VEHICLE-PEDESTRIAN IMPLICIT COORDINATION VIA UNCERTAINTY-AWARE PLANS

A Thesis

by

YA-CHUAN HSU

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Dylan A. Shell
Committee Members,	Swaminathan Gopalswamy
	Thomas R. Ioerger
Head of Department,	Scott Schaefer

December 2019

Major Subject: Computer Science

ABSTRACT

Drivers and other road users often encounter situations where priority is unclear or ambiguous, but must be resolved, for example, after arriving at an intersection nearly simultaneously. The participants in such scenarios reach agreement by communicating; while instinctive to humans, this is a significant challenge for autonomous vehicles. Currently, the nature of interaction for resolving ambiguous road situations between pedestrians and autonomous vehicles remains mostly in the realm of speculation, for which no direct means for expressing intent and acknowledgment has yet been established. This thesis approaches the challenge by contributing a model and approach for planning that can produce actions that are expressive and encode certain aspects of intent; the result is communicative in that vehicle-pedestrian coordination arises via a negotiation of intent in a prototypical unsignalized intersection crossing scenario. We deliberately construct a prototypical crossing setting with a vehicle and one pedestrian at an unsignalized intersection such that there is substantial ambiguity in crossing order. A decision-theoretic model is then used for capturing this scenario along with its ambiguity as uncertainty arising from non-determinism and partial observability. We solve the problem by first proposing a Markov decision process to express the interaction at the intersection. Next, we focus on the partial-observability and include it in the model to generate a sequence of vehicle actions by solving via a state-of-the-art online solver. We implement the approach on a self-driving Ford Lincoln MKZ platform and examine an experimental setting involving real-time interaction. The experiment shows that the method achieves safe and efficient navigation. We analyze the resulting policy in detail in simulation and examine the coupled behavior of the vehicle and pedestrian, interpreting evidence for implicit communication that emerges as the two resolve ambiguity to achieve safe and efficient navigation.

DEDICATION

To my lovely mother and father.

To my supportive grandparents.

To Ya-Han and Zhong-Jing.

ACKNOWLEDGMENTS

I would first like to express my sincere gratitude to Dr. Dylan Shell, who has been such a great advisor. He is a source of constant support and knowledge that guided me through questions about research, writing, and life as a graduate student. He consistently allowed me to work on my research but steered me in the right direction whenever he thought I needed it.

I would also like to thank Dr. Swaminathan Gopalswamy and Dr. Srikanth Saripalli for the various demos, presentations, and collaboration opportunities. Each and every event has served as an inspiration and source of motivation throughout my course of study at Texas A&M University.

Special thanks to members in CAST for their cooperation and feedback. And finally, last but by no means least, my fellow labmates and friends from Taiwan for their support, assistance and, of course, friendship.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a thesis committee consisting of Professor Dylan A. Shell (advisor) and Thomas R. Ioerger of the Department of Computer Science and Engineering and Professor Swaminathan Gopalswamy of the Department of Mechanical Engineering.

All work for the thesis was completed by the student in collaboration with Professor Swaminathan Gopalswamy of the Department of Mechanical Engineering and under the advisement of Professor Dylan A. Shell of the Department of Computer Science and Engineering.

Funding sources

Support for this research was provided in part by a grant from the U.S. Department of Transportation, University Transportation Centers Program to the Safety through Disruption University Transportation Center (451453-19C36).

The contents of this thesis reflect the views of the author, who is responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

With human drivers having been behind the wheel from the beginning of the early 20th century, the foundations of the social protocol that road users employ to mitigate contention is long-established. As applied control and systems engineering brings autonomous vehicles closer to day-to-day reality, the places where drivers once sat are starting to be seen empty. This replacement of human drivers may cost a price if we do not prudently supplant drivers who resolve social conundrums on the road.

Examples of current-generation autonomous vehicles lacking social competence are noticed from recent field studies [2, 3]. On the road, humans resolve ambiguities in traffic via social interaction, including in expressing intent. Common examples involve acknowledging or asserting the right of way, or communicating the intention to yield. In some cases, this protocol can even be so effective as to be almost entirely transparent. Pedestrians use minor signals and gaze to interact with car drivers [4]. Drivers demonstrate implicit communication actions, such as traveling at high speed as a signal to communicate their intention of not giving way to pedestrians at crossings [5]. The development of autonomous vehicles forces a careful reexamination of informal social protocol and communicative interactions. Several aspects that might previously have been understood merely as tacit parts of natural navigation need to be formulated precisely.

One place where informal protocols are particularly important, and in which autonomous vehicles must be competent participants, is when pedestrians wish to cross the road on which the vehicle is traveling. It is known that communicative interaction strengthens a pedestrian's confidence regarding when it is safe to cross a road [6]. As autonomous vehicles are starting to appear on the road, survey studies were carried out to gather pedestrian's experiences of encountering autonomous vehicles at crossings—some people expressed a great deal of concern when no human driver is seen [7]. In this work, we examine pedestrian and vehicle interaction by focusing on a scenario with an unsignalized intersection, as this is a representative circumstance in which communication is crucial. We will, later on, specifically explore some of the means to treat factors that cannot be sensed directly, such as crossing intention, but which is vital in reasoning about pedestrians' crossing.

We present a decision-theoretic model that expresses the interaction between a pedestrian and a vehicle. We aim to explore the possibility of mimicking human-controlled driving based on rewards that describe utilities of particular circumstances. Built on the results of our exploration, we then treat the ambiguity in crossing scenarios by considering aspects of pedestrian behaviour that are not directly observable as a form of uncertainty to be modelled. A plan that reasons over and manipulates this uncertainty is constructed. As we discuss in some detail, when the vehicle executes this plan, the result possesses some hallmarks of social competence, at least as applies to the small-scale scenario studied. Constructing plans that manage uncertainty, via information-gathering actions, or hedging outcomes to manage risk, does incur substantial computational costs, but recent improvements in both algorithms and hardware suggest this to be a promising direction.

We focus on approaches that are directly practicable. Our emphasis is on understanding minimal modifications to standard operating assumptions; thus, we use elements already commonplace, such as speeding up, slowing down and flashing headlights. This is appropriate in cases where pedestrians cannot determine whether the approaching vehicle is autonomous or not. More importantly, it gives us the opportunity to preserve *social* aspects so that existing understanding of pedestrians, including the role of implicit coordination and indirect signals, remains applicable without further presumptions.

1.1 Thesis Outline

In Chapter 2, we examine early studies on pedestrian crossing, which serve as background knowledge for designing our approach in this paper. In Chapter 3, we discuss work on communicative autonomous vehicles, including more recent studies of cooperative pedestrian interaction with autonomous vehicles. The chapter thereafter describes our problem in detail, which is then treated by models proposed in Chapter 5 and 7. Implementation and experiment design are also documented in both chapters, and we analyze the vehicle's behaviour individually in Chapter 6 and 8. Finally, Chapter 9 concludes the work with a highlight of our contributions.

2. BACKGROUND

We review the basic definitions and known influential factors of a pedestrian's crossing behaviour, followed by reviewing prior pedestrian and driver interaction studies as our foundation for later; ideally, precise formulating the crossing interaction model.

2.1 Pedestrian Behaviour at Crossings

Early studies of pedestrians crossing roads observed that pedestrians are primarily concerned with time-gaps [8]. According to the Highway Capacity Manual [9], a *critical gap* is defined to be below which a pedestrian will not attempt to begin crossing the street. Studies [8, 10] have identified that each person has their own critical gap, but which changes according to an oncoming vehicle's speed, and they will not cross if the vehicle is nearer than this threshold. Approximately $92\,\%$ of pedestrians cross a $7.0\,\mathrm{m}$ wide road when a vehicle is $7\,\mathrm{s}$ away, while $0\,\%$ cross when the time difference is less than $1.5\,\mathrm{s}$.

These primary studies were followed by research thoroughly evaluating factors that may influence peoples' critical gap. Brewer et al. [11] categorized pedestrian crossing maneuvers based on different traffic flow conditions and road geometry: 1) *single stage crossing* is when pedestrians cross the road in one crossing maneuver; 2) *two-stage* is when pedestrians cross up to the median first and then cross the far side subsequently; 3) *rolling* is when pedestrians search for gaps between a continuous flow of vehicles by adjusting the speed and direction of their movement. Harrell [12] analyzed crossing with variables for multiple parameters, including traffic volume, temperature, the width of the roadway, *etc.* He concluded that traffic volumes were found to obtain an inverse relationship with cautiousness. The model presented below includes a factor to represent a notion of caution.

Besides road conditions, it has been observed that personal characteristics pertaining to the specific pedestrian influence their critical gap. Studies have shown that gender affects the pedestrian's behaviour, most results indicating that males tend to take risky actions, whereas females often cross with greater caution [13, 14]. Age is another influence on pedestrian behaviour that has been studied. Oxley et al. [15] suggest that age-related perceptual and cognitive deficits affect one's crossing behaviour. Moreover, the complexity of the traffic has a larger effect on the behaviour of older pedestrians. For example, on two-way undivided roads, elderly people are frequently found crossing even when the traffic is already closing up [16].

2.2 Driver Behaviour at Crossings

Reciprocally, driver behaviour at crossings is also influenced by multiple aspects. These aspects include the group size of the crossing pedestrians, the distance of the pedestrian(s) to the crosswalk, and the size of the city [17]. Additionally, the current velocity of the vehicle affects the driver's decisions when approaching a crosswalk. If the vehicle is moving at high speed, drivers break significantly earlier [1]. An explanation is offered by the fact that drivers decelerate at a rate not exceeding $3.048\,\mathrm{m/s^2}$ for reasons of comfort [18]. The layout of the environment is another aspect that affects driver behaviour. At mid-road crosswalks with a curb extension, the vehicle has an average deceleration of $-1.92\,\mathrm{m/s^2}$. This is greater than an average deceleration rate of $-2.39\,\mathrm{m/s^2}$ occurring at mid-road crosswalks with advanced yield marks, but without curb extensions [19].

2.3 Markov Decision Processes

Formally, a Markov Decision Process (MDP) model [20] is defined by a tuple (S, A, T, R, γ), where S and A denote the model's state space and action space respectively. The transition function T is a conditional probability function $T(s, a, s') = p(s'|s, a)$ yielding the probability of transitioning from current state $s \in S$ to the next state $s' \in S$ when taking action $a \in A$. The process is decision-theoretic, being based on the rewards R, which describe the utility of particular circumstances, and which direct a maximizing agent to choose desirable actions. Immediate rewards are specified for each action taken in each state. The solution for an MDP, called a policy, $\pi : S \to A$ prescribes an action $a \in A$ for each state $s \in S$. More specifically, at each time step, an agent in state s performing action a will receive reward $R(s, a)$. The goal is to choose a policy that will maximize the agent's accumulated reward. We calculate the maximum accumulated reward with a discounted sum over a potentially infinite horizon $\sum_{t=0}^{\infty} \gamma^t R_{a_t}(s_t, s_{t+1})$, where $\gamma \in (0, 1)$ is a discount factor that models a preference for immediate rewards over future ones.

2.4 Partially Observable Markov Decision Process

Different from a Markov decision process, the state cannot be directly observed by the agent. The agent operates in a partially observable stochastic environment. Partially observable Markov decision process (POMDP) is formally expressed as a tuple (S, A, Z, T, O, R), where S is a set of states, A is a set of agent actions, and Z is a set of observations. When the agent takes an action $a \in A$, it moves to a new state $s' \in S$ with probability $T(s, a, s') = p(s'|s, a)$ and receives an observation $z \in Z$ with probability $O(s', a, z) = p(z|s', a)$. Therefore, without knowing the true state, the agent can maintain a belief, represented as a probability distribution, over the state based on the received observations. We solve a POMDP to obtain a policy π that maps from the belief space to the action space. In this work, we are interested in using an online POMDP planner, which uses forward search, from the current history or belief state, to form a local approximation to the optimal value function.

3. LITERATURE REVIEW

There is a rich literature in collision avoidance systems designed for autonomous robots operating in socially compliant settings [21], which robots are challenged to navigate around and avoid dynamic obstacles, human in particular, at shopping malls, airports and offices, where crowds of human exits. Among these studies, autonomous vehicles that operate on the road with pedestrians as interaction participants are less commonly found and are recently started to be appreciated. Therefore, we like to acknowledge the studies conducted on socially compliant robots collision avoidance systems in a different setting and then examine autonomous vehicles that are working towards being socially aware.

3.1 Autonomous Robots Navigating in Social Settings

Collision avoidance systems designed for autonomous vehicles operating in socially compliant settings can be categorized roughly into model-based and learning-based approaches. Our approach in this thesis falls into the scope of model-based methods, which typically introduce additional parameters to collision avoidance algorithms to account for social interactions [22, 23]. In existing works, people focus on modeling and replicating the detailed mechanism of social compliant, which remains hard to quantify as human behaviours are mainly stochastic. Thus, instead of focusing on modeling precise social information and feature-matching techniques, researchers have started to approach the problem from different angles.

With a different perspective on modeling, some studies have shown these approaches to be capable of handling socially competent environments while computing solutions efficiently. Chen et al. [24] believe the difficulty in quantifying sophisticated human social behaviour is in contrast with the intuitive evaluation of human behaviour is acceptable/reasonable. Thus, instead of modeling details of what to do, they model common social norms. Applying deep reinforcement learning, they developed a time-efficient navigation policy through pedestrian-rich environments. Another method for approaching social behaviour modeling is the work of Bandyopadhyay et al. [25], which combines the knowledge of the surrounding environment to predict pedestrian's heading destination and generate a conservative avoidance policy while maintaining uncertainty over pedestrians' heading directions. Data requirements are relaxed in learning methods through generalizing data format. Schneemann et al. [26] classified pedestrian's intention using a combination of an SVM along with a context-based feature descriptor. Okai et al. [27] use inverse reinforcement learning with a flexible graph-based representation to extend the degrees of social normativeness of the robot.

3.2 Cooperative Pedestrian Interaction with Autonomous Vehicles

There are many successful systems for autonomous driving [28, 29], but research study of close interactions between autonomous vehicles and pedestrians is more recent, which actual autonomous vehicle implementations that consider interaction are even less common. Most studies [1, 30] focus on reexamining driver-pedestrian interaction to provide reference data for future implementations of autonomous vehicles. The information gathered from these studies repeatedly point out that pedestrians use cues, such as eye-contact and facing direction, to indicate their next action, communicating intent.

A critical difficulty for autonomous vehicles driving amid pedestrians is to incorporate pedestrian intentions and behaviours into their decision making. People have modelled pedestrian intention using hidden Markov models and Gaussian processes. With a given pedestrian behaviour model, the simplest approach is to create a reactive system [31]. However, this ignores uncertainty inherent in making predictions, resulting in fast computation but sub-optimal solutions over time. Thus, different methods to reason about the prediction uncertainty during decision making have been proposed [32, 33]. The POMDP approach is general, assuming neither linear dynamics nor Gaussian noise [34]. Though POMDPs are widely known to be demanding computationally, steady improvements in efficiency have seen them being implemented in reasonably-scaled experiments [25]. This work leverages POMDPs to balance the uncertainties the pedestrian intention while having the vehicle operating safely and efficiently.

Among studies with autonomous vehicles that communicate their intentions (with regard to future actions) to nearby pedestrians, a variety of different methods have been explored, but no single solution outshines all others. Some studies state that showing physical information such as gap distance dominates the communication [35]. Others suggest that designing external interaction devices [36] can help boosts confidence in pedestrian decision-making. Several companies [37–39] have also proposed their own external hardware to interact with pedestrians. In contrast, we deliberately opt to use standard features only.

4. AUTONOMOUS VEHICLE AND PEDESTRIAN INTERACTION PROBLEM

4.1 Statement of Problem

We study a scenario where a pedestrian and an autonomous vehicle are both approaching the same segment of an unsignalized intersection, and there exists an ambiguity of who will cross the intersection first. In the case where human drives the vehicle, the vehicle and pedestrian interact via their respective choices of actions to efficiently and smoothly resolve this question as the situation unfolds. We pose the question of how to create an autonomous vehicle that can mimic the interactive behaviour of a human driver. A graphic representation of this scenario appears in Figure 4.1.

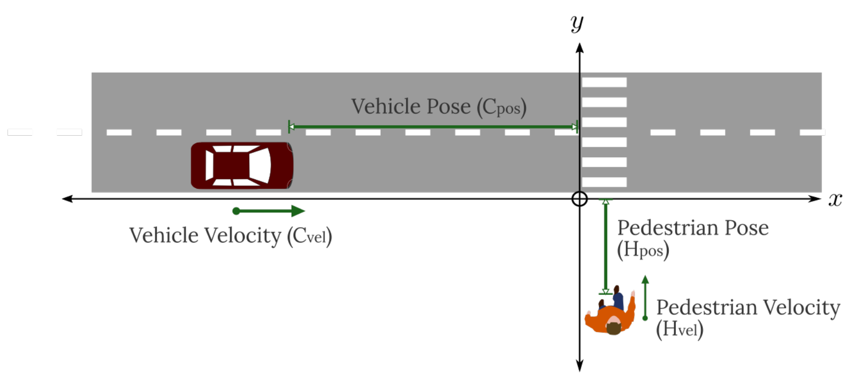

Figure 4.1: Bird's eye view of the unsignalized crossing.

4.2 Approaches

We aim to reproduce the dynamics of the social interaction mentioned above by designing the scenario as a decision-theoretical model. This provides a better understanding of both participants' dynamic process and the underlying incentives that lead the driver to perform socially; furthermore, we study the influential factors of pedestrian's crossing behaviour in order to formalize them such that we can create action selections for the vehicle which can maintain and even reduce the uncertainty in the scenario when necessary.

4.2.1 Markov decision process approach

From the perspective of implementing an autonomous vehicle, actions that are generated for solving the problem can only be sent to the vehicle. In our scenario, even though some actions such as honking a horn/hooter influences the pedestrian, there are limited to only creating an indirect influence. Considering the various factors that are involved in determining a pedestrian's crossing behaviour, we can state that the pedestrian's status becomes a stochastic outcome when given the vehicle executed an action.

We formulate the interaction as an MDP since the outcome of the vehicle's action creates a partly random result in the pedestrian's state. Our intention is for the vehicle to perform safely and efficiently in the situation. Solving for the MDP will choose actions that maximize the cumulated function of rewards designed based on our intentions. The uncertainty dynamics of the crossing interaction are maintained through this action selection process, which aims to reduce the probability of ending up in an undesired state, in other words, resulting in obtaining a lower accumulated reward.

4.2.2 Partially observable Markov decision process approach

Having understood the dynamic process, we now focus on the ambiguity of who will cross first. Unfortunately, the crossing order that the pedestrian has in mind cannot be observed directly by the vehicle. The vehicle is only able to gain this information regarding the pedestrian by integrating observations and using its model of the pedestrian's progress to learn more about the pedestrian's state. To model the pedestrian's evolving understanding of this crossing order question, we propose to use a concise representation, a single binary variable, which can be modelled as an unobservable state in a POMDP model. The vehicle maintains a belief (i.e., a distribution) over this binary variable while solving a policy for the model.

Based on the studies (in Section 2.1) of vehicle-pedestrian crossing interactions, the influence of a vehicle's action on a pedestrian's crossing behaviour and the replicate influence forms a complex relationship that couples the participants. Thus, this intertwined relationship can be used for resolving the crossing order ambiguity through observing and influencing each other's evolving understanding of the crossing order. From the vehicle's point of view, an appropriate choice of action helps ensure that it will make a sequence of observations that are informative. The vehicle is also able to manage its belief by initiating actions that create an interaction.

The preceding complexities are included in our model, along with one further important nuance: the vehicle is only disposed to gain information that is valuable for driving safely and efficiently through the intersection. We formulate an instance of a POMDP [40] that describes the impact of actions for the vehicle.

We leverage the solutions to POMDPs as they balance actions that gain information with ones that attain valuable rewards. The former actions are those that increase the vehicle's confidence in the pedestrian's crossing order understanding, including actions that influence the pedestrian's crossing behaviour. The latter would be actions that increase the safety and efficiency of crossing the intersection.

5. MARKOV DECISION PROCESS DESIGN*

We create a simulated autonomous vehicle that is capable of mimicking the dynamic behaviour of a human-controlled vehicle, including moderating its speed in a manner reflective of the presence of pedestrians at an unsignalized intersections. The autonomous vehicle is modelled as an MDP agent that needs to avoid collision with an approaching pedestrian who may opt to cross the intersection.

5.1 Basic Dynamic Model

First, we examine how a vehicle and a pedestrian can be modelled in a decoupled manner. Then the section which follows describes the details of how they are coupled together.

We express the dynamics of a pedestrian with a Markov chain. States represent discretized distances from the intersection. The velocity of the pedestrian determines the transition probabilities between the states. To calculate the transition probabilities, we first initialize some pedestrian with a given velocity and simulate the pedestrian moving forwards at this velocity. After several rounds of simulation, we average the times the pedestrian is located in each state and transform these numbers into probabilities. In this way, we create a unique Markov chain for every pedestrian velocity that would be used in our simulation (see Figure 5.1).

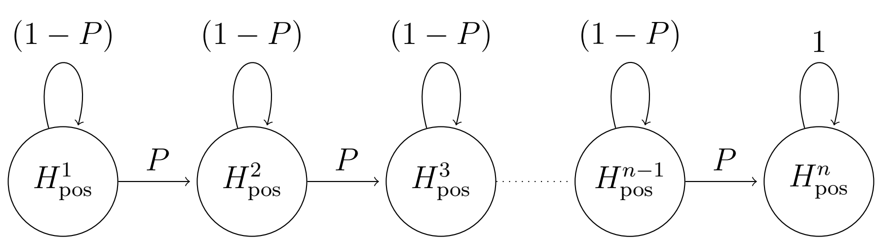

Figure 5.1: Pedestrian Dynamics Markov Chain. P is a number between 0 and 1 and represents the probability of transferring from one node to another

*Reprinted with permission from "An MDP Model of Vehicle-Pedestrian Interaction at an Unsignalized Intersection" by Ya-Chuan Hsu, Swaminathan Gopalswamy, Srikanth Saripalli, Dylan A. Shell, 2018. 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall), p. 1-6, copyright by © 2011 IEEE.

The vehicle's dynamics are expressed with the same method. But the states are the discretized distances from the vehicle to the intersection instead (see Figure 5.2).

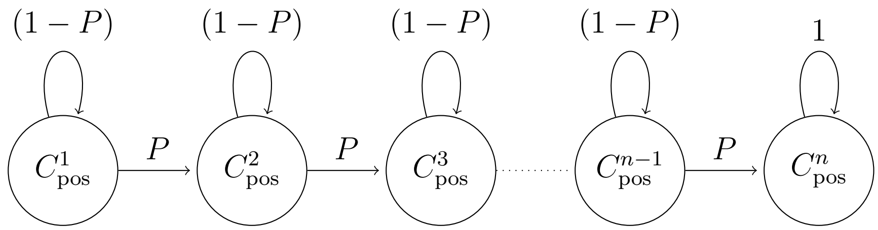

Figure 5.2: Vehicle Dynamics Markov Chain.

5.2 Markov Decision Process Model

We set the state space $\mathbb{S}$ of our MDP model as the joint dynamic state spaces of the vehicle and pedestrian. The set S is specified as a product: $(H_{\mathrm{pos}}, H_{\mathrm{vel}}, C_{\mathrm{pos}})$ (mnemonic: $H_{\square}$ for human and $C_{\square}$ for car). The pedestrian distance from the intersection is H_{pos}. When H_{pos} has a negative value, it represents a pedestrian that has not reached the crosswalk. The element H_{vel} represents the pedestrian speed in the direction of motion towards the intersection. Lastly, the state space S contains the vehicle's position C_{pos}, where C_{pos} is the vehicle's distance from the intersection. A negative value represents the fact that the vehicle has not yet traversed the crosswalk (see Figure 4.1). The states characterize variables that, though coarse, express details important for the interaction between the vehicle and the pedestrian.

The action space A encodes the control parameters of the autonomous vehicle. It is a set of different velocities C_{vel} that the vehicle can be commanded to drive at.

5.3 Reward Model

By adopting a decision-theoretic model, we are constructing an autonomous vehicle that acts as a rational agent to minimize its risk of collision with the pedestrian. Furthermore, we assume the autonomous vehicle wishes to cross the intersection as efficiently as possible. To model such behaviour within the MDP, we use the additive reward function $R(s, a) = R_{\mathrm{col}}(s) + R_{\mathrm{eff}}(a)$, where

$R_{\text{col}}(s)$ is the penalty imposed if at state $s \in S$, the autonomous vehicle and the pedestrian is both on the crosswalk and $R_{\text{eff}}(a)$ is the cost for the autonomous vehicle to perform action $a \in A$ from state $s \in S$.

6. MARKOV DECISION PROCESS EXPERIMENTS AND RESULTS

6.1 Markov Decision Process Approach Simulator

Given the problem modelled in the previous chapter, the motion strategy for the autonomous vehicle is generated by solving the associated MDP. Our software simulates the behaviour of an autonomous vehicle and pedestrian which may interact at an unsignalized intersection.

The simulation deals predominantly with the scenario in Figure 4.1. For the simulator to carry out an interaction, it requires two sources of input: a means by which the vehicle's speed is determined, and the same for the pedestrian. These sources, in general, can include prerecorded sequences, keyboard input, or other means. For the results we present, the vehicle's speed is always determined from a policy π^* that results from solving the MDP; determining the next action for the vehicle is then just a problem of evaluating $\pi^*\left[(H_{\mathrm{pos}}(t), H_{\mathrm{vel}}(t), C_{\mathrm{pos}}(t))\right]$. This requires that the pedestrian position, pedestrian speed, and vehicle position be mapped to the equivalence class of discrete states that describe that configuration:

$$\mathbf{State}^{\mathbf{Sim}}(t) \mapsto (H_{\mathrm{pos}}(t), H_{\mathrm{vel}}(t), C_{\mathrm{pos}}(t)).$$

The pedestrian's speed is determined in a different way. As we do not have control over the pedestrian, we construct separate models based on the knowledge of common pedestrian crossing behaviours. The model is varied with the experiments, so they will be described in detail next. With the scope of our MDP approach, it is important to note that we are interested in behaviour where the vehicle has an imperfect or inaccurate model. Thus, though evaluation uses a model of a pedestrian to generate that simulated pedestrian's action, the model provided as part of the MDP may differ markedly.

6.2 Experimental Setup

We consider two different choices of pedestrian characteristics, which each of these results in different dynamics for the pedestrian. Because the MDP expresses a coupling of the pedestrian and vehicle dynamics, each of these choices leads to a particular MDP.

Having solved each MDP, we then use different pedestrian behaviours to analyze the vehicle's reaction generated by the optimal policy. In all cases, of course, the vehicle is expected to perform a series of actions to cross safely and efficiently.

6.2.1 Setting parameters

The average velocity of a walking pedestrian is $1.4\,\mathrm{m/s}$ [41] and the fastest a pedestrian walks is $2.5\,\mathrm{m/s}$ [42]. We constructed Markov chains for pedestrians traveling at speed $0.0\,\mathrm{m/s}$ and the speeds between $1.4\,\mathrm{m/s}$ and $2.5\,\mathrm{m/s}$. For the vehicle, we set up three action choices, which are the vehicle traveling at a high speed of $10.0\,\mathrm{m/s}$, a slow speed of $5.0\,\mathrm{m/s}$, and a stop with the vehicle's velocity at $0.0\,\mathrm{m/s}$.

The design of the MDP, via a state space expressing the joint state of the vehicle and pedestrian, can capture important elements of the interaction between the two. We consider a state $(H_{\mathrm{pos}}, H_{\mathrm{vel}}, C_{\mathrm{pos}})$, where H_{pos} is a set from -6.0 to 6.0 with intervals of 0.5, C_{pos} is a set from -15.0 to 11.0 with intervals of 1.5, and H_{vel} is defined differently according to the pedestrian characteristic.

The two different choices of pedestrian characteristics are reckless and cautious. Our definition of the two different crossing dynamic behaviours are presented in detail as follow:

6.2.1.1 Reckless pedestrian crossing behaviour

Reckless pedestrian crossing behaviour is defined as a pedestrian that always aims to start crossing the intersection before the vehicle arrives at the crossing. When the pedestrian reaches $2\,\mathrm{m}$ from the intersection, it predicts the vehicle's arrival time according to the current vehicle's position and velocity (which we assume the pedestrian can approximately quantify). With the MDP model expressing the vehicle, we can obtain the vehicle's position information from the current state space and the vehicle's velocity is given by the action state the vehicle chooses to perform. If the vehicle appears to be arriving at the intersection before the pedestrian finishes crossing, the pedestrian will speed up in order to cross first. The action of speeding up may cause a collision, but the pedestrian does this nevertheless. This class of crossing behaviour can be seen on university campuses, especially periods between classes.

This description is shown quantitatively as

$$H_{\text{vel}}(x) = \begin{cases} 1.4 & \text{if } x > 2 \\ 2.5 \times e^{-0.289x} & \text{if } x \leq 2 \end{cases}, \tag{6.1}$$

where x is the time difference between the remaining time for the vehicle to arrive at the intersection and the remaining time for the pedestrian to finish crossing the intersection.

6.2.1.2 Cautious pedestrian crossing behaviour

Cautious pedestrians stop at the curb and wait for the vehicle to cross first when they determine that, continuing at their current speed, they cannot reach the other side of the road before the vehicle arrives.

Their velocity is given as

$$H_{\text{vel}}(d) = \begin{cases} 1.4 & \text{if } d < -1 \text{ and } d > 1 \\ 0 & \text{if } -1 \leq d \leq 1 \text{ and can't cross at } 1.4\,\mathrm{m/s} \\ 1.4 & \text{if } -1 \leq d \leq 1 \text{ and can cross at } 1.4\,\mathrm{m/s} \end{cases}, \tag{6.2}$$

where d is the distance between the pedestrian and crosswalk.

Using the reward model, $R(s,a) = R_{\text{col}}(s) + R_{\text{eff}}(a)$, described in Section 5.3, we define the collision zone to be the entire crosswalk area, assigning the $R_{\text{col}}(s)$ value in the states that cover these areas a reward of -1000. $R_{\text{eff}}(a)$ is set to encourage the autonomous vehicle to finish crossing the intersection as soon as possible; $R_{\text{eff}}(a)$ is set to 10 when the vehicle action a is travel in high speed, 5 when traveling in slow speed, and -10 when stopping/stopped.

6.3 Experimental Results and Analysis

The object of these experiments is to analyze the behaviour of the autonomous vehicle. We present quantitative data followed by an interpretation of the results.

We analyze the behaviour by observing the vehicle's speed profile, collision occurrence, and the accumulated rewards for each simulation. To generate the autonomous vehicle's behaviour, we first define the input pedestrian behaviour. In this experiment we simulate pedestrians moving at constant speed $1.0\,\mathrm{m/s}$, $1.4\,\mathrm{m/s}$, $2.0\,\mathrm{m/s}$ and $2.5\,\mathrm{m/s}$ and also simulate the reckless (Section 6.2.1.1) and cautious pedestrians (Section 6.2.1.2) as separate experimental conditions.

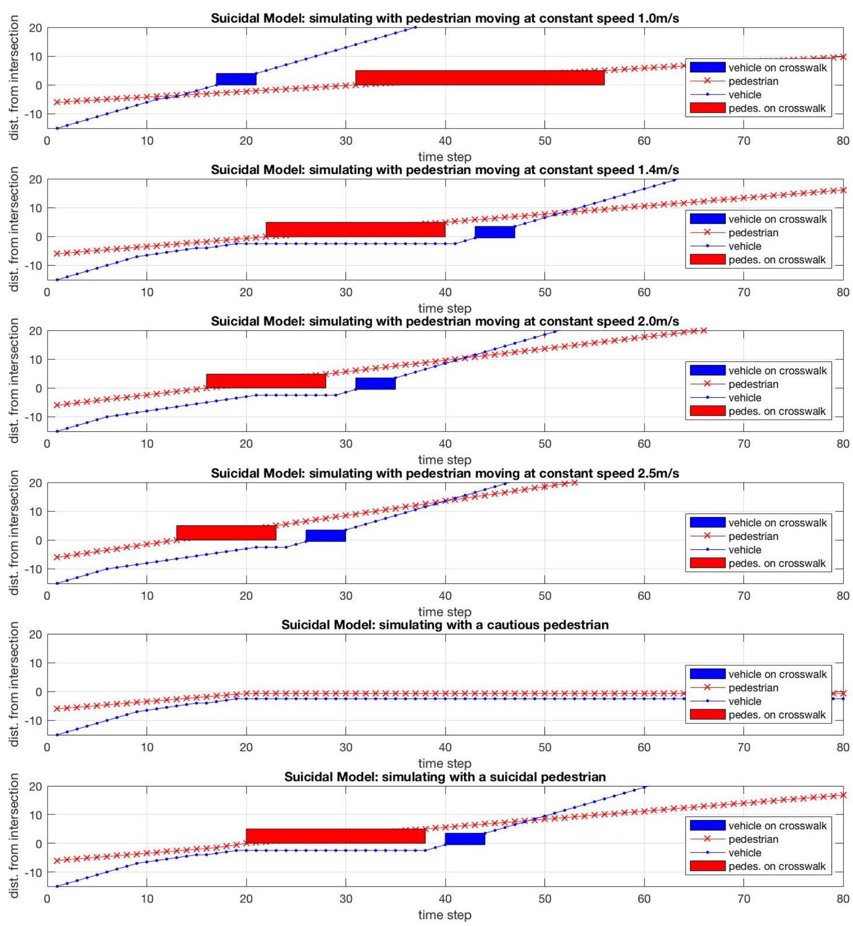

Figure 6.1: Reckless pedestrian-based MDP model: We construct a reckless pedestrian-based MDP model and examine the autonomous vehicle's behaviour by simulating the model with pedestrians moving at constant speed $1.0\,\mathrm{m/s}$, $1.4\,\mathrm{m/s}$, $2.0\,\mathrm{m/s}$ and $2.5\,\mathrm{m/s}$, a cautious pedestrian and a reckless pedestrian as separate experimental conditions.

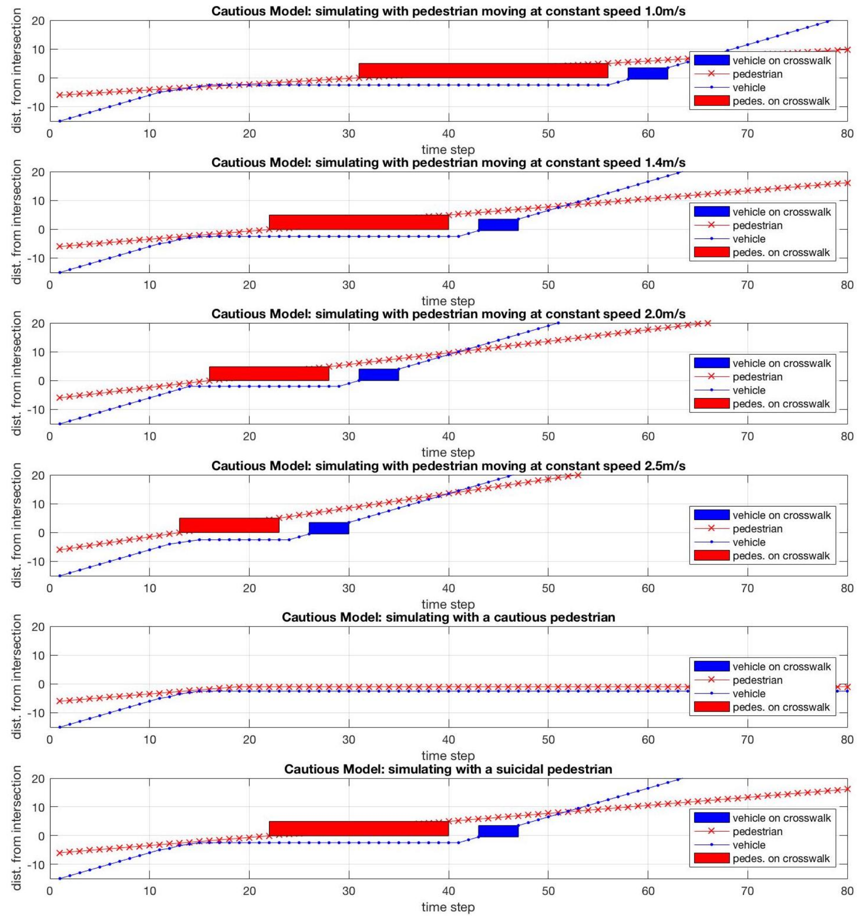

Figure 6.2: Cautious pedestrian-based MDP model: We construct a cautious pedestrian-based MDP model and examine the autonomous vehicle's behaviour by simulating the model with the same set of experimental conditions as in Figure 6.1.

The trajectories of both pedestrian (red) and autonomous vehicle (blue) is shown in Figs. 6.1 and 6.2. The horizontal axis is the time step (in units of $0.2\,\mathrm{s}$), and the vertical axis is the distance from the crosswalk (in meters). The slope of the lines represents the velocity. The boxes are a representation of the crosswalk position. As the crosswalk is $4\,\mathrm{m}$ wide and $5\,\mathrm{m}$ long, the blue and red boxes have different heights.

6.3.1 Vehicle's speed profile

The vehicle's characteristics are portrayed directly in its speed profile. The vehicle employing a model of a reckless pedestrian is more conservative, which is what one would expect: an optimal agent with a pessimistic model must be more risk-averse. This is reflected in the vehicle's trajectory— Figure 6.1 has shallow slopes, meaning that it slows down over a long period of time, initiating breaking sooner. On the other hand, the vehicle trajectory for the MDP solved under the assumption of a cautious pedestrian has steeper slopes and must perform a hard stop in some cases. For example, the second graph in Figure 6.2 shows a rapid stop being executed since the vehicle comes to a complete stop in a short time (around time steps 10–20). The vehicle's model assumes that the pedestrian will be cautious and can thus be counted on to stop before crossing, but the third graph is a pedestrian traveling at constant speed $1.4\,\mathrm{m/s}$ and not stopping, so the vehicle was forced to brake hard to avoid a collision.

6.3.2 Collision occurrences

The MDP is expected to provide a safe (collision-free) action choice in every state for the autonomous vehicle. We gave a large negative reward R_{col} in order to prevent the pedestrian and vehicle both being on the crosswalk simultaneously. Thus, as the boxes represent the crosswalk area for both pedestrian and vehicle, they should never overlap. By observing the distance between the boxes, we can see how conservative the vehicle is. When we lower the collision penalty, the distance reduces as the vehicle is more willing to risk colliding versus waiting for longer to lower the risk.

6.3.3 Unresolved turn taking

The magnitude of the penalty for collisions also affects the possibility of the vehicle and pedestrian approaching the crossing with their respective turn-taking still unresolved. In both Figure 6.1 and 6.2, the fifth plot shows an unresolved turn-taking situation. The definition of a cautious pedestrian is to wait for the vehicle to cross when it cannot reach the other side whilst simply maintaining the same speed. For instance, the pedestrian in both simulations had stopped to wait for the vehicle to cross first. Whereas our penalty for collision is large (at -1000) so the vehicle would wait for the pedestrian to cross first to prevent a collision. When the collision penalty is reduced to -160, this deadlock is resolved. The first plots in both Figure 6.3 and 6.4 safely control the vehicle (no

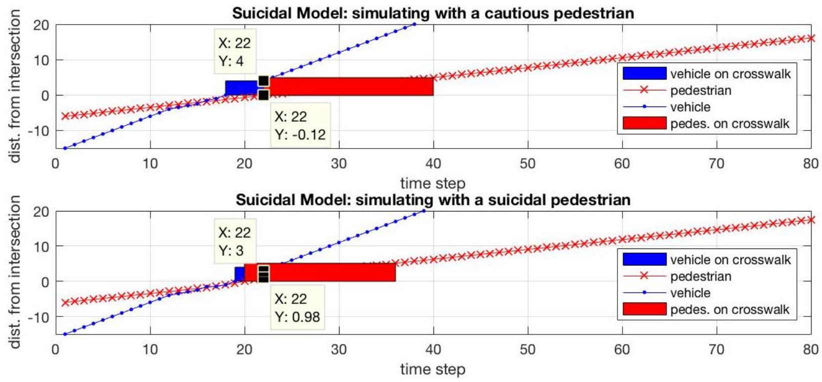

Figure 6.3: Reckless pedestrian-based MDP model with collision penalty as -160: the upper plot shows that lowering the collision penalty solves the unresolved turn taking issue. However, once the penalty lowers, the modelled vehicle collides with a reckless pedestrian, as shown in the lower plot.

boxes overlap) when there is a cautious pedestrian. Unfortunately, lowering the collision penalty now causes the vehicle to be on the crosswalk at the same time (boxes now overlap) as a reckless pedestrian.

6.3.4 Accumulated rewards

Accumulated rewards can be used to compare the performance between different MDP models. Given the same pedestrian behaviour as the simulation input, the model with the highest accumulated rewards is most suitable for purposes where the design matches real circumstances. For example, Table 6.1 shows that the reckless pedestrian-based model has a higher reward for simulations that have an interaction between the vehicle and pedestrian. This result fits with our goal of decreasing the amount of time the vehicle moves slowly. According to our reward definition, every time the vehicle stops, it gets a -10 penalty. Therefore, the simulations that have a shorter vehicle stopping period would have a higher accumulated reward.

6.4 Discussion

We have examined a decision-theoretic model for the interaction between a pedestrian and a vehicle. With the understanding of prior work conducted by psychologists examining similar experimental conditions, we were able to construct an MDP model that generates a sequence of

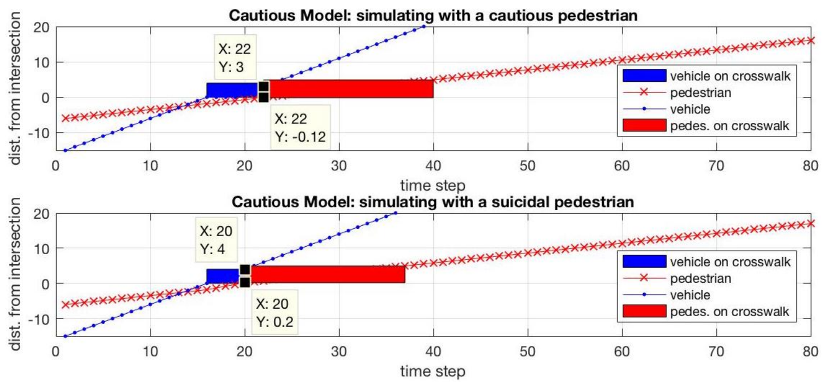

Figure 6.4: Cautious pedestrian-based MDP model with collision penalty as -160: The lowered collision penalty solves the unresolved turn taking problem, as shown in the upper plot. But lowering the collision penalty also causes the modelled vehicle to collide with a reckless pedestrian as shown in the lower plot.

Table 6.1: Accumulated rewards for each simulation

Accumulated Rewards						
	Const. Speed $1.0\,\mathrm{m/s}$	Const. Speed $1.4\,\mathrm{m/s}$	Const. Speed $2.0\,\mathrm{m/s}$	Const. Speed $2.5\,\mathrm{m/s}$	Cautious Pedestrian	Reckless Pedestrian
Reckless Pedestrian Based MDP ($R_{col} = -1000$)	165	485	765	865	-695	545
Cautious Pedestrian Based MDP ($R_{col} = -1000$)	900	455	690	795	-725	455

actions for the vehicle to execute such that it mimics a driver-pedestrian interaction at an unsignalized intersection. Furthermore, we were able to establish empirically that the crossing behaviour of the autonomous vehicle is safe and efficient. However, in our MDP approach, we assume the vehicle anticipates encountering a particular type of pedestrian crossing behaviour by encoding the behaviour inside the model. This raises two major issues for physical implementation when using our MDP approach: 1) the type of pedestrian crossing behaviour cannot be observed beforehand; therefore, it is difficult to encode the behaviour while modeling the MDP; 2) we can only divide types of pedestrian crossing behaviour so much that the uncertainty caused by neglecting the various influential factors is challenging to maintain. In the following chapter, we will build on the current understanding of the crossing interaction dynamics and propose an approach that aims to provide a solution to the above issues.

7. PARTIALLY OBSERVABLE MARKOV DECISION PROCESS DESIGN

There are various factors involved in pedestrian decision-making process at the time of crossing. Though current technologies encourage the installation of ever-richer sensors on robots, factors such as age, gender, culture, faith, and past experiences [43] are unlikely to be precisely sensed any time soon. Thus, it is crucial to have autonomous vehicles capable of balancing these unobservable factors as a type of uncertainty while crossing through an intersection safely and efficiently; and, if necessary, resolve the uncertainty through performing informative actions.

We solve the problem as a POMDP via a state-of-the-art online solver. Our POMDP model considers a state space $\mathbb{S}$ comprising states $S = \{S^H, S^C\}$, where S^H represents information about the pedestrian, including their position (H_{pos}), velocity (H_{vel}), characteristic (H_{chr}) and crossing intention (ξ); S^C represents information of the vehicle, including vehicle position (C_{pos}) and velocity (C_{vel}).

Next, we present the details of how we approach the abstract concept of human intention and capture the notion of beliefs and agreement over the crossing order and how we connect the pedestrian and vehicle's physical transitions, discussed in the previous MDP section, with the remaining pieces to construct the final POMDP model.

7.1 Mental States: Crossing Order

Let us denote the binary variable encoding crossing order at time t as ξ_t. We define $\xi_t \in \{0, 1\}$, where $\xi_t = 0$ means the pedestrian crosses first and $\xi_t = 1$ means the vehicle crosses first. The dynamics of ξ_t are based on domain knowledge (i.e., a time-gap–based decision), detailed in Section 7.2.1.

7.2 Pedestrian Dynamics

We restrict ourselves to a consideration of very basic motion: the pedestrian can either move along the crosswalk or pause. We will assume that the pedestrian can move at any reasonable speed, but, as clarified shortly, we treat speed in a particular way. As defined in the basic dynamic model section 5.1, we once again construct a Markov chain such that each node is a representation of the distance (discretized) from the crosswalk, and the transition probability between each physical state is calculated based on the speed the pedestrian is traveling. However, the speed is defined through a more complex model that takes crossing intentions (ξ_t) into consideration.

7.2.1 Determination of crossing intention (ξ_t)

Summing up the studies in Section 2.1, the influences on pedestrian crossing decision-making comprise two main factors: (i) Contextual factors include the position/velocity of the vehicle and the location of the crosswalk; (ii) Habitual factors include the pedestrian's traits and personal characteristics, like age and gender.

For contextual factors, we condense them into a notion of the 'level of peril' of the current world state. The level of peril is computed based on the time difference between the remaining time of the vehicle's arrival at the crossing point and the pedestrian's arrival at the crossing. The closer the vehicle appeals to be before the crossing while the pedestrian is about to cross, the higher level of peril (in other words: more dangerous) the pedestrian will feel to step off the curb and vice versa. The habitual factors determine how the pedestrian will act according to its sense of the level of peril. For example, a pedestrian who is more careful during crossings will tend to wait at the curb while the vehicle is seen to be beyond the 'average safe-crossing distance' from the intersection. Due to modeling purposes, we consider the extremes: a reckless ($H_{\text{chr}}^{\text{rkl}}$) and a cautious ($H_{\text{chr}}^{\text{cts}}$) pedestrian.

The final result is the making of a decision; we consider as having ξ_t take a value. Based on the factors described above, the dynamics of ξ_t can be expressed as

$$P(\xi_{t+1} = 0 |\, \xi_t = 0, S_t^{\mathrm{H}}, S_t^{\mathrm{C}}) = \begin{cases} 0.9 & \text{if vehicle slows down near the crosswalk} \\ & \text{and pedestrian is reckless} \\ 0.7 & \text{if vehicle slows down dramatically} \\ 0.5 & \text{if pedestrian is reckless} \\ 0.3 & \text{if vehicle slows down from far} \\ 1.0 & \text{otherwise,} \end{cases} \tag{7.1}$$

and

$$P(\xi_{t+1} = 1 |\, \xi_t = 1, S_t^{\mathrm{H}}, S_t^{\mathrm{C}}) = \begin{cases} 0.9 & \text{if vehicle speeds up near the crosswalk} \\ & \text{and pedestrian is cautious} \\ 0.5 & \text{if vehicle speeds up from far} \\ 1.0 & \text{otherwise,} \end{cases} \tag{7.2}$$

which here a reckless (or cautious) pedestrian is just one that has $H_{\mathrm{chr}} = H_{\mathrm{chr}}^{\mathrm{rkl}}$ (or $H_{\mathrm{chr}}^{\mathrm{cts}}$, respectively).

(The numbers above constitute examples, reflecting the general idea, if not the exact quantities used in our experiments. Precise values would be determined via psychological experiments and collection of data.)

7.2.2 Pedestrian locomotion

The pedestrian's motion depends on whether they currently intend to cross first or second. This is, of course, precisely the information in ξ_t. Hence the motion can be clearly defined into two cases expressed with functions $f_{\xi_t} : \mathbb{S} \to \mathbb{R}$ yielding velocities. The following definition is similar to Equation (6.1) and (6.2), but this time we define the intention as what determines pedestrian motion and pedestrian characteristics as part of an element that influences the intention (as mentioned in the previous section, Section 7.2.1).

7.2.2.1 *Pedestrian crosses first ($\xi_t = 0$)*

When the pedestrian decides to cross the intersection before the vehicle does, the pedestrian will attempt to travel at some speed to ensure it will cross first. We enforce some basic constraints: should the pedestrian reach the fastest walking speed, $2.5\,\mathrm{m/s}$, it remains moving at the highest speed it is capable of maintaining. If the speed needed to arrive in time is below average walking speed $1.4\,\mathrm{m/s}$, the pedestrian continues at a normal pace. Quantitatively, this is

$$f_0(\cdot) = \begin{cases} 1.4 & \text{if } o_{\Delta t} > 2 \\ 2.5 \times e^{-0.289 o_{\Delta t}} & \text{if } o_{\Delta t} \leq 2 \end{cases}, \tag{7.3}$$

where $o_{\Delta t}$, computed from $S_t \in \mathbb{S}$, is the time difference between the remaining time for the vehicle to arrive at the intersection and the remaining time for the pedestrian to finish crossing the intersection.

7.2.2.2 *Vehicle crosses first ($\xi_t = 1$)*

In this case, pedestrians will stop at the curb and wait for the vehicle to pass when they determine that, continuing at their current speed, they cannot reach the other side of the road before the vehicle arrives. Their velocity is given as

$$f_1(\cdot) = \begin{cases} 1.4 & \text{if } o_{\Delta d} < -1 \text{ and } o_{\Delta d} > 1 \\ 0 & \text{if } -1 \leq o_{\Delta d} \leq 1 \text{ and can not cross at } 1.4\,\mathrm{m/s} \\ 1.4 & \text{if } -1 \leq o_{\Delta d} \leq 1 \text{ and can cross at } 1.4\,\mathrm{m/s} \end{cases}, \tag{7.4}$$

where $o_{\Delta d}$, computed from $S_t \in \mathbb{S}$, represents the distance between the pedestrian and crosswalk. Once the pedestrian starts crossing the crosswalk, $o_{\Delta d}$ becomes a negative distance in our representation.

7.3 Vehicle Dynamics

The vehicle, unlike the pedestrian, has actions that we wish to determine. Hence, we model the vehicle's controls as actions of a decision process. The vehicle needs to avoid collision with the pedestrian, whose crossing behaviour is not perfectly known. The vehicle must deal with two forms of uncertainty: partial observability and stochasticity. By choosing actions, the vehicle seeks an optimal strategy through reasoning about the pedestrian's behaviour as expressed in the stochastic model.

7.3.1 Vehicle's motion model

Let C_{pos} be the state that represents the vehicle's distance from the crosswalk and state C_{vel} represent the vehicle's velocity. The evolving physical state of the vehicle is specified as ($C_{\text{pos},t}$, $C_{\text{vel},t}$) at time t. The vehicle is constrained to move in a fixed direction towards the crosswalk and its control is based on acceleration $a_t \in \{a_{\text{dec}}, 0.0, a_{\text{inc}}\}$, where $a_{\text{dec}} < 0.0$ and $a_{\text{inc}} > 0.0$. Given a_t, the new state of the vehicle is calculated as

$$\begin{aligned} C_{\text{pos},(t+1)} &= C_{\text{pos},t} + C_{\text{vel},t}, \\ C_{\text{vel},(t+1)} &= C_{\text{vel},t} + a_t. \end{aligned} \tag{7.5}$$

7.3.2 Vehicle-pedestrian interaction

The interactions between vehicle and pedestrian near the crossing point are embedded into transition functions. When the vehicle and the pedestrian are far from the crossing, they transition to their next state based on their individual dynamics. However, as modelled in Section 7.2.1, the pedestrian's crossing behaviour considers the vehicle position and velocity. Once the pedestrian is near the crosswalk, the behaviour of both the vehicle and the pedestrian are now tightly coupled: both their state transition probabilities are influenced by not only the vehicle's state but its action as well.

7.3.3 Sensors and observations

We assume that the vehicle is equipped with sensors capable of detecting the pedestrian and reporting his/her distance relative to the start of the pedestrian's crosswalk and velocity. These sensors produce data that has an error range which decreases as the vehicle gets closer to the pedestrian. Additionally, the vehicle is assumed to have sensors that return a (near) accurate value of the vehicle's velocity and distance to the start of the intersection. Taken together, this sensing equipment generates observations that we represent as a 4-tuple: $(H_{\text{pos}}, H_{\text{vel}}, C_{\text{pos}}, C_{\text{vel}})$.

7.3.4 Rewards

The reward model is simple: the primary objective of the vehicle is to minimize the risk of colliding with the pedestrian. Consequently, we assign a large penalty when both the vehicle and the pedestrian are on the crosswalk simultaneously. Additionally, to incentivize efficiency, the vehicle receives rewards for those states with a higher velocity. This reward model is similar to the one used in the Markov decision process modeling section. We emphasize that the vehicle is not specifically rewarded for knowing things about the pedestrian; any information of value is potentially a cause for the vehicle's action choices as it has implications for safe or/and efficient motion indirectly.

7.3.5 The vehicle's perspective on the crossing order

Unlike the pedestrian, who has a state ξ_t to represent whom he/she considers to be crossing first, the vehicle has no such explicit state. Instead, the POMDP maintains a distribution over the entire state space, i.e., a belief state. When all dimensions of the state other than ξ_t are marginalized out, what remains is a probability that represents the vehicle's estimate of the pedestrian's conception.

8. PARTIALLY OBSERVABLE MARKOV DECISION PROCESS EXPERIMENTS AND RESULTS

In this chapter, we show results from our implementation of the solver running on an autonomous vehicle and also summarize an extensive and carefully controlled evaluation conducted in a separately constructed simulation. The performance of the simulated autonomous vehicle is used to discuss and analyze the resulting behaviour in terms of the overall safety and further, for non-observable states, the dynamics of the vehicle's belief. Finally, we briefly discuss experiments where the vehicle may also generate actions that are explicitly communicative.

8.1 Experimental Setup and Details

We construct a continuous world describing a crossing scenario and employ an online POMDP solver that uses a belief-tree–based approach in which sampled scenarios produce nodes that are connected via edges to produce approximate policies, DESPOT [44, 45], to create a safe and efficient crossing policy for the vehicle. Both the simulator and solver are connected through the Robot Operating System (ROS) [46]. We implemented them as ROS nodes with inter-process communication handled by having them subscribed to one another.

8.1.1 Experimental flow

We treat the crossing scenario depicted in Figure 4.1. Each experimental trial begins with both the vehicle and the pedestrian in the simulator moving steadily towards the crossing. A trial concludes when the vehicle finishes crossing from one side to the other. The POMDP solver starts once the vehicle is $14\,\mathrm{m}$ or less away from the crosswalk. For every execution step, the solver considers the current state of the world, including both the vehicle and the pedestrian information, and outputs an acceleration value. The vehicle enacts the new acceleration and the aspects pertaining to the pedestrian evolve as per the transition model detailed earlier. The solver continues to output new acceleration values based on new inputs until the trial finishes.

8.1.2 Experiment parameters

To deliberately constructed the scenario which there is substantial ambiguity in crossing order, we position the pedestrian to start walking at a normal $1.4\,\mathrm{m/s}$ speed from $4.2\,\mathrm{m}$ before the crosswalk, and the vehicle starts $14\,\mathrm{m}$ away from the intersection moving at $3\,\mathrm{m/s}$ speed. The crosswalk is $4\,\mathrm{m}$ wide and $5\,\mathrm{m}$ long.

For the DESPOT solver, we used 500 sampled scenarios with the maximum depth of the belief tree as 100, and the discount factor set to 0.98. The solver was given $1\,\mathrm{s}$ to construct the search tree and choose an action. The experiments were executed on an Intel Core i7-6670HQ $2.6\,\mathrm{GHz}$ processor with $32\,\mathrm{GB}$ of RAM running Ubuntu 16.04.

8.2 Autonomous Vehicle Implementation

8.2.1 Autonomous vehicle experiment setup

We carried out experiments at our university autonomous vehicle testing ground. A virtual crossing was set up in the system that is $5.0\,\mathrm{m}\times4.0\,\mathrm{m}$ overseen by the camera mounted on a lamp post on the side of a roadway. (See Figure 8.1.)

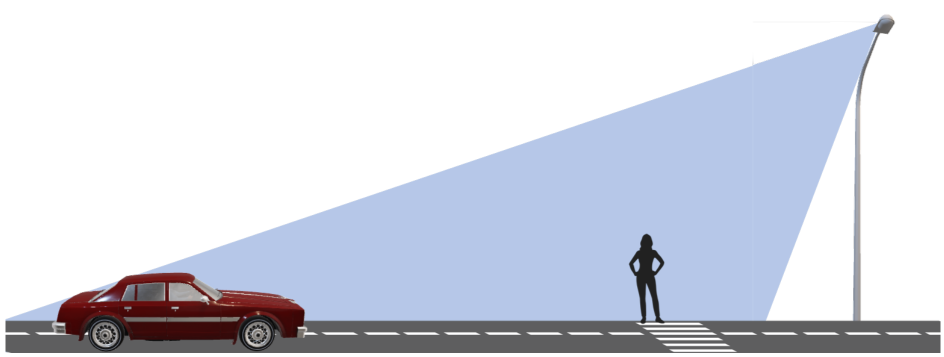

Figure 8.1: Cartoon depicting the experimental infrastructure.

Our test vehicle is a Ford Lincoln MKZ with auto-driving enabled. The auto-driving system is capable of following a pre-recorded path with GPS and vehicle orientation included by sending commands through the Robot Operating System (ROS) to the low-level vehicle controller to control the throttle, brake, and steering.

Our pedestrian is a manikin mounted on a pole installed on a remote control car. It is operated by people intending to either cross before the vehicle or after the vehicle has passed the crosswalk.

The sensors we use include a camera mounted on a lamp pole overseeing the crossing area and an RTK GNSS receiver, Piksi. With infrastructure enabled autonomy [47], we were able to have both sensors coordinated to provide the position of the vehicle and pedestrian relative from the base station. The velocity of the pedestrian is then calculated with a first-order filter. With the information gathered from the sensors, the planner uses them as observations to locate its current state in the belief state space and then outputs an acceleration value as the action for the vehicle to execute. (See Figure 8.2.)

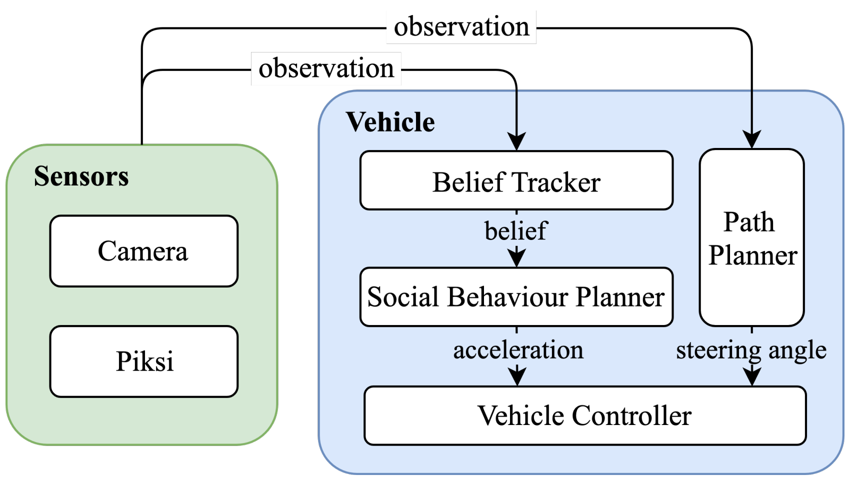

Figure 8.2: System architecture employed for the Lincoln MKZ.

Both pedestrian and vehicle's motion and sensor readings are imperfect owing to factors such as friction, bumps on the road, wind, sunlight, *etc.*, in the world affecting both agents. Seeking to balance between decision quality and computational expediency, observations of the pedestrian information are discretized (with resolution of $1.0\,\mathrm{m}$ for position and $1.0\,\mathrm{m/s}$ for speed) and the vehicle information is discretized (with resolution of $1.5\,\mathrm{m}$ for position and $1.0\,\mathrm{m/s}$ for speed).

The actions are sparse: $0.5\,\mathrm{m/s^2}$, $-1.0\,\mathrm{m/s^2}$, and $-2.0\,\mathrm{m/s^2}$. The planning horizon is 100 steps. The maximum planning time per step is $0.2\,\mathrm{s}$.

8.2.2 Results from the autonomous vehicle experiment

We present illustrative instances of the autonomous vehicle experiment results in Figure 8.3 and 8.4.

The vehicle is positioned at least $100.0\,\mathrm{m}$ away from the start of the crossing and approaches the crossing at a speed of $7.0\,\mathrm{m/s}$. Once the vehicle is as close as $70.0\,\mathrm{m}$, it receives input from the camera sensors for pedestrian detection. If the pedestrian is detected and determined to be approaching the crossing, the vehicle slows down to $3.0\,\mathrm{m/s}$; otherwise, it continues at the same speed, proceeding to approach and cross the crossing. Once the vehicle slows down and is within $14.0\,\mathrm{m}$ before the crossing, it activates the behaviour planner to interact with the pedestrian and begins to maintain a belief distribution over the pedestrian's crossing intention.

In the case of interacting with a reckless pedestrian, the vehicle slows down in advance (Figure 8.3a) with an initial belief that the pedestrian will cross first. While slowing down (Figure 8.3b), the vehicle's belief in the pedestrian being reckless increases which leads to it later picking up speed (Figure 8.3c) and crosses (Figure 8.3d).

As for the cautious pedestrian case, the vehicle slows down due to the initial belief the pedestrian intends to cross first (Figure 8.4a). Before it comes to a stop, the vehicle changes its belief distribution over the pedestrian's characteristic and intention as it now observes the pedestrian to be slowing down before the crossing (Figure 8.4b). The vehicle's acceleration changes from decelerating to maintaining speed as the majority of the weighting of belief shifts toward the pedestrian being cautious and who, thus, intends to let the vehicle cross first (Figure 8.4c). Then, passing by the cautious pedestrian slowly, the vehicle gradually gains speed (Figure 8.4d). Finally, the vehicle accelerates once past the pedestrian (Figure 8.4e).

These examples from our autonomous vehicle trials indicate that the planner produces effective crossing behaviour and that it performs well at managing uncertainty for states of the pedestrian that are not directly observable. The dynamic setting means it is impossible to repeat trials with

(a) (b) (c) (d)

Figure 8.3: The vehicle encounters a reckless pedestrian who speeds up to start crossing before vehicle arrives. Histogram indicates beliefs over pedestrian crossing intentions and characteristic: blue for *pedestrian crosses first*, green for *vehicle crosses first*, yellow for a pedestrian who is characteristically *reckless*, red for a *cautious* one. The length of the arrow above the vehicle expresses vehicle's velocity, in which, a circle indicates that the velocity is approximately zero. The color of the arrow describes the acceleration value: green for *accelerate*, yellow for *maintain*, and red for *decelerate*.

identical inputs, making it challenging to quantify performance or to interrogate particular aspects of the resulting behaviour. Consequently, we conducted a simulation study to probe the planner's behaviour in detail.

8.3 Simulated Experiments

We developed a custom simulator to model the crossing setting of Figure 4.1. It simulates pedestrian motions using the pedestrian crossing behaviour model described in Section 7.2. The vehicle is simulated to drive towards the crossing point its speed changing according to acceleration values produced as actions by DESPOT.

(a) (b) (c) (d) (e)

Figure 8.4: The vehicle encounters a cautious pedestrian who stops to wait for vehicle to cross. The histogram and arrow representations are the same design as used in Figure 8.3.

8.3.1 Simulation setup

The vehicle and pedestrian's position state space is generated by discretizing continuous space into intervals of 1.75 and 0.75. The velocity state for the vehicle contains values from 0.0 to 3.0 with an interval of 1.0. The pedestrian has three velocity states: $\{0.0, 1.4, 2.5\}$. We define the action space for of the POMDP model as $\{-1.0, 0.0, 0.5\}$, where each is an acceleration value that can be executed by the vehicle.

Each simulation trial begins with both the vehicle and the pedestrian in the simulator moving steadily towards the crossing. When the vehicle is $14.0\,\mathrm{m}$ or less away from the crosswalk, the simulator node sends a message to start the POMDP solver. For every execution step, the solver reads the current state of the simulator, including both the vehicle and the pedestrian information, and outputs an acceleration value. The simulator will transition to the next world state as the vehicle transitions based on the generated acceleration value as input and the pedestrian transitions based on its crossing behaviour model. The simulator continues to subscribe for new acceleration values until the trial finishes.

Pedestrian transitions at each step consider the vehicle's state as well as the pedestrian's characteristic. The vehicle's state represents contextual factors that influence the pedestrian's crossing behaviour, and the characteristic, which is set manually at the start of the simulation, represents the habitual factors. Both are needed to simulate the dynamics, which is achieved by sampling in proportion to the associated probabilities.

8.3.2 Analysis simulation results

We present the simulation results via detailed plots of a variety of variables as they evolve in time. Figure 8.5 and 8.6 depict runs of a cautious and a reckless pedestrian, respectively. The position of both pedestrian and vehicle are shown in the graph with the title 'continuous position.' The vertical axis of the graph is the distance from the crossing, where negative values represent positions that are before the crossing point.

8.3.2.1 Crossing safely

We can see that whether interacting with a reckless or a cautious pedestrian, the lines for the vehicle and the pedestrian positions are never seen to be between the crossing region, 0.0 to 4.0, simultaneously. This indicates that no collision occurs.

8.3.2.2 Beliefs over non-observable states

In our scenario, the key to communication is the inference of behaviour, which resolves to a question about the pedestrian's crossing decision (ξ_t). ξ_t is calculated based on the perilousness of the crossing for the pedestrian and also their habitual characteristics (H_{chr}). Since neither this characteristic nor ξ_t are observable, the vehicle's knowledge of these two elements is understood in terms of the belief state (or distribution) over both variables. The third plot in the results figures

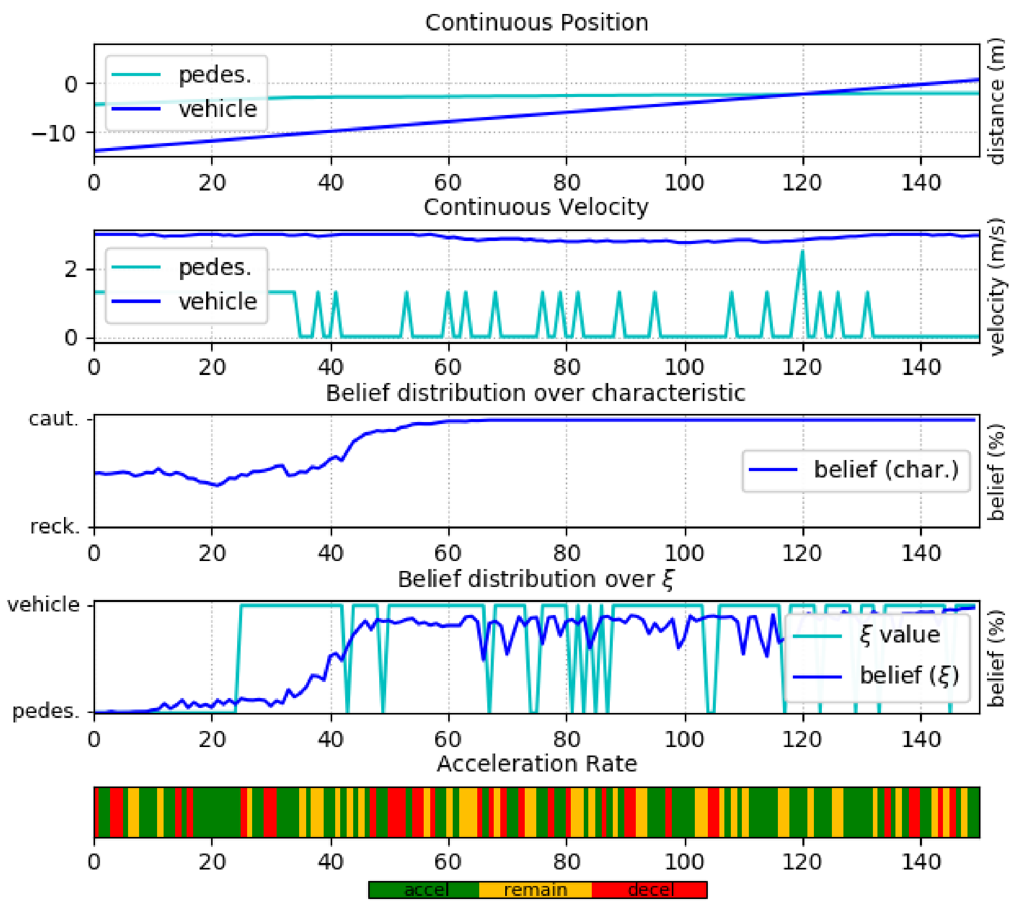

Figure 8.5: Simulation results showing a vehicle executing a policy, interacting with a cautious pedestrian.

shows how the belief of the characteristic converges to the correct trait. As for the belief distribution of ξ_t, it appears along with the actual pedestrian's ξ_t value for comparison in the fourth plot of the result figures. Notice that ξ_t changes, but the vehicle's belief distribution shows to track the changes in the pedestrian's ξ_t.

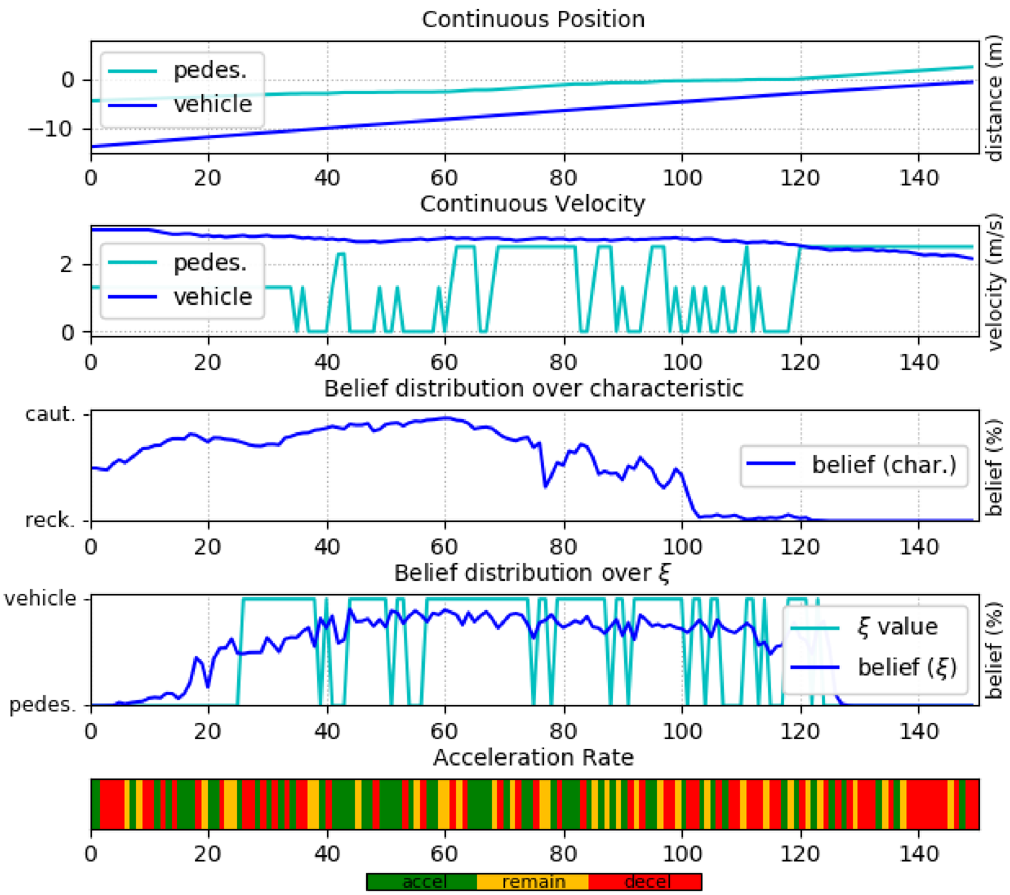

Figure 8.6: Simulation results showing a vehicle executing a policy, interacting with a reckless pedestrian.

8.3.2.3 Implicit communication — an interpretation

To help comprehend the results, we chose to compare the behaviour of simulated autonomous vehicle with human drivers under circumstances where they are uncertain of the pedestrian's sense of who should cross. Figure 8.7 is a summary of the behaviour of our simulated vehicle. We have redrawn this figure from [1], using their style of summarization, along with some modifications to improve clarity, but with numbers reporting results from our experiments. The percentages and speed values are based on data from simulations interacting with a reckless pedestrian. (One example of the strategy is shown in Figure 8.6.)

Schneemann and Gohl [1] report that human drivers resolve ambiguous situations by initially reducing their speed, and then decide to whether to speed up or come to a stop depending on the pedestrian's response to their speed reduction. We see that the vehicle's strategy is less conservative compared to human drivers. Figure 8.7 can be interpreted as the vehicle trying to gain efficiency rewards but also balancing uncertainty. Instead of slowing down to passively learn the pedestrian's crossing order decision, the vehicle remains at moderately high speeds, seemingly expressing its desire to cross first. This communicates with the pedestrian and the pedestrian's following movement can be explained as a reply to the crossing arrangement. In Figure 8.5, the cautious pedestrian is shown to slow down, giving the vehicle permission to cross first. In Figure 8.6, the reckless pedestrian accelerated to express disagreement on the vehicle crossing arrangement. Both the vehicle and the pedestrian continue to adapt their maneuvers thereafter in order to reach agreement on crossing order.

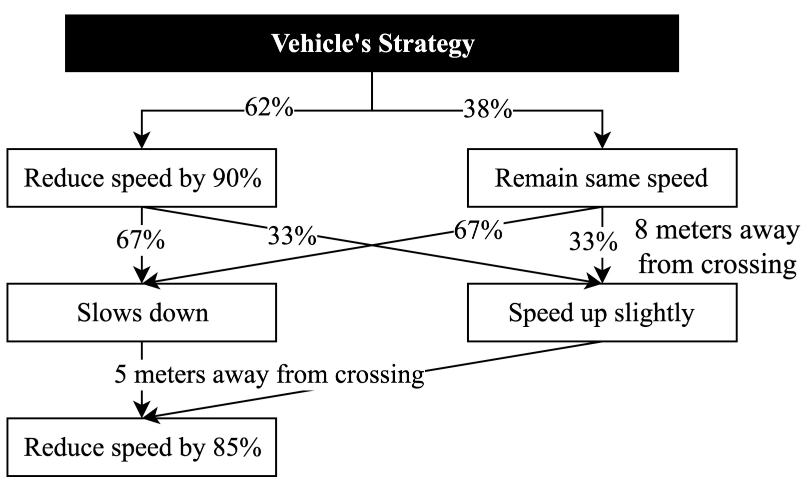

Figure 8.7: Strategies of the vehicle in ambiguous situations with a reckless pedestrian simulated for crossing (redrawn with modifications from [1]).

8.3.3 Explicit communication

We also conducted a simple experiment to analyze the value of communicating crossing order by creating an action that communicates ξ_t explicitly. In Figure 8.8, an action is added where the vehicle may flash its headlights. We model the pedestrian as understanding this action as indicating that the vehicle intends to let the pedestrian cross first. Additionally, to have the vehicle's policy be deliberate in choosing to communicate intent in establishing ξ_t, we penalize using the lights through modifying the reward model.

We conducted the experiment with a reckless pedestrian. The result, in Figure 8.8, shows that the vehicle opts to flash its lights (quite frequently) despite the negative reward incurred. Moreover, as we compare the third graph in Figure 8.6 and 8.8, it is clear that knowledge of the pedestrian's characteristic is recognized faster with explicit communication. The fourth graph in both figures shows that ξ_t also stabilizes sooner. And, as the ambiguity is resolved, the result is that both the vehicle and pedestrian cross the crosswalk more efficiently.

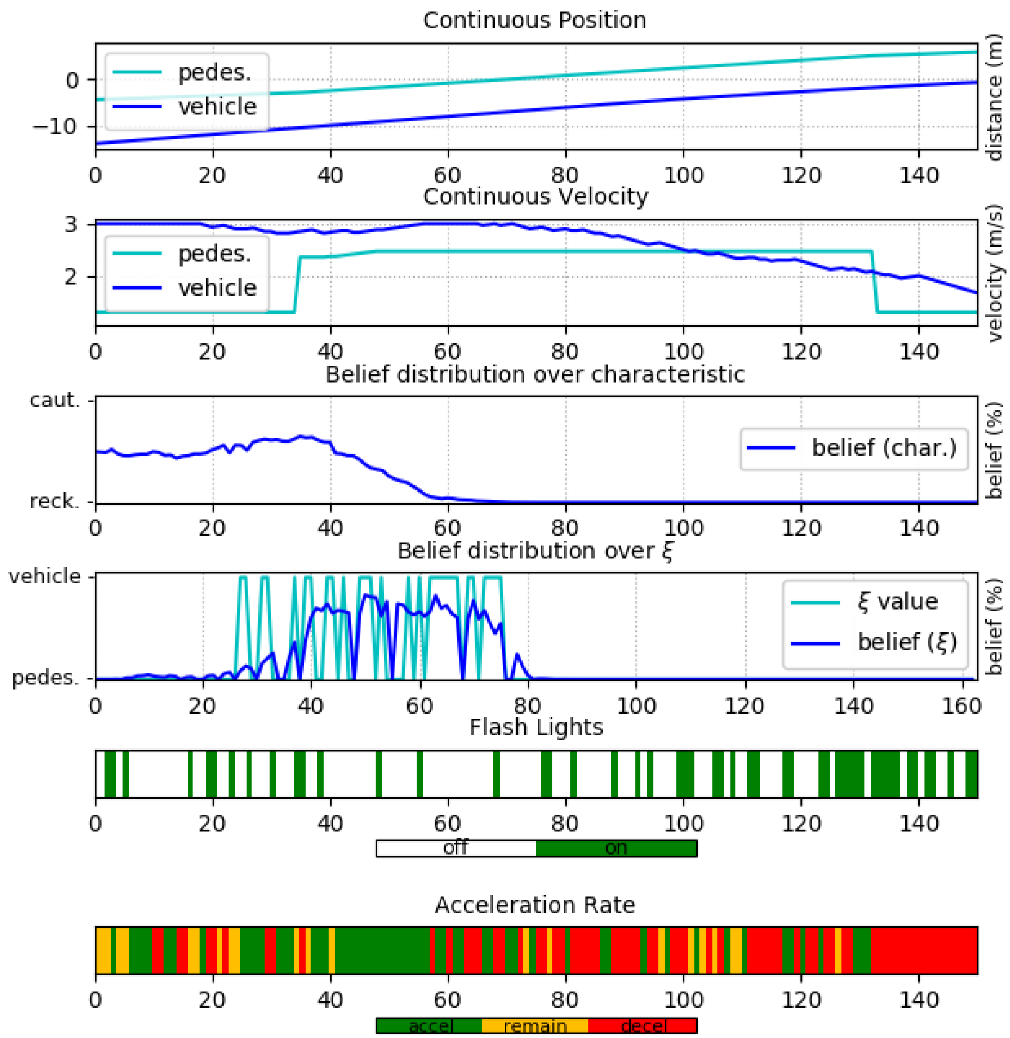

Figure 8.8: A reckless pedestrian interacts with a vehicle equipped to flash its headlights, communicating explicitly.

9. CONCLUSION

Social interaction is valuable in resolving ambiguity in traffic; it is necessary for autonomous vehicles if they are to operate harmoniously within the existing system, infrastructure, and norms. In this work, we proposed an MDP model to explore how velocity-based signaling affects vehicle-pedestrian interaction in a simple setting under a decision-theoretic model. The model shows to result in safe interact when pedestrians behave similarly to the model assumed within the MDP. Moreover, we planned under uncertainty arising from non-determinism and partial observability and examined how an uncertainty-aware planner can help an autonomous vehicle interact competently.

We consider the work's main contributions to be:

- *Reproduce crossing interaction via solving a Markov decision process:* We modelled the vehicle to perform socially at an unsignalized intersection based on reward functions describing the importance of safe and efficient crossing. Our results show that the vehicle is capable of selecting the sequence of actions that performs an efficient collision-free crossing. We can even observe how conservative the vehicle is by analyzing how long the vehicle waited to start crossing before or after the pedestrian crosses.

- *A minimal model of social ambiguity:* Though the sophisticated human social behaviour is challenging to quantify, in our scenario, we boiled several fairly complex and abstract concepts down to a single source of uncertainty, in a sort of concise, 'lumped parameter' model via the binary variable we term the pedestrian's crossing intention. Despite there being a collection of factors, our approach gives a single expression that can be interpreted in terms of probability.

- *Framing a practically solvable partially observable decision problem:* Exploiting the conciseness of the representation means the vehicle needs to maintain a low-dimensional distribution, making it practical to solve for a sequence of actions. Moreover, these actions include ones that elicit information and manage uncertainty. Initially, it was far from clear how causally cyclic nature of communication would be accommodated. The model we present expresses this satisfactorily, albeit indirectly via transition dynamics.

- *Implementation on autonomous vehicle:* We illustrate the feasibility of the approach by conducting a demonstration on a real vehicle. Our experiment results show a vehicle capable of resolving uncertainty in order to achieve efficiency. An examination of the vehicle's behaviour suggests that the strategy is less conservative than some driving behaviour, including some humans while trying to resolve ambiguous situations.

Our broader philosophy is that several aspects of social interaction are means for coping with uncertainty so that representing uncertainty explicitly and dealing with it efficiently, yields robots that are socially effective. Much further research remains to be conducted to better understand and realize social behaviour in vehicles.

REFERENCES

[1] F. Schneemann and I. Gohl, "Analyzing driver-pedestrian interaction at crosswalks: A contribution to autonomous driving in urban environments," in *2016 IEEE intelligent vehicles symposium (IV)*, pp. 38–43, IEEE, 2016.

[2] M. Richtel and C. Dougherty, "Google's driverless cars run into problem: Cars with drivers." https://www.nytimes.com/2015/09/02/technology/personaltech/google-says-its-not-the-driverless-cars-fault-its-other-drivers.html?_r=2, 2015. Accessed: 2019-07-14.

[3] S. E. Anthony, "The trollable self-driving car." https://slate.com/technology/2016/03/google-self-driving-cars-lack-a-humans-intuition-for-what-other-drivers-will-do.html. Accessed: 2019-07-14.

[4] N. Guéguen, S. Meineri, and C. Eyssartier, "A pedestrian's stare and drivers' stopping behavior: A field experiment at the pedestrian crossing," *Safety Science*, vol. 75, pp. 87–89, 2015.

[5] A. Varhelyi, "Drivers' speed behaviour at a zebra crossing: a case study," *Accident Anal. & Prev.*, vol. 30, no. 6, pp. 731–743, 1998.

[6] D. Rothenbücher, J. Li, D. Sirkin, B. Mok, and W. Ju, "Ghost driver: A field study investigating the interaction between pedestrians and driverless vehicles," in *2016 25th IEEE international symposium on robot and human interactive communication (RO-MAN)*, pp. 795–802, IEEE, 2016.

[7] B. Schoettle and M. Sivak, "A survey of public opinion about autonomous and self-driving vehicles in the us, the uk, and australia," tech. rep., University of Michigan, Ann Arbor, Transportation Research Institute, 2014.

[8] R. L. Moore, “Pedestrian choice and judgment,” *Journal of the Operational Research Society*, vol. 4, no. 1, pp. 3–10, 1953.

[9] W. Reilly, “Highway capacity manual 2000,” *Tr News*, no. 193, 1997.

[10] J. Cohen, E. Dearnaley, and C. Hansel, “The risk taken in crossing a road,” *Journal of the Operational Research Society*, vol. 6, no. 3, pp. 120–128, 1955.

[11] M. A. Brewer, K. Fitzpatrick, J. A. Whitacre, and D. Lord, “Exploration of pedestrian gap-acceptance behavior at selected locations,” *Transportation research record*, vol. 1982, no. 1, pp. 132–140, 2006.

[12] W. A. Harrell, “Factors influencing pedestrian cautiousness in crossing streets,” *The Journal of Social Psychology*, vol. 131, no. 3, pp. 367–372, 1991.

[13] F. M. Khan, M. Jawaid, H. Chotani, and S. Luby, “Pedestrian environment and behavior in karachi, pakistan,” *Accident Analysis & Prevention*, vol. 31, no. 4, pp. 335–339, 1999.

[14] G. Tiwari, S. Bangdiwala, A. Saraswat, and S. Gaurav, “Survival analysis: Pedestrian risk exposure at signalized intersections,” *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 10, no. 2, pp. 77–89, 2007.

[15] J. A. Oxley, E. Ihsen, B. N. Fildes, J. L. Charlton, and R. H. Day, “Crossing roads safely: an experimental study of age differences in gap selection by pedestrians,” *Accident Analysis & Prevention*, vol. 37, no. 5, pp. 962–971, 2005.

[16] J. Oxley, B. Fildes, E. Ihsen, J. Charlton, and R. Day, “Differences in traffic judgements between young and old adult pedestrians,” *Accident Analysis & Prevention*, vol. 29, no. 6, pp. 839–847, 1997.

[17] V. Himanen and R. Kulmala, “An application of logit models in analysing the behaviour of pedestrians and car drivers on pedestrian crossings,” *Accident Analysis & Prevention*, vol. 20, no. 3, pp. 187–197, 1988.

[18] B. J. Schroeder *et al.*, *A behavior-based methodology for evaluating pedestrian-vehicle interaction at crosswalks*. PhD dissertation, North Carolina State University, Raleigh, NC, 2008.

[19] F. Bella and M. Silvestri, "Effects of safety measures on driver's speed behavior at pedestrian crossings," *Accident Analysis & Prevention*, vol. 83, pp. 111–124, 2015.

[20] S. J. Russell and P. Norvig, *Artificial intelligence: a modern approach*. Prentice Hall Press, 3rd ed., 2009.

[21] J. Cheng, H. Cheng, M. Q.-H. Meng, and H. Zhang, "Autonomous navigation by mobile robots in human environments: a survey," in *2018 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pp. 1981–1986, IEEE, 2018.

[22] D. Helbing and P. Molnar, "Social force model for pedestrian dynamics," *Physical Review E*, vol. 51, no. 5, p. 4282, 1995.

[23] G. Ferrer, A. Garrell, and A. Sanfeliu, "Robot companion: A social-force based approach with human awareness-navigation in crowded environments," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1688–1694, IEEE, 2013.

[24] Y. F. Chen, M. Everett, M. Liu, and J. P. How, "Socially aware motion planning with deep reinforcement learning," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1343–1350, IEEE, 2017.

[25] T. Bandyopadhyay, C. Z. Jie, D. Hsu, M. H. Ang, D. Rus, and E. Frazzoli, "Intention-aware pedestrian avoidance," in *Experimental Robotics*, pp. 963–977, Springer, 2013.

[26] F. Schneemann and P. Heinemann, "Context-based detection of pedestrian crossing intention for autonomous driving in urban environments," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2243–2248, IEEE, 2016.

[27] B. Okal and K. O. Arras, "Learning socially normative robot navigation behaviors with bayesian inverse reinforcement learning," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2889–2895, IEEE, 2016.

[28] S. Thrun, M. Montemerlo, H. Dahlkamp, D. Stavens, A. Aron, J. Diebel, P. Fong, J. Gale, M. Halpenny, G. Hoffmann, *et al.*, "Stanley: The robot that won the darpa grand challenge," *Journal of Field Robotics*, vol. 23, no. 9, pp. 661–692, 2006.

[29] C. Urmson, J. Anhalt, D. Bagnell, C. Baker, R. Bittner, M. Clark, J. Dolan, D. Duggins, T. Galatali, C. Geyer, *et al.*, "Autonomous driving in urban environments: Boss and the urban challenge," *Journal of Field Robotics*, vol. 25, no. 8, pp. 425–466, 2008.

[30] A. Gorrini, G. Vizzari, and S. Bandini, "Towards modelling pedestrian-vehicle interactions: Empirical study on urban unsignalized intersection," *arXiv preprint arXiv:1610.07892*, 2016.

[31] J. Van Den Berg, S. J. Guy, M. Lin, and D. Manocha, "Reciprocal n-body collision avoidance," in *Robotics research*, pp. 3–19, Springer, 2011.

[32] D. Vasquez, T. Fraichard, and C. Laugier, "Growing hidden markov models: An incremental tool for learning and predicting human and vehicle motion," *IJRR*, vol. 28, no. 11-12, pp. 1486–1506, 2009.

[33] G. Ferrer and A. Sanfeliu, "Bayesian human motion intentionality prediction in urban environments," *Pattern Recognition Letters*, vol. 44, pp. 134–140, 2014.

[34] H. Bai, S. Cai, N. Ye, D. Hsu, and W. S. Lee, "Intention-aware online POMDP planning for autonomous driving in a crowd," in *International Conference on Robotics and Automation (ICRA)*, pp. 454–460, 2015.

[35] M. Clamann, M. Aubert, and M. L. Cummings, "Evaluation of vehicle-to-pedestrian communication displays for autonomous vehicles," tech. rep., Transportation Research Board, Washington, DC, USA, 2017.

[36] T. Lagstrom and V. M. Lundgren, "Avip-autonomous vehicles interaction with pedestrians," *Master of Science Thesis, Chalmers University of Technology*, 2015.

[37] C. P. Urmson, I. J. Mahon, D. A. Dolgov, and J. Zhu, "Pedestrian Notifications." https://patents.google.com/patent/US8954252B1/en. Accessed: 2019-10-24.

[38] Mercede-Benz, “Mercedes-Benz F 015 Luxury in Motion.” https://www.mercedes-benz.com/en/mercedes-benz/innovation/research-vehicle-f-015-luxury-in-motion/. Accessed: 2019-10-24.

[39] Nissan Motor Corporation, “Nissan IDS Concept: Nissan’s vision for the future of EVs and autonomous driving.” http://www.nissan-global.com/EN/NEWS/2015/{_}STORY/151028-01-e.html{%}5Cnhttp://bit.ly/1PGMJmw. Accessed: 2019-10-24.

[40] R. D. Smallwood and E. J. Sondik, “The optimal control of partially observable markov processes over a finite horizon,” *Operations Research*, vol. 21, no. 5, pp. 1071–1088, 1973.

[41] R. C. Browning, E. A. Baker, J. A. Herron, and R. Kram, “Effects of obesity and sex on the energetic cost and preferred speed of walking,” *Journal of Applied Physiology*, vol. 100, no. 2, pp. 390–398, 2006.

[42] B. M. Nigg, B. R. MacIntosh, and J. Mester, *Biomechanics and biology of movement*. Human Kinetics, 2000.

[43] A. Rasouli and J. K. Tsotsos, “Autonomous vehicles that interact with pedestrians: A survey of theory and practice,” *IEEE Transactions on Intelligent Transportation Systems*, 2019.

[44] A. Somani, N. Ye, D. Hsu, and W. S. Lee, “Despot: Online pomdp planning with regularization,” in *Advances in Neural Information Processing Systems*, pp. 1772–1780, 2013.

[45] N. Ye, A. Somani, D. Hsu, and W. S. Lee, “DESPOT: Online POMDP planning with regularization,” *JAIR*, vol. 58, pp. 231–266, 2017.

[46] M. Quigley, J. Faust, T. Foote, and J. Leibs, “Ros: an open-source robot operating system,” *ICRA Workshop on Open Source Software*, vol. 3, no. 3.2, p. 5, 2009.

[47] S. Gopalswamy and S. Rathinam, “Infrastructure enabled autonomy: A distributed intelligence architecture for autonomous vehicles,” in *2018 IEEE Intelligent Vehicles Symposium (IV)*, pp. 986–992, IEEE, 2018.