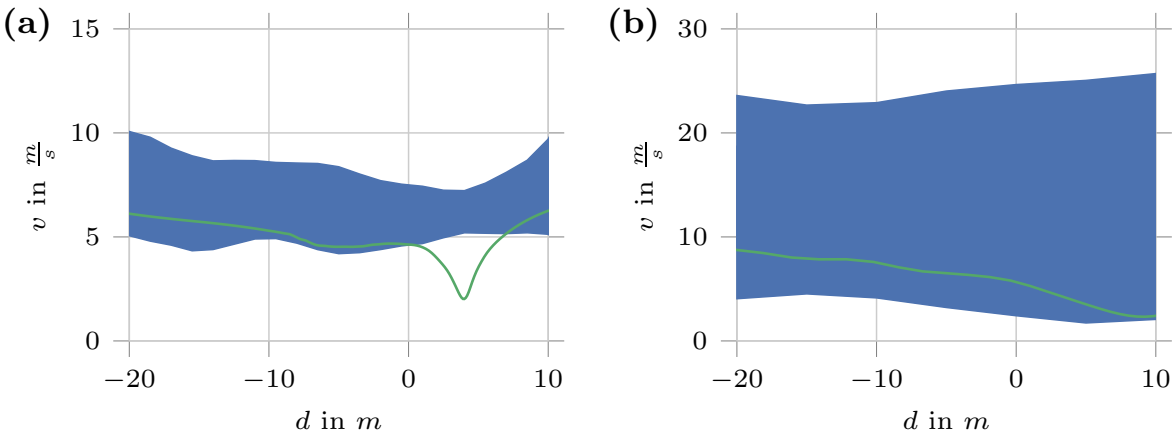
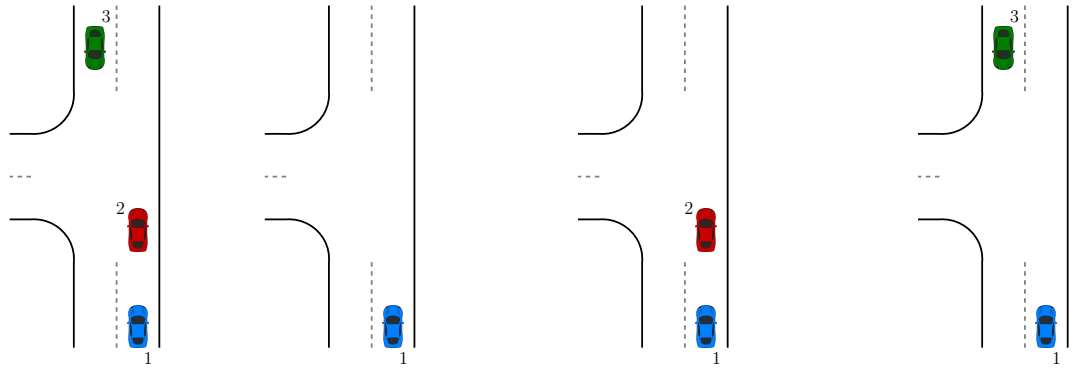


Figure 3.15: Prototypical velocity profiles in blue for (a) left turn maneuvers and (b) going straight maneuvers in free drive. The green velocity profile corresponds to a negative example, i.e. an influenced by preceding vehicle, in each case.

Fig. 3.15 (a) additionally shows a velocity profile which corresponds to a "left - influenced by preceding vehicle" maneuver, since it strongly deviates from the prototypical free drive velocity profile.
In the same manner, labels "straight - free drive" and "straight - influenced by preceding vehicle" maneuvers are obtained. Fig. 3.15 (b) shows a corresponding example for a "straight - influenced by preceding vehicle" compared to the prototypical velocity profile for a straight free drive.
The maneuver "left - turn shortly in front of an oncoming vehicle" is present if the subject vehicle passes the intersection slightly in front of the oncoming vehicle. This time frame is set to $3s$. The distinction between "approach slowly and yield to an oncoming vehicle" and "left - full stop and yield to an oncoming vehicle" rests upon the velocity of the subject vehicle dropping under a defined threshold of $2\frac{m}{s}$.
Fig. 3.14 (a) visualizes the spatial path based diversification of the possible maneuver alternatives.

The majority of defined possible maneuvers depend on the presence of either a preceding vehicle or an oncoming vehicle. Hence, they can obviously not be possible if one or both of the vehicles are not present in a current scene. This leads to a varying sample space of possible maneuvers. Therefore, the presented scenarios are well suited for evaluating the proposed approach as not all encountered scenes lead to $\Omega = \tilde{\Omega}$. Fig. 3.16 the different possible sample spaces.



(a) ${}^{\mathrm{I}}\Omega = \tilde{\Omega}$ **(b)** ${}^{\mathrm{II}}\Omega = \{b^1, b^6\}$ **(c)** ${}^{\mathrm{III}}\Omega = \{b^1, b^2, b^6, b^7\}$ **(d)** ${}^{\mathrm{IV}}\Omega = \{b^3, b^4, b^5, b^6\}$

Figure 3.16: Different possible scene layouts dependent on currently present traffic participants.

Evaluated Models

In order to obtain the desired probability distribution $p(b|\mathbf{x})$ using the proposed approach $\binom{|\tilde{\Omega}|}{2}$ prototypical binary classifiers are needed. The model based on pairwise probability coupling is from now on referred to as $\mathcal{M}_{\text{coup}}$.

This evaluation compares the proposed $\mathcal{M}_{coup}$. The following evaluation compares the proposed method with three other models for estimating $p(b|\mathbf{x})$.

The first comparison model is based on having specialized classifiers for all possibly occurring settings in Fig. 3.16. While this quickly becomes infeasible, it is still manageable for the restricted investigated test scenarios. The individual classifiers are summarized as model $\mathcal{M}_{\text{spec}}$.

The second comparison model is based on creating one classifier for the complete sample space $\tilde{\Omega}$ and normalizing over the possible events $b' \in \Omega$, as shown in (3.3). This model is referred to as $\mathcal{M}_{\text{norm}}$.

Additionally, a third model is created pursuing a generative approach. It is based on probabilistic state predictions and a subsequent comparison with the actual measured states, as discussed in the related work in Section 3.1.1. These approaches are truly flexible since they do not rely on factorizations of the generative joint probability distribution which are targeted at a fixed set of maneuvers, like e.g. Bayesian networks. There are two types of state of the art approaches for making probabilistic state predictions, learning based and model-based approaches, like discussed in Section 3.1.1. Since the currently available model based approaches

cannot cope with such complicated scenes, a learning based approach for making the needed probabilistic state predictions is chosen. For this purpose, nonlinear regression models for each maneuver based on commonly used Gaussian Processes ([110], [4]), with radial basis function kernels, are created. The likelihood $p(\mathbf{x}|b)$ of the observed subject vehicle's velocity v_1 belonging to maneuver b, given the remaining scene evidence $\mathbf{x}_{\setminus v_1}$ is determined by

$$p(\mathbf{x}|b) \propto p(v_1|\mathbf{x}_{\setminus v_1}, b).$$

Subsequently, the desired conditional probability distribution can be acquired using (3.2), assuming a uniform prior. This system is referred to as $\mathcal{M}_{\text{gen}}$.

The classifiers used to determine $p(b|\mathbf{x})$ of models $\mathcal{M}_{\text{norm}}$ and $\mathcal{M}_{\text{spec}}$, as well as the binary classifiers for determining the distributions $p^{k,l}(b|\mathbf{x})$ are based on probabilistic support vector machines with radial basis function kernels. The support vector machine is a discriminative kernel method. This makes it a suitable candidate for comparison with Gaussian processes, used for $\mathcal{M}_{\text{gen}}$. The kernels' length scales γ for the support vector machines as well as the Gaussian processes are set to the same value of 0.1 in order to improve comparability. Moreover, a neutral penalty parameter $C = 1$ is chosen for the support vector machines.

For the investigated, scenarios the distances to the intersection and velocities, depicted in Fig. 3.14 are expressive information. Hence, the information for each traffic participant is given by

$$\mathbf{x}_1 = [v_1, \Delta d_1]^T, \mathbf{x}_2 = [v_2, \Delta d_2]^T, \mathbf{x}_3 = [v_3, \Delta d_3]^T.$$

The feature vector for each model is a result of the maneuver(s) and the associated present traffic participants it is targeted at. E.g. the feature vector of $\mathcal{M}_{coup}$'s binary classifier discriminating between "left - free drive" and "left - influenced/preceding" is composed of $\mathbf{x}_1$ and $\mathbf{x}_2$. Another example is the regression model of $\mathcal{M}_{gen}$ estimating the velocity expected with maneuver "left - slow down/yield". Its feature vector is composed of $\mathbf{x}_1$ and $\mathbf{x}_3$. An exception is constituted by $\mathcal{M}_{\text{norm}}$. Its feature vector is the union of all possible features $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$. Missing values as a consequence of absent traffic participants are substituted by default values.

Results

On average, 486 cases could be extracted for each investigated *TTI*. Appendix A.1.2 depicts the occurrence distributions of the data set for all evaluated *TTI*s in detail[4]. The extracted data is split into training and test data using a 4-fold stratified cross validation.

Fig. 3.17 shows the resulting AUC curves for a *TTI* range from $0s$ to $2.5s$. The model based on the proposed pairwise probability coupling

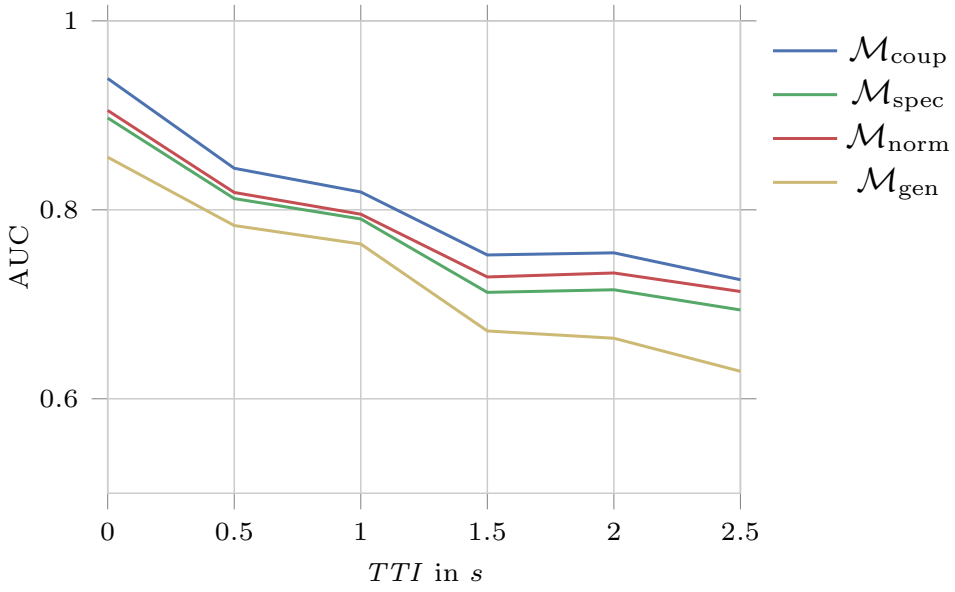


Figure 3.17: Area under the curve (AUC) for different *TTI*s.

$\mathcal{M}_{\text{coup}}$ shows the overall best performance. Model $\mathcal{M}_{\text{norm}}$ shows comparable performance in some areas but nevertheless does not reach the overall performance of $\mathcal{M}_{\text{coup}}$. It is expected that the performance deteriorates further for increasing $\tilde{\Omega}$ due to the increasing classifier complexity and the more prominent becoming normalization problem. The performance of $\mathcal{M}_{\text{spec}}$ does not reach the performance levels of the pairwise probability coupling approach. This observation strengthens the hypotheses stated in Section 3.4.1 that such highly specialized systems suffer from the therewith associated decreased number of training examples. In contrast, the proposed pairwise coupling approach acquires more training data and can benefit from utilizing the synergies between different, but partially similar scenes.

Fig. 3.18 illustrates this circumstance by showing exemplary learning

[4]Due to faulty traces or miss detections not every approach is usable for each investigated *TTI*; resulting in a slightly varying number of usable cases.

curves for a $TTI = 1.5s$ and the 4-fold cross validation of the evaluation in Fig. 3.17. Learning curves illustrate the performance of classification systems dependent on the number of training examples and thus are a good way of evaluating the adaptation of a classifier. The learning curve for $\mathcal{M}_{\text{coup}}$ converges to a stationary AUC which most likely corresponds to its optimum classification performance considering the classifier complexity and the available training data. Model $\mathcal{M}_{\text{spec}}$ still shows am increasing trend for the AUC for the maximum number of available training examples, indicating that the errors should further decrease for more training data. This emphasizes that specialized systems should generally perform at least on a similar level as the proposed coupling approach if a large enough training data set is available. However, this is rarely the case as collecting training data is associated with great effort and expenses.

Fig. 3.18 additionally indicates the $\mathcal{M}_{\text{norm}}$ model does also have a significant slope of its error compared to the coupling approach. This is reasonable, considering the demands on a classifier discriminating between a large number of possible classes.

The generative approach shows the poorest performance compared to the competing models adapted in a discriminative manner. This is a result of the individual Gaussian processes being trained to deliver accurate ve-

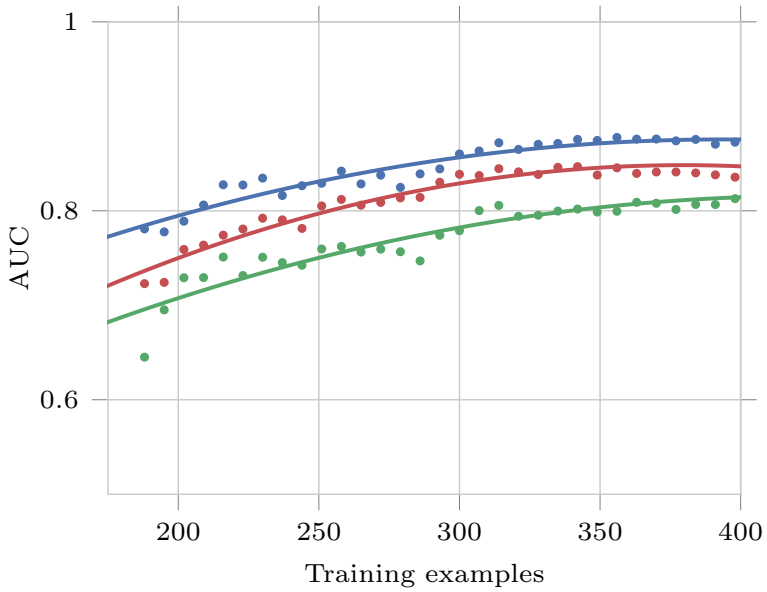


Figure 3.18: Learning curves for TTI = $1.5s$. The systems specific points correspond to the inverse accuracy for a certain number of training examples and folds. In order to outline the general trends more clearly a polynomial regression line is fitted.

locity predictions instead of discriminating between the maneuvers as good as possible. Furthermore, the evaluation confirms the findings reported in [53], [79] and [107], that discriminative maneuver estimation systems offer more reliable results compared to generative approaches.

3.5 Summary

This chapter stated the need for flexible estimation systems due to the extending application area of future advanced driver assistance systems and automated driving. Especially inner-city scenarios pose a challenge for these systems since they typically exhibit high variability. This high variability is the result of varying road layouts and constellations of traffic participants.

The literature review revealed that the majority of approaches does not tackle such problems. However, they can be categorized into either discriminative or generative approaches. Generative approaches tend to deal with the variability easier while comparative studies reveal that discriminative approaches tend to deliver superior performance on fixed problems. For this reason the chapter proposed a novel system architecture that is easily adaptable to changing environments, while providing the classification performance of purely discriminative approaches. The idea is based on partial, reusable classifiers combined in varying contexts using a probabilistic technique called pairwise probability coupling that is not dependent on a predefined set of maneuvers.

The experimental section included two evaluations. The first evaluation showed the reusability of the binary classifiers in varying contexts. The second evaluation showed that the proposed classification architecture is able to compete with specialized maneuver estimation systems on a complex intersection scenario with varying sets of possible maneuvers. In fact, it outperforms specialized systems due to its ability to use the available training data more efficiently and benefits from synergies between similar scenarios.

4 Interaction-aware Situation Recognition

Knowing how a scene evolves in the next couple of seconds is a prerequisite for future advanced driver assistance systems and autonomous vehicles. The previous chapter proposed a flexible framework for estimating the maneuvers of single vehicles. However, in complex scenarios the behavior of an intelligent vehicle might not just be influenced by one affecting traffic participant at a time but rather be influenced by several simultaneously. Fig. 4.1 shows a simple example clarifying this circumstance. The possible future behavior of the green vehicle, e.g. where and when it will safely merge onto the highway, strongly depends on which maneuvers the three gray cars in front are going to perform. Thus, possible combinations of their maneuvers and their occurrence probabilities are of great interest for the green vehicle. Possible maneuver combinations for example could be: the two cars on the most right lane making room for the entering vehicle or the two vehicles already on the highway are driving with constant velocity forcing the entering vehicle to stay on its. The exemplary maneuver combinations indicate that each traffic participant's maneuvers can in general not be assessed separately since their behaviors are strongly coupled with each other, i.e. the vehicles interact with each other. Interacting means that traffic participants take the maneuver alternatives of other vehicles and their likelihood explicitly into account. For the entering example in Fig. 4.1 this corresponds to the entering vehicles taking possible maneu-

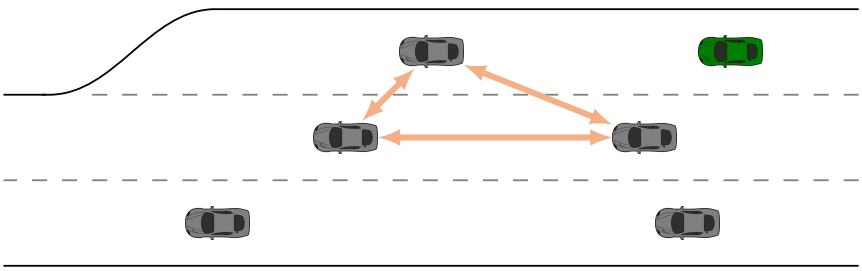


Figure 4.1: Interaction dominated highway scenario.

vers of the vehicles already on the highway into account. Vice versa, the vehicles on the highway take into account that vehicles on the entering will perform a merge at the same point. Respecting these interactions while performing situation assessment is referred to as *interaction-aware*.

Moreover, the problem of interaction-aware situation assessment does not only occur when confronted with more than two other dynamic traffic participants. It also arises in cases where the subject vehicle is part of an interaction and the prediction system does not know what the subject vehicle itself is planning to do, e.g. in warning systems where the vehicle is driven by a human. Examples for such systems are assistants for left turns in oncoming traffic or zebra crossing approaches, where there is typically an interaction between the driver and the pedestrian. For these reasons it is crucial for future ADAS and autonomous vehicles to assess scenes in a holistic and interaction-aware manner.

Thus, this chapter proposes a situation recognition framework for multiple, interacting traffic scenes. It is able to reconstruct a fully interaction-respecting probabilistic situation recognition, while relying on state-of-the-art single-entity-based maneuver estimations. The decomposition into single-entity-based maneuver estimations allows for tackling the combinatorics resulting from the possible maneuver combinations of involved traffic participants.

The remainder of this chapter is organized as follows. Section 4.1 reviews the current state-of-the-art and discusses general problems arising in the context of probabilistic situation assessment in generic traffic scenes with multiple, interacting traffic participants followed by an outline of the main contribution. Section 4.2 introduces an extension for state-of-the-art single traffic participant maneuver estimation, including the approach proposed in Chapter 3, enabling them to explicitly model interactions. These form the basis for the proposed interaction-aware situation assessment framework. Section 4.3 addresses the challenges arising when combining these interaction-aware maneuver predictions to assess situations with multiple traffic participants simultaneously. Section 5 propose a novel, general applicable, technique to overcome the addressed challenges. Furthermore, it gives an illustration on how to use it for situation assessment and evaluations on a real-world data set showing the benefits of the approach.
The main results of this chapter and Chapter 5 have been published in [64].

4.1 Problem Statement

The difficulties regarding probabilistic situation assessment with multiple dynamic traffic participants arise from the combinatorics regarding the sample space combinations of each traffic participants' possible maneuvers, [14]. This combinatorial increase of possible situation hypotheses leads to a high complexity of situation recognition systems. Furthermore, modeling the dependencies and interaction patterns between traffic participants is a challenging task. Both problems mutually reinforce each other, especially when considering the variability of traffic scenes. As stated in Chapter 3 traffic scenes undergo a high variability which results in varying sample spaces, i.e. varying possible maneuvers, for each traffic participant. Since each single maneuver sample space Ω_i is variable, the combined sample space $\Omega_{1:n}$ undergoes even stronger variations. Similarly, the number and characteristics of possible interaction patterns vary.

This section starts by reviewing related work, regarding assessing the maneuvers of several traffic participants simultaneously. Thereafter, the probability theoretical problems arising of systems confronted with interaction-aware situation recognition are discussed and the contribution is stated.

4.1.1 Related Work

The literature review in Section 3.1.1 focused on single-entity maneuver estimation systems. These approaches incorporate possible effecting dynamic scene entities by using relational features, e.g. relative velocities and distances. However, such features do not model possible interactions between traffic participants, i.e. the influence one vehicle's maneuver can have on another vehicle. The already discussed approach of [20] poses an exception. They incorporate the previously estimated lane change intention of a neighbouring vehicle into the maneuver estimation of the vehicle subject of maneuver estimation. However, the mutual influences between these vehicles are not considered.

This literature review focuses on approaches that model interactions between traffic participants and/or assess multiple traffic participants simultaneously.

The authors of the subsequent publications [91], [90], and [92] focus on modeling influences of other traffic participants by recognizing interrelated road users by so called *configurations*. While this approach is able

to find the cause for a certain vehicle's observed behavior it is limited to assessing each traffic participant individually. Moreover, they assume that dependencies between different traffic participants are strictly regulated and decomposable which might not always be the case when considering scenes where negotiation is an important factor, e.g. merging scenarios on highways.

Very few approaches have actually been proposed for predicting behaviors of multiple, possibly interacting traffic participants simultaneously. The situation recognition systems presented in [59], [72], [100], [43], [44], and [58] predict maneuvers of several traffic participants but do assume complete stochastic independence between the maneuvers of these traffic participants. They predict the maneuver of each traffic participant using systems like those discussed in Chapter 3 and simply multiply the obtained probabilities, justified by assuming stochastic independence. This represents a very strong assumption that does not hold in general traffic since it does not take any interactions into account at all.
Hence, to weaken these strong independence assumptions the authors of [58] and [72] propose to model interactions subsequently based on possible conflicting maneuvers. The idea is based on the assumption that drivers try to avoid risky situations. Thus, the occurrence probabilities of situations associated with a high collision probability are manually decreased. Decreasing these probability values artificially does not allow for sophisticated conclusions about which situations might appear instead. Moreover, such an approach can only capture the interactive character of scenes where a collision is inevitably going to happen.

The authors of [9] present an approach directly incorporating possible interactions between several traffic participants to estimate lane changes on highways. They use a game-theoretic approach to determine the most likely future motion of each traffic participant while taking the possible future evolution of other traffic participants' kinematic states into account. This calculated future motion is used as a predictive feature within a Bayesian network which estimates each vehicle's maneuver. Besides this interaction-aware predictive feature the maneuver prediction takes primarily lateral distances to lane markings into account. These features do only incorporate meaningful information if a lane change maneuver is already physically initialized and are therefore limited to short prediction horizons. Lateral features, e.g. intra-lane positioning, do not contain discriminative information for prediction horizons of more than three seconds. Moreover, while the predictive motion feature is calculated respecting possible interactions, the overall maneuver estimation

is performed separately for each vehicle without considering interactions. It is questionable if the full extent of interactions can be captured using one predictive feature. The authors pursue a similar strategy in their approach presented in [10].

The work presented in [74] indirectly incorporates interdependencies on behavior level via helper nodes. This means that each vehicle is represented by a dynamic Bayesian network where the individual behaviors do not influence each other directly. However, the individual dynamic Bayesian networks are coupled through an *expected* behavior for each vehicle, indirectly.

Fig. 4.2 shows an exemplary structure with two vehicles. It also reveals that the influence of the interaction-aware expected behavior takes effect in the subsequent timestep. While this approach incorporates interactions, the calculation of the expected behavior depends on vehicles with priority. The question remains how such an expected behavior can be determined if

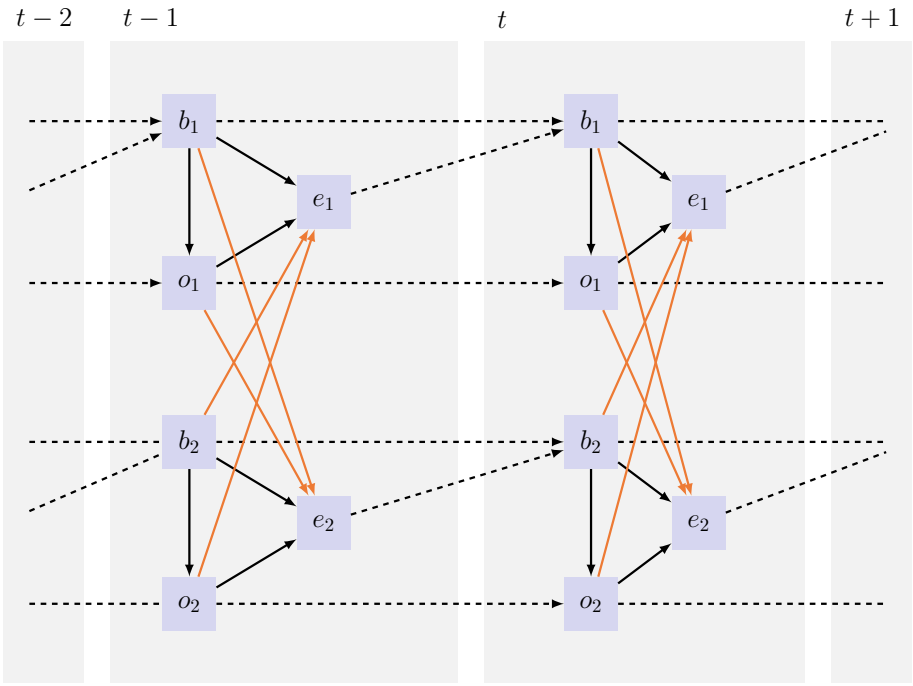


Figure 4.2: Shows a two vehicle example of the approach presented in [74]. Each vehicle is represented as one dynamic Bayesian network. The node b_i corresponds to possible behavior, o_i to the observations/measurements and e_i models the expected behavior variable. The dashed arrows correspond to connections between the individual timestemps, the red arrows incorporate the interaction between the vehicles.

the dependencies cannot be uniquely decomposed with priority rules (e.g. highway entering scenarios).

Therefore, it is not applicable universally. The authors of [71] propose a similar approach. They instantiate Bayesian networks dynamically based on an *Object-Oriented Probabilistic Relational Modeling Language* (OPRML) combined with their previously developed *Unified Traffic Situation Estimation Model* in [70]. The Unified Traffic Situation Estimation Model describes probabilistic dependencies reaching from low level measured states to trajectories up to drivable routes, represented by a separate Bayesian network for each present traffic participant. These individual networks are coupled using the OPRML for establishing connections between the individual networks and thus incorporating interactions. Fig. 4.3 shows the coupling of separate Bayesian networks for two vehicles. Due to the interdependent character of general interactions the coupled Bayesian networks develop cycles. The authors need to tackle these cycles since Bayesian networks cannot handle these types of interdependencies. Thus, they introduce helper nodes modeling how the state of a vehicle should be given a certain interaction. This is similar to the idea of expected behavior proposed in [74]. Again, it is unclear how such measures can be determined if an interaction is not beforehand regulated by traffic rules.

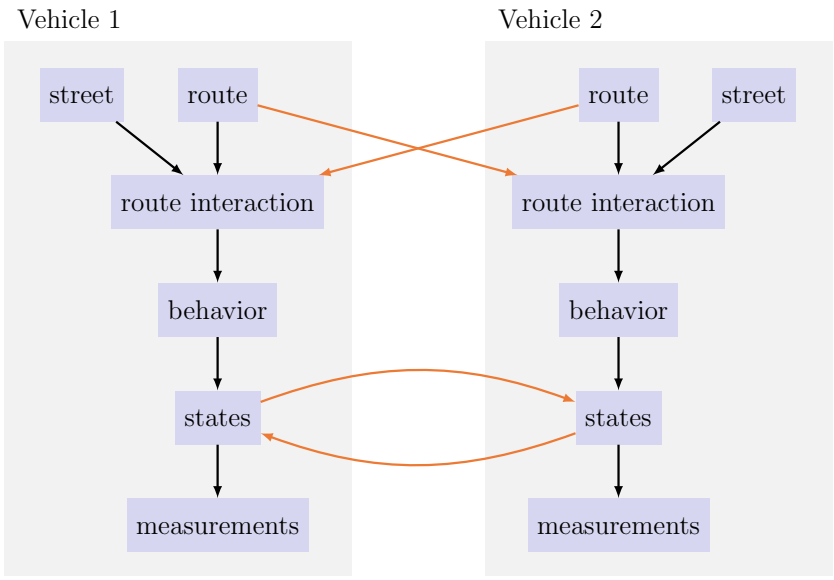


Figure 4.3: Outline of the approach proposed in [71]. Bayesian networks for two vehicles consisting of high level behavior abstractions down to low level sensory measurements are instantiated. The networks are coupled using the proposed OPRML to model possible interactions, indicated by the red arrows.

As a consequence, the system loses its ability to handle general interactions. Moreover, the approach is currently limited to model interactions of two vehicles at a time.

Discussion

In addition, a significant part of the approaches considering multiple traffic participants assumes stochastic independence. The few approaches that explicitly consider interactions between traffic participants do either rely on strong assumptions regarding the scope of these interactions or are dependent on decomposability of encountered scenes based on regulating traffic rules, e.g. priorities. Thus, an approach that generally tackles interaction-aware situation recognition, not relying on independence assumptions or distinct decomposable scenes, will be of great benefit. Moreover, the flexible applicability in generic traffic is not addressed by any of these approaches.

4.1.2 Probability Theoretical Considerations

The preceding literature review reveals that all presented approaches, without exception, rely on methods for decomposing the problem into smaller subparts. Therefore, the following discusses these decompositions from a probability theoretical point of view and outline their general benefits as well as their limitations.
Estimating the joint distribution $p(b_1, \ldots, b_n|\mathbf{x})$ directly is the most intuitive way of tackling multiple traffic participants situation recognition. However, it is striking that none of the discussed related works pursues such an approach. This is due to some some major limitations of modeling $p(b_1, \ldots, b_n|\mathbf{x})$. The following addresses these limitations.

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

Chapter 3 argued that making probabilistic maneuver estimations, i.e. estimating the posterior distribution $p(b|\mathbf{x})$, can be understood as probabilistic multiclass classification problem. In the same way making probabilistic situation recognition could be understood as a probabilistic classification problem, where each possible situation hypothesis, i.e. distinct maneuver combination, corresponds to one class of the classification algorithm. Accordingly, the posterior class probabilities correspond to the desired probability distribution $p(b_1, \ldots, b_n|\mathbf{x})$. However, for scenes with multiple

possibly interacting traffic participants, the number of possible situation hypotheses $|\Omega_{1:n}|$ increases combinatorially. Thus, the dimensionality of the classifier's target space increases, which in turn drastically increases the complexity of the classification algorithm. Moreover, the classifier needs to internally model the interdependencies between individual random variables. To overcome these two major issues large amounts of training data need to be available. However, recording data with the needed degree of detail is currently still expensive and annotating recorded scenes by human experts requires great efforts. Furthermore, complex classification algorithms pose higher computational requirements which in turn increases manufacturing costs.
For these reasons, directly estimating $p(b_1, \ldots, b_n|\mathbf{x})$ is hardly feasible and approaches tackling multi-entity traffic situations rely on different decompositions of the joint distribution $p(b_1, \ldots, b_n|\mathbf{x})$.

The addressed problems become even more prominent when the variability of traffic scenes is 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.

Decomposition of $p(b_1, \ldots, b_n|\mathbf{x})$ into Conditional Distributions

As stated above directly estimating $p(b_1, \ldots, b_n|\mathbf{x})$ is hardly feasible for complex traffic scenarios. Hence, decomposing the target joint distribution into conditional distributions is the current state of the art.

Assuming stochastic independence or conditional independence given $\mathbf{x}$ between traffic participants leads to the most trivial decomposition

$$p(b_1, \ldots, b_n|\mathbf{x}) = p(b_1|\mathbf{x})p(b_2|\mathbf{x}) \ldots p(b_n|\mathbf{x}),$$

where the individual conditionals correspond to single-entity maneuver estimations, e.g. those obtained from systems presented in Chapter 3. While this a preferable decomposition in terms of complexity, its applicability is limited to a small portion of traffic scenes due to its independence assumptions.

An interaction-respecting decomposition of the joint probability distribution can be obtained by recursively applying the chain rule of probabil-

ity, e.g.

$$\begin{aligned} p(b_1, b_2, \ldots, b_n|\mathbf{x}) =& p(b_1|\mathbf{x}, b_2, \ldots, b_n) p(b_2|\mathbf{x}, b_3, \ldots, b_n) \\ & \ldots p(b_{n-1}|\mathbf{x}, b_n) p(b_n|\mathbf{x}). \end{aligned} \tag{4.1}$$

The obtained decomposition is characterized by the decreasing number of regarded dependent entities for the conditional probability distributions

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

In interdependent cases, such as the entering scene of Fig. 4.1, decompositions of the chain rule of probability are still valid despite the lack of representing the dependencies between random variables in all conditional probability distributions, apparently. This is due to the reason that these conditionals are in fact marginal distributions, e.g. pointed out by

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

for $b_1 \not\perp b_2$. This is necessarily the case, since the chain rule of probability dictates an increasing reduction of represented random variables in contrast to the general underlying dependencies. The dependencies are incorporated through the marginal distributions overall random variables, nonetheless. Thus, these marginal conditionals cannot be obtained by single-entity maneuver prediction focusing on estimating the maneuver of one traffic participant at a time. At this point, it should be noted that the first factor in (4.1) forms an exception since all random variables are represented in this conditional but the distribution nevertheless corresponds to a single-entity maneuver estimation. These type of conditional distributions are referred to as *complete* conditional distribution. They have advantageous characteristics which the following sections will take advantage of throughout this chapter.
In contrast, the question of how to determine marginal conditional distributions leads to the same problems as already discussed for estimating $p(b_1, b_2, \ldots, b_n|\mathbf{x})$ directly. This is the case since the underlying target space of the marginal distributions (random variables left of the conditional) increases with each factor of (4.1) until the last marginal $p(b_n|\mathbf{x})$ needs to model the complete underlying sample space $\Omega_{1:n}$. Accordingly, the problems aggregate when confronted with varying sample spaces. For example, models for marginal probability distributions might in fact be marginals over not possible events or even not present traffic participants,

given a slightly altered layout of a scene. Therefore, they represent a completely different marginal distribution and can not be reused in different scene layouts.

4.1.3 Contribution

Future advanced driver assistance systems and autonomous vehicles need to be able to assess complex scenes consisting of several, possibly interacting traffic participants. This section pointed out that extending well-studied and numerously existing single-entity maneuver predictions to multi-maneuver estimations systems, i.e. predicting joint probabilities is hardly feasible. Moreover, the literature review reveals that current approaches focusing on several traffic participants simultaneously either rely on unique dependencies or neglect interactions completely. Hence, there is a strong demand for a general approach applicable in generic traffic that does not rely on any independence assumptions at all.
Therefore, this chapter proposes an approach that decomposes the problem by making fully interaction-respecting maneuver predictions for each traffic participant separately. Subsequently, these are used for reconstructing the fully interaction-aware joint probability $p(b_1, \ldots, b_n|\mathbf{x})$. The decomposition is advantageous, as the high target space dimensionality is partitioned into several partial classifiers with reduced complexity. The partial classifiers are easily reusable in other scenarios since each of them relies on fewer assumptions which tackles the variability of traffic scenes. Additionally, there are no special requirements for the individual interaction-aware maneuver predictions and thus well-studied approaches discussed in Section 3.1.1 as well as the flexible maneuver prediction framework proposed in Chapter 3 are applicable. Evaluations on a real-world data set show the benefits of the proposed approach.

4.2 Interaction-respecting Maneuver Predictions

The probabilistic single-entity-based approaches, discussed in Section 3.1.1, are estimating conditional probability distributions of the form

$$p(b_i|\mathbf{x}), \tag{4.2}$$

where the behavior b_i is conditioned on the currently observed scene evidence $\mathbf{x}$. This scene evidence includes relational features, which is a direct

way to reflect the influence of other potentially present traffic participants on participant i's behavior.
In general, the behavior of subject i will not only depend on the current states of effecting traffic participants but also on their possible future behaviors. In these cases, the posterior distribution $p(b_i|\mathbf{x})$ in fact models a marginal distribution along with its associated disadvantages, as outlined in Section 4.1.2. Another possibility is to add the behavioral alternatives of all affecting traffic participants explicitly to the conditional of (4.2), leading to

$$p(b_i|\mathbf{x}, b_1, \ldots, b_{i-1}, b_{i+1}, \ldots b_n). \tag{4.3}$$

These fully interaction-respecting conditional probability distributions are called *complete* conditionals. They include all random variables the target joint distribution includes. However, the target space dimensionality is reduced to $|\Omega_i|$, as $n-1$ random variables are shifted to the input space of the classification algorithm. This is an important difference compared to marginal distributions modeling the complete joint distribution internally. In fact the concrete behavior information of all effecting traffic participants leads to a better discriminability of the behavioral options of traffic participant i. This is due to the reason that the knowledge of other traffic participants' behaviors, incorporated as evidence in the conditional of (4.3), allows for drawing conclusion on a subject i's behavior. Fig. 4.4 shows a simple example where the knowledge of traffic participant 2's behavior helps with correctly estimating vehicle 1's (blue) behavior. E.g. knowing that red vehicle 2 is going to perform a left turn increases the probability of the blue vehicle being influenced by a preceding vehicle significantly.

The following investigation of a scene setting similar to the one evaluated in Section 3.4.2 further highlights this circumstance. The modified data set excludes cases with an oncoming vehicle or cases where no additional vehicle at all is present. This leads to cases with the same setting as in Fig. 4.4. Tab. 4.1 lists the remaining maneuver alternatives for vehicle 1 and vehicle 2. The red vehicle has only two possible maneuvers consisting of two path options in free drive due to missing impairments since only scenes with no other additional vehicles are investigated.

To show the benefits of explicitly representing other traffic participants' behaviors in complete conditionals, two models for estimating the maneuver of vehicle 1 are trained. Expressive information for discriminating the different maneuver alternatives of such settings consist of distances to the intersection and the current velocities as pointed out in Section 3.4.2.

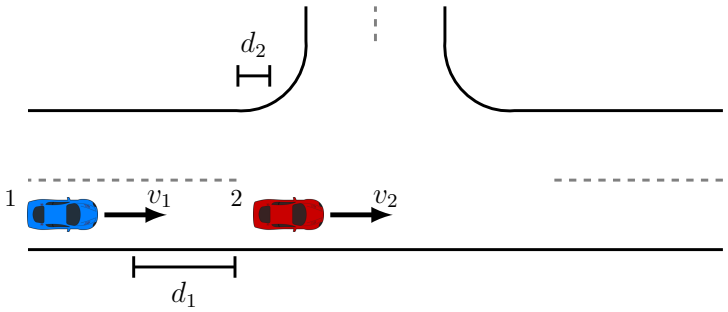


Figure 4.4: shows the investigated setting. Two vehicles on the same lane approaching an intersection. The preceding red vehicle possibly influences the blue vehicle due to slow turning maneuvers.

Hence, for this scenario the relevant scene information consists of the velocities of vehicles 1 and 2, as well as their distances to the intersection

$$\mathbf{x} = [v_1, d_1, v_2, d_2]^T. \tag{4.4}$$

The first model estimates the conditional distribution

$$\mathcal{M}_{1,\perp} : p(b_1|\mathbf{x}),$$

without adding the behavior information of vehicle 2, i.e. it is a marginal over vehicle's events. The second model estimates the complete conditional distribution

$$\mathcal{M}_{1,\not\perp} : p(b_1|\mathbf{x}, b_2), \tag{4.5}$$

incorporating knowledge of the maneuver vehicle 2 performs as a categorical variable into the feature space of the model estimating (4.5). Both models are based on a logistic regression as underlying classification algorithm.

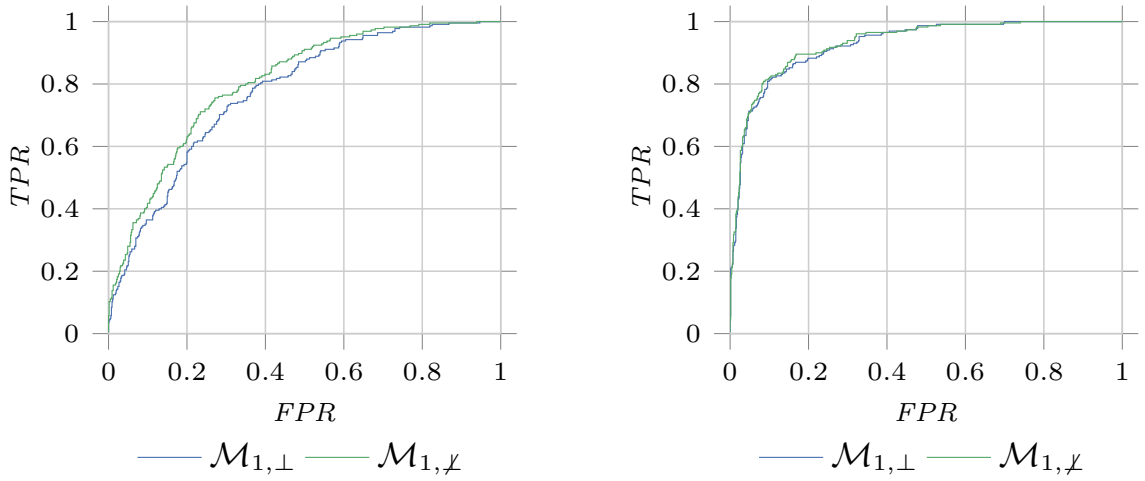
Ω_1	maneuver
b_1^1	left - free drive
b_1^2	left - influenced/prec. vehicle
b_1^3	straight - free drive
b_1^4	straight - influenced/prec. vehicle

(a) sample space of vehicle 1

Ω_2	maneuver
b_2^1	left - free drive
b_2^2	straight - free drive

(b) sample space of vehicle 2

Table 4.1: Different possible events/maneuvers for the subject vehicle.



(a) $TTI = 3s$ **(b)** $TTI = 0.5s$

Figure 4.5: Shows the two ROC curves for $\mathcal{M}_{1,\perp}$ and $\mathcal{M}_{1,\not\perp}$ for TTI's each obtainded from a stratified 4-fold cross-validation. Appendix A.2 shows the detailed situation occurrences in the data set.

Fig. 4.5 shows the ROC curves for two different predicted TTIs. Adding the information b_2 into the conditional results in an immediate performance gain for the longer prediction horizon of $TTI = 3s$, without any changes to the underlying classification algorithm. These results are expectable considering the obvious influence b_2 onto b_1.
For short prediction horizons, $TTI = 0.5s$, the ROC curves show that the performance differences vanish. This indicates that for short prediction horizons the importance of interactions are negligible due to the reactive character of short term predictions.

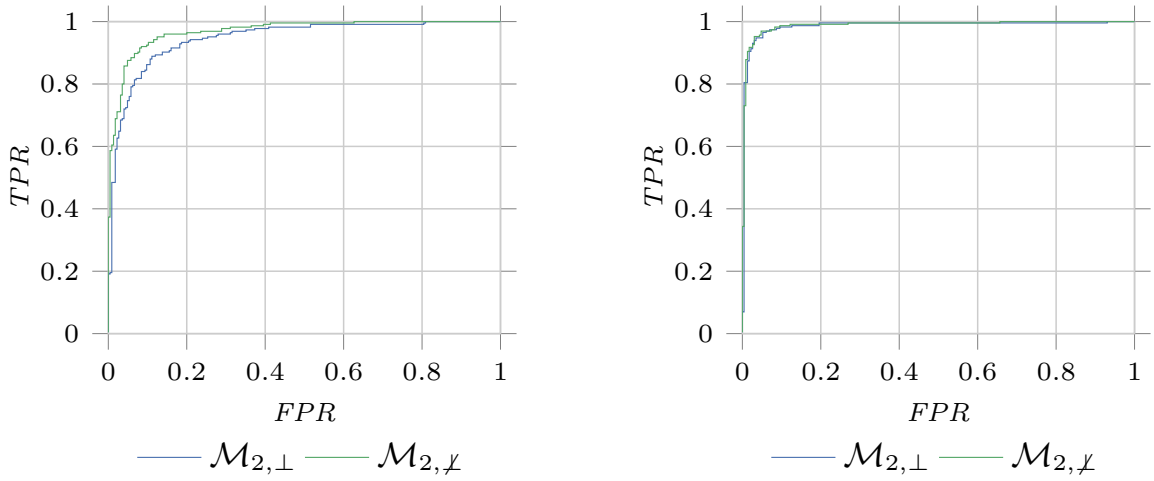
The following evaluation shows that also the estimation of vehicle 2's behavior benefits from the knowledge which maneuver vehicle 1 is going to perform. Two models for estimating the maneuvers of vehicle 2 are created. The first estimates the conditional distribution

$$\mathcal{M}_{2,\perp} : p(b_2|\mathbf{x}),$$

while the second model explicitly takes the possible influence of vehicle 1 into account by estimating the complete conditional distribution

$$\mathcal{M}_{2,\not\perp} : p(b_2|\mathbf{x}, b_1).$$

Both models are build upon a logistic regression with the same feature vector (4.4) The model estimating complete conditional has an additional



(a) $TTI = 3s$ **(b)** $TTI = 0.5s$

Figure 4.6: Shows the two ROC curves for $\mathcal{M}_{2,\perp}$ and $\mathcal{M}_{2,\not\perp}$ for TTI's each obtainded from a stratified 4-fold cross-validation. Appendix A.3 shows the detailed situation occurrences in the data set.

categorical variable for incorporating b_1.

Fig. 4.6 shows the ROC curves for two different predicted TTIs. Again, a clear performance increase can be seen for the complete conditional for the longer prediction horizon of $TTI = 3s$. This is a remarkable result considering that a maneuver of b_1 can hardly influence the pursued path of the preceding vehicle 2. However, knowing for example that vehicle 1 performs a b_1^4: "straight - influenced by preceding vehicle" maneuvers allows for inferring that vehicle b_2 will most likely perform a left turn. This is due to the symmetry of stochastic dependencies. Section 5.4 discusses this in more detail. For the short prediction horizon of $TTI = 0.5s$, the performance increase vanishes.

To sum it up, complete conditional distributions do in fact increase discriminability while being completely interaction-aware. Moreover, the target space dimensionality is decreased drastically compared to modeling joint distributions. Evidences expressing which maneuver the vehicles, represented in the particular conditionals, are going to perform are not available during run-time and need to be predicted at the same time[1].

[1] Although such information might be available in the case Vehicle-To-X communication is comprehensively introduced.

However, generally the maneuvers of several traffic participants are of interest, and thus they need to be predicted either way. During the training phases of models that estimate the complete conditional distributions, this information is in fact available, since the maneuvers have already been observed.

4.2.1 Complete Conditional Distributions and Generic Traffic

This section addresses the possibilities to adapt and reuse models that determine the needed complete conditional distributions for interaction-aware situation recognition to variable traffic scenes.
Sections 3.1.2 and 4.1 stated that having specialized models for each possibly encountered traffic scene is not feasible. Thus, the ability to adapt complete conditional distribution models to varying traffic scenes is crucial. In fact, the design of these models allows for an easy adaption.
Fig. 4.7 (a) illustrates a scene that clarifies this circumstance. Assuming that the maneuver alternatives consist of the three path alternatives, without considering any further typical temporal behavior patterns, the sample spaces for both vehicles are identical ${}^{a}\Omega_1 = {}^{a}\Omega_2 = \{\text{right}, \text{straight}, \text{left}\}$. Like outlined in the example of Section 4.2, probabilistic classification methods that incorporate the maneuver information of other vehicles as categorical feature can be used to obtain the needed complete conditional distributions $p(b_1|\mathbf{x}, b_2)$ and $p(b_2|\mathbf{x}, b_1)$. Thus, a probabilistic model $\mathcal{M}_{1,\not\perp}$ determines $p(b_1|\mathbf{x}, b_2)$ and second probabilistic model $\mathcal{M}_{2,\not\perp}$ determines $p(b_2|\mathbf{x}, b_1)$.

Fig. 4.7 (b) depicts a slightly different scene. The green vehicle cannot perform a straight going maneuver due to a one way street. Hence, for this second scene the sample space of the green vehicle 2 is re-

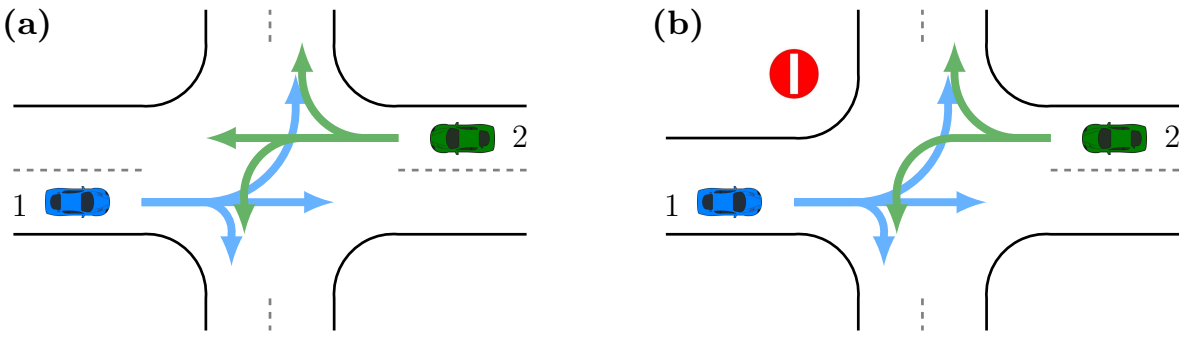


Figure 4.7: two slightly different scenes leading to different sample spaces.

duced to ${}^{b}\Omega_2 = \{\text{right}, \text{left}\}$ while the sample space of the blue vehicle 1 remains the same ${}^{b}\Omega_1 = {}^{a}\Omega_1 = \{\text{right}, \text{straight}, \text{left}\}$. These altered sample spaces do affect the conditional distributions $p(b_1|\mathbf{x}, b_2)$ and $p(\{b_2 \in {}^{b}\Omega_2 : B_2 = b_2|\mathbf{x}, b_1)$ (from now on abbreviated as $p(\bar{b}_2|\mathbf{x}, b_1)$) describe the outcome of the scene. Nevertheless, the same models $\mathcal{M}_{1,\not\perp}$ and $\mathcal{M}_{2,\not\perp}$ can be used to determine the second scene's complete conditional distributions. The changed sample space ${}^{b}\Omega_2$ only affects the conditional of $p(b_1|\mathbf{x}, b_2)$. The information of the conditional is part of the feature vector for model $\mathcal{M}_{1,\not\perp}$ and the complete conditional $p(b_1|\mathbf{x}, b_2)$ for the reduced sample space can be obtained by simply not making predictions for $b_2 = \text{straight}$.
This circumstance does not apply for the complete conditional of $p(\bar{b}_2|\mathbf{x}, b_1)$ since the changed sample does affect the target variable $\bar{b}_2$. However, varying sample spaces have been in the focus of Chapter 3. Hence, choosing the variable discriminative maneuver estimation framework for building model $\mathcal{M}_{2,\not\perp}$ allows for adapting its target space to the reduced sample space of ${}^{b}\Omega_2$. In this example it becomes possible since $\mathcal{M}_{2,\not\perp}$ consists of three binary classifiers, each of them discerning between one of the possible pairs of ${}^{a}\Omega_2$. Hence, $\mathcal{M}_{2,\not\perp}$ includes a binary *OvO* classifier targeted at exactly the reduced sample space of ${}^{b}\Omega_2$. This particular classifier is able to determine the complete conditional distribution of $p(\hat{b}_2|\mathbf{x}, b_1)$.

Thus, changed sample spaces for random variables in the conditional of complete conditional distributions do not pose a problem for the reusability of maneuver estimation systems. Moreover, the flexible discriminative maneuver estimation framework of Chapter 3 solves the problem of changing sample spaces regarding random variables in the target space of complete conditional distributions. For these reasons complete conditional distributions are adaptable to variable traffic scenes.

4.3 Reconstructing Joint Probabilities

The question remains of how the proposed complete conditionals along with their associate benefits can be used to construct a joint probability

distribution remains since a straightforward multiplication

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

does not provide a valid decomposition.

For many practical problems, to infer a probability distribution of interest exactly is not directly possible or intractable. A common method for obtaining information about complicated distributions is to use sampling methods [16, 48]. However, due to the complexity of many target distributions drawing samples from such distributions directly is not possible.

A popular sampling method, the *Gibbs sampling*, overcomes this problem by successively drawing samples from complete conditional distributions. 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 [55]. Chapter 5 will discuss the formal prerequisites in more detail.

Gibbs sampling starts with a randomly initialized value for each random variable b_i

$$p(b_1^{(0)}, b_1^{(0)}, \ldots, b_n^{(0)}|\mathbf{x}).$$

Starting from this initial, random state of the joint distribution a complete cycle of samples for each random variable is drawn for $\tau = 1, ..., n_\tau$ overall samples:

$$\begin{aligned} b_1^{(\tau+1)} &\sim p(b_1|\mathbf{x}, b_2^{(\tau)}, ..., b_n^{(\tau)}) \\ b_2^{(\tau+1)} &\sim p(b_2|\mathbf{x}, b_1^{(\tau+1)}, b_3^{(\tau)} ..., b_n^{(\tau)}) \\ &\vdots \\ b_{n-1}^{(\tau+1)} &\sim p(b_{n-1}|\mathbf{x}, b_1^{(\tau+1)}, b_2^{(\tau+1)} ..., b_{n-2}^{(\tau+1)}, b_n^{(\tau)}) \\ b_n^{(\tau+1)} &\sim p(b_n|\mathbf{x}, b_1^{(\tau+1)}, b_2^{(\tau+1)} ..., b_{n-1}^{(\tau+1)}), \end{aligned} \tag{4.6}$$

resulting in a new sample

$$p(b_1^{(\tau+1)}, b_2^{(\tau+1)}, \ldots, b_n^{(\tau+1)}|\mathbf{x}),$$

as stated in [16], where the upper index notation $b_i^{(\tau)}$ depicts a concrete event drawn at iteration τ. The successively drawn samples result in a Markov chain. This chain can be used to obtain the sufficient statistics of the joint probability distribution $p(b_1, \ldots, b_n|\mathbf{x})$, e.g. the mode over situation hypotheses samples. However, an approximation of the joint distribution $p(b_1, b_2, \ldots, b_n|\mathbf{x})$ is of interest for probabilistic situation recognition. The following metric

$$\hat{p}(\mathbf{b}_{1:n}^{\mathbf{i}}|\mathbf{x}) = \frac{|\{p(\mathbf{b}_{1:n}^{(\tau)}|\mathbf{x}) \; \forall \, \tau \in 1, \ldots, n_\tau \mid \mathbf{b}_{1:n}^{(\tau)} = \mathbf{b}_{1:n}^{\mathbf{i}}\}|}{n_\tau}, \quad \forall \mathbf{i} = \mathbf{1}, \ldots |\boldsymbol{\Omega}_{\mathbf{1:n}}| \tag{4.7}$$

approximates the target joint distribution using the resultant Markov chain. It determines the distribution of the successive situation hypotheses samples generated by the Gibbs sampler. For a sufficient large number of drawn samples, the sample distribution $\hat{p}(\mathbf{b}_{1:n}|\mathbf{x})$ converges towards the underlying joint distribution $p(\mathbf{b}_{1:n}|\mathbf{x})$. An estimate about what a sufficient large number n_τ is is hard to determine and represents a major issue for the application in automotive situation recognition systems. Available convergence diagnostics show only restricted reliability [30] or rely on several parallel instances of generate Markov chains which do further boost the already high computational demands.

Another issue arises since the complete conditionals are given in analytical form, i.e. as probability mass function outputted by the underlying classification algorithms. Thus, in an interim step multinomial distributions based on the probability values of the complete conditionals need to draw the needed samples for the Gibbs sampling.

Summing up, using a Gibbs sampler for reconstructing joint distribution from complete conditional distributions, obtained from maneuver prediction systems, is associated with several issues. For this reason, Chapter 5 proposes a new analytical solution for determining the joint distribution directly, while taking advantage of the functional form (pmf) of the complete conditional distributions determined by probabilistic classification algorithms.

4.4 Application to Situation Recognition

This section gives an overview on how to perform a probabilistic situation assessment in generic traffic scenes using the complete conditional

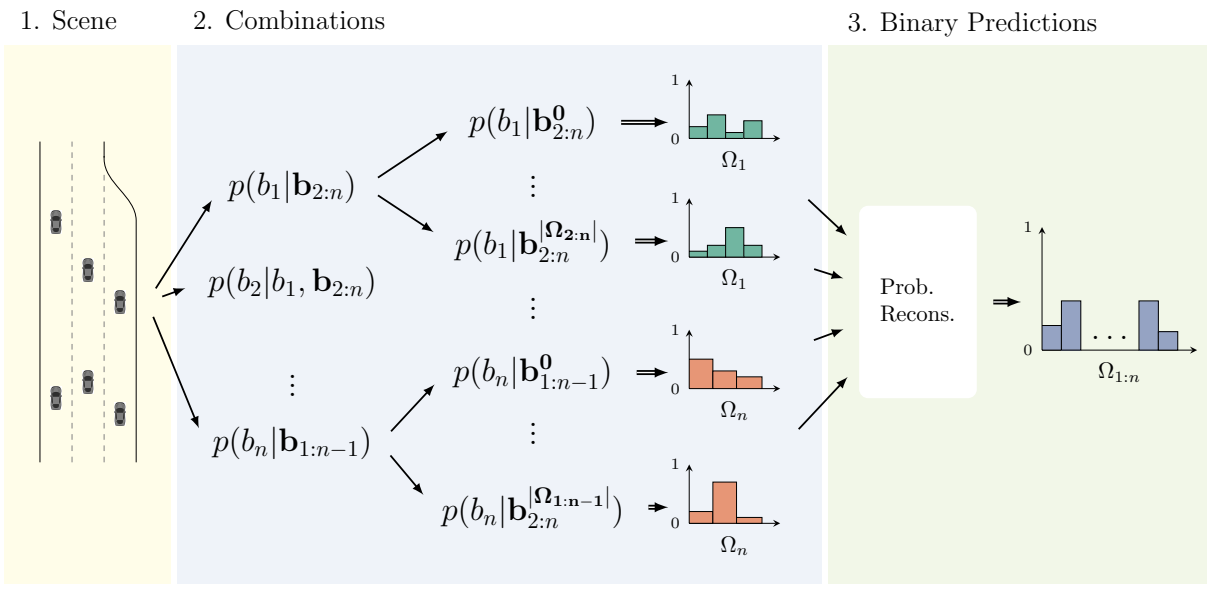


Figure 4.8: Overview of the proposed framework, detailed in Section 4.4.

distribution-based interaction-aware joint probability reconstruction. Fig. 4.8 illustrates the entire procedure. **First**, the set of possible situation hypotheses is instantiated based on the n present traffic participants and their possible maneuver options Ω_i. The complete sample space $\Omega_{1:n}$ is the result of a Cartesian product of all individual sample spaces $\Omega_i, i = 1, \ldots, n$. **Second**, each dynamic scene entity i is assessed separately. A single-entity-based prediction model is selected for obtaining the respective complete conditionals $p(b_i|\mathbf{x}, \mathbf{b}_{1:i-1}, \mathbf{b}_{1:i+1})$. 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 predicted. The corresponding classifiers treat the event combinations of the conditionals as input features along the scene evidence $\mathbf{x}$. **Third**, based on the obtained complete conditionals the desired joint probability distribution is reconstructed either by using the discussed sampling technique of Section 4.3 or one of the analytic solutions Chapter 5 presents. The result is the fully interaction-aware joint probability distribution over all situation hypotheses.

4.5 Summary

This chapter stated that there is a strong demand for a general situation recognition approach that is able to deal with multiple, interacting traffic participants. The literature review revealed that the few approaches that tackle the problem of assessing several traffic participants simulta-

neously either rely on unique dependencies like priority rules or neglect interactions completely. The probability theoretical discussion regarding situation recognition of multiple traffic participants concluded that there is a need for decomposing the problem into smaller subproblems due to the combinatorial increase of situation hypotheses. Subsequently, Section 4.2 outlined that state-of-the-art single-entity maneuver predictions can be extended to obtain complete conditional distributions. These are completely interaction-aware and form the basis for reconstructing a full joint distribution by applying discrete sampling techniques as outlined in Section 4.3. Section 4.4 summarized the complete process of interaction-aware situation recognition with multiple traffic participants. However, sampling techniques are problematic in terms of computational costs and the reliability of their results. Hence, the next chapter outlines procedures for reconstructing the target joint distribution analytically. Moreover, it conducts experiments on a complex highway scenario that outline the benefits of the proposed framework.

5 Analytic Reconstruction of Discrete Joint Probabilities

This chapter outlines how complete conditional distributions, obtained from probabilistic maneuver predictions for single-entities based on all other affecting traffic participants, can be used to reconstruct the interdependent joint probability distribution over all random variables without relying on costly sampling techniques.
Starting point is a possible factorization of the joint probability distribution, e.g.

$$p(b_1, \ldots, b_n) = p(b_1|b_2, \ldots, b_n)p(b_2|b_3, \ldots, b_n) \ldots p(b_n).$$

The underlying idea is to make use of every complete conditional probability distribution by restoring every factor to containing all random variables. These *remarginalized* distributions are the starting point for a substitution containing only the respective complete conditional and the desired joint distribution.

This chapter starts by discussing general considerations for reconstructing joint distributions from complete conditional distributions in Section 5.1. Section 5.2 outlines the basis for the developed reconstruction methods on a two random variable example while Section 5.3 generalizes the method to n random variables. Subsequently, Section 5.4 addresses the incorporation of stochastic independencies into the reconstruction process. and concludes by presenting experimental results on a complex, interactive highway entering that outlines the benefits of the interaction-aware situation recognition framework of Chapter 4 and the theoretical reconstruction methods developed in this chapter.

5.1 Compatibility of Complete Conditionals

Reconstructing joint probabilities from complete conditional distributions is a challenging task. In general, there exists no factorization of the joint

distribution, i.e.

$$p(b_1, \ldots, b_n) \neq \prod_{i=1}^{n} p(b_i | \mathbf{b}_{1:i-1}, \mathbf{b}_{i+1:n}),$$

such that all complete conditionals directly specify the target joint distribution without any other partial joint or marginal probability occurring. By applying the chain rule of probability, factorizations of the joint distribution can be obtained that contain one complete conditional respectively, e.g.

$$p(b_1, \ldots, b_n) = p(b_i | \mathbf{b}_{1:i-1}, \mathbf{b}_{i+1:n}) p(\mathbf{b}_{1:i-1}, \mathbf{b}_{i+1:n}). \tag{5.1}$$

If it is possible to express the unknown marginal probability $p(\mathbf{b}_{1:i-1}, \mathbf{b}_{i+1:n})$ through all $n-1$ remaining complete conditional distributions, the desired joint distribution ($p(b_1, \ldots, b_n)$) can be calculated directly. Despite, $n-1$ other starting factorizations of (5.1) can be chosen to calculate the target joint distribution that must all yield the same result. As a consequence, it is quite obvious that there must be strong restrictions to the functional form of the complete conditionals in order to yield the same joint distribution, regardless of the chosen factorization.
In other words, a reconstructed joint distribution from an arbitrary starting factorization ($\tilde{p}(b_1, \ldots, b_n)$) must yield all complete conditionals it has been reconstructed from, i.e.

$$p(b_i | \mathbf{b}_{1:i-1}, \mathbf{b}_{i+1:n}) \stackrel{!}{=} \frac{\tilde{p}(b_1, \ldots, b_n)}{\sum_{b_i} \tilde{p}(b_1, \ldots, b_n)}, \ \forall i \in 1, \ldots, n. \tag{5.2}$$

Complete conditional distributions specifying a unique joint distribution that satisfy (5.2), regardless of the chosen starting factorization, are said to be *compatible*.

Deciding whether a set of complete conditionals is compatible is not straightforwardly possible. In [7], necessary and sufficient conditions for compatibility are discussed. The authors focus on problems with finite, discrete sample spaces and two random variables. Compatibility for two dimensional distributions means that

$$p(b_1, b_2) = p(b_1 | b_2) p(b_2) \stackrel{!}{=} p(b_2 | b_1) p(b_1), \tag{5.3}$$

applies. The authors propose the construction of a $|\Omega_1| \times |\Omega_2|$ matrix $\mathbf{C}$ consisting of the elements

$$c^{k,l} = \frac{p(b_1^k | b_2^l)}{p(b_2^l | b_1^k)}.$$

The complete conditionals are compatible if

$$c^{k,l} \sum_{k,l} c^{k,l} = \sum_{l} c^{k,l} \sum_{k} c^{k,l}, \forall k, l.$$

In [6, 7], the authors prove that compatibility for strictly positive complete conditional distributions $p(b_1|b_2)$ and $p(b_2|b_1)$ is also guaranteed if $\mathbf{C}$ is of *rank* 1.

Compatibility and Sampling Techniques

The problem of determining whether complete conditionals are compatible is not just a problem of reconstructing joint probabilities analytically; this also effects sampling based methods, e.g. Gibbs sampling. Applying a Gibbs sampler to a set of incompatibile distributions yields different results for different ordering of drawn samples [48].
However, for the following reason this circumstance is rarely addressed in literature focusing on sampling techniques. Sampling techniques, as the name suggests, are based on drawing samples of distributions in order to obtain the desired target distributions. Thereby, it is unlikely that the distributions the samples are drawn from are present in their functional form but rather a population result of experiments representing the underlying target distribution. Hence, it is assumed that the target joint distribution exists, and thus specifies complete conditionals in the form of (5.2).
This is in contrast to the idea of having functional models for complete conditional distributions and reconstructing a possible joint distribution.

5.2 Two Random Variables

The following considers two random variables b_1 and b_2 with their sample spaces Ω_1 and Ω_2. Starting point is one of two possible factorizations from the chain rule of probability

$$p(b_1, b_2) = p(b_1|b_2)p(b_2). \tag{5.4}$$

Taking a closer look at (5.4) reveals that it already contains a complete conditional $p(b_1|b_2)$. Remarginalizing $p(b_2)$ over b_1 results in

$$p(b_1, b_2) = p(b_1|b_2) \sum_{b_1' \in \Omega_1} p(b_1', b_2),$$

and offers the possibility to incorporate the remaining complete conditional probability.

This is achieved by using the other possible factorization of the joint probability distribution which leads to

$$p(b_1, b_2) = p(b_1|b_2) \sum_{b_1' \in \Omega_1} p(b_2|b_1')p(b_1'),$$

where the second complete conditional distribution $p(b_2|b_1')$ is incorporated, while introducing $p(b_1')$. Remarginalizing over b_2 results in

$$p(b_1, b_2) = p(b_1|b_2) \sum_{b_1' \in \Omega_1} p(b_2|b_1') \sum_{b_2' \in \Omega_2} p(b_1', b_2'), \tag{5.5}$$

containing only the complete conditional distributions $p(b_1|b_2)$ and $p(b_2|b_1)$ as well as all unknown concrete probability values of $p(b_1, b_2)$. Equating (5.4) with (5.5) reveals the substitution for the incomplete marginal $p(b_2)$

$$p(b_2) = \sum_{b_1' \in \Omega_1} p(b_2|b_1') \sum_{b_2' \in \Omega_2} p(b_1', b_2').$$

Coefficient Matrix

Applying a concrete event combination $b_1^f \in \Omega_1$, $b_2^h \in \Omega_2$ of random variables b_1, b_2 for (5.5) leads to

$$p(b_1^f, b_2^h) = p(b_1^f|b_2^h) \sum_{b_1' \in \Omega_1} p(b_2^h|b_1') \sum_{b_2' \in \Omega_2} p(b_1', b_2').$$

Writing the sums out in full yields a better understanding of the underlying regularities of the derived equation

$$\begin{aligned} p(b_1^f, b_2^h) = p(b_1^f|b_2^h) \Big(& p(b_2^h|b_1^1) \big(p(b_1^1, b_2^1) + \dots p(b_1^1, b_2^{|\Omega_2|}) \big) \\ & + \dots + p(b_2^h|b_1^{|\Omega_1|}) \big(p(b_1^{|\Omega_1|}, b_2^1) + \dots p(b_1^{|\Omega_1|}, b_2^{|\Omega_2|}) \big) \Big). \end{aligned} \tag{5.6}$$

It is composed of products of the two complete conditionals values, respectively for b_1 and b_2 multiplied with the sum of a subset of the unknown joint probabilities.

The column vector $\mathbf{p}$ summarizes all joint probabilities $p(b_1, b_2)$, thus $\mathbf{p}$ is of size $(m = |\Omega_{1:2}|, 1)$. The order of the singular elements are given by the Cartesian product $\Omega_1 \times \Omega_2$

$$\begin{aligned}\mathbf{p} =&[p(b_1^1, b_2^1), \ldots, p(b_1^1, b_2^{|\Omega_2|}),\\ &p(b_1^2, b_2^1), \ldots p(b_1^2, b_2^{|\Omega_2|}),\\ &\ldots,\\ &p(b_1^{|\Omega_1|}, b_2^1), \ldots p(b_1^{|\Omega_1|}, b_2^{|\Omega_2|})].\end{aligned}$$

This allows for writing (5.6) in a vector form

$$p(b_1^f, b_2^h) = (\mathbf{a}^{f,h})^T \mathbf{p}, \tag{5.7}$$

using the dot product with

$$\begin{aligned}(\mathbf{a}^{f,h})^T = [&\overbrace{p(b_1^f|b_2^h)p(b_2^h|b_1^1), \ldots, p(b_1^f|b_2^h)p(b_2^h|b_1^1)}^{|\Omega_2|-\text{times}},\\ &\ldots,\\ &\underbrace{p(b_1^f|b_2^h)p(b_2^h|b_1^{|\Omega_1|}), \ldots, p(b_1^f|b_2^h)p(b_2^h|b_1^{|\Omega_1|})}_{|\Omega_2|-\text{times}}]^T.\end{aligned}$$

It should be noted that $(\mathbf{a}^{f,h})^T$ is composed of $|\Omega_1|$ blocks, where each block is composed of $|\Omega_2|$-times repeated, identical elements $p(b_1^f|b_2^h)p(b_2^h|b_1^k)$. k is fixed for each block and increases with each block, i.e. $k \in 1, \ldots, |\Omega_1|$ in the same succession as the events in b_1 in $\mathbf{p}$.

A linear equation system can be created using (5.7) for all $\Omega_{1:2}$ combinations

$$\mathbf{p} = \begin{pmatrix} (\mathbf{a}^{1,1})^T \\ \vdots \\ (\mathbf{a}^{1,|\Omega_2|})^T \\ (\mathbf{a}^{2,1} \\ \vdots \\ (\mathbf{a}^{2,|\Omega_2|})^T \\ \vdots \\ (\mathbf{a}^{|\Omega_1|,|\Omega_2|})^T \end{pmatrix} \mathbf{p} = \mathbf{A}\mathbf{p} \tag{5.8}$$

where $\mathbf{A}$ is a $(|\Omega_{1:2}| \times |\Omega_{1:2}|)$ coefficient matrix with elements $a^{o,u}$. The index o indicates the corresponding row and u the respective column.

5.2.1 Analogy to Discrete Markov Chains

Revisiting equation (5.8) reveals the similarity of determining $\mathbf{p}$ to finding the stationary distribution of a discrete Markov chain on a finite state space. Section 3.2.3 discussed the problem of verifying the existence and finding the stationary distribution of discrete Markov chains. Hence, the following shows that determining the joint distribution can as well be transferred to finding the stationary distribution of a discrete Markov chain.

Finite State Space and Stochastic Matrices

For the purpose of interpreting (5.8) as a discrete, homogeneous Markov chain, its state space needs to be finite and $\mathbf{A}$ needs to be a stochastic matrix. Each possible state of the Markov chain corresponds to one event combination of the sample space. Hence, the Markov chain's state space is identical to the sample space $\Omega_{1:2}$ and is thus finite and fixed.
In order for the coefficient matrix $\mathbf{A}$ to be a stochastic matrix, the following two properties have to be satisfied. The first property

$$0 \leq a^{o,u} \leq 1,\ \forall o, \forall u \in 1, \ldots, |\Omega_{1:2}|, \tag{5.9}$$

is satisfied as each element $a^{o,u}$ is the product of two conditional probabilities values. The second property

$$\sum_{o=1}^{|\Omega_{1:2}|} a^{o,u} = 1,$$

states that all elements of each column of $\mathbf{A}$ must sum up to 1. The column sum of an arbitrary column u, corresponding to an event b^k of $\mathbf{A}$ is given by

$$\sum_{o=1}^{|\Omega_{1:2}|} a^{o,u} = \sum_{f=1}^{|\Omega_1|} \sum_{h=1}^{|\Omega_2|} p(b_2^f | b_2^h) p(b_2^h | b_1^k).$$

The event b^k is fixed for all rows and determined by the column u, thus $p(b_2^h|b_1^k)$ is independent of the running index f and is rearranged to

$$\sum_{o=1}^{|\Omega_{1:2}|} a^{o,u} = \sum_{h=1}^{|\Omega_2|} p(b_2^h | b_1^k) \underbrace{\sum_{f=1}^{|\Omega_1|} p(b_2^f | b_2^h)}_{=1},$$

where the sum over Ω_1 for a fixed h is identical to 1. The remaining sum of

$$\sum_{o=1}^{|\Omega_{1:2}|} a^{o,u} = \underbrace{\sum_{h=1}^{|\Omega_2|} p(b_2^h|b_1^k)}_{=1},$$

does also sum up to one for arbitrary u. Therefore, it is guaranteed that each column of $\mathbf{A}$ sums up to 1, which in turn proofs that $\mathbf{A}$ is a stochastic matrix.
Hence, the equivalent Markov chain is given by

$$\hat{\mathbf{p}}^{(\tau+1)} = \mathbf{A}\hat{\mathbf{p}}^{(\tau)},$$

at each iteration step τ.

Existence of the Stationary Distribution

In order for

$$\lim_{\tau\to\infty} \hat{\mathbf{p}}^{(\tau)} = \mathbf{p}$$

to actually converge to the stationary distribution and for (5.8) to hold true, $\mathbf{A}$ needs to be irreducible and aperiodic, as Section 3.2.3 outlined. Both requirements are satisfied by strengthening (5.9) to

$$0 < a^{o,u} < 1,\ \forall o, \forall u \in 1, \ldots, |\Omega_{1:2}|.$$

The requirement (5.9) is fulfilled since the elements of the coefficient matrix $\mathbf{A}$ are products of two concrete conditional probability values. Demanding strict positivity for all elements of $\mathbf{A}$ is equivalent to demanding strict positivity for all complete conditional distributions

$$p(b_i|b_{1:i-1,i+1:n}) > 0. \tag{5.10}$$

This constraint can easily be fulfilled in practical implementations by slightly adjusting the complete conditional distributions, violating (5.10), without significantly altering the target distribution $\mathbf{p}$.

Determining the Stationary Distribution

After proving the existence of a unique stationary distribution for $\mathbf{p} = \mathbf{A}\mathbf{p}$, the actual distribution needs to be calculated. Section 3.2.3 introduced

three ways for doing this. Either the Markov chain is calculated directly (3.20), an eigenvector problem is formulated (3.21), or an inhomogeneous linear equation system (3.22) is solved.

A **fourth** option, apart from methods directly related to Markov chains, arises in the context of formal comparability in Section 5.1. This option for calculating the stationary distribution is based on the definition for compatibility given in (5.3) allowing for a direct calculation of the desired joint target distribution by rearranging (5.3) to

$$\frac{p(b_1|b_2)}{p(b_2|b_1)} = \frac{p(b_1)}{p(b_2)}.$$

Marginalizing over b_1 leads to

$$\begin{aligned} \sum_{b_1' \in \Omega_1} \frac{p(b_1'|b_2)}{p(b_2|b_1')} &= \sum_{b_1' \in \Omega_1} \frac{p(b_1')}{p(b_2)}, \\ \sum_{b_1' \in \Omega_1} \frac{p(b_1'|b_2)}{p(b_2|b_1')} &= \frac{1}{p(b_2)}, \end{aligned} \tag{5.11}$$

yielding a direct way to calculate the marginal $p(b_2)$ with

$$p(b_2) = \frac{1}{\sum_{b_1' \in \Omega_1} \frac{p(b_1'|b_2)}{p(b_2|b_1')}}. \tag{5.12}$$

Based on this the target joint distribution can be directly calculated

$$p(b_1, b_2) = p(b_1|b_2)p(b_2) = \frac{p(b_1|b_2)}{\sum_{b_1' \in \Omega_1} \frac{p(b_1'|b_2)}{p(b_2|b_1')}}.$$

Similarly, the target joint distribution can be calculated with the decomposition $p(b_1, b_2) = p(b_2|b_1)p(b_1)$ while determining the marginal $p(b_1)$ by marginalizing (5.11) with respect to b_2.

5.3 n Random Variable

After outlining the general idea of reconstructing joint distributions from complete conditional distributions on two random variables this section introduces two novel methods for reconstructing joint distributions over more than two random variables.

5.3.1 Linear Equation System

The discussed two random variable example can be straightforwardly extended to the general n random variable case. The same procedure of successively incorporating all b_i's complete conditionals can be applied $n-1$ times, e.g.

$$\begin{aligned} p(b_1,\ldots,b_n) &= p(b_1|\mathbf{b}_{2:n})p(\mathbf{b}_{2:n}) \\ &= p(b_1|\mathbf{b}_{2:n})\sum_{b'_1} p(b_2|b'_1,\mathbf{b}_{3:n})p(b'_1,\mathbf{b}_{3:n}) \\ &= p(b_1|\mathbf{b}_{2:n})\sum_{b'_1} p(b_2|b'_1,\mathbf{b}_{3:n})\sum_{b'_2} p(b_3|\mathbf{b}'_{1:2},\mathbf{b}_{4:n})p(\mathbf{b}'_{1:2},\mathbf{b}_{4:n}) \\ &= \ldots, \end{aligned}$$

ultimately leading to the general analytic reconstruction formula

$$\begin{aligned} p(b_1,\ldots,b_n) = \; & p(b_1|\mathbf{b}_{2:n})\sum_{b'_1} p(b_2|b'_1,\mathbf{b}_{3:n})\sum_{b'_2} p(b_3|\mathbf{b}'_{1:2},\mathbf{b}_{4:n}) \\ \ldots & \sum_{b'_{n-2}} p(b_{n-1}|\mathbf{b}'_{1:n-2},b_n)\sum_{b'_{n-1}} p(b_n|\mathbf{b}'_{1:n-1})\sum_{b_n} p(b'_1,\ldots,b'_n). \end{aligned} \tag{5.13}$$

It can be seen that (5.13) only contains the known complete conditional distributions and the target joint distribution itself, exactly like the two random variable case in Section 5.2. Thus, an equation system can be set up in order to determine the target joint distribution. For this purpose a concrete event combination for $p(\mathbf{b}^{\mathbf{k}})$ is applied, leading to

$$\begin{aligned} p(\mathbf{b}^{\mathbf{k}}_{1:n}) = & p(b^{\mathbf{k}}_1|\mathbf{b}^{\mathbf{k}}_{2:n})\sum_{b'_1} p(b^{\mathbf{k}}_2|b'_1,\mathbf{b}^{\mathbf{k}}_{2:n}) \\ & \ldots\sum_{b'_{n-1}} p(b^{\mathbf{k}}_n|\mathbf{b}'_{1:n-1})\sum_{b_n} p(b'_1,\ldots,b'_n) \end{aligned} \tag{5.14}$$

which can again be condensed to a dot product of

$$p(\mathbf{b}^{\mathbf{k}}_{1:n}) = (\mathbf{a}^{\mathbf{k}})^T\mathbf{p}, \tag{5.15}$$

of a coefficient row vector $(\mathbf{a}^{\mathbf{k}})^T$ for each event combination $\mathbf{k}$. The multi-index notation, introduced in Section 2.1, for $\mathbf{k} = \mathbf{1},\ldots,|\mathbf{\Omega_{1:n}}|$ represents

all $|\Omega_{1:n}|$ possible event combinations. Accordingly, the joint probability vector $\mathbf{p}$ contains discrete probability values for these m event combinations.
Taking a closer look at (5.14) reveals that all discrete probability values of $p(\mathbf{b}^{\mathbf{k}}_{1:n})$ that only differ in the last event $b^{\mathbf{k}}_n$ are multiplied with the same coefficient. This is due to the fact that the summation over the complete conditionals in (5.14) is not influenced by the marginalization of b'_n, as only the last factor, i.e the target joint distribution itself, is subject to the marginalization over b_n. The coefficient vector for event combination $\mathbf{k}$ is given by

$$\begin{aligned}(\mathbf{a}^{\mathbf{k}})^T = &\overbrace{[p(b^{\mathbf{k}}_1|\mathbf{b}^{\mathbf{k}}_{2:n})p(b^{\mathbf{k}}_2|b^{\mathbf{1}}_1,\mathbf{b}^{\mathbf{k}}_{2:n})\dots p(b^{\mathbf{k}}_n|\mathbf{b}^{\mathbf{1}}_{1:n-1})}^{|\Omega_n|-\text{times}},\\ &\dots,\\ &\underbrace{p(b^{\mathbf{k}}_1|\mathbf{b}^{\mathbf{k}}_{2:n})p(b^{\mathbf{k}}_2|b^{|\mathbf{\Omega_{1:n}}|}_1,\mathbf{b}^{\mathbf{k}}_{2:n})\dots p(b^{\mathbf{k}}_n|\mathbf{b}^{|\mathbf{\Omega_{1:n}}|}_{1:n-1})}_{|\Omega_n|-\text{times}}].\end{aligned} \tag{5.16}$$

The complete reconstruction matrix for $p(b_1,\dots,b_n)$ can be obtained by setting up all m equations (5.15) for each possible event combination $\mathbf{k}$, leading to

$$\mathbf{p} = \begin{pmatrix}(\mathbf{a}^{\mathbf{1}})^T\\ \vdots\\ (\mathbf{a}^{|\mathbf{\Omega_{1:n}}|})^T\end{pmatrix}\mathbf{p} = \mathbf{A}\mathbf{p}, \tag{5.17}$$

where each element of $\mathbf{A}$ is given by

$$a^{\mathbf{k},\mathbf{l}} = p(b^{\mathbf{k}}_1|\mathbf{b}^{\mathbf{k}}_{2:n})p(b^{\mathbf{k}}_2|b^{\mathbf{l}}_1,\mathbf{b}^{\mathbf{k}}_{2:n})\dots p(b^{\mathbf{k}}_n|\mathbf{b}^{\mathbf{l}}_{1:n-1}),$$

where the multi-indices $\mathbf{k}$, $\mathbf{l}$ index the row corresponding to the event combination given by $\mathbf{k}$ and the column corresponding to the event combination given by $\mathbf{l}$.

Equation (5.17) has the same structure, i.e being equivalent to a discrete Markov chain, as the equation system obtained for the two random variable case in Section 5.2. Hence, the same requirements as stated in Section 5.2 hold for $\mathbf{A}$ to specify a unique stationary distribution.

Stochastic Matrix & Stationary Distribution

A prerequisite for (5.17) to yield a stationary distribution is that $\mathbf{A}$ needs to be a stochastic matrix. As a reminder, this can be guaranteed if the $\mathbf{A}$'s column and/or row sum is identical to 1.

In the following a proof for the columns of $\mathbf{A}$ summing up to 1, i.e

$$\sum_{\mathbf{k}} a^{\mathbf{k},\mathbf{l}} = 1,\ \forall \mathbf{l} \in \mathbf{1}, \ldots, |\mathbf{\Omega_{1:n}}|,$$

is provided. For an arbitrary column $\mathbf{l}$ the column sum over all $\mathbf{k}$ rows based on (5.16) is given by

$$\sum_{b_1,\ldots,b_n} p(b_1|\mathbf{b}_{2:n}) p(b_2|b_1^{\mathbf{l}}, \mathbf{b}_{3:n}) \ldots p(b_n|\mathbf{b}_{1:n-1}^{\mathbf{l}}).$$

For convenience of the subsequent steps, the order of the factors is reversed resulting in

$$\sum_{b_1,\ldots,b_n} p(b_n|\mathbf{b}_{1:n-1}^{\mathbf{l}}) \ldots p(b_2|b_1^{\mathbf{l}}, \mathbf{b}_{3:n}) p(b_1|\mathbf{b}_{2:n}). \tag{5.18}$$

It can be seen that only the last factor of (5.18) is affected by the summation over the sample space of b_1, since all other occurrences of b_1 are specified by the particular column $\mathbf{l}$ for the remaining factors. Thus, (5.18) can be rewritten as

$$\sum_{b_2,\ldots,b_n} p(b_n|\mathbf{b}_{1:n-1}^{\mathbf{l}}) \ldots p(b_2|b_1^{\mathbf{l}}, \mathbf{b}_{3:n}) \underbrace{\sum_{b_1} p(b_1|\mathbf{b}_{2:n})}_{=1},$$

where it is obvious that the summation over the complete conditional $p(b_1|\mathbf{b}_{2:n})$ is identical to 1, regardless of possible event combinations of $\mathbf{b}_{2:n}$. Having removed the factor of $p(b_1|\mathbf{b}_{2:n})$ offers the possibility to apply the same permutation of sums with regard to to the complete conditional $p(b_2|b_1^1, \mathbf{b}_{3:n})$ yielding

$$\sum_{b_3,\ldots,b_n} p(b_n|\mathbf{b}_{1:n-1}^{\mathbf{l}}) \ldots \underbrace{\sum_{b_2} p(b_2|b_1^{\mathbf{l}}, \mathbf{b}_{3:n})}_{=1}.$$

This procedure can be repeated until only the summation of the first factor of (5.18) is left,

$$\sum_{b_n} p(b_n|\mathbf{b}_{1:n-1}^{\mathbf{l}}) = 1,$$

which is again identical to one, and thus proofs that all columns of $\mathbf{A}$ sum up to one.

Besides of $\mathbf{A}$ being stochastic, all its elements need to be in the range

$$0 < a^{\mathbf{k},\mathbf{l}} < 1,\ \forall \mathbf{k}, \forall \mathbf{l} \in \mathbf{1}, \ldots |\mathbf{\Omega_{1:n}}|$$

in order to have a unique, stationary distribution as stated in Section 3.2.3. This requirement is satisfied while demanding strict positivity of the complete conditional distributions since each element $a^{\mathbf{k},\mathbf{l}}$ is a product of discrete probability values of complete conditional distributions.

Thus, the equation system of (5.17) satisfies all prerequisites for obtaining a stationary distribution $\mathbf{p}$. It can be calculated using one of the three presented methods for determining the stationary distribution of discrete Markov chains, presented in Section 3.2.3.

5.3.2 Recursive Backwards Reconstruction

This section proposes a general procedure to break down the n random variable case into a successive solution of linear equation systems based on two random variables at a time.
The derivation starts with the same strategy as discussed in Section 5.2

$$\begin{aligned} p(b_1, b_2, \ldots, b_n) =& p(b_1|b_2, \ldots b_n)p(b_2, \ldots b_n) \\ =& p(b_1|b_2, \ldots b_n) \sum_{b'_1} p(b'_1, b_2, \ldots b_n), \end{aligned}$$

where one possible factorization of the joint probability distribution is chosen in order to incorporate the first complete conditional distribution. Afterwards the introduced incomplete marginal is restored to all n random variables. However, instead of using a factorization to directly incorporating a second complete conditional, e.g. $p(b'_2, |b_1, b_3, \ldots b_n)$, another decomposition is chosen yielding

$$p(b_1, b_2, \ldots, b_n) = p(b_1|b_2, \ldots b_n) \sum_{b'_1} p(b_2, \ldots, b_n|b'_1)p(b'_1),$$

and remarginalizing the introduced $p(b'_1)$ results in

$$\begin{aligned} p(b_1, b_2, \ldots, b_n) =& p(b_1|b_2, \ldots b_n) \sum_{b'_1} \underbrace{p(b_2, \ldots, b_n|b'_1)}_{\text{unknown joint}} \\ & \sum_{b'_2, \ldots, b'_n} p(b'_1, b'_2, \ldots, b'_n). \end{aligned} \tag{5.19}$$

Thus, (5.19) has the same structure as the two random variable case, derived in (5.5). Hence, the formal proof of Section 3.2.3 can be applied guaranteeing that the stationary distribution $p(b_1, \ldots, b_n)$ exists. Moreover, the presented methods for determining the stationary distribution can also be applied. This presupposes that the partial joint $p(b_2, \ldots, b_n|b'_1)$ is known. Nevertheless, as Section 4.1.2 stresses, a large number of random variables in the distribution's target space increases the required complexity of the underlying classification algorithm. Additionally, the number of needed models to cope with variable traffic scenes rapidly increases.

However, the same strategy as in (5.19) can be used to express $p(b_2, \ldots, b_n|b'_1)$ for a fixed event b'_1 in the conditional

$$\begin{aligned} p(b_2, \ldots, b_n|b'_1) =& p(b_2|b_3, \ldots, b_n, b'_1)p(b_3, \ldots b_n|b'_1) \\ =& p(b_2|b_3, \ldots, b_n, b'_1) \sum_{b'_2} p(b_3, \ldots, b_n|b'_1, b'_2)p(b'_2|b'_1) \\ =& p(b_2|b_3, \ldots, b_n, b'_1) \sum_{b'_2} \underbrace{p(b_3, \ldots, b_n|b'_1, b'_2)}_{\text{unknown joint}} \\ & \sum_{b'_3, \ldots, b'_n} p(b'_2, b'_3, \ldots, b'_n|b'_1), \end{aligned} \tag{5.20}$$

leading again to a structure identical to the linear case. It should be noted that (5.20) needs to be calculated for each $b'_1 \in \Omega_1$ to have all of the needed partial joints in (5.19). Though, (5.20) makes use of another unknown partial joint $p(b_3, \ldots, b_n|b'_1, b'_2)$, which needs to be determined for all $|\Omega_1||\Omega_2|$ possible event combinations. Compared to the first linear equation system, the number of random variables on the left of the conditional has further decreased by one for this partial joint.

Consequently, repeating this procedure of formulating linear equation systems for unknown partial joint probabilities leads to a point where only two random variables b_{n-1} and b_n remain on the left side of the conditional, i.e. $p(b_{n-1}, b_n|b'_1, \ldots b'_{n-2})$. The fixed event combinations for random variables $1, \ldots, n-2$ on the right of the conditional are abbreviated

by the multivariate $\mathbf{b}_{1:n-2} = [b'_1, \ldots, b'_{n-2}]$, leading to

$$\begin{aligned} p(b_{n-1}, b_n|\mathbf{b}_{1:n-2}) =& p(b_{n-1}|b_n, \mathbf{b}_{1:n-2})p(b_n|\mathbf{b}_{1:n-2}) \\ =& p(b_{n-1}|b_n, \mathbf{b}_{1:n-2}) \sum_{b'_{n-1}} p(b_n|b'_{n-1}, \mathbf{b}_{1:n-2}) \\ & p(b'_{n-1}|\mathbf{b}_{1:n-2}) \\ =& p(b_{n-1}|b_n, \mathbf{b}_{1:n-2}) \sum_{b'_{n-1}} p(b_n|b'_{n-1}, \mathbf{b}_{1:n-2}) \\ & \sum_{b'_n} p(b'_n, b'_{n-1}|\mathbf{b}_{1:n-2}), \end{aligned}$$

only containing complete conditionals and the partial joint probability distribution $p(b_n, b'_{n-1}|\mathbf{b}_{1:n-2})$ itself. Hence, $p(b_{n-1}, b_n|\mathbf{b}_{1:n-2})$ can be directly determined for all event combinations of $\mathbf{b}_{1:n-2}$ with the methods discussed in Section 3.2.3. Subsequently, this determined partial joint distribution is used to solve the linear equation system for determining $p(b_{n-2}, b_{n-1}, b_n|\mathbf{b}_{1:n-3})$ which again is used to determine $p(b_{n-3}, b_{n-2}, b_{n-1}, b_n|\mathbf{b}_{1:n-4})$, etc. This recursive procedure is repeated until the target joint distribution can be determined using (5.19).

After this exemplary outline of how a recursive strategy is able to determine the target joint distribution, the following formally outlines the approach. For the general n random variable case $r = 1, \ldots, n-1$ recursion steps are necessary, where each recursion step r determines the partial joint distribution

$$p(\mathbf{b}_{r:n}|\mathbf{b}_{1:r-1}). \tag{5.21}$$

This is possible by calculating (5.21) for each event combination $\mathbf{b}_{1:r-1} \in \Omega_{1:r-1}$, separately, by determining the stationary distribution of

$$\mathbf{p}_r^{\mathbf{s}} = \begin{pmatrix} p(\mathbf{b}_{r:n}^{\mathbf{1}}|\mathbf{b}_{1:r-1}^{\mathbf{s}}) \\ \vdots \\ p(\mathbf{b}_{r:n}^{\mathbf{\Omega_{r:n}}}|\mathbf{b}_{1:r-1}^{\mathbf{s}}), \end{pmatrix}$$

where the multi-index $\mathbf{s} = \mathbf{1}, \ldots, |\mathbf{\Omega_{1:r-1}}|$ indexes all event combinations of $\Omega_{1:r-1}$, with

$$\mathbf{p}_r^{\mathbf{s}} = \mathbf{A}_r^{\mathbf{s}} \mathbf{p}_r^{\mathbf{s}}.$$

$\mathbf{A}_r^{\mathbf{s}}$ denotes the coefficient matrix and is identically constructed like outlined in Section 5.2 for the two random variable case using the complete

conditional

$$p(b_r|\mathbf{b}_{r+1:n}, \mathbf{b}^s_{1:r-1})$$

and the partial joint

$$p(\mathbf{b}_{r+1:n}|b_r, \mathbf{b}^{\mathbf{s}}_{1:r-1}), \tag{5.22}$$

resulting from the preceding recursion step $r+1$. In the deepest possible case of $r = n-1$ the partial joint of (5.22) is identical to a complete conditional distribution. Fig. 5.1 outlines the recursive approach.

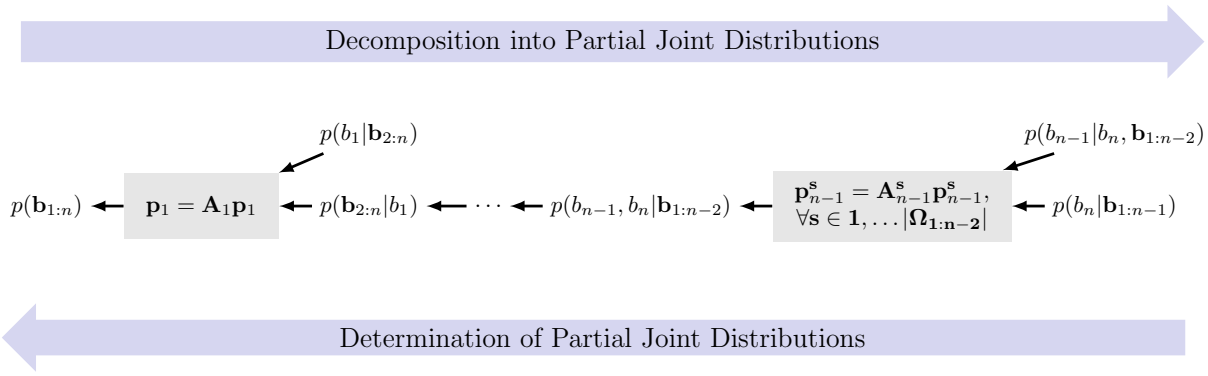


Figure 5.1: Outline of the recursive approach for reconstructing the target joint distributions. At each recursion depth r, $|\Omega_{r:n}|$ linear equation systems determine the partial joint distributions until the target distribution $p(\mathbf{b}_{1:n}) = \mathbf{p}_1$ is obtained.

Stochastic Matrices and Stationary Distributions

The following transfers the requirements assuring the existence of a unique stationary distribution from the two random variable case to the proposed recursive procedure. Section 3.2.3 discussed requirements for the existence of a unique stationary distribution although they can only be used directly for the deepest recursion step ($r = n-1$). That is because all remaining recursion depths' coefficient matrices are based on one complete conditional and one partial joint, i.e. the stationary distribution from the previous recursion step, e.g. (5.19) and (5.20).

The existence as well as the possibility to calculate a unique stationary distribution for the linear case is based on the following two requirements, already stated in Section 3.2.3. **First**, the coefficient matrix $\mathbf{A}^{\mathbf{s}}_r$ needs to be a stochastic matrix, i.e. its columns or rows (or both) sum up to 1. **Second**, all elements of $\mathbf{A}^{\mathbf{s}}_r$ of need to satisfy $0 < a^{o,u} < 1, \forall o, u \in 1, \ldots, |\Omega_{r:n}|$. As a consequence, to guarantee the existence of $p(b_1, \ldots, b_n)$

calculated by the proposed recursive procedure, these requirements need to be satisfied each recursion step.

Section 3.2.3 conducted the general proof that the column sum of $\mathbf{A}$ sums always up to one, and thus $\mathbf{A}$ is a stochastic matrix. Taking a closer look at how the coefficient matrix is constructed, e.g. (5.20), it becomes obvious that each coefficient is the product of a complete conditional and a partial joint distribution instead of the product of two complete conditionals. However, this difference does not alter with the columns summing up to 1 as the following shows. The sum of an arbitrary column o of $\mathbf{A}_r^{\mathbf{s}}$ with its fixed characteristic b_r^k is given by

$$\sum_{b'_r \in \Omega_1} \sum_{\mathbf{b}'_{r+1:n} \in \Omega_{r+1:n}} p(b'_r | \mathbf{b}'_{r+1:n}, \mathbf{b}^{\mathbf{s}}_{1:r-1}) p(\mathbf{b}'_{r+1:n} | b_r^k, \mathbf{b}^{\mathbf{s}}_{1:r-1}),$$

which is identical to 1 as

$$\begin{aligned} &\sum_{\mathbf{b}'_{r+1:n} \in \Omega_{r+1:n}} p(\mathbf{b}'_{r+1:n} | b_r^k, \mathbf{b}^{\mathbf{s}}_{1:r-1}) \underbrace{\sum_{b'_r \in \Omega_1} p(b'_r | \mathbf{b}'_{r+1:n}, \mathbf{b}^{\mathbf{s}}_{1:r-1})}_{=1} \\ &= \sum_{\mathbf{b}'_{r+1:n} \in \Omega_{r+1:n}} p(\mathbf{b}'_{r+1:n} | b_r^k, \mathbf{b}^{\mathbf{s}}_{1:r-1}) = 1 \end{aligned}$$

yields. Hence, the **first** requirement for an arbitrary recursion depth r is met.

Additionally, the **second** requirement demanding that all elements of $\mathbf{A}_r^{\mathbf{s}}$ of need to satisfy $0 < a^{o,u} < 1, \forall o, u \in 1, \ldots, |\Omega_{r:n}|, \forall r$. Each coefficient in $\mathbf{A}_r^{\mathbf{s}}, \forall r < n-1$ is a product of a complete conditional and a partial joint distribution, obtained from the previous recursion depth. The strict positivity of complete conditionals has already been guaranteed in Section 3.2.3. Hence, showing that the partial joints $p(\mathbf{b}_{r+1:n} | b_r, \mathbf{b}^{\mathbf{s}}_{1:r-1})$ obtained by each recursion step are also strictly positive fulfills the second requirement.

Equation (3.20) states the possibility of obtaining the stationary distribution through actually calculating the discrete Markov chain

$$\mathbf{p}_r^{\mathbf{s}} = \lim_{\alpha \to \infty} (\mathbf{A}_r^{\mathbf{s}})^{\alpha} \hat{\mathbf{p}}_r^{\mathbf{s}^{(0)}},$$

for large exponents α. Since the stationary distribution $\mathbf{p}_r^{\mathbf{s}}$ is independent of the starting distribution $\hat{\mathbf{p}}_r^{\mathbf{s}^{(0)}}$ the columns of $(\mathbf{A}_r^{\mathbf{s}})^{\alpha}$, for a sufficient large

exponent α, are identical to the actually desired stationary distribution $\mathbf{p}_r^{\mathbf{s}}$, i.e.

$$(\mathbf{A}_r^{\mathbf{s}})^{\alpha} = (\mathbf{p}_r^{\mathbf{s}}, \ldots, \mathbf{p}_r^{\mathbf{s}}).$$

Since the deepest recursion step $r = n - 1$ is identical to the two random variable case of Section 3.2.3, it satisfies $0 < a^{o,u} < 1, \forall o, u \in 1, \ldots, |\Omega_{n-1:n}|$. Consequently, $\mathbf{p}_{n-1}^{\mathbf{s}}$ is also strictly positive since multiplying strictly positive matrices ($(\mathbf{A}_{n-1}^{\mathbf{s}})^{\alpha}$) results in a strictly positive matrix itself. Furthermore, $\mathbf{p}^{\mathbf{s}}$ is a probability distribution ($\sum \mathbf{p}^{\mathbf{s}} = 1$) and thus each element $\mathbf{p}^{\mathbf{s}}$ is truly smaller than 1.
Hence, the elements of $\mathbf{A}_{n-2}^{\mathbf{s}}$ are also in the range $(0, 1)$ since its coefficients are products of a complete conditional distribution and $\mathbf{p}_{n-1}^{\mathbf{s}}$, which in turn results in $\mathbf{p}_{n-2}^{\mathbf{s}}$ being in the range $(0, 1)$, for $r = n-2$. This procedure can be applied to all recursion steps r. For this reason, the coefficient matrix and its corresponding stationary distribution of each recursion step are always strictly positive smaller 1.

Consequently, both stated requirements are satisfied in each recursion step and a unique stationary distribution, identical to the desired target joint distribution $p(b_1, \ldots, b_n)$, exists and can be calculated.

Total Linear Equation Systems

The recursive reconstruction of joint probabilities with more than two random variables only based on complete conditionals can be performed by $r = 1, \ldots, n-1$ recursion steps. The previous section indicated that in each recursion step a certain number of linear equation systems needs to be solved. This number corresponds to the cardinality of possible event combinations of the random variables in the conditional of the partial joints, e.g. for $p(b_{n-1}, b_n|\mathbf{b}_{1:n-2})$ for each of $|\Omega_1| \ldots |\Omega_{n-2}|$ possible event combinations one linear equation system needs to be solved. This results in the following succession of number of equation systems to be solved per recursive depth r:

$$\begin{aligned}
&r = 1 \rightarrow 1,\\
&r = 2 \rightarrow |\Omega_1|,\\
&r = 3 \rightarrow |\Omega_1||\Omega_2|,\\
&\quad\vdots\\
&r = n-1 \rightarrow |\Omega_1||\Omega_2| \ldots |\Omega_{n-2}|.
\end{aligned}$$

As a result in order to reconstruct $p(b_1, \ldots, b_n)$ completely

$$1 + \sum_{r=2}^{n-1} |\Omega_{1:r-1}| \tag{5.23}$$

linear equation systems need to be solved in total.

While this number quickly increases with large sample spaces and random variables n, the solution of each linear equation system is in general based on highly optimized algorithms, as discussed in Section 3.2.3 and can therefore still be applied for large overall sample spaces $\Omega_{1:n}$.

5.3.3 Recursive Forward Reconstruction

The newly developed method for recursively calculating the joint distribution is characterized by first determining the partial joint distributions on the deepest recursion depth $r = n - 1$. Starting from there, the degree of the partial joints is successively increased by conducting the recursion steps in reversed order $r = n - 2, \ldots, 2, 1$. Hence, the joint distribution is reconstructed backwards.
The authors of [12] presented an alternative approach that determines the missing incomplete conditionals of a general factorization. Their approach makes use of the possibility to directly calculate incomplete distributions from their corresponding complete conditionals. The basic concept is related to the direct marginal determination, discussed in Section 3.2.3. The following illustrates the application of the forward recursion presented in [12] on a $n = 3$ random variable example. Starting with the factorization given in lexicographical order

$$p(b_1, b_2, b_3) = \underbrace{p(b_1|b_2, b_3)}_{r=1} \underbrace{p(b_2|b_3)}_{r=2} \underbrace{p(b_3)}_{r=3}. \tag{5.24}$$

It consists of $r = 1, \ldots, 3$ factors that correspond to the needed recursion steps. The factor $p(b_1|b_2, b_3)$ for recursion step $r = 1$ is equivalent to the first of the three given complete conditional distributions. The incomplete conditional $p(b_1|b_2)$ can be calculated directly, starting with two possible factorizations of $p(b_1, b_2|b_3)$

$$p(b_1, b_2|b_3) = p(b_1|b_2, b_3)p(b_2|b_3) = p(b_2|b_1, b_3)p(b_1|b_3). \tag{5.25}$$

Subsequently, marginalizing b_1 yields

$$\sum_{b_1' \in \Omega_1} p(b_2|b_3) \frac{p(b_1'|b_2, b_3)}{p(b_2|b_1', b_3)} = 1, \tag{5.26}$$

where the unknown incomplete can be directly determined by

$$p(b_2|b_3) = \frac{1}{\sum_{b'_1 \in \Omega_1} \frac{p(b'_1|b_2,b_3)}{p(b_2|b'_1,b_3)}}. \tag{5.27}$$

The unknown incomplete of $r = 3$ can be calculated in the same way with

$$p(b_3) = \Big(\sum_{b'_2 \in \Omega_2} \frac{p(b'_2|b_3)}{p(b_3|b'_2)} \Big)^{-1}.$$

However, it is based on the unknown conditional $p(b_3|b_2)$ distribution. It can be determined in the same way as done for $p(b_2|b_3)$, outlined in equations (5.25), (5.26) and (5.27). Thus, the missing distribution can be determined using

$$p(b_3|b_2) = \Big(\sum_{b'_1 \in \Omega_1} \frac{p(b'_1|b_2,b_3)}{p(b_3|b'_1,b_2)} \Big)^{-1}.$$

Finally, all needed unknown conditional distributions are determined and the joint distribution is calculated according to (5.24).

The authors of [12] generalized this procedure for n random variables. Assuming that the factorization is given in a lexicographical order, i.e. each incomplete conditional's random variable with the smallest superscript is given by the current recursion depth r, the missing incomplete conditional distributions can be calculated according to

$$p(b_i|\mathbf{b}_{r:i-1}, \mathbf{b}_{i+1:n}) = \Big(\sum_{b'_{r-1} \in \Omega_{r-1}} \frac{p(b'_{r-1}|\mathbf{b}_{r:n})}{p(b_i|b'_{r-1}, \mathbf{b}_{r:i-1}, \mathbf{b}_{i+1:n})} \Big)^{-1}, \forall i = q, \ldots, n. \tag{5.28}$$

As indicated before, the factor for recursion step $r = 1$ is already given by a complete conditional distribution.

Equation (5.28) emphasizes the forward character of the recursion scheme. The dimension of the conditional distributions on the right hand side of (5.28) surpasses the dimension of the target distribution on the left-hand side by one. Thus, the missing incompletes need to be calculated in the order of $r = 2, \ldots, n$.

In general, at each recursion step $r > 2$, all combinations of remaining random variables $\mathbf{b}_{r:n}$ need to be determined. This corresponds to determining $n - r + 1$ distributions at each recursion step, and thus overall

$$1 + \sum_{r=2}^{n} (n - r + 1), \tag{5.29}$$

distributions need to be calculated.

Comparison of Backward and Forward Recursion

Taking a closer look at the overall number of unknown distributions that need to be determined in (5.23) and (5.29) for both discussed methods suggests that the forward recursion of [12] scales more favorably, although these two numbers do not take the size and the effort for determining these distributions into account. Moreover, the proposed backwards recursion can be calculated using the same principals of the marginalization calculation (5.28) or (5.12) to omit solving linear equation systems at each recursion step.

The following models the needed computational effort for both methods explicitly. For the comparison of the computational effort of both methods the cardinality of sample spaces is assumed to be identical, hence, $|\Omega|_i = |\Omega|, \forall i \in 1, \ldots, n$.

Section 5.3.3 stated that the forward recursion needs to determine

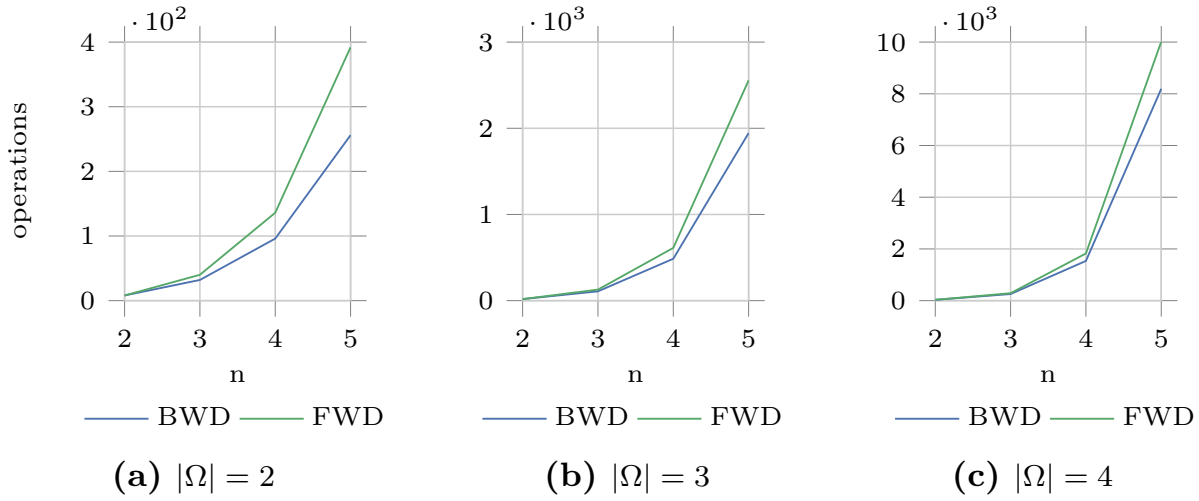
$$1 + \sum_{r=2}^{n} (n - r + 1),$$

marginal distributions where each marginal distribution has $|\Omega|^{n-r+1}$ elements that need to be determined. Taking a closer look at (5.28) reveals that each determination of a marginal distribution needs $|\Omega| - 1$ summations and $|\Omega|$ divisions in the denominator as well as one additional operation to divide nominator and denominator. Thus, one marginal calculation according to (5.28) needs $2|\Omega|$ operations.
The overall effort for determining all needed distributions for the forward recursion sums up to

$$\sum_{r=2}^{n} 2|\Omega|(n - r + 1)|\Omega|^{n-r+1} = \sum_{r=2}^{n} 2(n - r + 1)|\Omega|^{n-r+2}. \tag{5.30}$$

For the proposed backward algorithm one marginal distribution with $|\Omega|^{n-1}$ elements at each recursion depth $r > 1$ needs to be calculated. The effort for the determining marginal distributions is the same as for the determining marginal distributions for the forward recursion and is given by $2|\Omega|$. Thus, the overall effort for determining all needed marginal distributions for the backward recursion sums up to

$$\sum_{r=2}^{n} 2|\Omega||\Omega|^{n-1} = \sum_{r=2}^{n} 2|\Omega|^{n}. \tag{5.31}$$



(a) $|\Omega| = 2$ **(b)** $|\Omega| = 3$ **(c)** $|\Omega| = 4$

Figure 5.2: Detailed differences of the computational effort for the forward and the proposed backward recursion.

Taking a closer look at the computational efforts (5.30) and (5.31) shows that both scale in the same order $\mathcal{O}(|\Omega|^n)$. Fig. 5.2 shows the detailed differences in computational effort for exemplary samples space sizes $|\Omega|$ and different number of random variables n. The comparison points out that the proposed backward recursions effort for determining the missing marginal distributions scales more favorably than the forward recursion.

The newly proposed recursion is an alternative to the forward recursion presented in [12] that scales with same order of magnitude. Moreover, the partial joint probability of its interim results pose interesting distributions that can be subject of further investigations. Section 6 introduces a hypotheses selection scheme based on these partial joint distributions.

5.3.4 Robustness of Reconstruction Methods

This section compares the robustness of the discussed methods for analytically reconstructing joint distributions to errors in the conditional distributions. These errors can be a result of imperfectly learned or derived heuristic models that determine the conditional distributions. They potentially affect the reconstruction process.

The following evaluations compare four different ways to reconstruct joint distributions from erroneous conditional distributions. All of them do not rely on any independence assumptions. The first model uses one

decomposition based on the chain rule of probability

$$\mathcal{M}_{\text{decomp}} = \prod_{i=1}^{n} p(y_i|\mathbf{y}_{i+1:n}),$$

and serves as reference. It depends on marginal distributions. The other three models reconstruct the target joint distribution using only complete conditional distributions. They only differ in their reconstruction method. Model $\mathcal{M}_{\text{LGS}}$ directly solves the linear equation system based on the derived reconstruction matrix in (5.17). As an alternative, model $\mathcal{M}_{\text{BWD, LGS}}$ uses the proposed recursive method of Section 5.2.1, solving smaller linear equation systems at each recursion depth r. Both methods iteratively incorporate all $!n$ possible factorizations of the joint distribution.
The last evaluated model $\mathcal{M}_{\text{marg}}$ makes use of the same recursive decomposition as $\mathcal{M}_{\text{BWD, LGS}}$ but calculates the interim distributions according to (5.12).

Randomly instantiated joint distributions $p(\mathbf{y}_{1:n})$ with a sample space sizes $|\Omega_i|^n$ provide the basis of the comparative evaluation. The required conditionals can be determined using the chain-rule of probability. The resultant conditional distributions form the basis for the reconstruction methods. However, to simulate errors in the conditional distributions, they are superimposed with a random process according to

$$\check{p}(y_i^l|\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) = p(y_i^l|\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) + \epsilon, \; \forall l \in 1, \ldots, |\Omega_i|, \qquad (5.32)$$

where ϵ represents a random number in the interval $[0, 0.05]$ drawn from a uniform density.
The same random process is used to affect the marginal distributions $\check{p}(y_i|\mathbf{y}_{i+1:n})$ with errors in the same magnitude of ϵ. The addition of a positive error term preserves the non-negativity of the distributions. However, they will no longer sum up to one. To restore this property the altered distributions are normalized afterwards.
The four models reconstruct a joint distribution $\hat{p}(\mathbf{y}_{1:n})$ that approximates the starting distribution $p(\mathbf{y}_{1:n})$.
A popular measure for quantifying the dissimilarity of probability distributions is the Kullback-Leibler divergence, [16]. It is therefore well suited to quantify the difference between the concise starting joint distribution $p(\mathbf{y}_{1:n})$ and the reconstructed joint distribution $\hat{p}(\mathbf{y}_{1:n})$, based on the erroneous conditional distributions.

The Kullback-Leibler divergence for discrete distributions is given by

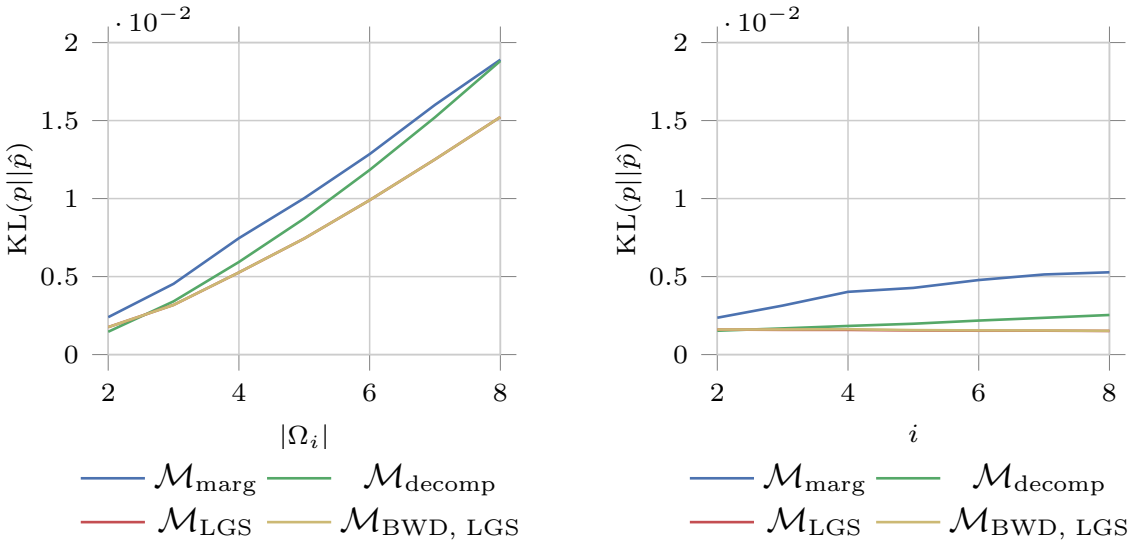
$$\mathrm{KL}(p|\hat{p}) = \sum_{\mathbf{y}_{1:n} \in \Omega_{1:n}} p(\mathbf{y}_{1:n}) \ln \left(\frac{p(\mathbf{y}_{1:n})}{\hat{p}(\mathbf{y}_{1:n})} \right).$$

The process of randomly initiating joint distributions, calculating their conditionals, artificially adding noise to them, reconstruct the joint distribution and determine the resultant Kullback-Leibler divergence is repeated for different numbers of random variables i and sample space sizes $|\Omega_i|$. For each distinct combination $(i, |\Omega_i|)$ 4000 randomly chosen joint distributions are reconstructed.

Fig. 5.3 (a) shows the development of the $\mathrm{KL}(p|\hat{p})$ for increasing sample space sizes $|\Omega_i|$ while the number of random variables $i = 2$ remains constant. The curves show that the reconstruction method based on solving linear equation systems ($\mathcal{M}_{\mathrm{LGS}}, \mathcal{M}_{\mathrm{BWD,\,LGS}}$) cope better with erroneous conditional distributions than models $\mathcal{M}_{\mathrm{decomp}}$ and $\mathcal{M}_{\mathrm{marg}}$. Moreover, the evaluation outlines that there is no significant difference between reconstructing the joint distribution with one equation systems ($\mathcal{M}_{\mathrm{LGS}}$) or several smaller equation systems ($\mathcal{M}_{\mathrm{BWD,\,LGS}}$). Fig. 5.3 (b) shows similar results for an increasing number of random variables i while leaving the sample space size $|\Omega_i| = 2$ constant. Moreover, the comparison of both figures outlines that the increase in sample space size has a more severe impact than an increase in the number of random variables. E.g. the overall sample space size $|\Omega|_{1:n}$ is identical for $|\Omega_i| = 8$ in (a) and $i = 6$ in (b), while the dissimilarity for the larger sample space size in (a) is larger.

This observation can be explained by the stronger relative influence of the error term ϵ for larger sample space sizes. The error term is added for each event of y_i in (5.32). However, for increasing sample space sizes the probability values tend to take on smaller values while ϵ remains constant. This leads to an increasing distortion of the conditional distributions. Increasing the number of random variables while leaving the sample space size constant is not subject to this effect and the overall error remains on a significantly lower overall level. In fact the reconstruction methods based on solving linear equation systems are almost not effected by the increasing number of random variables.

The reason for them to perform better than $\mathcal{M}_{\mathrm{decomp}}$ when confronted with erroneous conditional distributions is due their averaging characteristic. Taking a closer look at the coefficient matrices $\mathbf{A}$ of (5.8) and (5.17) reveals each reconstructed probability value is a sum over all probability values where all $!n$ possible decompositions of the factorization of $p(\mathbf{y}_{1:n})$



(a) Increasing $|\Omega_i|$, constant i **(b)** Increasing i, constant Ω_i

Figure 5.3: (a) Shows the mean Kullback-Leiber divergence for an increasing sample space sizes $|\Omega_i|$ and for a constant $i = 2$ for 4000 runs. (b) Shows the mean Kullback-Leiber divergence for an increasing number of random variables i and a constant $|\Omega_i| = 2$ for 4000 runs. The curves of models $\mathcal{M}_{\text{LGS}}$ and $\mathcal{M}_{\text{BWD, LGS}}$ are overlapping

are incorporated. This leads to a beneficial averaging effect. In contrast, model $\mathcal{M}_{\text{marg}}$ suffers from equating different factorizations ((5.11)), since for erroneous conditional distributions this equation does not hold true. This distorts the resultant distribution. Moreover, it does not necessarily sum up to one. Not taking into account this possibility has an drastic effect on the KL$(p|\hat{p})$. Therefore, in the above evaluations this has already been compensated.

These results demonstrate that the proposed reconstruction methods, based on solving linear equation systems are beneficial when confronted with erroneous conditional distributions. However, directly calculating the missing marginal distributions as being done for $\mathcal{M}_{\text{marg}}$ scales more favorably. Hence, the different reconstruction methods offer advantages either in robustness or computational complexity.

5.4 Stochastic Independencies

The presented methods for reconstructing joint probabilities from complete conditional distributions are based on completely interdependent random

variables. While it is the most general assumption it is nevertheless possible and quite likely that settings are encountered where distinct stochastic independencies exist. This means that the complete conditional distributions do not have all remaining random variables $\mathbf{b}_{1:i-1}$ and $\mathbf{b}_{i+1:n}$ in their conditional, but a subset of it. Therefore, this section discusses possibilities to incorporate such a priori knowledge into the reconstruction process. This can be achieved by replacing each of the random variables in the Markov blanket of random variable i. By doing that, each present complete conditional $p(b_i|\mathbf{b}_{1:i-1}, \mathbf{b}_{i+1:n})$, in any of the presented methods for reconstruction joint probability distributions, can be replaced by a conditional only depending on its Markov blanket $MB(b_i)$. This can be done, since

$$p(b_i|\mathbf{b}_{1:i-1}, \mathbf{b}_{i+1:n}) = p(b_i|MB(b_i)).$$

The set of random variables of $\mathbf{b}_{1:i-1}, \mathbf{b}_{i+1:n}$ that specify the Markov blanket of b_i depends on the presence and type of dependencies between b_i and all remaining elements $\mathbf{b}_{1:i-1}, \mathbf{b}_{i+1:n}$.

A convenient way of finding a random variable's Markov blanket is to derive it from the causal graph of random variables $1, \ldots, n$. A causal graph illustrates the causal dependencies between random variables [87]. Fig. 5.4 shows two different scene layouts with different causal dependencies.

The authors of [96] state the relationship of casual graph structures and resultant probabilistic independencies, i.e. how to determine Markov blankets. All random variables that have bidirectional dependencies with random variable b_i belong to the Markov blanket $MB(b_i)$. In the case of unidirectional dependencies two types of connections between two random variables are possible. If b_i is causally dependent on some variable $b_{\setminus i}$, $b_{\setminus i}$ is said to be its parent. The other way around, if some variable $b_{\setminus i}$ is

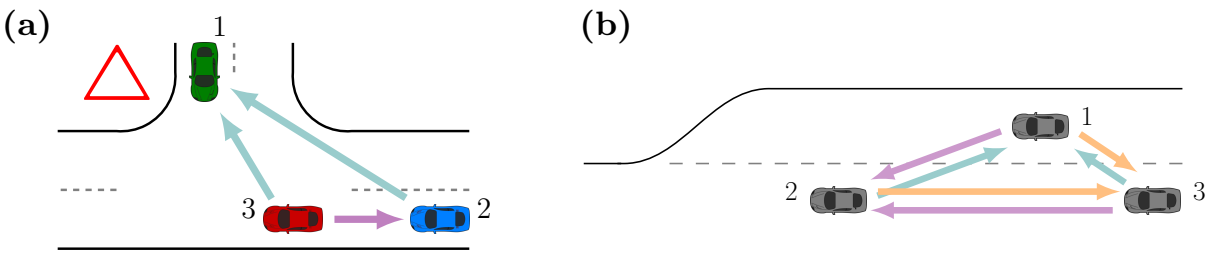


Figure 5.4: (a) Example where the causal dependencies are distinctly decomposable. The green car's behavior depends on the two vehicles on the road with priority. The blue vehicle depends on the behavior of the preceding vehicle, while the red car is independent from the two other vehicles. Part (b) shows a highway entering scene an example for an interdependent case.

causally dependent on b_i, $b_{\setminus i}$ is said to be its child. The Markov blanket of $MB(b_i)$ is composed of b_i's parents, its children and its children's parents in the case of directed dependencies [16]. At first glance, it is surprising that the conditional of complete conditional cannot simply be replaced by its parents. This is due to the fact that even though b_i is not causally dependent on its children, the knowledge of their state allows for conclusion on the distribution of b_i, i.e. the symmetry of stochastic dependencies. However, this is only possible if also the state of b_i's children's parents is considered, since the state of b_i's children is equally dependent on their other parents. This phenomenon is also referred to as *explaining away*, [88].

Fig. 5.5 shows two artificially created causal dependencies graphs. The example emphasizes the influence of different causal dependencies

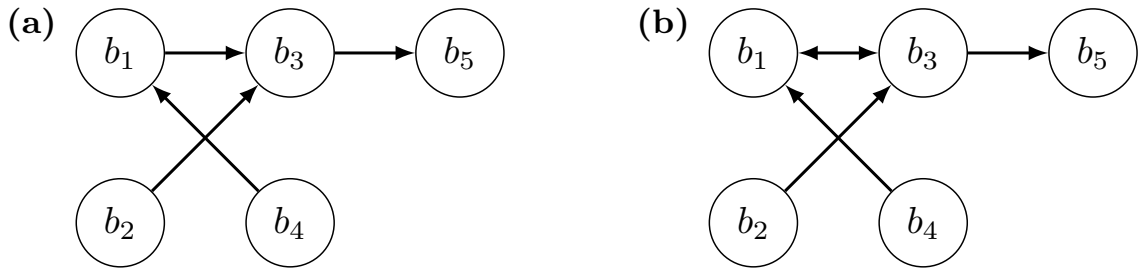


Figure 5.5: Two different set of dependencies displayed by a probabilistic graphical model.

between random variables on their corresponding Markov blanket. The models only differ in the type of connection between b_1 and b_3. For model (a) the $MB(b_1) = \{b_2, b_3, b_4\}$, whereas the Markov blanket of b_1 for model (b) is composed of $MB(b_1) = \{b_3, b_4\}$. This is a quite remarkable result as an additional dependency actually reduces the cardinality of the variables b_1 is conditioned on.

Regardless of the dependencies, the proposed reconstruction methods can be straightforwardly used by simply replacing the complete conditionals with the conditionals only dependent on the corresponding Markov blanket. The proposed framework uses learned models for determining the complete conditional distribution of random variable b_i. Hence, these models can directly be based on $MB(b_i)$ during the design process.

5.5 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, the following evaluation compares the performance of the proposed fully interaction-respecting approach compared to a set of different possibilities to model joint distributions on these scenarios.
The used data is extracted from the NGSIM real-world data set introduced in detail in Section 2.3.1. 1070 lane change scenarios are extracted.

5.5.1 Complex Highway 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. 5.6 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 this group as well as for assistance systems in vehicles which are part of the group.

According to Section 4, three possible situation hypotheses are defined 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. The scene allows for three expressive longitudinal behavior patterns in relation to the expected longitudinal position obtained by a kinematic extrapolation for a prediction horizon of Δt

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

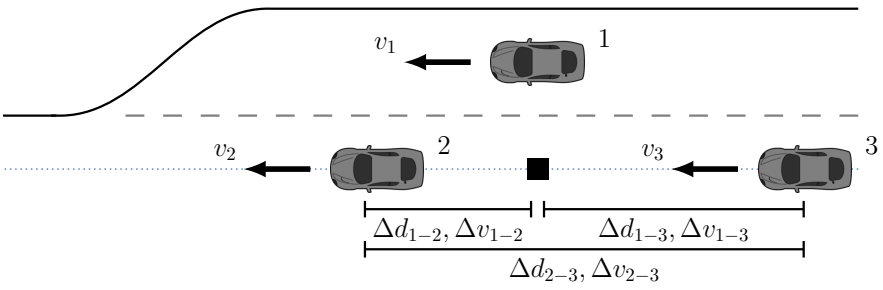


Figure 5.6: Investigated scenario consistent of three scene entities. Additionally, the extracted relative features are shown.

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 sample spaces Ω_2 and Ω_3 are identically composed of the events

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

are illustrated in 5.7 (a).

Regarding vehicle 1, it is of interest in front of which vehicle (2 or 3) it is going to merge or if no merge is performed during the observed time horizon Δt. Thus, the sample space Ω_1 is composed of

$$\Omega_1 = \{\text{betw}, \text{front}, \text{no}\},$$

and is illustrated in Fig. 5.7 (b). Merging scenarios where vehicle 1 enters the highway behind vehicle 3 are not considered, i.e. took not place during the time horizon of $\Delta t = 5s$ The set of possible hypotheses $\Omega_{1:3}$ is instantiated by all possible event combinations $\Omega_{1:3} = \Omega_1 \times \Omega_2 \times \Omega_3$ resulting in $|\Omega_{1:3}| = 27$ situation hypotheses.

Cases are extracted for vehicles on the on-ramp of the Interstate-80 data for a prediction horizon of Δt before a lane change (t_{ch}) or no lane change ($t_{\text{no,ch}}$). The point in time where 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 $\Omega_{1:3}$. Appendix A.4 shows the detailed distribution over the situation occurrences.

Evaluated Models

This evaluation benchmarks the proposed approach for reconstructing the joint probability distribution $p(b_1, b_2, b_3|\mathbf{x})$. It uses the backwards recursion discussed in Section 5.3.2 and is referred to as $\mathcal{M}_{\text{recons}}$. The conditional 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 the base for the reconstruction process and are learned from the extracted

data. Section 5.5.1 discusses the learning process. The joint probability distribution is obtained by calculating the partial joint distribution $p(b_2, b_3|\mathbf{x}, b_1)\forall b_1 \in \Omega_1$, based on the complete conditionals $p(b_2|\mathbf{x}, b_1, b_3)$ and $p(b_3|\mathbf{x}, b_1, b_2)$. Afterwards the target joint distribution results from combining $p(b_1|\mathbf{x}, b_2, b_3)$ and the calculated partial joints $p(b_2, b_3|\mathbf{x}, b_1)$ in the previous step.

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

$$\mathcal{M}_{\text{direct}} : p(b_1, b_2, b_3|\mathbf{x}), \tag{5.33}$$

is trained. 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 4.1.2.

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 modeling only subsets of the present dependencies between the three vehicles.
First, a model is created that uses a decomposition obtained by the chain rule of probability:

$$\mathcal{M}_{\text{decomp}} : p(b_1|\mathbf{x}, b_2, b_3)p(b_3|\mathbf{x}, b_2)p(b_2|\mathbf{x}). \tag{5.34}$$

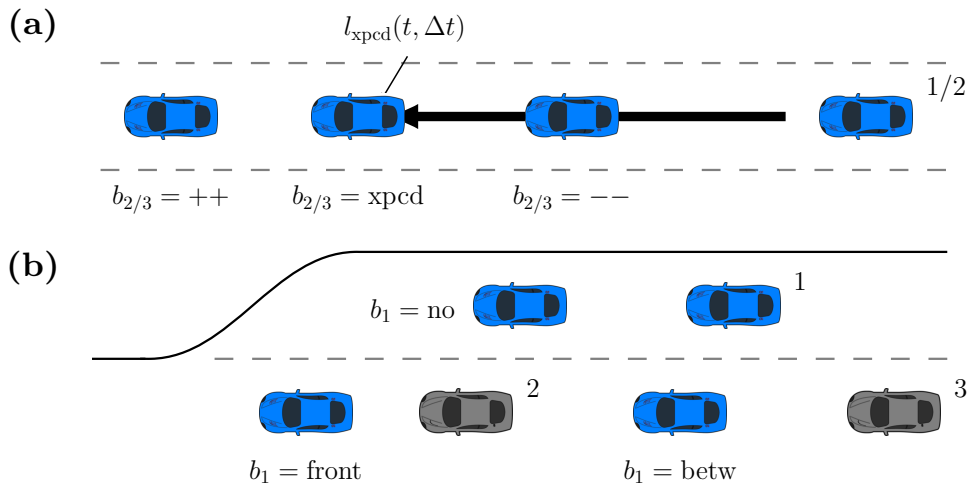


Figure 5.7: (a) the three possible behavior patterns/events for vehicle 2 and 3. The kinematic extrapolated position $l_{\text{xpcd}}(t, \Delta t)$ is indicated by the square. Possible behavior patterns/events based on the actual observed $l(t + \Delta t)$ are indicated by the transparent blue vehicles. (b) the three possible events of vehicle 1. The transparent blue vehicles indicate each possibly occurring behavior pattern/event at time $t + \Delta t$.

The chosen decomposition focuses on explicitly modeling the most prominent dependencies. Thus, the dependency of vehicle 1 on vehicles 2 and 3 on the on-ramp, as well as the dependency of vehicle 3 of its direct predecessor, i.e. vehicle 2. Thus, $p(b_3|\mathbf{x}, b_2)$ and $p(b_2|\mathbf{x})$ correspond to marginal distributions rather than single-entity maneuver predictions.

Second, a model making the assumption that the scene entities behave completely stochastic independent ($b_1 \perp b_2 \perp b_3$) from each other is evaluated

$$\mathcal{M}_{\text{indep}} : p(b_1|\mathbf{x})p(b_2|\mathbf{x})p(b_3|\mathbf{x}), \tag{5.35}$$

where no direct dependencies between the vehicles are respected. This represents a strong assumption, since it can not be assumed that vehicles do not take possible future maneuvers of their surrounding vehicles into account in such a scenario.

Learning (Conditional) Probability Distributions

For each of the conditional probability distributions present in models $\mathcal{M}_{\text{recons}}$, (5.34) and (5.35), a probabilistic multiclass classifier, where each possible class corresponds to a possible event of the particular sample space Ω_i, is trained. 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 (5.33), consists of $|\Omega_{1:3}| = 27$ classes.

A logistic regression is a suitable candidate for determining 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 [85]. 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 $b_2 \perp b_1 \perp b_3$, nevertheless uses relative

kinematic features dependent on scene entities 1 and 3. The common scene information is given by

$$\mathbf{x} = [v_1, v_2, v_3, \Delta d_{1-2}, \Delta d_{1-3}, \Delta d_{2-3}, \Delta v_{1-2}, \Delta v_{1-3}, \Delta v_{2-3}]^T \quad (5.36)$$

where $\Delta d_{1-2}, \Delta d_{1-3}, \Delta d_{2-3}$ correspond to the relative longitudinal distances (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. 5.6 illustrates the construction of the used features.[1] 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. A.4 into account, all models are trained weighting the importance of each case with the inverse of its overall situation hypothesis prototype occurrence.

Evaluation

The evaluation is performed using a 4-fold stratified cross validation extracted observations. The *ROC* curve is a suitable metric to compare the different models, as it allows for comparing the quality of the obtained joint probability distributions. Moreover, it also takes 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. For this general evaluation no specific application is assumed and hence no minimum reliability (TPR) or maximum number of false warnings (FPR) specified. Thus, the *ROC* curve and its respective *AUC* provide the most objective metrics. Section 2.4 introduced these metrics in detail.

The *ROC* performance for a prediction horizon $\Delta t = 5s$ is evaluated. Fig. 5.8 shows the corresponding *ROC* curves and Tab. 5.1 contains detailed information regarding the *AUC*s for all investigated models.

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

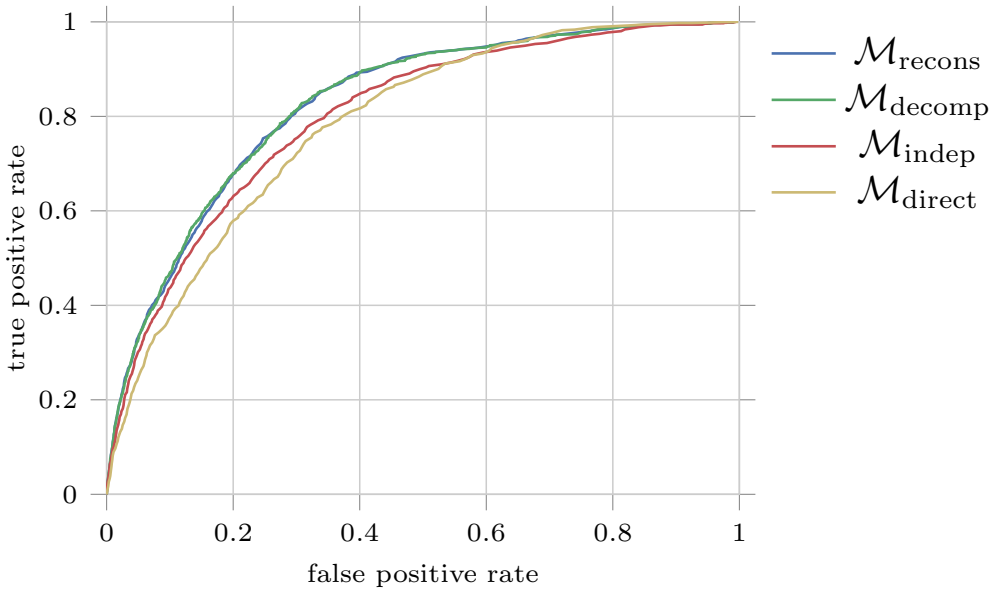


Figure 5.8: *ROC* for the different investigated models for estimating $p(b_1, b_2, b_3|\mathbf{x})$

The gap between the performance of the proposed model $\mathcal{M}_{\text{recons}}$ and $\mathcal{M}_{\text{direct}}$ is a result of the high complexity of its target space.The classifier for determining $p(b_1, b_2, b_3|\mathbf{x})$ needs to directly discriminate between 27 classes (situation hypotheses). This is a drastic increase in target space dimensionality compared to the models for obtaining the complete conditionals (3 classes each) of $\mathcal{M}_{\text{recons}}$ and explains the performance differences. In theory, these systems should perform on similar levels for sufficiently large data sets and a more complex classification algorithm for model $\mathcal{M}_{\text{direct}}$. However, models directly estimating $p(b_1, b_2, b_3|\mathbf{x})$ are not suited for variable traffic.

Fig. 5.8 shows a clear performance advantage of model $\mathcal{M}_{\text{recons}}$ compared to the model $\mathcal{M}_{\text{indep}}$ assuming stochastic independence. This result underlines that it is beneficial to explicitly model interactions in traffic scenarios with more than one vehicle and that the assumption of stochastic independence needs to be verified beforehand.

Moreover, the *ROC* performance of the proposed model $\mathcal{M}_{\text{recons}}$ and $\mathcal{M}_{\text{decomp}}$ are difficult to separate based on the curves in Fig. 5.8. This is not unexpected considering that both models respect interactions and decompose the determination of the joint probability distribution into smaller subproblems. Moreover, the probability decomposition of directly $\mathcal{M}_{\text{decomp}}$ models the most prominent interaction patterns of the scene. This means that for most scenarios the vehicle on the on-ramp is influ-

	$\Omega_{1:3}$		$\Omega_{1:3}^{\not}$	
	AUC	rel. increase	AUC	rel. increase
$\mathcal{M}_{\text{recons}}$	82.7	-	84.0	-
$\mathcal{M}_{\text{direct}}$	78.4	5.36	81.5	3.07
$\mathcal{M}_{\text{decomp}}$	82.7	0.00	82.3	1.32
$\mathcal{M}_{\text{indep}}$	80.1	3.12	81.5	3.07

Table 5.1: Relative performance increase of model $\mathcal{M}_{\text{recons}}$ compared to the three benchmark models in %.

enced by the vehicles on the entering lane and the vehicle in the rear is influenced mostly be the vehicle in front. However, it is expected that the proposed model performs better in more complex scenarios with more situation hypotheses where it is more challenging to learn the marginal distributions of model $\mathcal{M}_{\text{decomp}}$ and interactions have no preferred direction.
For this reason, Tab. 5.1 additionally shows the results on a subset of the situation hypotheses $\Omega_{1:3}^{\not}$ that show interactions patterns that are not directly modeled by $\mathcal{M}_{\text{decomp}}$. The subset $\Omega_{1:3}^{\not}$ consists of the 6 hypotheses $\{2, 3, 6, 9, 11, 20\}$. 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. Hence, the direction of this interaction is not directly modeled by the marginal distribution $p(b_2)$ the maneuver of vehicle 2'. Likewise, hypotheses $\{3, 6, 9\}$ are characterized by a merging maneuver in front of vehicle 3, which in turn needs to perform a deceleration maneuver. This interaction is only modeled indirectly in the marginal of $p(b_3|\mathbf{x}, b_2)$. Consequently, the comparison of Tab. 5.1 for the subset $\Omega_{1:3}^{\not}$ shows that the performance of $\mathcal{M}_{\text{decomp}}$ decreases compared to the proposed $\mathcal{M}_{\text{recons}}$ that models interactions directly in both directions in its complete conditional distributions.

Thus, the evaluation on the complex highway scenario shows that it is beneficial to model interactions and to decompose the recognition task into smaller sub problems. Moreover, the proposed universally applicable reconstruction method performs at least on the same or a higher level compared to dedicated joint probability decompositions while having the advantage to be adaptable to variable traffic scenes.

5.5.2 Comparison to Gibbs Sampling

Section 4.3 stated that the reconstruction of joint probabilities using sampling techniques is associated with several issues, the most prominent being the high computational demands and the difficulty to estimate convergence progress. The following evaluation concentrates on the concrete performance differences and compares the computational effort of the proposed analytic reconstruction methods to Gibbs sampling with different sample sizes. The evaluated setting, learning algorithms, and cross-validations are identical to those in the evaluation of Section 5.5.1.

The implemented Gibbs sampler follows the procedure of (4.6). Its resulting AUC is calculated for a prediction horizon of $\Delta t = 5s$ with an increasing number performed iterations T. In each iteration τ one sample for each of the 3 random variables is drawn. The statistics of the resulting Markov chain are calculated according to (4.7). 50 runs are conducted for each sample size to obtain statistically reliable results on the 1070 investigated scenarios.

Fig. 5.9 shows the AUC histograms for different sample sizes with 50 runs each. For small sample sizes the resultant AUCs do not reach the baseline performance of the analytic reconstruction schemes and are comparable to models that assume stochastic independence (see Tab. 5.1 for details). Moreover, they show a large deviations over 50 performed runs.

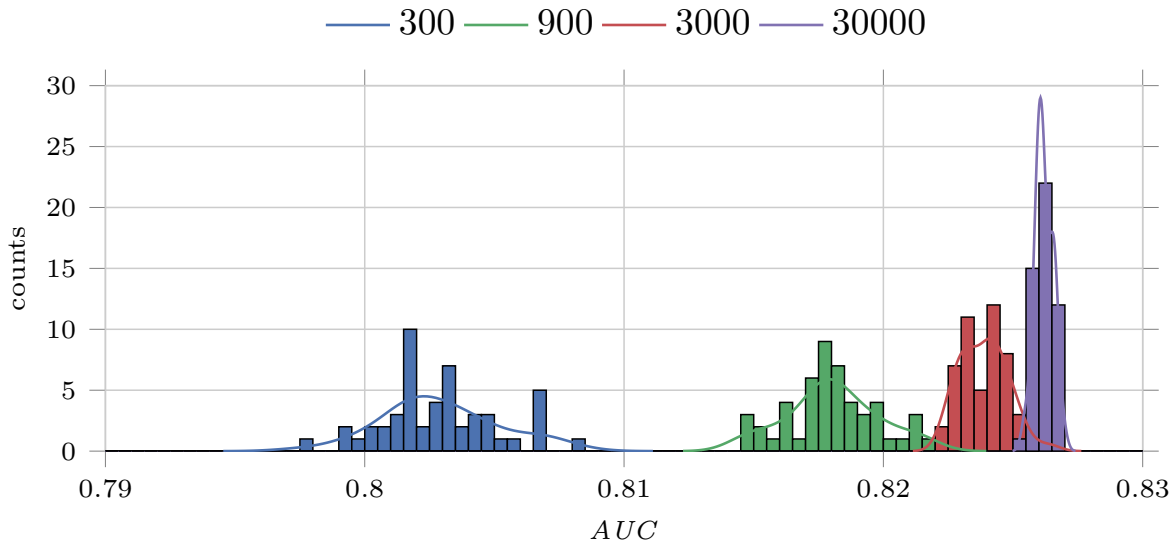


Figure 5.9: Histograms regarding resultant AUC for the conducted 50 runs for each investigated sample size.

Tab. 5.3 shows detailed mean and standard deviation of the AUCs for successive sample sizes. Starting with 30000 drawn samples the results of Gibbs sampling and the analytic reconstruction become comparable. However, drawing 30000 samples is time consuming and nevertheless no guarantee to obtain reliable results.

samples	300	900	3000	30000	base line
AUC	80.3 ± 0.22	81.8 ± 0.17	82.3 ± 0.09	82.6 ± 0.03	82.6

Table 5.2: Mean and standard deviations regarding resultant AUC for 50 runs for each investigated sample size.

Tab. 5.3 compares calculation times of the reconstruction based on one linear equation systems (Section 5.3.1), the recursive backward calculation (Section 5.3.2) as well as Gibbs sampling for different sample sizes. All implementations and calculations are done in *python* on one core (Intel Xeon @3.7Ghz) of a consumer grade PC.

Tab. 5.3 reveals that Gibbs sampling takes significantly more time compared to both of the proposed reconstruction methods. For a sample size of 30000, most likely resulting in a competitive performance, the calculation time increases by over 544 times. Summing up, sampling techniques

	Analytic		Gibbs sampler			
	one	recursive	300	900	3000	30000
time	$1.21ms$	$1.53ms$	$7.89ms$	$20.7ms$	$65.3ms$	$659ms$
relative increase	1	1.264	6.52	17.10	53.96	544.62

Table 5.3: Average calculation times for 100 repeated calculations for one joint probability reconstruction of the evaluation in Section 5.5.1 in ms. The second row shows the increase in calculation time in relation to the reconstruction based on one linear equation system.

for reconstruction joint probabilities can yield comparable results, however, the computational effort is much higher. Thus, the proposed analytic reconstruction schemes are preferable over sampling techniques for the application in situation recognition systems.

5.5.3 Combination with Pairwise Probability Coupling

This evaluation investigates the combination of the variable discriminative maneuver estimation framework of Chapter 3 with the interaction-aware situation reconstruction of this chapter.
Section 4.2.1 stated that decomposing the joint probability distribution $p(b_1, \ldots, b_n)$, i.e. the target interaction-aware situation recognition, into complete conditional distributions offers the advantage of being easily adaptable to changing traffic scenes if the underlying probabilistic classification models use the pairwise probability coupling approach of Chapter 3. For this reason, the following conducts mostly the identical evaluation as in Section 5.5.1, but instead of training a multiclass logistic regression that directly estimates the maneuvers of each vehicle a set of binary *OvO* classifiers are trained and combined using pairwise probability coupling. This model is referred to as $\mathcal{M}_{\text{pair}}$

The three possible maneuvers $\Omega_1 = \{\text{betw}, \text{front}, \text{no}\}$ of vehicle 1 on the entering lane lead to three possible binary maneuver combinations $\Omega_1^{1,2} = \{\text{betw}, \text{front}\}$, $\Omega_1^{1,3} = \{\text{betw}, \text{no}\}$ and $\Omega_1^{2,3} = \{\text{front}, \text{no}\}$. Thus, a logistic regression for each binary combination is trained that estimates the binomial distributions $p^{1,2}(b_1|b_2, b_3, \mathbf{x})$, $p^{1,3}(b_1|b_2, b_3, \mathbf{x})$ and $p^{2,3}(b_1|b_2, b_3, \mathbf{x})$. The feature vector $\mathbf{x}$ is identical to (5.36) and the incorporation of other vehicles' behaviors via categorical variables is the same as in Section 5.5.1.
In the same way, models for the maneuver combinations $\Omega_2^{1,2} = \Omega_3^{1,2} = \{--, \text{xpcd}\}$, $\Omega_2^{1,3} = \Omega_3^{1,3} = \{--, ++\}$ and $\Omega_2^{1,2} = \Omega_3^{1,2} = \{\text{xpcd}, ++\}$ of vehicles 2 and 3 are trained.
Hence, there are three binary classifiers for each traffic participant that are trained completely separately without incorporating knowledge on what situation hypotheses are possible in the overall scene.

During run-time, two subsequent calculations determine the target joint probability distribution $p(b_1, b_2, b_3|\mathbf{x})$. Fig. 5.10 shows the architecture of the recognition framework. First, based on each traffic participants' possible maneuver alternatives and their pairwise combinations, the corresponding binary classifiers predict the binomial distributions that form the basis for applying pairwise probability coupling. The results are the complete conditional distributions $p(b_1|b_2, b_3, \mathbf{x})$, $p(b_2|b_1, b_3, \mathbf{x})$ and $p(b_3|b_1, b_2, \mathbf{x})$ of each traffic participant. Subsequently, these are used to reconstruct the target joint probability distribution $p(b_1, b_2, b_3|\mathbf{x})$ using the analytic reconstruction.

The model $\mathcal{M}_{\text{pair}}$ based on the interaction-aware reconstruction framework and the variable discriminative maneuver estimation framework performs with an $AUC = 82.6\%$ in the same evaluations of the experiment conducted in Section 5.5.1. Thus, it shows nearly the identical performance as $\mathcal{M}_{\text{recons}}$ ($AUC = 82.7\%$), only neglectable differences occur. This is a remarkable result considering that $\mathcal{M}_{\text{pair}}$ reconstructs the target distribution $p(b_1, b_2, b_3|\mathbf{x})$ starting with separate maneuver predictions discerning between two possible maneuver alternatives at a time. Consequently, the performance of $\mathcal{M}_{\text{pair}}$ surpasses the performance of models $\mathcal{M}_{\text{indep}}$ and $\mathcal{M}_{\text{direct}}$ due to reduced model complexities and the absence of independence assumptions.

These results demonstrate the possibility to combine both frameworks to obtain a completely adaptable situation recognition framework without the need for any independence assumptions.

5.6 Summary

This chapter introduced two novel analytic methods for reconstructing interaction-aware joint probabilities based on complete conditional distributions. The first method is based on solving linear equation systems while the second method is based on solving subproblems in a recursive manner. The analogy to discrete Markov chains provided the foundation

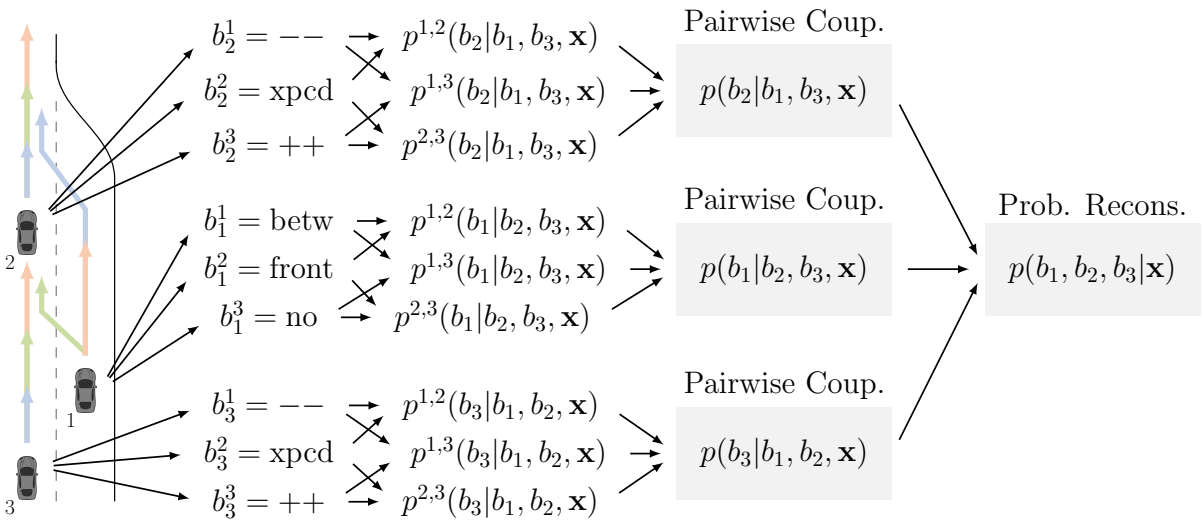


Figure 5.10: Procedure for using the variable discriminative maneuver estimation framework with the interaction-aware joint probability reconstruction framework on the complex highway entering scene.

for proving the existence of the stationary distribution for both methods. Additionally, the chapter discussed strategies for incorporating actually present stochastic independencies in the analytic reconstruction process and outlined that the novelly proposed reconstruction methods are superior in terms of robustness.

The complete framework for recognizing situations with multiple, interacting traffic participants was tested on a complex highway scene. The evaluation showed the benefits of the reduced model complexities of the entity based recognition system compared to learning a direct fully-interaction-respecting situation recognition system and that ignoring interactions leads to decreased recognition performance. Moreover, a comparative experiment benchmarked the proposed analytic reconstruction methods with state of the art sampling techniques showing that sampling techniques suffer from the needed sample sizes and the associated computational effort. A final evaluation investigated the impact of using the variable discriminative maneuver estimation framework of Chapter 3 to determine the needed complete conditionals. The results were almost identical compared to learning these complete conditional distributions directly using multiclass classifiers, and thus demonstrate the possibility to combine both frameworks to obtain a completely adaptable situation recognition framework.

6 Situation Hypotheses Reduction

This chapter addresses the combinatorial aspect of situation assessment. Due to the large variability of traffic scenes, especially inner-city traffic, the number of situation hypotheses increases quickly. This drastic increase challenges the situation assessment itself as well as subsequent systems that depend on the results of the situation assessment. For this reason, there is a strong need for reducing the number of explicitly investigated situation hypotheses.

The remainder of this chapter is organized as follows. Section 6.1 discusses the problems of situation assessment for large sample spaces and states the contribution of this chapter. Section 6.2 proposes the approach for reducing the number of situation hypotheses by making use of the joint probability reconstruction methods of Chapter 5. Subsequently, Section 6.3 discusses a detailed qualitative example for applying the hypotheses pooling technique and addresses its computational complexity. The chapter closes with an evaluation investigating possibilities to apply the proposed hypotheses pooling to the complex entering scenario in Section 6.4.

6.1 Problem Statement

Chapters 3 and 5 outlined problems occurring in generic traffic due to the large variability of possible encountered scenes. The associated large number of specialized models and the combinatorial increasing classifiers' target spaces was tackled by decomposing the recognition task into adaptable single-entity maneuver predictions with a subsequent reconstruction of the full joint probability distribution. This approach drastically reduced both the number of needed classifiers as well as their complexity.

However, even a fixed but complex scene can easily lead to several hundred situation hypotheses. Fig. 6.1 shows a scene with three traffic participants approaching a 4-way intersection. Assuming that each traffic participant has 3 path alternatives, with 2 characteristic temporal be-

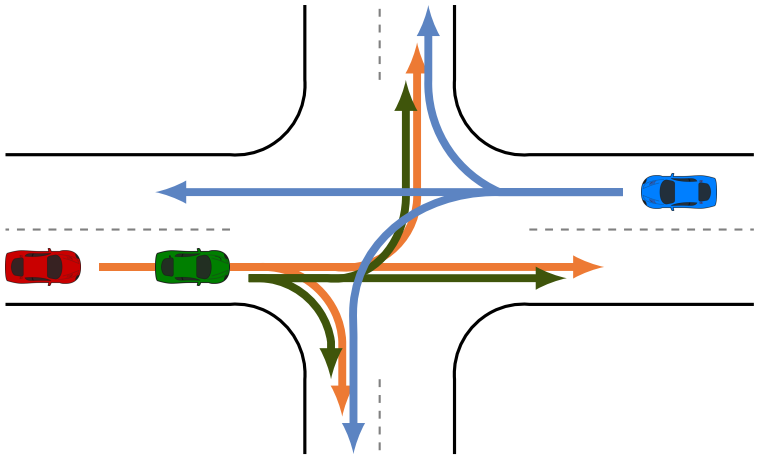


Figure 6.1: Traffic scene consisting of 3 vehicles at a 4-way intersection crossing. Each vehicle has 3 path alternatives, assuming 2 typical longitudinal behavior patterns leads to 6 possible maneuvers for each vehicle. All possible maneuver combinations result $6^3 = 216$ situation hypotheses.

havior patterns each, already leads to 216 different situation hypotheses (maneuver combinations). Considering that the scene of Fig. 6.1 is still rather manageable compared to possibly encountered constellations in a crowded inner-city scene with additional traffic participants, like pedestrians or cyclists, highlights the problem.
The large sample spaces of complex scenes lead to two main challenges. First, while the methods presented in the previous chapters break down the complexity to small subproblems, the large sample space needs to be dealt with at some point during the situation recognition, e.g. the coefficient matrix of Section 5.3.1 or Section 5.3.2. This becomes computationally demanding for large $|\Omega_{1:n}|$. The second problem arises in combination with subsequent systems that rely on the results of the situation assessment step, e.g. warning systems, automated braking or fully autonomous functions. A large number of situation hypotheses pose an immense flood of information that needs to be dealt with by those systems.

6.1.1 Contribution

The literature reviews in Section 3.1.1 and Section 4.1.1 revealed that existing approaches do not tackle general complex traffic scenarios with several interdependent traffic participants. Thus, to the author's knowledge there is no previous work that explicitly addresses the problems arising from the large number of investigated situation hypotheses in complex traffic

scenes.
For this reason, this work proposes an approach that tackles the combinatorics associated with situation recognition with several traffic participants based on the reconstruction process of Section 5. A key aspect of the approach is that it does not change the underlying probability distribution. This is important since it is theoretically possible to exclude uninteresting maneuvers from a traffic participant's sample space directly, e.g. using the variable discriminative maneuver estimation framework of Chapter 3, to tackle the combinatorics. However, such an approach completely changes the resulting probability distribution. This is due to the reason that the excluded maneuvers might in fact be very likely but are not represented in the resulting probability distribution anymore. Hence, excluding them completely results in a probability distribution that contains the relative occurrence probabilities of the remaining hypotheses, not their absolute occurrence probabilities.

The next section proposes an approach that summarizes several maneuver into groups and performs the joint probability reconstruction on these groups instead. This ensures that probability values of maneuvers that are not grouped remain unchanged while the number of investigated situation hypotheses is efficiently reduced as a real-world evaluation shows.

6.2 Hypotheses Pooling

The proposed approach for reducing the size of the sample space bases on *pooling* several maneuvers (events) of one traffic participant into a new dummy maneuver (event), e.g.

$$\{b_i^2, b_i^3, b_i^5\} = \mathcal{I}_i.$$

Such a dummy event comprises several maneuvers, where each individual maneuver's probability of occurrence is not of interest but rather that of the group as a whole. Section 6.2.1 discusses reasonable criteria for deciding when it is appropriate to apply hypotheses pooling.

The conditional distributions the joint probability reconstruction rests upon need to represent the distributions regarding these artificially created events $\mathcal{I}$, i.e

$$p(B_i = \mathcal{I}_i | \mathbf{b}_{1:i-1}, \mathbf{b}_{1:i+1}, \mathbf{x})$$

and

$$p(\mathbf{b}_{1:i-1}, \mathbf{b}_{1:i+1} | B_i = \mathcal{I}_i, \mathbf{x}).$$

Section 6.2.2 explains the details on how to calculate these distribution based on the original complete conditional distributions.

6.2.1 Pooling Criteria

The reasons for pooling different maneuvers can be manifold and depend on the area of application of the situation assessment.
This section gives two examples for meaningful and universally usable pooling techniques.

Minimum Likelihood Pooling

One possibility is to group very unlikely maneuver alternatives into dummy events based on the single-entity maneuver estimations' results. This means that several unlikely maneuvers are combined, each of them being almost irrelevant, e.g. being lower than chance level

$$p(b_i|\mathbf{b}_{\setminus i}, \mathbf{x}) < \frac{1}{|\Omega_i|}, \forall b_i \in \Omega_i.$$

Simply neglecting these unlikely maneuvers would lead to two main problems.

First, completely discarding the probabilities associated with the maneuvers in $\mathcal{I}_i$ leads to unnormalized complete conditional distributions. i.e

$$\sum_{b_i \in \Omega_i \setminus \mathcal{I}_i} p(b_i|\mathbf{b}_{1:i-1}, \mathbf{b}_{1:i+1}) \neq 1. \tag{6.1}$$

This property plays an essential role in proving the existence of a unique reconstructed joint probability of Section 3.2.3. Thus, the successful reconstructing of the target joint probability is no longer guaranteed.

Second, even if it is made possible to approximate the target joint distribution despite of (6.1) either by renormalizing (6.1) or formulating an optimization problem based on (5.17), the resulting joint distribution is distorted and does not represent the same distribution than without the hypotheses reduction. This is due to the reason that the individual events in $\mathcal{I}_i$ may be unlikely but their cumulated probability $\sum_{b_i \in \Omega_i} p(b_i|\mathbf{b}_{1:i-1}, \mathbf{b}_{1:i+1})$ may represent a considerable part of their distribution $p(b_i|\mathbf{b}_{1:i-1}, \mathbf{b}_{1:i+1})$. Hence, the individual probability values actually represent relative rather than absolute occurrence probabilities.

Equivalent Impact Pooling

It is also possible to group maneuvers together which lead to similar implications for the situations assessment's overlying advanced driver assistance system or automated driving function. These are typically maneuvers associated with a low empirical risk that do not have any or only minor implications on a vehicle's behavior at all, e.g. turning maneuvers at an intersection that do not overlap the paths of other vehicles.
The complex highway entering scene of Section 5.5.1 displays another possible application of equivalent impact pooling. For a safety system assisting the entering of vehicle 1, it is important to know whether vehicle 3 is going to perform an unexpected acceleration maneuver (event: $b_3 = ++$) to detect conflicting merging trajectories. In contrast, the detailed assessment of vehicle 3 performing an expected ($b_3 = \text{xpcd}$) or decelerated ($b_3 = --$) maneuver is not crucial for the safety of the entering maneuver since these maneuvers do not reduce the anticipated gap in front of vehicle 3. Thus, they are pooled to one dummy event $\mathcal{I}_3 = \{\text{xpcd}, --\}$

These two examples for pooling strategies are not mutually exclusive. They can be combined by introducing several dummy events $\mathcal{I}_i^c$ for different categories c, e.g. different associated risk categories or likelihood levels.

6.2.2 Complete Conditional Recalculation

The pooling of several events of a random variable b_i changes its sample space Ω_i which affects the complete conditional distributions in two different ways

$$p(b_i|b_j, \mathbf{b}_{\setminus\{i,j\}}), \tag{6.2}$$

and

$$p(b_j|b_i, \mathbf{b}_{\setminus\{i,j\}}), \tag{6.3}$$

where $\mathbf{b}_{\setminus\{i,j\}}$ are all remaining $n-2$ random variables for convenient notation. Additionally, for this formal outline of the procedure, the scene evidence $\mathbf{x}$ is omitted to improve readability. The concrete probability values of these two distributions do not change for all event combinations regarding $b_i \notin \mathcal{I}_i$. The probability value of (6.2) for an event pooled by the dummy event $\mathcal{I}_i$ occurring is identical to the summation of the individual

event probabilities

$$p(B_i = \mathcal{I}_i | b_j, \mathbf{b}_{\setminus\{i,j\}}) = \sum_{b'_i \in \mathcal{I}_i} p(b'_i | b_j, \mathbf{b}_{\setminus\{i,j\}}). \tag{6.4}$$

The calculation of (6.3), where the pooled dummy event is part of the conditional, is based on the calculation of cumulative conditional distribution

$$p(b_j | B_i = \mathcal{I}_i, \mathbf{b}_{\setminus\{i,j\}}) = \frac{\sum_{b'_i \in \mathcal{I}_i} p(b_j, b'_i | \mathbf{b}_{\setminus\{i,j\}})}{\sum_{b'_i \in \mathcal{I}_i} p(b'_i | \mathbf{b}_{\setminus\{i,j\}})},$$

as given in [61]. Neither the partial joint $p(b_i, b_j | \mathbf{b}_{\setminus\{i,j\}})$ nor the marginal $p(b_i | \mathbf{b}_{\setminus\{i,j\}})$ are known during the reconstruction approach based on complete conditional distributions.
However, decomposing the partial joint distribution of the nominator with the chain rule of probability yields

$$p(b_j | B_i = \mathcal{I}_i, \mathbf{b}_{\setminus\{i,j\}}) = \frac{\sum_{b_i \in \mathcal{I}_i} p(b_j | b'_i, \mathbf{b}_{\setminus\{i,j\}}) p(b'_i | \mathbf{b}_{\setminus\{i,j\}})}{\sum_{b_i \in \mathcal{I}_i} p(b'_i | \mathbf{b}_{\setminus\{i,j\}})}. \tag{6.5}$$

A closer look at (6.5) reveals that denominator is equal for all events of b_j and thus corresponds to a normalization constant for the distribution $p(b_j | B_i = \mathcal{I}_i, \mathbf{b}_{\setminus\{i,j\}})$. Hence, (6.5) is proportional to

$$p(b_j | B_i = \mathcal{I}_i, \mathbf{b}_{\setminus\{i,j\}}) \propto \sum_{b_i \in \mathcal{I}_i} p(b_j | b'_i, \mathbf{b}_{\setminus\{i,j\}}) p(b'_i | \mathbf{b}_{\setminus\{i,j\}}).$$

The complete conditional $p(b_j | b_i, \mathbf{b}_{\setminus\{i,j\}})$ is a direct result from the single-entity maneuver prediction. However, this does not apply for the marginal distribution $p(b_i | \mathbf{b}_{\setminus\{i,j\}})$ but it can be determined by

$$p(b_i | \mathbf{b}_{\setminus\{i,j\}}) = \left(\sum_{b'_j \in \Omega_j} \frac{p(b_i | b'_j, \mathbf{b}_{\setminus\{i,j\}})}{p(b'_j | b_i, \mathbf{b}_{\setminus\{i,j\}})} \right)^{-1}, \tag{6.6}$$

using the principle from (5.12) to only rely on the available complete conditional distribution itself. Accordingly, applying equation (6.6) leads to

$$p(b_j | B_i = \mathcal{I}_i, \mathbf{b}_{\setminus\{i,j\}}) \propto \sum_{b_i \in \mathcal{I}_i} p(b_j | b'_i, \mathbf{b}_{\setminus\{i,j\}}) \left(\sum_{b'_j \in \Omega_j} \frac{p(b'_i | b'_j, \mathbf{b}_{\setminus\{i,j\}})}{p(b'_j | b'_i, \mathbf{b}_{\setminus\{i,j\}})} \right)^{-1} \tag{6.7}$$

for calculating the second complete conditional distribution over the reduced sample space Ω_i. Thus, all needed cumulative distributions are entirely specified by available complete conditional distributions.

Dependent on the pooling criteria for reducing the sample space the calculation of (6.7) can be simplified by neglecting the influence of the marginal $p(b_i|\mathbf{b}_{\setminus\{i,j\}})$ without noticeably affecting the reconstructed joint distribution. This is a valid approximation assuming that probabilities of $p(b_i|\mathbf{b}_{\setminus\{i,j\}}), \forall b_i \in \mathcal{I}_i$ are in a similar range, e.g. as for the case of *minimum likelihood pooling*. Applying this approximation simplifies (6.7) to

$$p(b_j|B_i = \mathcal{I}_i, \mathbf{b}_{\setminus\{i,j\}}) \propto \sum_{b_i' \in \mathcal{I}_i} p(b_j|b_i', \mathbf{b}_{\setminus\{i,j\}}). \tag{6.8}$$

6.2.3 Recursive Hypotheses Pooling

Section (5.3.2) proposed a recursive algorithm for reconstructing joint probabilities from complete conditional distributions. The reconstruction bases on calculating interim partial joint probability distributions. These partial joints offer additional possibilities to apply hypotheses pooling for reducing the sample space even further. Applying the hypotheses pooling in combination with the backward recursion allows to not only pool single maneuvers but also pool maneuver combinations over several traffic participants. By considering maneuver combinations, the *minimum likelihood pooling* criterion extends to combining maneuver combinations that are unlikely given the reconstructed partial joint distributions at each recursion step. This allows to reduce the number of hypotheses drastically Fig. 6.2 shows the simplified procedure and highlights that the hypotheses pooling is applicable at any interim step. The procedure is as follows. At each recursion step, the complete conditionals (or partial joint distributions at later recursion steps) are checked whether certain maneuver (or maneuver combinations from partial joint distributions over several traffic participants) should be pooled. If pooling is a applied the complete conditionals/partial joint distributions are recalculated. The following example further clarifies the procedure.

The suggested maneuver pooling for the complex highway entering scene can be extended to pooling maneuver combinations for interim partial joints during the reconstruction process. Assuming the reconstruction process starts with calculating $p(b_1, b_3|b_2, \mathbf{x})$ from combining $p(b_1|b_2, b_3, \mathbf{x})$ and $p(b_3|b_1, b_2, \mathbf{x})$. The example in Section 6.2.1 suggested to reduce the sample space for traffic participant 3 to $\Omega_3 = \{++, \mathcal{I}_3\}$ with $\mathcal{I}_3 = \{\text{xpcd}, --\}$

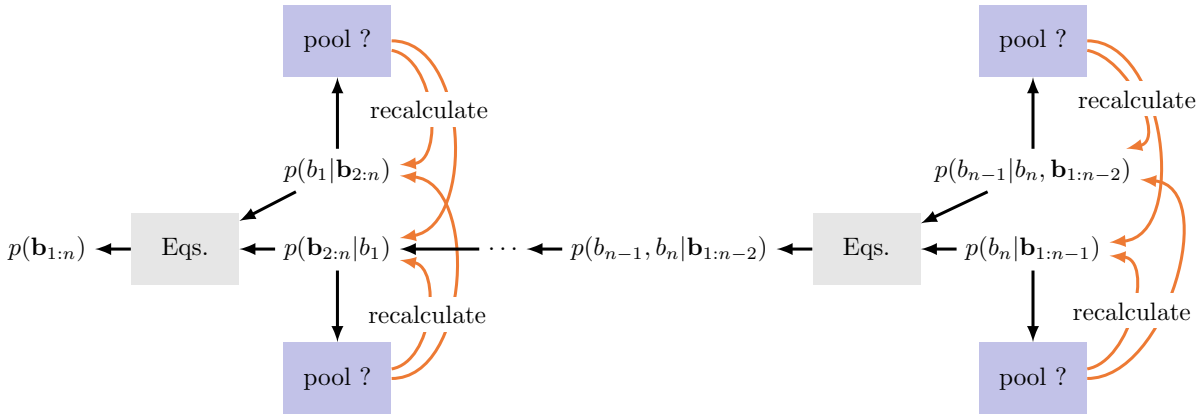


Figure 6.2: Procedure for applying recursive hypotheses pooling based on the recursive backwards calculation.

when the goal of the situation assessment is to assist a save entering of vehicle 1. This reduces the combined sample space $\Omega_{1\times3}$ from a cardinality of 9 to 6. The obtained partial joint distribution $p(b_1, b_3|b_2, \mathbf{x})$ allows for further application of the *equivalent impact pooling*. For example combining all hypotheses associated where no change maneuver occurs at all leads to $\mathcal{I}^{\text{no}}_{1\times3} = \{(\text{no}, \mathcal{I}_3), (\text{no}, ++)\}$. Moreover, an additional dummy category can be introduced summarizing maneuver combinations where vehicle 1 is going to merge in front of vehicle 2 which makes the detailed maneuver of vehicle 3 irrelevant for the safety of vehicle 1, i.e. $\mathcal{I}^{\text{front}}_{1\times3} = \{(\text{front}, \mathcal{I}_3), (\text{front}, ++)\}$. This reduction leads to a sample space cardinality of $|\Omega_{1\times3}| = 4$ instead of 9, originally. Hence, the full reconstructed joint probability distribution contains only $\Omega_{1\times3} \times |\Omega_2| = 12$ hypotheses instead of 27. These capture all relevant information for the safe highway entering of vehicle 1.

6.2.4 Partial Joints Recalculation

The conditional distributions' recalculation for pooling partial joint events is analogous to that for one dimensional target spaces of Section 6.2.2. For a convenient notation of the recalculation equations it is assumed that the random variables are numerated according to their backwards reconstruction ordering. In the multivariate case of random variable $\mathbf{b}_{m:n}$ the dummy event $\mathcal{I}_{m:n}$ contains several event combinations $\mathbf{b}^{\mathbf{j}}_{m:n}$.

The calculation of the partial joint distribution, having the random variables affected by the pooling in the target space, is analogously to (6.4)

and results in a summation of

$$p(\mathbf{b}_{m:n} = \mathcal{I}_{m:n}|b_{m-1}, \mathbf{b}_{1:m-2}) = \sum_{\mathbf{b}'_{m:n} \in \mathcal{I}_{m:n}} p(\mathbf{b}'_{m:n}|b_{m-1}, \mathbf{b}_{1:m-2}). \quad (6.9)$$

The cumulative complete conditional distributions, having the multivariate random variable $\mathbf{B}_{m:n}$ in their conditionals, are determined by

$$p(b_{m-1}|\mathbf{b}_{m:n} = \mathcal{I}_{m:n}, \mathbf{b}_{1:m-2}) \propto \sum_{\mathbf{b}'_{m:n} \in \mathcal{I}_{m:n}} p(b_{m-1}|\mathbf{b}'_{m:n}, \mathbf{b}_{1:m-2}) \left(\sum_{\mathbf{b}'_{m-1} \in \Omega_{m-1}} \frac{p(\mathbf{b}'_{m:n}|b'_{m-1}, \mathbf{b}_{1:m-2})}{p(b'_{m-1}|\mathbf{b}'_{m:n}, \mathbf{b}_{1:m-2})} \right)^{-1}, \quad (6.10)$$

or

$$p(b_{m-1}|\mathbf{b}_{m:n} = \mathcal{I}_{m:n}, \mathbf{b}_{1:m-2}) \propto \sum_{\mathbf{b}'_{m:n} \in \mathcal{I}_{m:n}} p(b_{m-1}|\mathbf{b}'_{m:n}, \mathbf{b}_{1:m-2}), \quad (6.11)$$

in case of assuming nearly uniform distributed values

$$p(\mathbf{b}_{m:n}, \mathbf{b}_{1:m-2}), \forall \mathbf{b}_{m:n} \in \mathcal{I}_{m:n}.$$

6.3 Qualitative Calculation Example

This section illustrates the details of using the recursive hypotheses pooling on an artificial 3 random variable example. The backward recursion for three random variables has two depths. At first, recursion depth $r = 2$, the partial joint $p(b_2, b_3|b_1)$ conditioned on b_1 is calculated while the final step ($r = 1$) determines the full joint by combining $p(b_2, b_3|b_1)$ with $p(b_1|b_2, b_3)$. The hypotheses pooling is performed for maneuver probabilities that do not exceed the chance level.

Fig. 6.3 shows the complete conditional distributions for reconstructing the joint probability $p(b_1, b_2, b_3)$. Without using any situation hypotheses reduction technique the majority of occurrence probabilities are close to zero. Fig. 6.4 shows the detailed joint probability distribution $p(b_1, b_2, b_3)$.

The first step of the recursive hypotheses pooling is to select maneuvers of the complete conditional distributions that are used to reconstruct the partial joint $p(b_2, b_3|b_1)$. This example pools maneuvers, on the basis of the likelihood pooling, that are below chance level

$$\mathcal{I}_2 = \{p(b_2|b_1, b_3) < \frac{1}{|\Omega_2|}, \forall b_1 \in \Omega_1 \wedge \forall b_3 \in \Omega_3\},$$

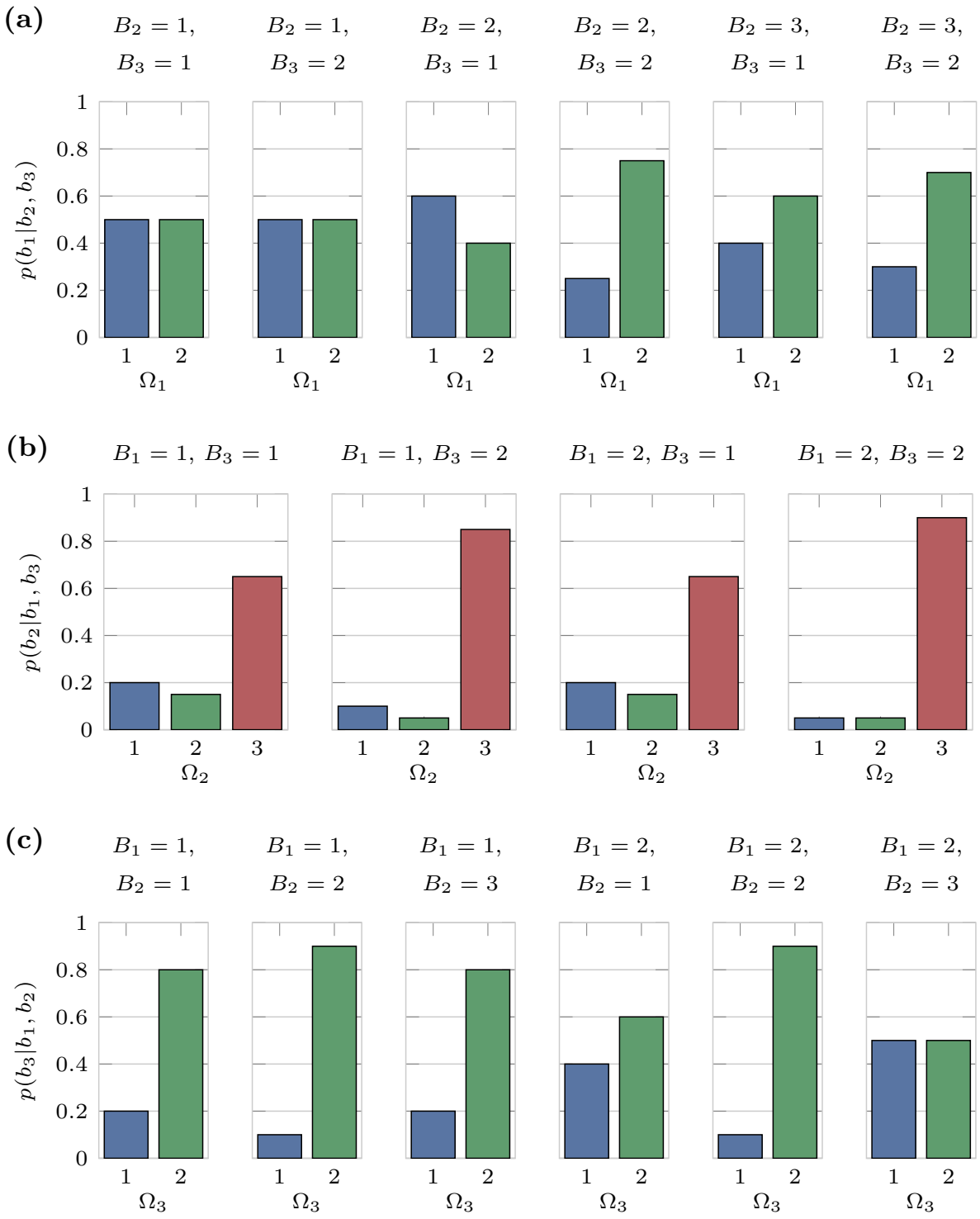


Figure 6.3: Complete conditional distributions for each event combination in their conditional. (a) $p(b_1|b_2, b_3)$, (b) $p(b_2|b_1, b_3)$ and (c) $p(b_3|b_2, b_3)$

which results in $\mathcal{I}_2 = \{1, 2\}$ for pooling. Pooling maneuvers of b_3 is not possible since a sample space cardinality of two events does not allow for reasonable maneuver pooling. Thus, $\mathcal{I}_3 = \emptyset$. Hence, the complete

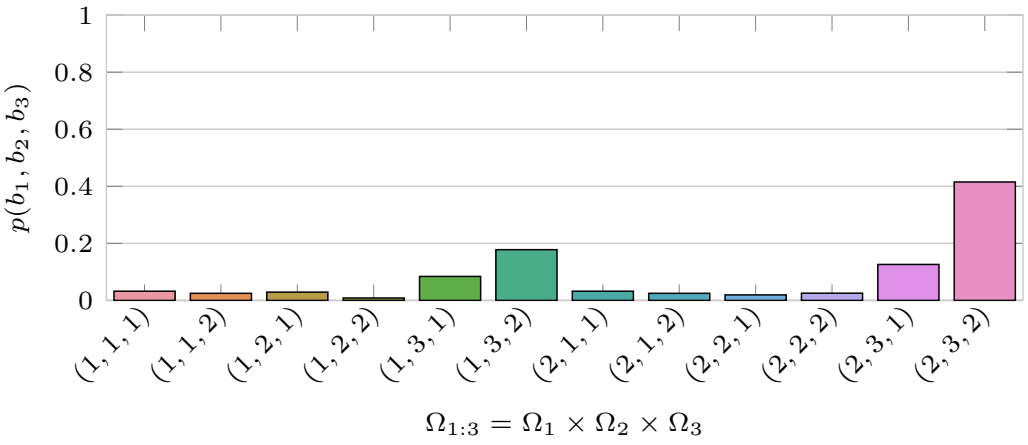


Figure 6.4: Reconstructed joint probability $p(b_1, b_2, b_3)$ from the complete conditional distributions of Fig. 6.3.

conditionals of $p(b_2|b_1, b_3)$ and $p(b_3|b_2, b_1)$ need to be recalculated on the reduced sample space of $\bar{\Omega}_2 = \{1, \mathcal{I}_2\}$ while Ω_3 remains unchanged. The cumulative complete conditional $p(\bar{b}_2 = \mathcal{I}_2|b_1, b_3)$ is determined using (6.4) as well as $p(b_3|\bar{b}_2 = \mathcal{I}_2, b_1)$ using (6.8). Fig. 6.5 shows the recalculated cumulative complete conditional distributions.

With the calculation of the partial joint $p(\bar{b}_2,b_3|b_1)$, based on the newly obtained $p(\bar{b}_2|b_1, b_3)$ and $p(b_3|\bar{b}_2, b_1)$ the first reconstruction step is finalized. Fig. 6.6 displays the resultant partial joint distribution.

The last reconstruction step combines the partial joint distribution $p(\bar{b}_2, b_3|b_1)$ with the up to now considered complete conditional distribution of $p(b_1|\bar{b}_2, b_3)$. The calculation starts with identifying possible maneuver (or maneuver combinations in the case for $p(\bar{b}_2, b_3|b_1)$) for pooling. The distribution of $p(b_1|b_2, b_3)$ does not offer the possibility to combine maneuver alternatives on the condition of being under chance level for all $b_2 \in \Omega_2$ and $b_3 \in \Omega_3$. However, the partial joint distribution $p(\bar{b}_2, b_3|b_1)$ allows for combining the event combinations $\mathcal{I}_{2:3} = \{(\mathcal{I}_2, 1), (\mathcal{I}_2, 2), (3, 1)\}$. Hence, the complete conditionals of $p(b_1|\bar{\mathbf{b}}_{2:3})$ and $p(\bar{\mathbf{b}}_{2:3}|b_1)$ need to be recalculated on the reduced sample space of $\bar{\Omega}_{2:3} = \{(3, 2), \mathcal{I}_{2:3}\}$ while the sample space Ω_1 remains unchanged. The cumulative complete conditional $p(\bar{\mathbf{b}}_{2:3}|b_1)$ is determined using (6.9) as well as $p(b_1|\bar{\mathbf{b}}_{2:3})$ using (6.11).

Fig. 6.7 shows the resulting cumulative conditional distributions. With the calculation of the target joint distribution $p(b_1, \bar{\mathbf{b}}_{2:3})$, based on the newly obtained $p(\bar{\mathbf{b}}_{2:3}|b_1)$ and $p(b_1|\bar{\mathbf{b}}_{2:3})$ the second reconstruction step is finalized.

Fig. 6.8 shows the resulting target joint distribution while using a chance

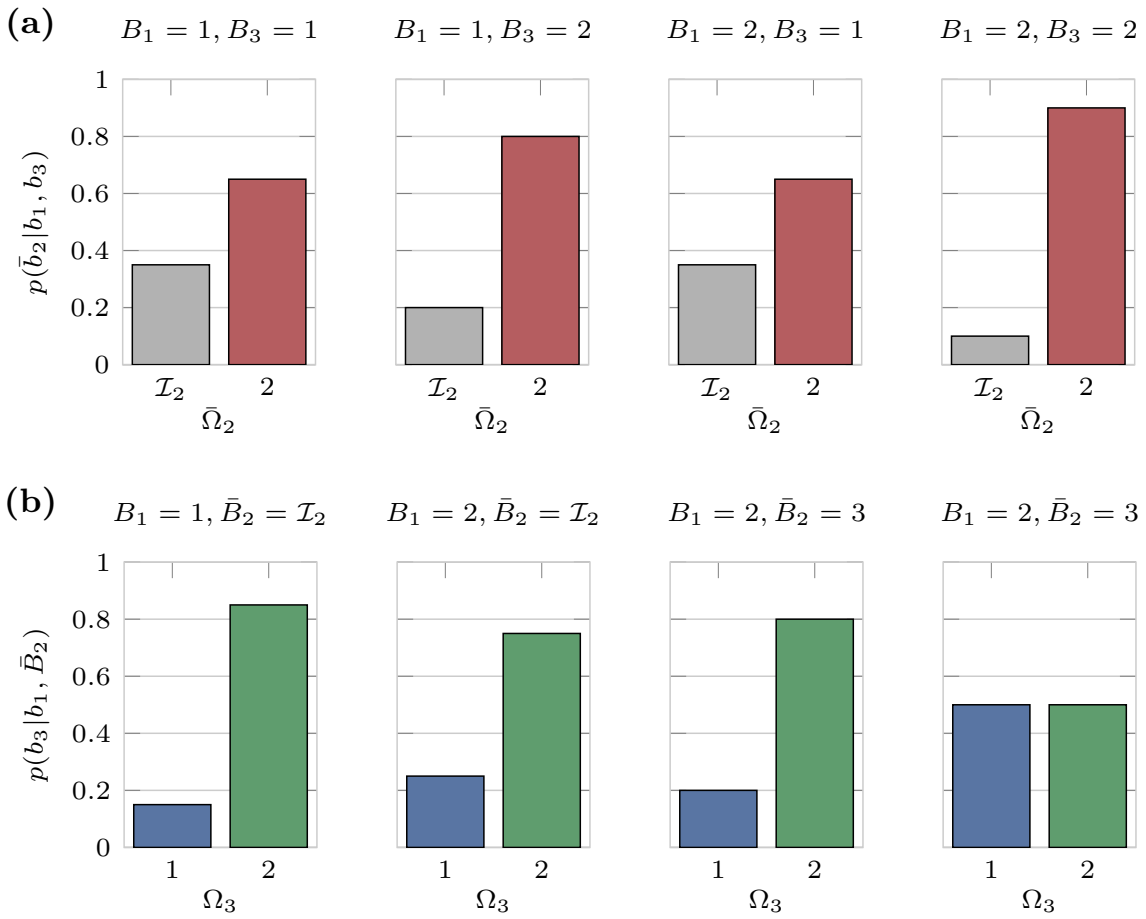


Figure 6.5: Recalculated conditional distributions for the pooled events $\mathcal{I}_2 = \{1, 3\}$. (a) $p(\bar{b}_2|b_1, b_3)$ and (b) $p(b_3|b_1, \bar{b}_2)$.

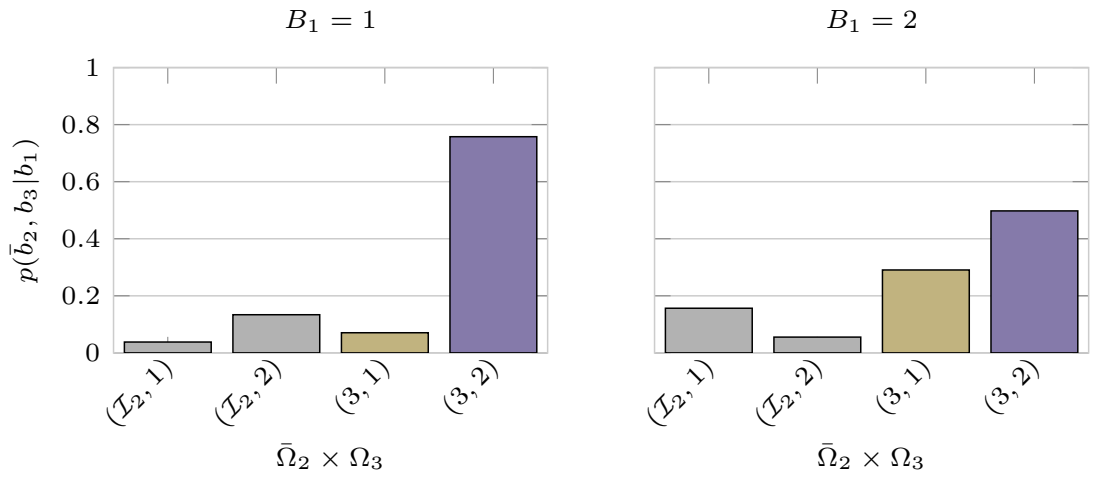


Figure 6.6: Reconstructed partial joint distributions $p(\bar{b}_2, b_3|b_1)$ for all event combinations of $b_1 \in \Omega_1$.

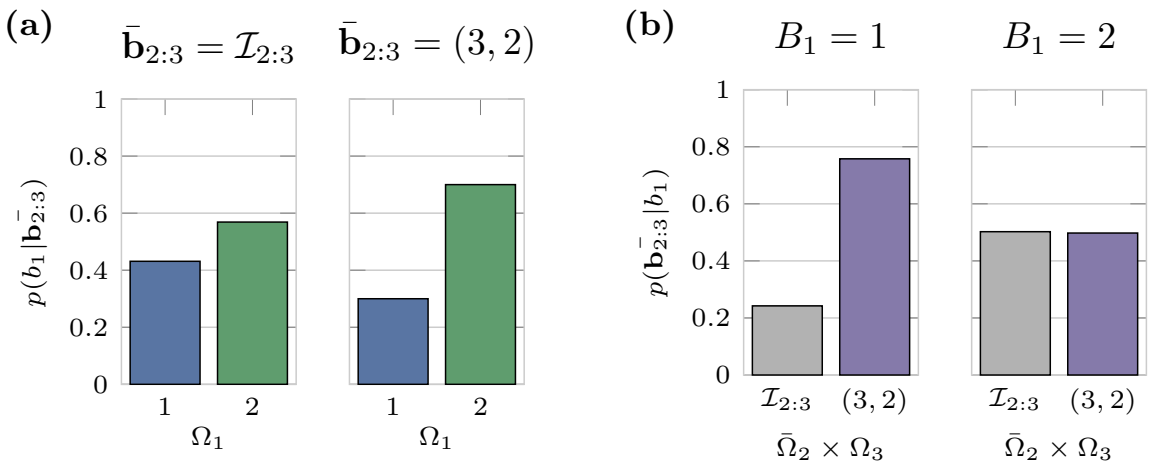


Figure 6.7: Recalculated conditional and partial joint distributions for pooled event combinations $\mathcal{I}_{2:3} = \{(\mathcal{I}_2, 1), (\mathcal{I}_2, 2), (3, 1)\}$. (a) $p(b_1|\bar{\mathbf{b}}_{2:3})$ and (b) $p(\bar{\mathbf{b}}_{2:3}|b_1)$.

level based hypotheses pooling during the reconstruction process. The sample space of $|\Omega_{1:3}| = 12$ has been reduced to a pooled sample space of cardinality $|\bar{\Omega}_{1:3}| = 4$ without changing the probability values of the most likely events ($\bar{\mathbf{b}}_{1:3} = (1, 3, 2)$ and $\bar{\mathbf{b}}_{1:3} = (2, 3, 2)$) of the reconstructed joint distribution of Fig. 6.4.

The event probability values effected by the reduction can be uniformly spread over the original sample space leading to an approximated full joint distribution. Fig. 6.9 shows the resulting distribution and illustrates that the probability values of the two most likely events $((1, 3, 2), (2, 3, 2))$ in the fully reconstructed distribution of Fig. 6.4 remain unchanged. Thus, the sample space has been drastically reduced from a cardinality of $|\Omega_{1:3}| = 12$ to $|\bar{\Omega}_{1:3}| = 4$ while not changing the relevant characteristics of the probability distribution. A smaller pooling threshold would have summarized fewer events and event combinations, preserving more of the original probability values but also concomitantly increasing the resultant sample space size.

Computational efficiency

Each reconstruction step leads to solving a linear equation systems of type (3.14). Using a Gaussian elimination algorithm corresponds to a computational complexity of $\mathcal{O}(m^3)$, where m corresponds to the sample space size. The detailed number of operations (multiplication, addition, divi-

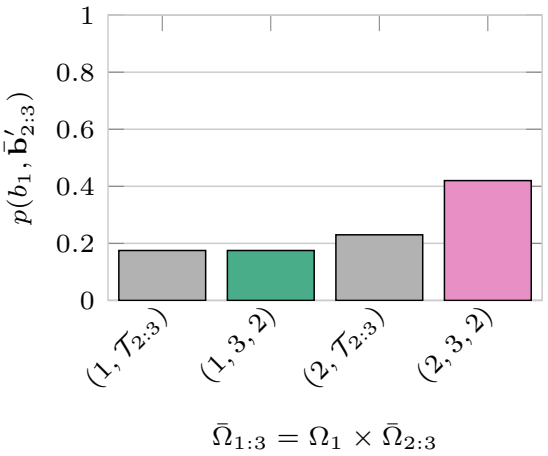


Figure 6.8: Final reconstructed joint probability distribution $p(b_1, \bar{\mathbf{b}}_{2:3})$ on the reduced sample space $\bar{\Omega}_{1:3}$

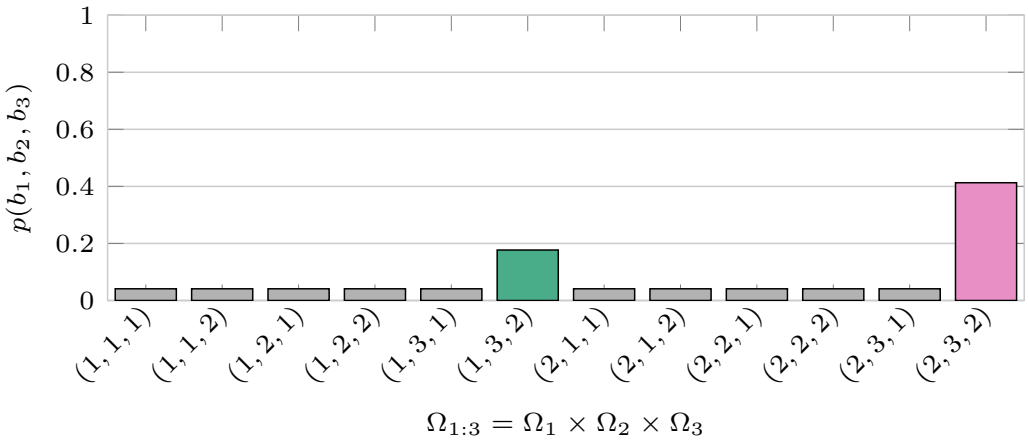


Figure 6.9: Restored joint probability distribution by uniformly distributing the pooled situation hypotheses probability values of distribution $p(b_1, \bar{\mathbf{b}}_{2:3})$.

sion, subtraction) approximate to $\frac{2m^3}{3}$ according to [37]. Thus, solving the complete sample space ($|\Omega_{1:3}| = 12$) without any reduction leads to 1152 operations.
The recalculation of conditional distributions, either (6.2) or (6.8), needs $|\mathcal{I}_i||\Omega_{\setminus i}|$ operations. In the first recursion step $|\Omega_1|$ linear equation systems with $\frac{2|\bar{\Omega}_{2\times 3}|^3}{3} \approx 42$ operations to solve each linear equation system and 16 operations to setup the linear equation systems. Additionally, recalculating the complete conditionals takes 12 operations. The complete effort for the first recursion step amounts to 128 operations. The second recursion step consists of solving a linear equation system with $\frac{2|\bar{\Omega}_{1:3}|^3}{3} \approx 42$ operations required with 16 additional multiplications to set up the linear

equation systems. Additionally, the overhead for recalculating the partial joint $p(\bar{b}_2, b_3|b_1)$ and complete conditional $p(b_1|\bar{b}_2, b_3)$ is 16 operations. Hence, the complete effort for the backwards recursion amounts to 202 overall operations. This is significantly lower than the 1152 operations without any hypotheses pooling techniques and corresponds to a reduction of the computational effort of 82%.
The reduction in computational effort when relying on the more efficient, yet also more incompatibility-sensitive, direct solution of the reconstruction problem (see Section 5.2.1) still accounts to 15% compared to the 1152 operations needed when not using any reduction technique.

Thus, the proposed hypotheses pooling techniques yield a significant reduction in target sample space cardinality and a concomitant reduction of computational effort. The potential for reducing the computational complexity decreases further with increasing number of random variables n and sample space cardinality.

6.4 Experimental Results

This section investigates the impact of applying recursive reconstruction process with a minimum likelihood pooling criteria to the complex entering scene evaluated in Section 5.5.1. The objective is to investigate the performance differences when applying hypotheses pooling to unlikely situations with varying minimum likelihood thresholds. The thresholds for pooling hypotheses result from multiplying the particular chance level at each recursion step $\frac{1}{|\Omega|}$ with a factor β in the range $(0, 1)$.
To calculate the AUC on the complete original sample space of $|\Omega|_{1:3}$, the pooled situation hypotheses are uniformly distributed in the same way as in Fig. 6.9 of the qualitative example. By doing so, the probability values of the pooled situation hypotheses do represent approximations of those resultant from a full reconstruction over the complete sample space. However, it is expected that this has only a minor impact on the overall quality of the reconstructed joint distribution since the individual probability values of the not pooled situation hypotheses are unaffected.

Tab. 6.1 shows the performance results for applying the hypotheses pooling with different prefactors β. The table depicts different βs associated with their AUC, their mean sample space cardinality $|\Omega_{1:3}|$, as well as their standard deviation σ. The standard deviation quantifies how much the uniformly distributed pooled probability values differ from

β	Eq. (6.11) (approx.) AUC	σ	$\lvert\bar{\Omega}_{1:3}\rvert$	Eq. (6.10) AUC	σ	$\lvert\bar{\Omega}_{1:3}\rvert$
0	82.6	±0.00	27.0	82.6	±0.00	27.0
0.2	82.4	±0.05	23.3	82.4	±0.05	23.3
0.4	82.0	±0.22	20.4	82.1	±0.23	20.4
0.6	81.4	±0.50	17.2	81.6	±0.52	17.3
0.8	79.9	±0.95	13.2	80.2	±0.98	13.2
1	77.2	±1.52	9.6	77.1	±1.62	9.6

Table 6.1: Table showing the impact of different prefactors β on the AUC performance, the standard deviation σ and the mean sample space size $|\bar{\Omega}_{1:3}|$. Additionally, the table opposes the performance differences in using the approximation of (6.11) and the full recalculation of (6.10).

those calculated by the full reconstruction process. Moreover, Tab. 6.1 shows the differences of using the approximation of (6.8) compared to the formal recalculation of (6.7).

These results support the hypothesis that the approximation errors made by recalculating the pooled conditional distributions according to (6.8) do not significantly influence the quality of the reconstructed joint probability distribution. Additionally, the AUC decreases as well as the standard deviation σ increases for larger pooling thresholds. The reason for this is as follows. Larger pooling thresholds promote the pooling of comparitively likely situation hypotheses (almost at chance level). Thus, if the situation assessment is slightly off and the actual occurring situation hypotheses gets pooled, its probability value is underestimated as a consequence of distributing the probability values over all pooled situation hypotheses, uniformly. In similar way, hypotheses' probability values that would have been actually below the level of the uniform distributed values are overrated and can lead to false positives. However, the AUC decreases only for larger thresholds ($\beta > 0.6$) to levels comparable to those of the alternative models evaluated in Section 5.5.1. Thus, the hypotheses pooling provides noticeable reduction on computational complexity while still outperforming competing approaches. For example $\beta = 0.2$ leads to a AUC decrease of 0.2% while at the same time reduces the sample space on average from 27 to 23.33. This corresponds to a reduction of the needed operations for solving the linear equations systems by at least 35%, only

taking the effort for solving the linear equation system at the last recursion step $r = 1$ into account.

6.5 Summary

This chapter stated the need for approaches that are able to tackle the combinatorics associated with recognizing situations consisting of several traffic participants. The method suggests to combine maneuvers that are probably not interesting for the overall situation assessment by calculating the cumulative distribution of these events. By doing so, no maneuvers are neglected completely and the characteristic of the resulting probability distribution is preserved. This is achieved by recalculating the complete conditional distributions for the combined maneuver events and reconstructing the joint distribution afterwards. This leads to a reconstructed joint probability distribution where the probability values for the hypotheses that are not affected by the hypotheses reduction are not altered. The potential for decreasing the computational effort can be further increased when applying the recursive reconstruction scheme of Chapter 5.
After the formal outline of the procedure, the chapter provided a detailed qualitative example on how to apply the hypotheses pooling followed by a evaluation that highlighted the potential for complexity reduction on the complex real-world highway-entering scenario.

7 Conclusion

The main purpose of this thesis was to tackle challenges regarding complex situation recognition for the application in next generation ADAS and automated driving functions. This chapter summarizes the main results of the thesis and outlines possible directions for future research activities.

7.1 Summary

Future ADAS and automated driving promise to increase safety while simultaneously relieving stress on the driver. This work stated that to fulfill these promises, situation recognition systems need to be applicable to a large variety of complex traffic scenes. This lead to the three research questions that have been investigated throughout this work:

1. How can situation recognition systems be adapted to the currently encountered scene?

2. How can situation recognition systems be designed so that they are able to deal with several interacting traffic participants?

3. How can situation recognition systems scale to large number of possible situation classes?

Before the thesis at hand investigated these questions in detail, Chapter 2 gave an introduction to elementary terms and their notation. The usage and notation of probability theoretical concepts is generally affected by many misconceptions. For this reason Chapter 2 provided a common starting point. Similarly, the terms *maneuver* and *situation* are used with a variety of different meanings in the ITS community. In this work a maneuver is as a typical saptio-temporal behavior pattern of traffic participants and a situation is a combination of these typical behavior patterns over several traffic participants. After clarifying the usage and notation of these terms the chapter gave a brief description of the real-world data sets used to evaluate the approaches of the following chapters.

The question of how situation recognition systems can be adapted to a currently encountered scene was addressed in detail in Chapter 3. The literature review revealed that the majority of current maneuver estimations system can be categorized into two groups, *generative* and *discriminative* approaches. Moreover, it was found that generative approaches tend to deal easier with varying traffic scenes due to their probability theoretical foundation while discriminative approaches tend to provide superior recognition performances. Thus, Chapter 3 investigated possibilities to combine the advantageous of both categories and proposed a discriminative maneuver estimation framework that is fully adaptable to varying traffic scenes. The approach is based on a concept called pairwise probability coupling that allows to combine binary classifiers online. Experimental results showed that the proposed approach is not only able to construct a potent maneuver estimation for completely unseen scenes, it also outperforms specialized maneuver estimation systems on complex traffic scenes undergoing a large variability due to its efficient usage of available training data.

The following two chapters addressed the question of how situation recognition systems can be designed to deal with several interacting traffic participants without losing the ability to adapt to varying scene layouts. Chapter 4 started by revealing that very few approaches have been presented that actually consider situations consisting of several traffic participants. Furthermore, the majority of these approaches rely on the decomposability of the encountered scene into concrete independence assumptions or are restricted to specific scenes. For this reason, the chapter proposed to decompose the situation recognition into assessing each present traffic participant individually while taking possible future maneuvers of the other traffic participants into account without relying in any independence or infrastructure assumptions at all. These individual maneuver estimations can be easily adapted to changing traffic scenes, especially if the maneuver prediction framework of Chapter 3 is utilized. The resultant *complete conditional distributions* formed the basis for restoring the joint probability distribution of the holistic situation recognition when combined with probabilistic sampling techniques. However, in general these pose problems for the application in intelligent vehicles since they are time consuming and not deterministic.

Hence, Chapter 5 proposed two analytic reconstruction methods for restoring the joint probability distribution over all traffic participants. The methods centered around solving linear equation systems either directly or in recursive manner. They can be applied to a variety of general discrete

probabilistic sampling problems apart from the automotive area. Moreover, the chapter proved the existence of a stationary distribution for both methods. The proposed interaction-aware situation recognition framework showed its benefits on a complex highway scene with multiple interacting traffic participants in comparison with a system that does not decompose the scene into single maneuver predictions and a system assuming stochastic independence. Additionally, a comparison of Gibbs sampling with the proposed analytic reconstruction methods outlined their benefits in terms of computational effort and reliability of the results. A final evaluation outlined that the variable maneuver estimation framework of Chapter 3 can be used as basis for the interaction-aware situation recognition framework enabling it to deal with varying scene layouts.

Chapter 6 addressed the question of how to deal with the rapidly increasing number of situation hypotheses in the presence of several traffic participants and proposed a method for reducing it. The approach centered around calculating the cumulative distribution of uninteresting or similar maneuver combinations before restoring the joint probability distribution. The procedure can either be applied directly or in a recursive manner which extends the potential for reducing situation hypotheses even further. A real-world evaluation showed that the presented approach has great potential for reducing the number of investigated situation hypotheses while having only minor implications on the reconstruction quality.

Altogether, possible solutions to the stated research questions have been outlined in the main chapters. This thesis provides options for adapting situation recognition systems to changing environments, proposed a framework to deal with multiple interacting traffic participants, and investigated possibilities for scaling the proposed methods to large number of situation hypotheses. The conducted real-world evaluations encourage the feasibility of the proposed approaches.

7.2 Future Research

Situation recognition in generic traffic is a challenging task. Despite the promising solutions that were developed throughout this thesis there are still opportunities for improvement that are beyond the scope of this work.

The incorporation of **temporal dependencies** has not been investi-

gated as the scope of this thesis was not to push a specific application to its absolute performance maximum. Instead it provides universally applicable approaches to tackle general problems associated with situation recognition. Nevertheless, it can be expected that the incorporation of temporal dependencies, i.e. the recent history regarding vehicle states, if available, can improve situation recognition systems. These can be incorporated in two ways. The first possibility is to use classification methods that can deal with time-series data like long-short term memories [49]. A second possibility is to concatenate the reconstructed joint probabilities of successive time steps with recursive Bayesian filters [8]. Both options seem promising and need further investigation.

Rating the **importance of interactions** will be of great benefit. This work provided a framework that does not rely on any independence assumptions at all, however, there will be definitely scenes encountered where dependencies between traffic participants' behaviors are less important (or completely neglectable) compared to others. Recognizing these independencies and incorporating them into the reconstruction of joint probabilities within the situation recognition considerably reduces the inference task.

An important part for developing situation recognition systems as well as testing them is **data acquisition**. The acquisition of sufficiently large data sets for designing and testing situation recognition systems consumes a considerable amount of time and requires a comprehensive sensory set up. For this reason, possibilities need to be investigated that use artificially created data, e.g. from realistic simulation environments like SUMO [68]. This includes testing the generalization capabilities of systems trained on artificial data sets to their application in real-world scenarios. The author and his colleagues pursued such an approach in [63] with promising results. Another possibility is to take advantage of increasingly naturalistic driving studies like UDRIVE [34] or SHRP2 [24]. However, when considering such data sources the question of how to make use of these large amounts of data arises. The majority of related work and the classification algorithms used throughout this thesis are based on supervised learning techniques. Labeling large data sets takes great effort for human experts. Thus, developing automatic label generators or using unsupervised learning algorithms pose interesting future research fields.

In any case, this work provides a solid foundation to tackle remaining challenges and to thrive the development of situation recognition systems that are able to deal with generic traffic scenes while providing reliable estimation results.

A Experimental Data

A.1 Maneuver Estimation

A.1.1 Prototypical Velocity Profile Calculation

Labeling intersection approaches with possibly affecting traffic participants as a "straight - free drive" and "left - free drive" relies on the similarity to velocity profiles that were observed in completely undisturbed intersection approaches. To determine this similarity a probabilistic, prototypical velocity profile for the "free" intersection approaches is calculated for left and straight going paths.

To accomplish this, all $n^{\mathcal{T}}$ recorded velocity profiles, for one path alternative in free drive, are resampled for equidistant distances. These distances are chosen from $-10m$ before and $15m$ after the intersection in $1m$ steps and are indicated by $\hat{d} = -10m, -9m, \ldots, 14m, 15m$. The resampling is accomplished by using linear interpolation and yields $g = 1, \ldots, n_{\mathcal{T}}$ velocity profiles of the form $\mathcal{T}_g = \{(\hat{v}_{1,\hat{d}})_{\hat{d}=-10m,\ldots,15m}\}$. For each discretized distance the mean

$$\mu_{\hat{d}} = \frac{1}{n_g} \sum_{g=1}^{n_{\mathcal{T}}} \hat{v}_{1,\hat{d},g}, \;\; \forall \hat{d} = -10m, \ldots, 15m,$$

and the variance

$$\sigma^2_{\hat{d}} = \frac{1}{n_{\mathcal{T}}} \sum_{g=1}^{n_{\mathcal{T}}} (\hat{v}_{1,\hat{d},g} - \mu_{\hat{d}})^2, \;\; \forall \hat{d} = -10m, \ldots, 15m,$$

are calculated. These two values describe a prototypical velocity for a certain distance $\hat{d}$ in free drive.

For a new trajectory's velocity profile with velocity $\hat{v}_{1,\hat{d},i}$ that needs to be labeled, the probability of being an approach in free drive is calculated according to

$$p(B_1 = \text{"free drive"} | \hat{v}_{1,\hat{d}}, \mu_{\hat{d}}, \sigma^2_{\hat{d}}) = \frac{1}{25} \sum_{\hat{d}=-10m}^{15m} \frac{\mathcal{N}(\hat{v}_{1,\hat{d}} | \mu_{\hat{d}}, \sigma^2_{\hat{d}})}{\mathcal{N}(\mu_{\hat{d}} | \mu_{\hat{d}}, \sigma^2_{\hat{d}})},$$

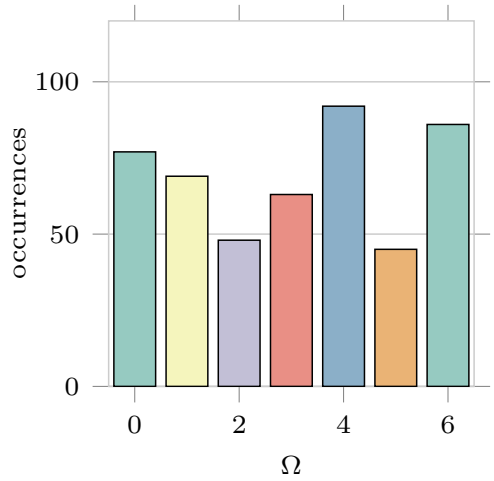
bar key/	maneuver
0	left - turn/ in front
1	left - slow down/yield
2	left - full stop/yield
3	left - influenced/preceding
4	left - free drive
5	straight - free drive
6	straight - influenced/preceding

Table A.1: Shows the mapping between maneuvers and the bars present in Fig. A.1

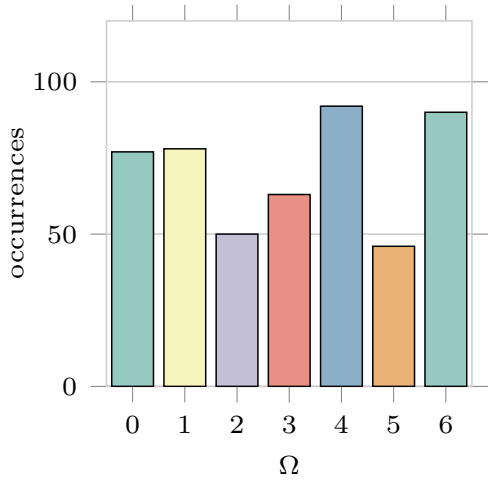
where $\mathcal{N}$ refers to a Gaussian distribution. The intersection approach is labeled as a free drive if the probability exceeds

$$p(B_1 = \text{“free drive”}|\hat{v}_{1,\hat{d},g}, \mu_{\hat{d}}, \sigma^2_{\hat{d}}) \geq 0.5.$$

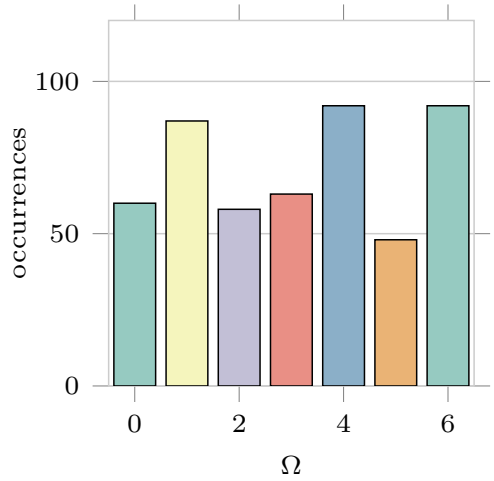
A.1.2 Detailed Situation Distribution



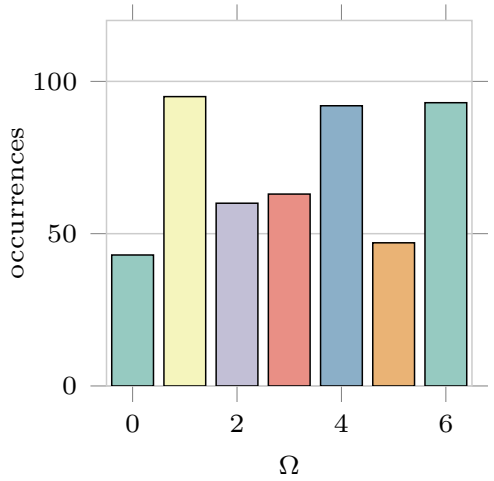
(a) $TTI = 0s$, 480 cases extracted



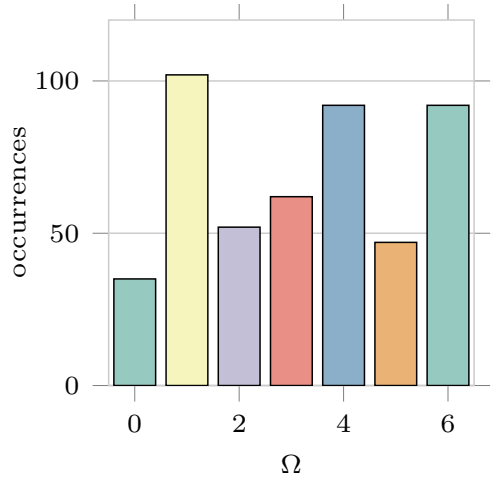
(b) $TTI = 0.5s$, 489 cases extracted



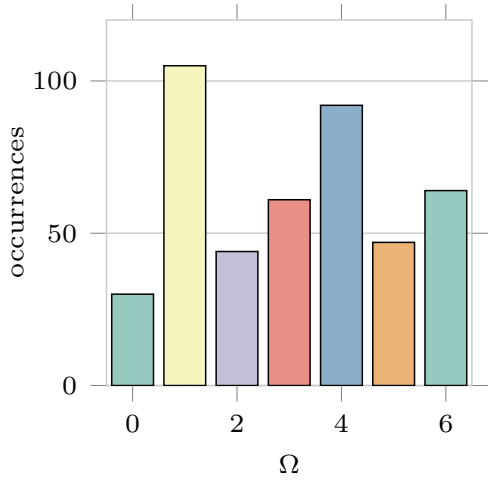
(c) $TTI = 1s$, 500 cases extracted



(d) $TTI = 1.5s$, 493 cases extracted



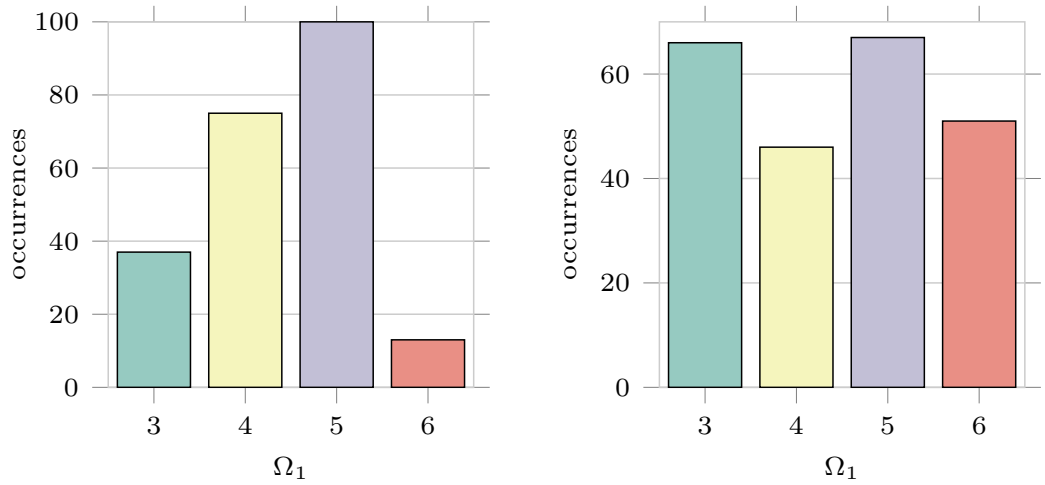
(e) $TTI = 2s$, 482 cases extracted



(f) $TTI = 2.5s$, 473 cases extracted

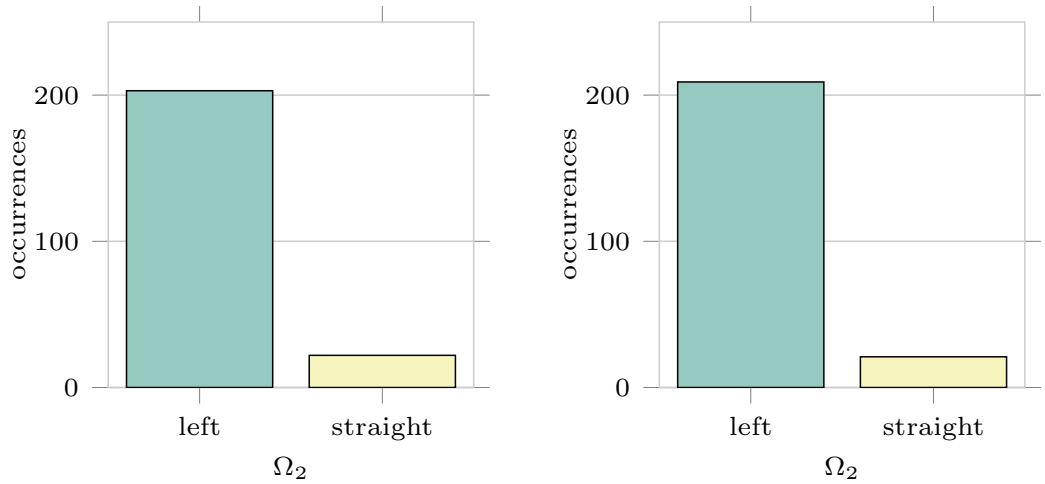
Figure A.1: Number of occurrences according to the 7 investigated maneuvers for different Time To Intersections (TTI). Tab. A.1 shows the bars' membership to their maneuvers.

A.2 Interaction-aware Maneuver Predictions



(a) $TTI = 3s$, 225 cases extracted **(b)** $TTI = 0.5s$, 230 cases extracted

Figure A.2: Number of occurrences according to the 4 investigated maneuvers of vehicle 1 for different Time To Intersections (TTI). the bars' membership to their maneuvers are listed in Tab. A.1.



(a) $TTI = 3s$, 225 cases extracted **(b)** $TTI = 0.5s$, 230 cases extracted

Figure A.3: Number of occurrences according to the 2 investigated maneuvers of vehicle 2 for different Time To Intersections (TTI).

A.3 Interaction-aware Situation Recognition

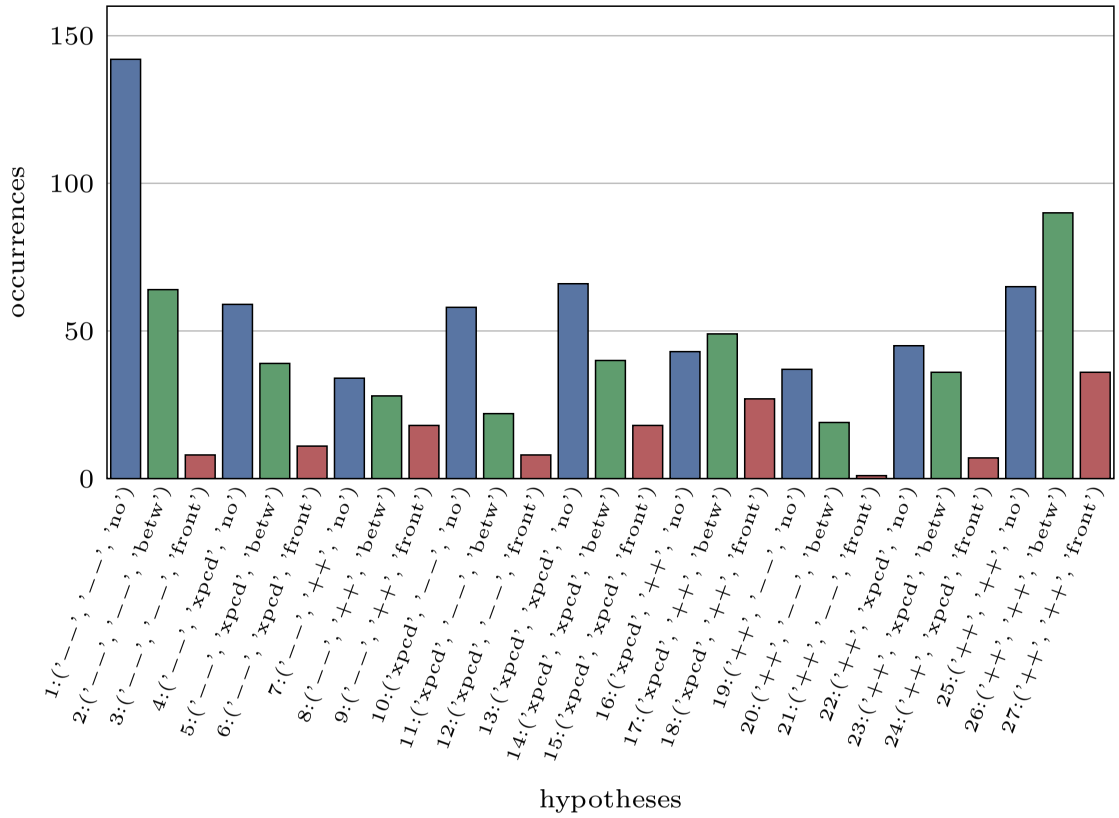


Figure A.4: Number of occurrences according to the 27 situation hypotheses. The hypotheses labels indicate one possible event combination over all three random variables in the form (b_3, b_2, b_1).

Bibliography

[1] Google Maps. https://goo.gl/PRkgme, 2017-01-11.

[2] Google Trends for "self-driving car". https://trends.google.ca/trends/explore?date=all&q=self%20driving%20car, 2017-01-11.

[3] Georges S. Aoude, Vishnu R. Desaraju, Lauren H. Stephens, and Jonathan P. How. Behavior classification algorithms at intersections and validation using naturalistic data. In *IEEE Intelligent Vehicles Symposium Proceedings*, 2011.

[4] Alexandre Armand, David Filliat, and Javier Ibañez-Guzman. Modelling stop intersection approaches using Gaussian processes. In *16th IEEE International Conference on Intelligent Transportation Systems*, 2013.

[5] Alexandre Armand, David Filliat, and Javier Ibañez-Guzman. Ontology-based context awareness for driving assistance systems. In *IEEE Intelligent Vehicles Symposium Proceedings*, 2014.

[6] Barry C. Arnold, Enrique Castillo, and José Maria Sarabia. Compatibility of partial or complete conditional probability specifications. *Journal of Statistical Planning and Inference*, 123(1), 2004.

[7] Barry C. Arnold and S. James Press. Compatible Conditional Distributions. *Journal of the American Statistical Association*, 84(405), 1989.

[8] M. Sanjeev Arulampalam, Simon Maskell, Neil Gordon, and Tim Clapp. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on signal processing*, 50(2), 2002.

[9] Mohammad Bahram, Constantin Hubmann, Andreas Lawitzky, Michael Aeberhard, and Dirk Wollherr. A Combined Model- and

Learning-Based Framework for Interaction-Aware Maneuver Prediction. *IEEE Transactions on Intelligent Transportation Systems*, 17(6), 2016.

[10] Mohammad Bahram, Andreas Lawitzky, Jasper Friedrichs, Michael Aeberhard, and Dirk Wollherr. A Game Theoretic Approach to Replanning-aware Interactive Scene Prediction and Planning. *IEEE Transactions on Vehicular Technology*, PP(99), 2016.

[11] David Barber. *Bayesian Reasoning and Machine Learning*. Cambridge University Press, 2015.

[12] Francesco Bartolucci and Julian Besag. A recursive algorithm for Markov random fields. *Biometrika*, 89, 2002.

[13] Asher Bender, James R. Ward, Stewart Worrall, and Eduardo M. Nebot. Predicting Driver Intent from Models of Naturalistic Driving. In *18th IEEE International Conference on Intelligent Transportation Systems*, 2015.

[14] Philipp Bender, Omer Sahin Tas, Julius Ziegler, and Christoph Stiller. The combinatorial aspect of motion planning: Maneuver variants in structured environments. *IEEE Intelligent Vehicles Symposium Proceedings*, 2015.

[15] Massimo Bertozzi, Alberto Broggi, Allesandro Coati, and Rean Isabella Fredriga. A 13,000 km Intercontinental Trip with Driverless Vehicles: The VIAC Experiment. *IEEE Intelligent Transportation System Magazine*, 5, 2013.

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

[17] Christopher M. Bishop and Julia Lasserre. Generative or discriminative? getting the best of both worlds. *Bayesian Statistics*, 8, 2007.

[18] Sarah Bonnin, Franz Kummert, and Jens Schmüdderich. A Generic Concept of a System for Predicting Driving Behaviors. In *15th IEEE International Conference on Intelligent Transportation Systems*, 2012.

[19] Sarah Bonnin, Thomas H. Weisswange, Franz Kummert, and Jens Schmüdderich. Accurate behavior prediction on highways based on a systematic combination of classifiers. In *IEEE Intelligent Vehicles Symposium Proceedings*, 2013.

[20] Sarah Bonnin, Thomas H. Weisswange, Franz Kummert, and Jens Schmüdderich. General Behavior Prediction by a Combination of Scenario-Specific Models. 15(4), 2014.

[21] Guillaume Bouchard and Bill Triggs. The tradeoff between generative and discriminative classifiers. In *16th IASC International Symposium on Computational Statistics*, 2004.

[22] Martin Buczko and Volker Willert. How to Distinguish Inliers from Outliers in Visual Odometry for High-speed Automotive Applications. In *IEEE Intelligent Vehicles Symposium Proceedings*, 2016.

[23] Lukáš Burget, Oldřich Plchot, Sandro Cumani, Ondřej Glembek, Pavel Matějka, and Niko Brümmer. Discriminatively trained probabilistic linear discriminant analysis for speaker verification. In *2011 IEEE international conference on acoustics, speech and signal processing*, 2011.

[24] Kenneth L. Campbell. The SHRP2 naturalistic driving study: Addressing driver performance and behavior in traffic safety. *TR News*, (282), 2012.

[25] Jorge Cano, Jordanka Kovaceva, Magdalena Lindman, and Matthias Bränström. Automatic Incident Detection and Classification at Intersections. Technical report, Volvo Car Corporation, 2009.

[26] Claas Rodemerk, Hermann Winner, and Robert Kastner. Predicting the driver's turn intentions at urban intersections using context-based indicators. In *IEEE Intelligent Vehicles Symposium Proceedings*, 2015.

[27] Claas Rodemerk, Robert Kastner, and Hermann Winner. Manöverprädiktion an innerstädtischen Knotenpunkten durch Exklusion alternativer Manöveroptionen. In *10. Workshop Fahrerassistenzsysteme*, 2015.

[28] David D. Clarke, Patrick J. Ward, and Jean Jones. Overtaking road-accidents: Differences in manoeuvre as a function of driver age. *Accident Analysis & Prevention*, 30(4), 1998.

[29] Michael Collins. Discrimative Training Methods for Hidden Markov Models: Theory and Experiments with Perceptron Algorithms. *Proceedings of the Conference on Empirical Methods in NLP*, 2002.

[30] Mary K. Cowles and Bradley P. Carlin. Markov chain Monte Carlo Convergence Diagnostics: A Compartive Review. *Journal of the American Statistical Association*, 91, 1996.

[31] Florian Damerow, Stefan Klingelschmitt, and Julian Eggert. Spatio-Temporal Trajectory Similarity and its Application to Predicting Lack of Interaction in Traffic Situations. In *19th IEEE International Conference on Intelligent Transportation Systems*, 2016.

[32] Deutscher Verkehrssicherheitsrat. Was leisten fahrerassistenzsysteme? Technical report, 2010.

[33] Pedro Domingos. A few useful things to know about machine learning. *Communications of the ACM*, 55(10), 2012.

[34] Rob Eenink, Yvonne Barnard, Martin Baumann, Xavier Augros, and Fabian Utesch. UDRIVE: the European naturalistic driving study. In *Proceedings of Transport Research Arena*, 2014.

[35] Julian Eggert, Florian Damerow, and Stefan Klingelschmitt. The Foresighted Driver Model. In *IEEE Intelligent Vehicles Symposium Proceedings*, 2015.

[36] Julian Eggert, Stefan Klingelschmitt, and Florian Damerow. The Foresighted Driver: Future ADAS Based on Generalized Predictive Risk Estimation. In *FAST-zero 2015 Symposium*, 2015.

[37] Richard W. Farebrother. *Linear Least Squares Computations*. CRC Press, 1988.

[38] Tom Fawcett. An Introduction to ROC Analysis. *Pattern Recognition Letters*, 27(8), 2006.

[39] Richard M. Feldman and Ciriaco Valdez-Flores. *Applied Probability and Stochastic Processes*. Springer, 2009.

[40] FHWA. Home of the Next Generation SIMulation Community. http://ngsim-community.org, 2015-11-18.

[41] Michaël Garcia Ortiz, Jannik Fritsch, Franz Kummert, and Alexander Gepperth. Behavior prediction at multiple time-scales in inner-city scenarios. In *IEEE Intelligent Vehicles Symposium Proceedings*, 2011.

[42] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite. In *Conference on Computer Vision and Pattern Recognition*, 2012.

[43] Tobias Gindele, Sebastian Brechtel, and Rüdiger Dillmann. A Probabilistic Model for Estimating Driver Behaviors and Vehicle Trajectories in Traffic Environments. In *13th International IEEE Conference on Intelligent Transportation Systems*, 2010.

[44] Tobias Gindele, Sebastian Brechtel, and Rüdiger Dillmann. Learning context sensitive behavior models from observations for predicting traffic situations. In *16th IEEE International Conference on Intelligent Transportation Systems*, 2013.

[45] Regine Graf, Hendrik Deusch, Florian Seeliger, Martin Fritzsche, and Klaus Dietmayer. A learning concept for behavior prediction at intersections. In *IEEE Intelligent Vehicles Symposium Proceedings*, 2014.

[46] Daniel Grossman and Pedro Domingos. Learning Bayesian network classifiers by maximizing conditional likelihood. In *International Conference on Machine Learning*, 2004.

[47] Mark Harris. How Google's Autonomous Car Passes the First U.S. State Self-Driving Test, 2014. http://spectrum.ieee.org/transportation/advanced-cars/how-googles-autonomous-car-passed-the-first-us-state-selfdriving-test.

[48] David Heckermann, David Maxwell Chickering, Christopher Meek, Robert Rounthwaite, and Carl Kadie. Dendency Networks for Inference, Collaborative Filtering, and Data Visualization. *Journal of Machine Learning Research*, 2000.

[49] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8), 1997.

[50] Kaizhu Huang. Discriminative Naive Bayesian Classifiers. *Department of Computer Science and Engineering, The Chinese University Hong Kong*, 2003.

[51] Michael Hülsen, J. Marius Zöllner, and Christian Weiss. Traffic intersection situation description ontology for advanced driver assistance. In *IEEE Intelligent Vehicles Symposium Proceedings*, 2011.

[52] Ashesh Jain, Hema S. Koppula, Bharad Raghavan, Shane Soh, and Ashutosh Saxena. Car that Knows Before You Do: Anticipating Maneuvers via Learning Temporal Driving Models. *arXiv:1504.02789*, 2015.

[53] Ashesh Jain, Hema S. Koppula, Shane Soh, Bharad Raghavan, Avi Singh, and Ashutosh Saxena. Brain4Cars: Car That Knows Before You Do via Sensory-Fusion Deep Learning Architecture. *arXiv:1601.00740*, 2016.

[54] Tony Jebara. *Discriminative, Generative and Imitative Learning.* PhD thesis, 2002.

[55] Michael Johannes and Nicholas Polson. MCMC Methods for Continuous-Time Financial Econometrics. *Handbook of Financial Econometrics*, 2, 2010.

[56] Björn Johansson, Johan Wiklund, Per-Erik Forssén, and Gösta Granlund. Combining shadow detection and simulation for estimation of vehicle size and position. *Pattern Recognition Letters*, 30(8), 2009.

[57] Jonas Firl. *Probabilistic Maneuver Recognition in Traffic Scenarios.* PhD thesis, 2014.

[58] Eugen Käfer, Christoph Hermes, Christian Wöhler, Helge Ritter, and Franz Kummert. Recognition of situation classes at road intersections. In *IEEE International Conference on Robotics and Automation*, 2010.

[59] Dietmar Kasper, Galia Weidl, Thao Dang, Gabi Breuel, Andreas Tamke, and Wolfgang Rosenstiel. Object-Oriented Bayesian Networks for Detection of Lane Change Maneuvers. *Intelligent Transportation Systems Magazine, IEEE*, 4(3), 2012.

[60] Aaron Kessler. Elon Musk Says Self-Driving Tesla Cars Will Be in the U.S. by Summer, 2015.

https://www.nytimes.com/2015/03/20/business/elon-musk-says-self-driving-tesla-cars-will-be-in-the-us-by-summer.html?hpw&rref=automobiles&action=click&pgtype=Homepage&module=well-region&region=bottom-well&WT.nav=bottom-well&_r=1, 2017-06-02.

[61] Nick Kingsbury. *Joint and Conditional cdfs and pdfs.* OpenStax CNX, 2005.

[62] Aldebaro Klautau, Nikola Jevtic, and Alon Orlitsky. Discriminative Gaussian mixture models: A comparison with kernel classifiers. In *International Conference on Machine Learning*, 2003.

[63] Stefan Klingelschmitt, Florian Damerow, and Julian Eggert. Managing the Complexity of Inner-City Scenes : An Efficient Situation Hypotheses Selection Scheme. In *IEEE Intelligent Vehicles Symposium Proceedings*, 2015.

[64] Stefan Klingelschmitt, Florian Damerow, Volker Willert, and Julian Eggert. Probabilistic Situation Assessment Framework for Multiple, Interacting Traffic Participants in Generic Traffic Scenes. In *IEEE Intelligent Vehicles Symposium Proceedings*, 2016.

[65] Stefan Klingelschmitt and Julian Eggert. Using Context Information and Probabilistic Classification for Making Extended Long-Term Trajectory Predictions. In *18th IEEE International Conference on Intelligent Transportation Systems*, 2015.

[66] Stefan Klingelschmitt, Matthias Platho, Horst-Michael Groß, Volker Willert, and Julian Eggert. Combining Behavior and Situation Information for Reliably Estimating Multiple Intentions. In *IEEE Intelligent Vehicles Symposium Proceedings*, 2014.

[67] Stefan Klingelschmitt, Volker Willert, and Julian Eggert. Probabilistic, Discriminative Maneuver Estimation in Generic Traffic Scenes using Pairwise Probability Coupling. In *19th IEEE International Conference on Intelligent Transportation Systems*, 2016.

[68] Daniel Krajzewicz, Jakob Erdmann, Michael Behrisch, and Bieker Laura. Recent Development and Applications of SUMO - Simulation of URBAN MObility. *International Journal on Advances in Systems and Measurements*, 2012.

[69] Tobias Kühnl and Jannik Fritsch. Visio-spatial road boundary detection for unmarked urban and rural roads. In *IEEE Intelligent Vehicles Symposium Proceedings*, 2014.

[70] Florian Kuhnt, Ralf Kohlhaas, Thomas Schamm, and J. Marius Zöllner. Towards a Unified Traffic Situation Estimation Model – Street-dependent Behaviour and Motion Models. In *IEEE Intelligent Vehicles Symposium Proceedings*, 2015.

[71] Florian Kuhnt, Jens Schulz, Thomas Schamm, and J. Marius Zöllner. Understanding Interactions between Traffic Participants based on Learned Behaviors. 2016.

[72] Andreas Lawitzky, Daniel Althoff, Christoph F. Passenberg, Georg Tanzmeister, Dirk Wollherr, and Martin Buss. Interactive scene prediction for automotive applications. In *IEEE Intelligent Vehicles Symposium Proceedings*, 2013.

[73] Stéphanie Lefèvre, Christian Laugier, and Javier Ibañez-Guzmán. Exploiting map information for driver intention estimation at road intersections. In *IEEE Intelligent Vehicles Symposium Proceedings*, 2011.

[74] Stéphanie Lefèvre, Christian Laugier, and Javier Ibañez-Guzmán. Risk assessment at road intersections: Comparing intention and expectation. In *IEEE Intelligent Vehicles Symposium Proceedings*, 2012.

[75] Stéphanie Lefèvre, Dizan Vasquez, and Christian Laugier. A survey on motion prediction and risk assessment for intelligent vehicles. *ROBOMECH Journal*, 1(1), 2014.

[76] Anders Lie and Claes Tingvall. Government status report. In *Government Status Report*. Swedish National Road Administration, 2009.

[77] Martin Liebner, Felix Klanner, Michael Baumann, Christian Ruhhammer, and Christoph Stiller. Velocity-Based Driver Intent Inference at Urban Intersections in the Presence of Preceding Vehicles. *Intelligent Transportation Systems Magazine, IEEE*, 5(2), 2013.

[78] Martin Liebner, Christian Ruhhammer, Felix Klanner, and Christoph Stiller. Generic driver intent inference based on parametric models. In *16th IEEE International Conference on Intelligent Transportation Systems*, 2013.

[79] Hiren M. Mandalia and Dario D. Salvucci. Using Support Vector Machines for Lane-Change Detection. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 49(22), 2005.

[80] Sean Meyn and Richard L. Tweedie. *Markov Chains and Stochastic Stability*. Springer-Verlag, 2012.

[81] Jonathan Milgram, Mohamed Cheriet, and Robert Sabourin. "One Against One" or "One Against All": Which One is Better for Handwriting Recognition with SVMs? *Tenth International Workshop on Frontiers in Handwriting Recognition*, 2006.

[82] Brendan Morris, Anup Doshi, and Mohan M. Trivedi. Lane Change Intent Prediction for Driver Assistance: Design and On-Road Evaluation. In *IEEE Intelligent Vehicles Symposium Proceedings*, 2011.

[83] Kevin P. Murphy, Yair Weiss, and Michael I. Jordan. Loopy belief propagation for approximate inference: An empirical study. In *Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence*, 1999.

[84] Andrew Y. Ng and Michael I. Jordan. On Discriminative vs. Generative classifiers: A comparison of logistic regression and naive Bayes. In *Neural Information Processing Systems Proceedings*, 2001.

[85] Alexandru Niculescu-Mizil and Rich Caruana. Predicting Good Probabilities with Supervised Learning. In *International Conference on Machine Learning*, 2005.

[86] Jean-Marc Nigro, Sophie Loriette-Rougegrez, and Michèle Rombaut. Driving situation recognition with uncertainty management and rule-based systems. *Engineering Applications of Artificial Intelligence*, 15(3–4), 2002.

[87] Judea Pearl. *Causality*. Cambridge University Press, 2009.

[88] Judea Pearl. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann, 2014.

[89] Dominik Petrich, Thao Dang, Gabi Breuel, and Christoph Stiller. Assessing map-based maneuver hypotheses using probabilistic methods and evidence theory. In *IEEE 17th International Conference on Intelligent Transportation Systems*, 2014.

[90] Matthias Platho and Julian Eggert. Deciding what to inspect first: Incremental situation assessment based on information gain. In *15th IEEE International Conference on Intelligent Transportation Systems*, 2012.

[91] Matthias Platho, Horst-Michael Groß, and Julian Eggert. Traffic situation assessment by recognizing interrelated road users. In *15th IEEE International Conference on Intelligent Transportation Systems*, 2012.

[92] Matthias Platho, Horst-Michael Groß, and Julian Eggert. Learning Driving Situations and Behavior Models from Data. In *16th IEEE International Conference on Intelligent Transportation Systems*, 2013.

[93] John Platt. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Advances in large margin classifiers*, 10(3), 1999.

[94] Lawrence R. Rabiner. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE*, 77(2), 1989.

[95] Gareth O. Roberts and Jeffrey S. Rosenthal. General state space Markov chains and MCMC algorithms. *Probability Surveys*, 1, 2004.

[96] Richard Scheins. An introduction to Causal Inference. In *Causality in Crisis?* 1997.

[97] Julian Schlechtriemen, Florian Wirthmueller, Andreas Wedel, Gabi Breuel, and Klaus-Dieter Kuhnert. When will it change the lane? A probabilistic regression approach for rarely occurring events. In *IEEE Intelligent Vehicles Symposium Proceedings*, 2015.

[98] Jens Schmüdderich, Rebhan Sven, Thomas H. Weisswange, Marcus Kleinehagenbrock, Robert Kastner, Mirmichi Nishigaki, Hiroyuki Kamiya, Naoki Mori, Shunsuke Kusuhara, and Shimosuke Ishida. A novel approach to driver behavior prediction using scene context and

physical evidence for intelligent Adaptive Cruise Control (i-ACC). In *FAST-zero 2015 Symposium*, 2015.

[99] Matthias Schreier. *Bayesian Environment Representation, Prediction, and Criticality Assessment for Driver Assistance Systems*. PhD thesis, 2015.

[100] Matthias Schreier, Volker Willert, and Jürgen Adamy. Bayesian, maneuver-based, long-term trajectory prediction and criticality assessment for driver assistance systems. In *IEEE 17th International Conference on Intelligent Transportation Systems*, 2014.

[101] Statistisches Bundesamt. Unfallentwicklung auf deutschen Straßen 2012. Technical report, 2013.

[102] Statistisches Bundesamt. Fachserie 8 Reihe 7 - Verkhersunfälle 2014. Technical report, 2015.

[103] Statistisches Bundesamt. Datenhandbuch zum Mirkozensus Scientific Use File 2016. Technical report, 2016.

[104] Thomas Streubel. *Situation Assessment at Intersections for Driver Assistance and Automated Vehicle Control*. PhD thesis, 2016.

[105] Thomas Streubel and Karl Heinz Hoffmann. Predicting of Driver Intended Path at Intersections. In *IEEE Intelligent Vehicles Symposium Proceedings*, 2014.

[106] Marvin Struwe, Stephan Hasler, and Ute Bauer-Wersing. A Two-Stage Classifier Architecture for Detecting Objects under Real-World Occlusion Patterns. In *24th International Conference on Artificial Neural Networks*, 2014.

[107] Bo Tang, Salman Khokhar, and Rakesh Gupta. Turn Prediction at Generalized Intersections. In *IEEE Intelligent Vehicles Symposium Proceedings*, 2015.

[108] Quan Tran and Jonas Firl. A probabilistic discriminative approach for situation recognition in traffic scenarios. In *IEEE Intelligent Vehicles Symposium Proceedings*, 2012.

[109] Quan Tran and Jonas Firl. Modelling of traffic situations at urban intersections with probabilistic non-parametric regression. In *IEEE Intelligent Vehicles Symposium Proceedings*, 2013.

[110] Quan Tran and Jonas Firl. Online maneuver recognition and multimodal trajectory prediction for intersection assistance using nonparametric regression. In *IEEE Intelligent Vehicles Symposium Proceedings*, 2014.

[111] Simon Ulbrich, Till Menzel, Andreas Reschka, Fabian Schuldt, and Markus Maurer. Defining and Substantiating the Terms Scene, Situation, and Scenario for Automated Driving. In *18th IEEE International Conference on Intelligent Transportation Systems*, 2015.

[112] Christoph Urmson. Just press go: designing a self-driving vehicle, 2014. https://googleblog.blogspot.de/2014/05/just-press-go-designing-self-driving.html, 2017-06-02.

[113] Vladimir N. Vapnik. *Statistical Learning Theory.* Wiley New York, 1998.

[114] Erik Ward and John Folkesson. Multi-classification of Driver Intentions in Yielding Scenarios. In *18th IEEE International Conference on Intelligent Transportation Systems*, 2015.

[115] Galia Weidl, Anders L. Madsen, Dietmar Kasper, and Gabi Breuel. Optimizing Bayesian networks for recognition of driving maneuvers to meet the automotive requirements. In *IEEE International Symposium on Intelligent Control*, 2014.

[116] Jürgen Wiest, Matthias Hoffken, Ulrich Kresel, and Klaus Dietmayer. Probabilistic trajectory prediction with Gaussian mixture models. In *IEEE Intelligent Vehicles Symposium Proceedings*, 2012.

[117] Jürgen Wiest, Matthias Karg, Felix Kunz, Stephan Reuter, Ulrich Kresel, and Klaus Dietmayer. A probabilistic maneuver prediction framework for self-learning vehicles with application to intersections. In *IEEE Intelligent Vehicles Symposium*, 2015.

[118] Hermann Winner, Stephan Hakuli, Felix Lotz, and Christian Singer. *Handbuch Fahrerassistenzsysteme, Grundlagen, Komponenten und Systeme für aktive Sicherheit und Komfort.* Springer Viewig, 2015.

[119] Ting-fan Wu and Ruby C. Weng. Probability Estimates for Multiclass Classification by Pairwise Coupling. *Journal of Machine Learning Research*, 5, 2004.

Index

Curriculum Vitae

History	
starting 08.2017	Data Scientist at BI X (Boehringer Ingelheim's digital Lab)
2013 – 2017	Research associate at Institute of Automatic Control and Mechatronics, Control Methods and Robotics Lab at Technische Universität Darmstadt in cooperation with the Honda Research Institute Europe GmbH in Offenbach
2011 – 2013	Master of Science in Mechatronics, Technische Universität Darmstadt, specialization: simulation and control of mechatronic systems
2007 – 2011	Bachelor of Science in Mechatronics, Technische Universität Darmstadt
2007	Abitur, Gymnasium am Römerkastell in Alzey