# Causal Discovery and Inference for Autonomous-Agent Interactions

DPhil in Engineering Science


**Rhys Peter Matthew Howard**

**Lincoln College**

**University of Oxford**

*Supervised by Professor Lars Kunze and Professor Nick Hawes*

April 2025

## DECLARATION

This thesis is submitted to the Department of Engineering Science, University of Oxford, in fulfilment of the requirements for the degree of Doctor of Philosophy. I declare that this thesis is entirely my own work, and except where otherwise stated, describes my own research.

Rhys Peter Matthew Howard

Lincoln College

3rd September 2025

To my parents Dawn E Howard and Simon P Howard; my sister Bella M Howard;

my grandmothers Joan H Evans and Pamela Howard;

and my dear friends Kyle A R Allen-Taylor and Tom Enright.

Your unending support gave me the strength to keep going even when the going got tough.

## ABSTRACT

Everyday humans interact with others and generally understand how their actions impact them, even when not sharing a common goal. This is particularly crucial in high-risk domains where a failure to comprehend how to react or how others might react could result in injury or death. Thus it is important to develop such functionality for the autonomous agents that increasingly operate around people. This thesis focuses upon the expanding field of autonomous vehicles as agent interactions are prevalent there, yet explainability and transparency are also important. Given the inherently causal nature of agent interactions and that explanations can utilise counterfactual inference this work builds upon causality literature. Four main chapters of content provide the primary contributions of this thesis in addition to a literature review of the aforementioned research areas. The first chapter introduces causality theory and applies it to the autonomous-vehicle domain, before benchmarking causal-discovery methods on real-world autonomous-vehicle data to identify challenges for existing techniques. The second chapter combines an action-based theory of mind and counterfactual inference in order to produce SimCARSv1, which outperforms the methods of the first chapter. The third chapter builds upon the previous by tackling challenges associated with representing a system of interacting autonomous agents via structural causal models. The final chapter concludes by merging the contributions from all previous chapters, along with integrating a means by which to estimate the instantaneous reward parameters of agents. This culminates in the creation of SimCARSv2, which offers similar quantitative performance to SimCARSv1, but with greater expressiveness due to its structural-causal-model-based architecture. The sum of these contributions represents an important step towards bridging the promising fields of causality and autonomous vehicles, with the goal of building autonomous-agent technologies that can interact with humans safely.

## PREFACE

Just under five years ago I began the journey that would culminate in the creation of this thesis. At that time I knew that I wanted to carry out research that would give robots — and by extension autonomous agents — the ability to better understand the world and interactions around them. At the time I was not exactly sure what exact avenue this would take me down. However, after being taken on by Lars Kunze and introduced the the field of causality matters quickly fell into place. This was simultaneously challenging and rewarding, as although causality was already quite established within big-data domains, there was comparatively little when it came to autonomous-agent domains — particularly when considering interactions between agents. Conducting research within such a frontier field led to many instances of self doubt, and this combined with other life events pushed me to nearly quit several times along the way. Nonetheless I am truly glad that I stuck it through, not just for the academic achievement of it, but as a personal milestone of perseverance and growth. However, as proud as I am in having made it to this point where I am putting the finishing touches on the thesis as I write this, the most valuable outcome of these last five years has to be the experience and insight I have gained as a result of pursuing this. With that in mind, I would like to think of this thesis not just as a finale to the DPhil, but also as the beginning of all that comes next. I can only hope that whatever I go on to do now will make all those who have supported me thus far proud.

Rhys Peter Matthew Howard

Lincoln College

8th April 2025

## ACKNOWLEDGEMENTS

First and foremost I must thank Lars Kunze for making this entire endeavour possible by taking me on as a DPhil student. Although joining me at a later point in my DPhil journey, I am also thankful for Nick Hawes stepping up to act as one of my supervisors, not to mention introducing me to Lars in the first place. I must also thank the Engineering Science department of the University of Oxford for providing the bulk of the funding that has enabled me to spend the past several years of my life carrying out this research. I also thank both Lincoln College and the RobotCycle Project for having assisted me in covering the costs of conferences thus allowing me to engage with my fellow researchers in the field. As for my colleagues: Efimia Panagiotaki, Jumana Baghabrah, and Ricardo Cannizzaro have all been a massive source of help and encouragement during the course of the DPhil and have my thanks. Ricardo in particular has been a rock of support over the years, not to mention having directly collaborated with myself on several occasions. In addition to those already mentioned would like to also thank Ben Hardin, Carl Hentges, Divya Thuremella, Jianhao Yuan, and Michael Groom for aiding in the write-up of various publications by providing peer-review. Lastly I would like to thank the administration team at the Oxford Robotics Institute, for being vital in making various arrangements, and just generally being very helpful.

# Contents

# List of Tables

# List of Figures

# List of Acronyms

**AI** Artificial Intelligence

**XAI** Explainable Artificial Intelligence

**RQ** Research Question

**1D** One-Dimensional

**2D** Two-Dimensional

**3D** Three-Dimensional

**AES** Autonomous Embodied System

**AV** Autonomous Vehicle

**ROS** Robot Operating System

**RL** Reinforcement Learning

**IRL** Inverse Reinforcement Learning

**DAG** Directed Acyclical Graph

**SCM** Structural Causal Model

**MDP** Markov Decision Process

**POMDP** Partially-Observable Markov Decision Process

**PWGC** Pair-Wise Granger Causality

**MVGC** Multi-Variate Granger Causality

**TCDF** Temporal Causal Discovery Framework

**AD-DSTCN** Attention-based Dilated Depth-wise Separable Temporal Convolutional Network

**NN** Neural Network

**ABCNN** Attention-based Convolutional Neural Network

**NAVAR** Neural Additive Vector Auto-Regression

**MLP** Multi-Layer Perceptron

**LiNGAM** Linear Non-Gaussian Acyclical Model

**DirectLiNGAM** Direct Linear Non-Gaussian Acyclical Model

**VAR-LiNGAM** Vector Auto-Regressive Linear Non-Gaussian Acyclical Model

**TiMINo** Time-series Models with Independent Noise

**PC** Peter-Clark algorithm

**MEC** Markov Equivalence Class

**PCMCI** Peter-Clark algorithm with Momentary Conditional Independence

**oCSE** Optimal Causation Entropy

**FCI** Fast Causal Inference

**tsFCI** Time-Series Fast Causal Inference

**CD-NOD** Constraint-based causal Discovery from Nonstationary / heterogeneous Data

**NOTEARS** Non-combinatorial Optimisation via Trace Exponential and Augment Lagrangian for Structure learning

**DYNOTEARS** Dynamic Non-combinatorial Optimisation via Trace Exponential and Augment Lagrangian for Structure learning

**NRI** Neural Relational Inference

**GRI** Grounded Relational Inference

**NS-DR** Neuro-Symbolic Dynamic Reasoning

**V-CDN** Visual Causal Discovery Network

**TP** True Positive

**TN** True Negative

**FP** False Positive

**FN** False Negative

**ROC** Receiver Operating Characteristic

**MTL** Maximum Time Lag

**TTC** Time To Collision

**CCT** Cumulative Collision Time

**PTS** Previous Time Step

**TSSP** Time-Step-Size Product

**TSSQ** Time-Step-Size Quotient

**CTD** Current Time Difference

**RCS** Retrospective Causal Stationarity

**ADAS** Advanced Driver-Assistance System

**IDM** Intelligent Driver Model

**CARLA** Car Learning to Act

**FWD** Front-Wheel Drive

**LiDAR** Light Detection and Ranging

**SimCARSv1** Simulation-based Causal Analysis and Reasoning System Version 1

**SimCARSv2** Simulation-based Causal Analysis and Reasoning System Version 2

# List of Symbols

| | |
|---|---|
| $i,\ j,\ k$ | Index Counters |
| $u,\ v,\ w,\ x,\ y,\ z$ | Generic Variables |
| $f(\cdot)$ | Generic Function / Structural Equation |
| $G$ | Graph / Directed Acyclical Graph |
| $N$ | Nodes / Vertices / Variables |
| $E$ | Edges / Links |
| $Pa(\cdot)$ | Node Parent Function |
| $An(\cdot)$ | Node Ancestor Function |
| $M$ | Structural Causal Model |
| $U$ | Exogenous Variables |
| $V$ | Endogenous Variables |
| $F$ | Structural Equations |
| $P(\cdot)$ | Probability Distribution |
| $T$ | Set of Time Steps |
| $\mathbb{N}$ | Set of Natural Numbers |

| | |
|---|---|
| $t$ | Time Step |
| $M^T$ | Full-Time-Graph Structural Causal Model |
| $U^T$ | Time-Indexed Exogenous Variables |
| $V^T$ | Time-Indexed Endogenous Variables |
| $E^T$ | Full-Time-Graph Edges |
| $F^T$ | Time-Indexed Structural Equations |
| $Pa^T(\cdot)$ | Full-Time-Graph Node Parent Function |
| $An^T(\cdot)$ | Full-Time-Graph Node Ancestor Function |
| $\mathcal{G}^S$ | Causal-Summary Graph |
| $N^S$ | Causal-Summary-Graph Nodes |
| $E^S$ | Causal-Summary-Graph Edges |
| $\tau$ | Maximum Time Lag / Sliding Time-Window Size |
| $L^\tau$ | Relative Time Lags |
| $M^\tau$ | Window-Time-Graph Structural Causal Model |
| $U^\tau$ | Relative-Time-Lag-Indexed Exogenous Variables |
| $V^\tau$ | Relative-Time-Lag-Indexed Endogenous Variables |
| $E^\tau$ | Window-Time-Graph Edges |
| $F^\tau$ | Relative-Time-Lag-Indexed Structural Equations |
| $Pa^\tau(\cdot)$ | Window-Time-Graph Node Parent Function |

| | |
|---|---|
| $An^{\tau}(\cdot)$ | Window-Time-Graph Node Ancestor Function |
| $\mathbf{c0}, \mathbf{c1}, \ldots$ | Convoy Vehicle Agents |
| $\mathbf{i0}, \mathbf{i1}, \ldots$ | Independent Vehicle Agents |
| $\mathbb{R}$ | Set of Real Numbers |
| $\mathbf{W}$ | Linear-Coefficient Array / Matrix |
| $\mathbf{I}$ | Identity Matrix |
| $\varepsilon$ | Additive Noise |
| $F^{\varepsilon}$ | Structural Equations with Explicit Noise |
| $\ell(\cdot)$ | Loss Metric |
| $\beta_{\mathbf{W}}$ | Coefficient-Array Regularisation Constant |
| $\|\cdot\|_1$ | Manhattan Distance |
| $t_{max}$ | Maximum Time |
| $\hat{\mathcal{G}}^S$ | Discovered Causal-Summary Graph |
| $\hat{N}^S$ | Discovered Casual-Summary-Graph Nodes |
| $\hat{E}^S$ | Discovered Causal-Summary-Graph Edges |
| $\|TP\|$ | True-Positive Count |
| $\|FP\|$ | False-Positive Count |
| $\|FN\|$ | False-Negative Count |
| $\|TN\|$ | True-Negative Count |

| | |
|---|---|
| $\|TP\|^O$ | Orientation-Based True-Positive Count |
| $\|FP\|^O$ | Orientation-Based False-Positive Count |
| $\|FN\|^O$ | Orientation-Based False-Negative Count |
| $\|TN\|^O$ | Orientation-Based True-Negative Count |
| $\|TP\|^A$ | Adjacency-Based True-Positive Count |
| $\|FP\|^A$ | Adjacency-Based False-Positive Count |
| $\|FN\|^A$ | Adjacency-Based False-Negative Count |
| $\|TN\|^A$ | Adjacency-Based True-Negative Count |
| $\lambda_\alpha$ | Significance Alpha |
| $L^A$ | Markov Relative Time Lags |
| $M^A$ | Autonomous-Agent Structural Causal Model |
| $U^A$ | Autonomous-Agent Exogenous Variables |
| $V^A$ | Autonomous-Agent Endogenous Variables |
| $V_S$ | State Variables |
| $V_A$ | Action Variables |
| $V_D$ | Driving-Actuation Variables |
| $V^O$ | Other Endogenous Variables |
| $E^A$ | Autonomous-Agent Graph Edges |
| $F^A$ | Autonomous-Agent Structural Equations |

| | |
|---|---|
| $\mathcal{G}^{BI}$ | Behavioural-Interaction Graph |
| $N^{BI}$ | Behavioural-Interaction-Graph Nodes / Autonomous-Agent Actions |
| $E^{BI}$ | Behavioural-Interaction-Graph Edges |
| $a$ | Composite-Action Variable Value |
| $g$ | Goal Variable Value |
| $s$ | State Variable Value |
| $T^{g,sp,A}$ | Speed-Goal Start Times |
| $T^{g,sp,\Omega}$ | Speed-Goal Target Times |
| $d$ | Driving-Actuation Variable Value |
| $d_a$ | Acceleration Variable Value |
| $\lambda_{d,a}$ | Speed-Goal Acceleration Actuation Threshold |
| $\boldsymbol{g}_{sp}$ | Speed Goals |
| $s_{sp}$ | Speed Variable Value |
| $\lambda_{\delta a,sp,t}$ | Speed-Goal Duration Threshold |
| $\lambda_{\delta a,sp,sp}$ | Speed-Goal Speed-Difference Threshold |
| $\boldsymbol{g}_{la}$ | Lane Goals |
| $s_{la}$ | Lane Variable Value |
| $\mathcal{T}^X$ | Data Type / Domain of Variable $X$ |
| $\lambda_{\delta a,la,t}$ | Pre-Lane-Change Lane-Goal-Start-Time Threshold |

| | |
|---|---|
| $\lambda'_{\delta a,la,t}$ | Post-Lane-Change Lane-Goal-Target-Time Threshold |
| $A$ | Composite Actions |
| $a_{\mathcal{C}}$ | (Potentially) Causing Composite Action |
| $a_{\mathcal{A}}$ | (Potentially) Affected Composite Action |
| $\mathcal{C}$ | (Potentially) Causing Agent |
| $\mathcal{A}$ | (Potentially) Affected Agent |
| $\delta_t$ | Time-Step Size |
| $st_{la}(\cdot)$ | Lane-Curvature Steering Function |
| $st_{mid}(\cdot)$ | Mid-Point Steering Function |
| $s_{pos}$ | Position Variable Value |
| $s_{dir}$ | Direction Variable Value |
| $l_{mid}(\cdot)$ | Lane Mid-Point Function |
| $[\cdot]^X$ | Counterfactual Context Under $X$ |
| $r(\cdot)$ | Reward Function |
| $r_{ttc}(\cdot)$ | Time-To-Collision Sub-Reward Function |
| $r_{cct}(\cdot)$ | Cumulative-Collision-Time Sub-Reward Function |
| $r_s(\cdot)$ | Speed Sub-Reward Function |
| $s_{ttc}$ | Time-To-Collision Variable Value |
| $s_{cct}$ | Cumulative-Collision-Time Variable Value |

| | |
|---|---|
| $\exp(\cdot)$ | Exponential Function |
| $\delta_r$ | Combined Reward-Based Conditional-Optimality Metric |
| $\delta_r^+$ | Positive Reward-Based Conditional-Optimality Metric |
| $\delta_r^-$ | Negative Reward-Based Conditional-Optimality Metric |
| $\beta_h$ | Simulation Horizon |
| $\lambda_{\delta r}$ | Reward-Based Conditional-Optimality Threshold |
| $\xi^{\text{reward}}$ | Reward-Based Causal-Link Test |
| $\mathbb{B}$ | Binary Set |
| $ag(\cdot)$ | Agency Function |
| $\xi^{\text{active}}$ | Agency-Based Active-Interaction Test |
| $\xi^{\text{passi.}}$ | Agency-Based Passive-Interaction Test |
| $\xi^{\text{facil.}}$ | Agency-Based Facilitation-Interaction Test |
| $\xi^{\text{m.e.m.}}$ | Agency-Based Mutual-Effect-Motive-Interaction Test |
| $\xi^{\text{agency}}$ | Agency-Based Causal-Link Test |
| $\xi^{\text{hybrid}}$ | Hybrid Causal-Link Test |
| $\mathfrak{M}$ | Generic Monad |
| $\mathfrak{M}(\cdot)$ | Generic-Monad Type Constructor |
| $u_{\mathfrak{M}}(\cdot)$ | Generic-Monad Unit / Return Function |
| $\mathcal{T}$ | Generic Data-Type / Domain |

| | |
|---|---|
| $b_{\mathfrak{M}}(\cdot)$ / $\triangleright_{\mathfrak{M}}$ | Generic-Monad Bind Function |
| $C$ | Variable Context |
| $C_T$ | Current-Time Meta-Variable |
| $C_{\delta t}$ | Time-Step-Size Meta-Variable |
| $\mathfrak{S}$ | State Monad |
| $\mathcal{S}$ | State-Monad State Domain |
| $V^{\delta}$ | Previous-Time-Step Variable |
| $f_{V^{\delta}}$ | Previous-Time-Step-Variable Structural Equation |
| $V_{Pa}$ | Parent Variable |
| $\delta_{-}$ | Time-Step-Decrement Monadic Value |
| $\delta_{+}$ | Time-Step-Increment Monadic Value |
| $\emptyset$ | Empty Set / Value |
| $s_{\delta sp}$ | Speed-Difference Variable Value |
| $V_{S,sp}$ | Speed Variable |
| $V_{S,\delta sp}$ | Speed-Difference Variable |
| $V_{D,a}$ | Acceleration Variable |
| $V^{\cdot\delta t}$ | Time-Step-Size-Product Variable |
| $V^{\div\delta t}$ | Time-Step-Size-Quotient Variable |
| $f_{V^{\cdot\delta t}}$ | Time-Step-Size-Product-Variable Structural Equation |

| | |
|---|---|
| $f_{V^{\div \delta t}}$ | Time-Step-Size-Quotient-Variable Structural Equation |
| $V^{\delta T}$ | Current-Time-Difference Variable |
| $f_{V^{\delta T}}$ | Current-Time-Difference-Variable Structural Equation |
| $V^{T?}$ | Time-Conditional Variable |
| $f_{V^{T?}}$ | Time-Conditional-Variable Structural Equation |
| $\theta_T$ | Time-Conditional Divergence-Point Parameter |
| $V^B$ | Buffer Variable |
| $\Phi_{V^B}$ | Buffer-Variable Time-Indexed Dictionary |
| $T_{V^B}$ | Set of Buffer-Variable Dictionary Time Steps |
| $\phi_{V^B}(\cdot)$ | Buffer-Variable Dictionary Time-Indexing Function |
| $f_{V^B}$ | Buffer-Variable Structural Equation |
| $\mathfrak{B}$ | Buffer Monad |
| $\mathrm{UP}_X$ | Buffer-Monad Data Constructor (Update Dictionary Based Upon $X$) |
| $U^S$ | Socket Variables |
| $U^H$ | Hidden Exogenous Variables |
| $T_{<t}$ | Set of Time Steps Before $t$ |
| $U^{S,\emptyset}$ | Empty-Set Degenerate-Distribution Socket Variable |
| $V^{\cup}$ | Union Variable |
| $\theta_T^A$ | Input-Set-Source Introduction Time |

| | |
|---|---|
| $\theta_T^{\Omega}$ | Input-Set-Source Removal Time |
| $\mathbf{p0}, \mathbf{p1}, \ldots$ | Pedestrian Agents |
| $s_o$ | Outcome Variable Value |
| $s_{lt}$ | Lane-Transition Variable Value |
| $s_{dh}$ | Distance-Headway Variable Value |
| $s_{ef}$ | Maximum-Environmental-Force-Magnitude Variable Value |
| $s_{ad}$ | Action-Done Variable Value |
| $r^*(\cdot)$ | Weighted Reward Function |
| $\mathbf{r}(\cdot)$ | Composite Reward Function |
| $\boldsymbol{\gamma}$ | Reward-Profile Weight Vector |
| $\gamma$ | Reward-Profile Weight |
| $r_0(\cdot)$ | Lane-Transition Sub-Reward Function |
| $r_1(\cdot)$ | Distance-Headway Sub-Reward Function |
| $r_2(\cdot)$ | Speed-Maximisation Sub-Reward Function |
| $r_3(\cdot)$ | Speed-Minimisation Sub-Reward Function |
| $r_4(\cdot)$ | Crash-Avoidance Sub-Reward Function |
| $\sigma(\cdot)$ | Sigmoid Function |
| $\lambda_{dh}$ | Distance-Headway Braking-Time Limit |
| $\lambda_{sp}$ | Speed Limit |

| | |
|---|---|
| $\lambda_{ef}$ | Maximum-Environmental-Force-Magnitude Threshold |
| $\hat{a}$ | Hypothetical Composite-Action Value |
| $\hat{s}_o$ | Hypothetical Outcome Value |
| $\zeta_o(\cdot)$ | Outcome-Distance Function |
| $\alpha_o$ | Overall Outcome-Distance Scaling Factor |
| $\alpha_{lt}$ | Lane-Transition Outcome-Distance Scaling Factor |
| $\alpha_{sp}$ | Speed Outcome-Distance Scaling Factor |
| $\alpha_{dh}$ | Distance-Headway Outcome-Distance Scaling Factor |
| $\alpha_{ef}$ | Maximum-Environmental-Force-Magnitude Outcome-Distance Scaling Factor |
| $\alpha_{ad}$ | Action-Done Outcome-Distance Scaling Factor |
| $\boldsymbol{\gamma}_{\mathcal{A}}$ | Reward Profile of (Potentially) Affected Composite Action |
| $\hat{a}^{\neg\mathcal{C}}$ | Hypothetical Action in the Absence of a (Potentially) Causing Action |
| $\zeta_a(\cdot)$ | Action-Distance Function |
| $\alpha_a$ | Overall Action-Distance Scaling Factor |
| $\alpha_{la}$ | Lane-Goal-Value Action-Distance Scaling Factor |
| $\delta^K$ | Kronecker Delta |
| $\lambda_{\zeta a}$ | Action-Distance Conditional-Optimality Threshold |

# Chapter 1

# Introduction

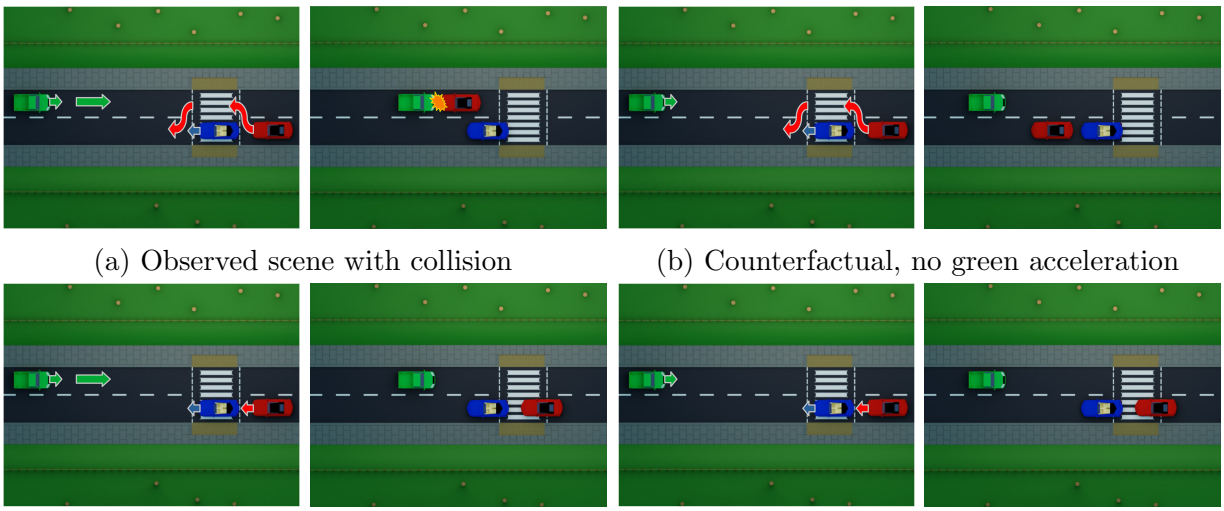
Autonomous systems are becoming increasingly prevalent in the day-to-day lives of people. Such systems ought to understand cause and effect in relation to their behaviour and the behaviour of others. The field of Pearlean causality offers a strong basis with which to approach this task as it has existing techniques aimed at identifying causal relationships between variables, not to mention many techniques that can be applied to established causal models.

## 1.1 Motivations

Before outlining the contents of this thesis it is first necessary to establish what challenges it aims to address and why it is important that these challenges are tackled.

### 1.1.1 Why Model Autonomous-Agent Interactions?

While there are many reasons one may want to model behavioural interactions this work is primarily interested in doing so to carry out post-hoc analysis of said interactions. Specifically, this work focuses upon high-risk domains where autonomous agents present potential harm to people. Understanding the series of events that occurred and how agents interacted is of crucial importance, as there is substantially more ambiguity over who or what might be to blame for a negative outcome than with non-autonomous systems. Thus the first instance where behavioural-interaction modelling is useful is in determining culpability within such circumstances. The second instance is using such investigations to inform the future development of laws, regulations, and technologies to increase the safety of autonomous systems. All of these present critical objectives for any society that aims to develop autonomous technologies responsibly.



(a) Observed scene with collision

(b) Counterfactual, no green acceleration

(c) Counterfactual, no red overtake

(d) Counterfactual, neither action

Figure 1.1: Illustration of example fault-attribution scenario. Subfigures show the initial scene with the plan on the left and the final outcome on the right. (© 2025 IEEE)

### 1.1.2 Which Scenarios Benefit from Modelling Autonomous-Agent Interactions?

While the general motivations behind wanting to model agent interactions have been discussed, the following example scenarios offer a better insight into domains where this work could be applied, as well as greater insight as to how such use-cases could be approached.

**Autonomous-Vehicle Crash** In the scenario depicted in Figure 1.1 a red vehicle attempts to overtake a blue vehicle, meanwhile a green vehicle approaching from the opposite direction speeds up. The combination of these actions results in a collision between the red and green vehicles. In this scenario one can imagine that one or more AVs are present, or that at least some of the vehicles are outfitted with an array of high-fidelity sensors. Under such circumstances it would be desirable to try and approximate some level of culpability for each agent in an objective manner. Now legally speaking the exact proportion of blame one can attribute to each of the agents will of course depend upon the temporal ordering of actions, the laws of the jurisdiction in which the crash occurred, and other matters of context. However, one can attempt to establish via causal techniques some level of agent action attribution. One can approach this by tying measures of causal necessity to the 'but-for' test as laid out in the Model Penal Code [1] in the manner described by Pearl and Mackenzie [2]. This also bears resemblance to the form of 'actual causality' between events as described by Halpern [3], albeit in the naive form without considering counterfactual contingencies over variables.

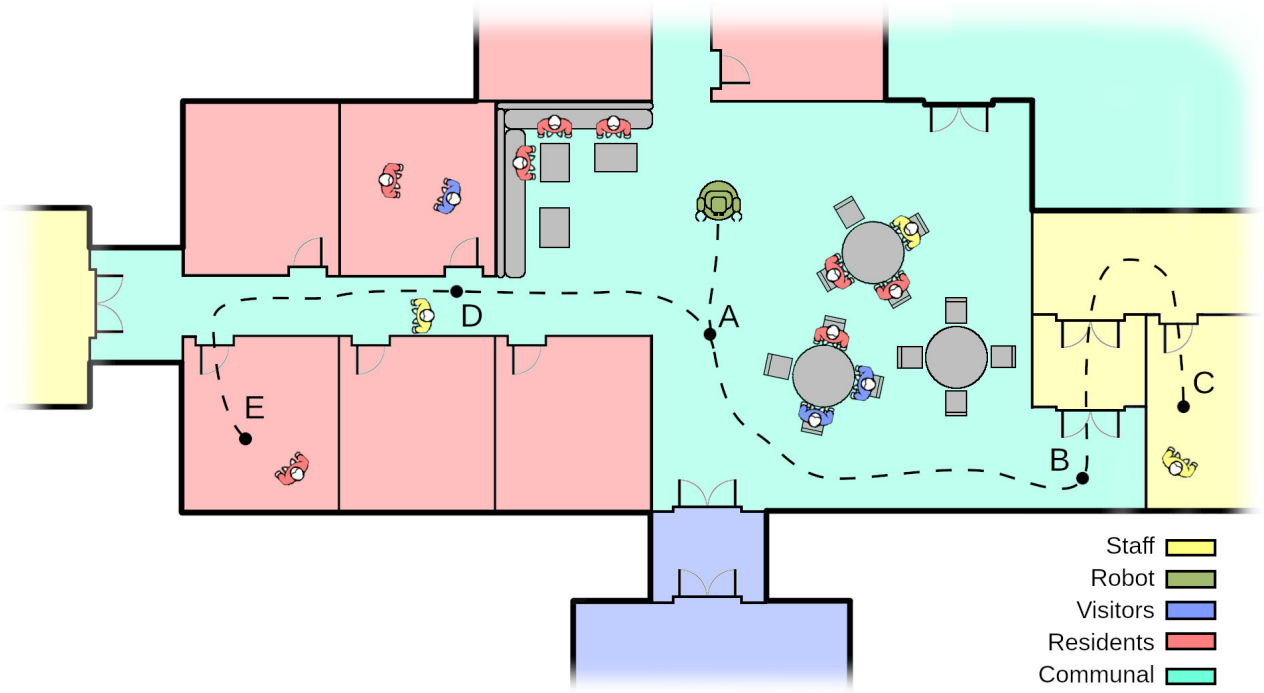



Figure 1.2: Illustration of example self-diagnostic scenario set in a retirement home.

**Socially-Aware Robots** This scene shown in Figure 1.2 depicts a green support robot and a yellow staff member, among others, in a retirement-home setting. They cross paths in a tight corridor at vertex D. The robot's default behaviour is to always give way to people, however the staff member, in an attempt to be polite, attempts to give way to the robot. This leads to the agents being at an impasse, that is both awkward and time wasting. The robot should ideally be able to flag the anomalous duration of time that has passed during the impasse and begin self-diagnosis in order to deduce what went wrong from a causal perspective. In contrast to the AV scenario described above the aim here is less to establish culpability and more to better inform the robot on its future behaviour. Understanding the causal links inherent in agent interactions is important for autonomous agents (e.g. robots) acting in such roles to build a model of social cues that does not rely upon training for specific scenarios. Instead it is hoped that through a combination of causal modelling and theory of mind that one can build autonomous agents that gradually develop a grasp on how their actions affect others.

### 1.1.3 Why Utilise Pearlean Causality?

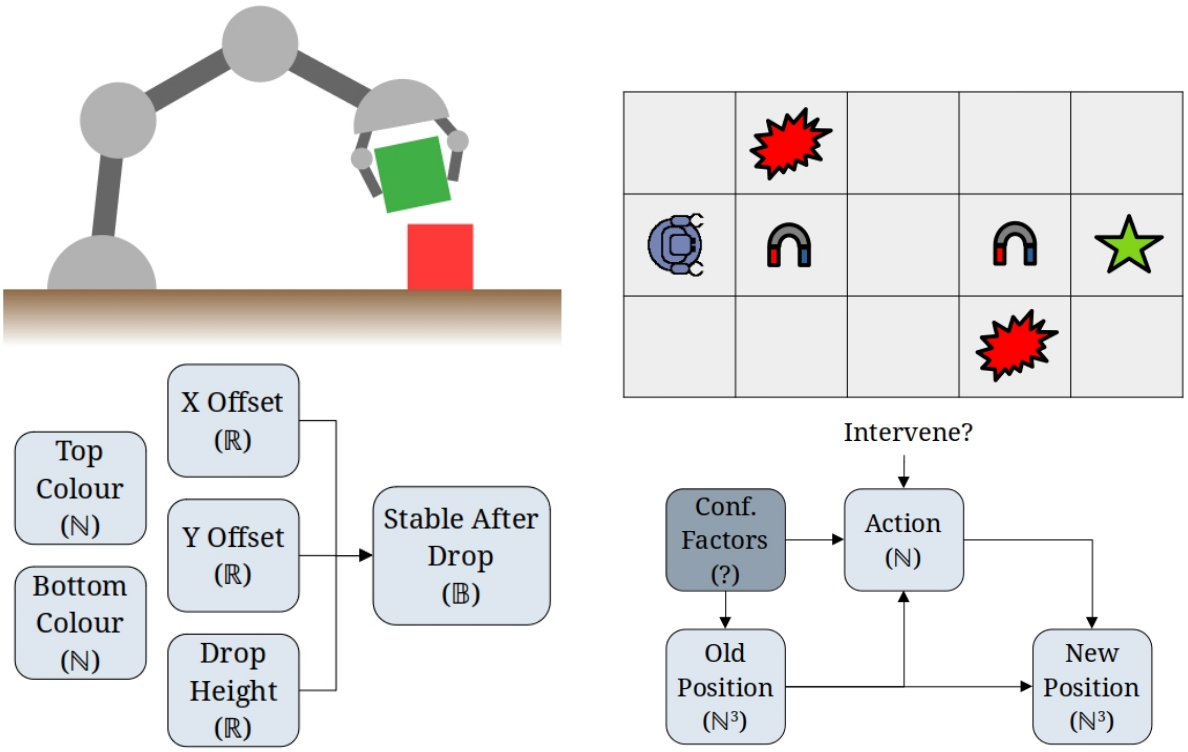
Causality has been tied to counterfactual reasoning since Pearl [4] helped to establish many of the formalisms that are used today, in addition to describing the concept of the ladder of

causation. This concept ascribes counterfactual reasoning to the highest rung of a series of steps increasing cognitive capabilities, and identifies causality as a necessary component in order to carry out said reasoning. In turn counterfactual reasoning has been found to tie in with how humans perceive and reason about the world [5, 6]. Importantly this means that such conceptions of causality are inherently tied to explainability, or in other words answering 'Why?' questions about the world and the events that occur within it.

Beyond that, models such as Structural Causal Models (SCMs) offer several beneficial traits that align with responsible development philosophies. Key among these is that SCMs are highly transparent, and clearly delineate modelled factors and those that are not modelled via endogenous and exogenous variables respectively. This allows one to understand exactly how an SCM arrives at the results it does, offering a level of grounding that cannot be matched by black-box methodologies. In a similar thread recent work has shown that many interpretations of accountability are rooted in causality [7, 8]. This implies that not only are causal systems desirable for interpretation purposes, but are also critical for making such systems accountable.

All of these traits make these causal models preferable over other contemporary approaches to modelling agent interactions. For example, a Neural Network (NN) based method might be able to weave a compelling and even correct story regarding how a particular outcome came about. However, unless it can be understood how it arrived at that story — which in many cases is difficult [9] — the output may be of little use, depending on context. This is of primary importance when it comes to potential litigation, where grounded evidence is of paramount importance. However, this could also be beneficial to systems which might be able to learn from such meta-knowledge. As such, it is with all these factors in mind that this thesis explores this particular research avenue into causality.

In terms of practical examples of Pearlean causal models, two example applications within autonomous-agent domains are depicted in Figure 1.3. The first of these, depicted in Figure 1.3a, is based upon the work of Diehl and Ramirez-Amaro [10]. In this work they learn a causal Bayesian model that describes the causal relationships between the input parameters of a block pick and place task and whether the resulting stack of blocks is stable. By identifying which input parameters influence the outcome of the task they can more efficiently identify the causes




(a) An example of causal Bayesian network usage. Based off work by Diehl and Ramirez-Amaro [10].

(b) An example of SCM usage as part of an Markov Decision Process (MDP) formulation. Based off work by Cannizzaro and Kunze [11].

Figure 1.3: Two examples of causal model usage within autonomous-agent domains. Descriptions of their usage are given in Section 1.1.3.

behind task failures by attempting to search for the closest input parameter configuration that would have succeeded in each case. The second example, depicted in Figure 1.3b, is based upon the work of Cannizzaro and Kunze [11]. In their work they utilise a hybrid SCM-MDP model[1] to capture a mobile robot operating in a grid world attempting to reach a goal — indicated by a green star. Within this grid world there are both hazards for the robot — indicated via red explosions — and areas of magnetic interference — shown as magnets. The former of these can easily be accounted for as moving into a hazardous square would result in a large negative reward. However, the areas of magnetic interference have a confounding effect upon both the senses and actuation of the robot. In the aforementioned work, the knowledge of potential confounding factors allows the proposed method to take pre-emptive action by intervening upon actions in order to minimise the influence of the aforementioned confounders and ultimately maximise reward. For example, in the grid world shown, the proposed method is able to correctly judge

[1]The work in question utilises a Partially-Observable Markov Decision Process (POMDP) formulation, but for the sake of the example's simplicity an MDP formulation was given.

whether it is worth heading directly for the goal or whether it would be more beneficial to zig-zag and avoid the magnetic interference. Both of these examples illustrate that the utilisation of Pearlean causality is both useful and varied in its applications.

## 1.2 Outline

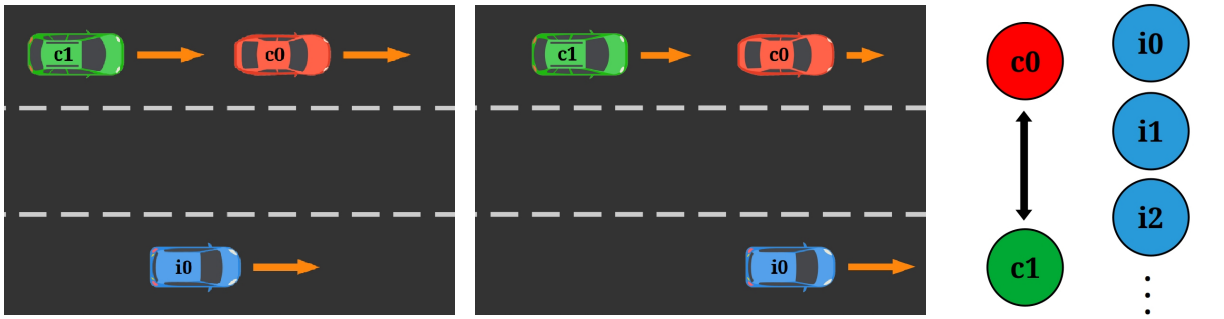
With the motivations behind the work established, it is now possible to provide a high-level outline of the work, including how the work is situated against contemporary work in literature, the choice of application domain, and the research questions the thesis aims to explore.

### 1.2.1 Situating the Contributions of This Work

It is little wonder that causality has found itself the subject of interest of several fields, including Artificial Intelligence (AI) and robotics as of late [12, 13]. However, much of this interest has been regarding the application of causality to augment existing methodologies [14], establish models from an egocentric perspective [15], or to avoid confounding influences affecting controllers / planners [11]. While works applying causality to Explainable Artificial Intelligence (XAI) exist [16], they remain focused on an egocentric perspective, as mentioned above. This is not necessarily a negative, but it does mean that explanations from such methods are inherently subjective, and while this has its own uses, this work aims to approach the task from an objective viewpoint. This work is the first to its author's knowledge, to utilise causality for the non-egocentric modelling of agent behavioural interactions to the extent considered here.

### 1.2.2 Relation of This Work to the Autonomous-Vehicle Domain

In examining the utilisation of causal reasoning to model agent behavioural interactions AVs make for a particularly motivating case. They have a substantial amount of investment currently, yet can also pose a significant risk to human life. This in many ways reflects the early airline industry, and there is work suggesting autonomous systems record their data — similar to aircraft — for post-hoc analysis [17]. Additionally from a practicality standpoint the AV domain offers several advantages. There is plenty of data relevant to the domain available due to substantial industry support and from a technical perspective road driving offers a good balance between chaotic social behaviour and structured rule following (e.g. lanes markings, a highway code, etc.). Thus most of this work concerns itself with the AV domain.




(a) Initial State (b) Final State (c) Causal Adjacency

Figure 1.4: Various illustrations of the two-vehicle-convoy scenario.

### 1.2.3 Research Questions

Each segment of work leading up to this thesis was carried out with the aim of tackling several Research Questions (RQs). Naturally the direction of the latter RQs was partially informed by the outcomes of the RQs that came beforehand. These are documented as follows:

RQ-1. ***'Can contemporary causal-discovery approaches identify autonomous-agent behavioural interactions?'***

This is a question of how one might be able to formulate a causal model to represent agent interactions, whether it is possible to utilise existing methodologies to identify them, and if it is not possible, what are the challenges preventing it.

- **Example Case:** Consider the scenario depicted in Figure 1.4 which depicts two vehicles occupying a single lane within close proximity to one another. Naturally the movement of one vehicle has a causal effect upon the movement of the other vehicle as they each make an effort to avoid collision while also attempting to take advantage of free space where they can. The causal relationship between the agents can be seen in the causal-adjacency graph of the figure. This question asks if it is possible to deploy traditional temporal causal-discovery techniques to identify the causal relationship between the agents autonomously.

RQ-2. ***'Can the nature of autonomous agents be exploited to overcome the challenges associated with identifying autonomous-agent behavioural interactions?'***

Based upon RQ-1, several difficulties were identified with real-world agent data that made utilising existing temporal causal-discovery methods difficult. Therefore this question aims to explore whether any traits specific to autonomous agents can be exploited in order to overcome these difficulties and identify behavioural interactions.

- **Example Case:** Once again consider the scenario depicted in Figure 1.4. One can extend the scenario it by specifying that the direction of causality switches back and forth, and the time lags involved in causal relationships are inconsistent. As a result traditional temporal causal-discovery methods cannot be applied here. However, autonomous agents possess a certain level of intelligence which one can make assumptions about. For example, it can reasonably be assumed that autonomous agents typically act to maximise some conception of utility or reward. As such this question asks whether it is possible in such a situation to utilise what is known regarding autonomous agents to still discover the causal relationship between the agents despite the conditions of the scenario.

RQ-3. ***'How can SCMs be extended in order to better represent a system of multiple interacting autonomous agents?'***

The research contributions of RQ-2 significantly improved the ability to reliably identify interactions between agents, but in doing so there was a departure from a fully SCM-based architecture. This prevents one from exploiting the beneficial traits discussed in Section 1.1.3 as well as limiting the counterfactual queries one can make. As such this question aims to identify how SCMs can be modified and extended in order to support the advances made under RQ-2 while retaining an underlying SCM format.

- **Example Case:** So far the research questions have focused upon the identification of causal relationships in the scenario depicted in Figure 1.4 via either traditional temporal causal discovery or the application of theory of mind. However, one could instead attempt to represent the AV scenario itself as an SCM based upon established

low-level knowledge of agent and physics models in literature. With an SCM in place one could then attempt to infer high-level causal relationships between agents by carrying out counterfactual inferences directly on the SCM. Despite this, SCMs have traditionally been used directly with tabulated data and thus are not immediately well suited to represent such a scenario. Thus this question asks how one might be able to extend SCMs to better capture a scenario involving autonomous agents.

RQ-4. ***'How can autonomous-agent motivations be factored into the identification of behavioural interactions?'***

With the SCM-based architecture developed under RQ-3, it was possible to return to an open issue identified as an outcome of RQ-2, namely the relatively poor performance of the methods considering numerical agent reward. One possibility is that this poor performance was due to a simple product of reward metrics being used across all agents / actions. Thus this question explores how one can capture variations in agent motivations, and how this information can then be integrated within the identification of agent interactions.

- **Example Case:** Returning to the scenario depicted Figure 1.4 now with the possibility of utilising an SCM-based architecture, it becomes desirable to address the shortcomings of the previous theory-of-mind-based approach. One such shortcoming relates to the manner in which the previous approach captured the motivations of agents via a fixed composite reward model. In addition to motivations varying between agents and across time, the comparatively poor performance of the variant utilising the previous composite reward model indicates its unsuitability. To once again extend the scenario described for Figure 1.4, consider if the two agents are engaged in road rage, acting aggressively and otherwise going against typical driver behaviour. If a reward model is fixed and assumes all drivers follow a sensible pattern of driving behaviour then it will be unable to model agent cognition effectively in situations which deviate from this, such as in the previously described case. Taking this all into account, this question asks how best to estimate the instantaneous motivations of the agents within a scenario, and how this information can be incorporated to address the shortcomings of the reward-based variant of the previous approach.

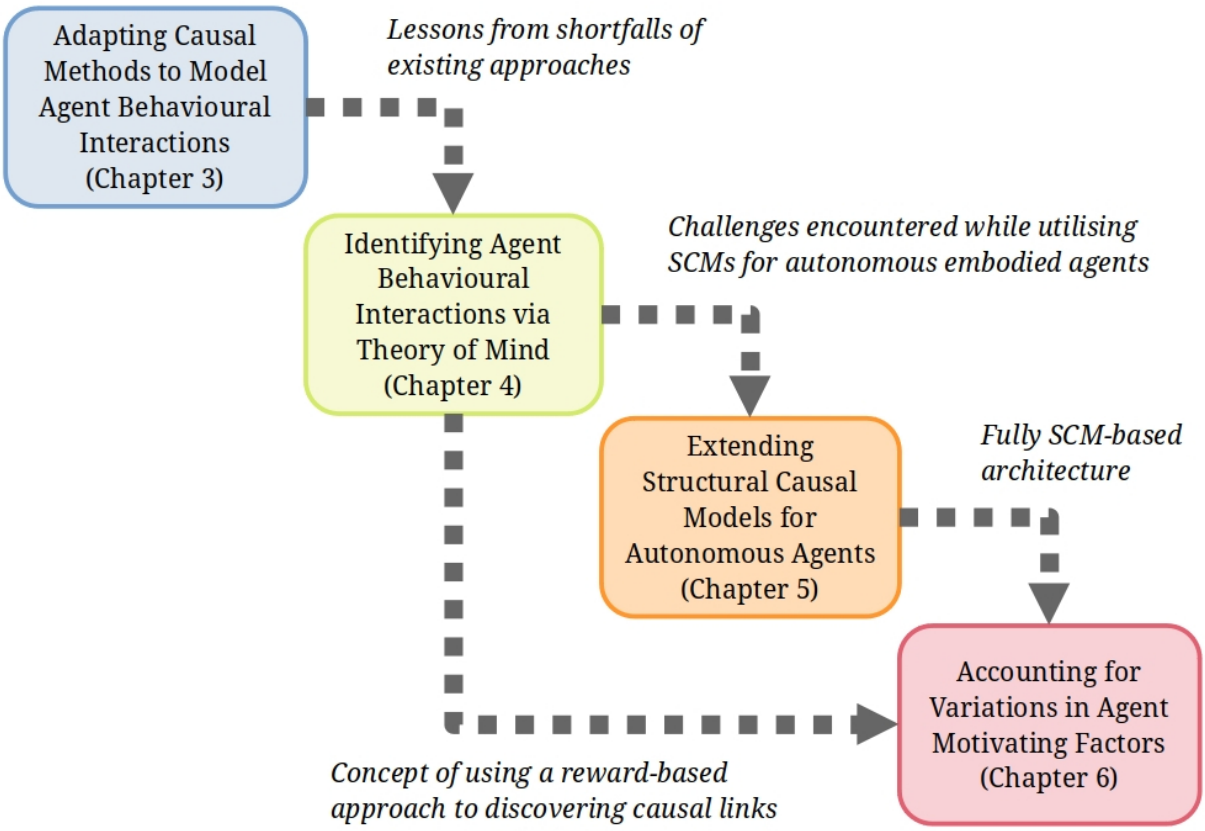



Figure 1.5: An illustration of how the insights and contributions from each of the main-content chapters guided the development of subsequent chapters.

## 1.3 Thesis Layout

With the premise of this thesis established, the structure of the rest of the chapters can be laid out. Furthermore, an illustration of how each of the main-content chapters relate to one another is given in Figure 1.5. Each of the main-content chapters between the literature review and conclusion aims to tackle one of the research questions documented above and is associated with a publication, as laid out below in Section 1.4.

**Literature Review (Chapter 2)** This work begins with a literature review that explores the state of the field, particularly within areas relating to causal reasoning, in order to contextualise the rest of the work moving forwards.

**Adapting Causal Methods to Model Agent Behavioural Interactions (Chapter 3)** The first main-content chapter of the thesis, this chapter aims to explore the readiness of the field of temporal causal discovery to be applied to autonomous-agent data — in doing so addressing

RQ-1. An overview of the formalisms associated with causality is presented, a scenario in the AV domain is introduced, and a multitude of contemporary methodologies are described before benchmark experiment data is provided and discussed.

**Identifying Agent Behavioural Interactions via Theory of Mind (Chapter 4)** This chapter builds upon the findings of Chapter 3 by both framing the behavioural-interaction identification task in terms of discrete actions rather than continuous variables, and approaching the task from a theory-of-mind perspective — both with the goal of answering RQ-2. After presenting a counterfactual simulation method grounded in a theory of mind, the method is evaluated quantitatively and qualitatively against the results of Chapter 3.

**Extending Structural Causal Models for Autonomous Agents (Chapter 5)** This chapter explores the limitations against being able to utilise causal modelling effectively for autonomous embodied systems from a practical standpoint. While the theory-of-mind counterfactual simulation approach presented in Chapter 4 represented a substantial step in the right direction, it was ultimately more closely grounded in game theory than true causality. Namely, by intervening upon both potential causes and effects, any causal ancestry between the two is severed. This approach was largely the result of practical limitations at the time, limitations that are overcome in this chapter by extending causal model formalisms to better support autonomous embodied systems — in line with RQ-3.

**Accounting for Variations in Agent Motivating Factors (Chapter 6)** Having introduced new formalisms with which to model autonomous embodied systems in Chapter 5, it became possible to utilise these formalisms with the theory-of-mind behavioural-interaction identification strategy established in Chapter 4. Furthermore, by utilising the AV case-study architecture from Chapter 5 it became possible to utilise a planner during counterfactual inference. As part of this planner integration a means of approximating agent motivations via an Inverse Reinforcement Learning (IRL) inspired technique is described — aiming to tackle RQ-4. Finally this second iteration of techniques grounded in theory of mind is evaluated both quantitatively and qualitatively and compared against the first iteration.

**Conclusion (Chapter 7)** A conclusion to the thesis that summarises the contributions of the work, along with any significant findings, and discusses these with consideration for the future of research in this field.

## 1.4 Publications

In the course of carrying out the work described in this thesis there has been the opportunity to publish several papers, a summary of which is presented here.

### 1.4.1 Core Publications

These publications cover material which corresponds directly to the work described within this thesis. In the interest of transparency and proper attribution, these are listed as follows:

- The work corresponding to Chapter 3 was published in [18]:

  Rhys Peter Matthew Howard and Lars Kunze. 'Evaluating Temporal Observation-Based Causal Discovery Techniques Applied to Road Driver Behaviour'. In: *Proceedings of the Second Conference on Causal Learning and Reasoning*. Ed. by Mihaela van der Schaar, Cheng Zhang and Dominik Janzing. Vol. 213. Proceedings of Machine Learning Research. PMLR, Apr. 2023, pp. 473—498.

- The work corresponding to Chapter 4 was published in [19]:

  Rhys Peter Matthew Howard and Lars Kunze. 'Simulation-Based Counterfactual Causal Discovery on Real World Driver Behaviour'. In: *2023 IEEE Intelligent Vehicles Symposium (IV)*. 2023. DOI: 10.1109/IV55152.2023.10186705.

- The work corresponding to Chapter 5 was published in [20]:

  Rhys Peter Matthew Howard and Lars Kunze. 'Extending Structural Causal Models for Autonomous Vehicles to Simplify Temporal System Construction & Enable Dynamic Interactions Between Agents'. In: *Proceedings of the Fourth Conference on Causal Learning and Reasoning*. Ed. by Biwei Huang and Mathias Drton. Vol. 275. Proceedings of Machine Learning Research. PMLR, May 2025, pp. 1477–1505.

- The work corresponding to Chapter 6 was published in [21]:

  Rhys Peter Matthew Howard, Nick Hawes and Lars Kunze. 'Generating Causal Explanations of Vehicular Agent Behavioural Interactions with Learnt Reward Profiles'. In: *2025 IEEE International Conference on Robotics and Automation (ICRA)*. 2025, pp. 10416–10423. DOI: 10.1109/ICRA55743.2025.11127745

### 1.4.2 Adjacent Publications

These publications heavily relate to the themes and research fields of this thesis, yet do not correspond directly to the content discussed in this work. For the purpose of providing the reader research directly adjacent to this thesis, they are as follows:

- The author co-supervised two masters students and assisted in the development of their work, culminating in the publication of [22] and [23]:

  Richa Nahata, Daniel Omeiza, Rhys P M Howard and Lars Kunze. 'Assessing and Explaining Collision Risk in Dynamic Environments for Autonomous Driving Safety'. In: *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*. 2021, pp. 223–230. DOI: 10.1109/ITSC48978.2021.9564966.

  and

  Enrik Maci, Rhys P M Howard and Lars Kunze. 'Generating and Explaining Corner Cases Using Learnt Probabilistic Lane Graphs'. In: *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*. 2023, pp. 4201–4208. DOI: 10.1109/ITSC57777.2023.10422229.

- The author collaborated with colleagues from within the Oxford Robotics Institute, as well as academics from the University of York to work on the application of causality to drones operating in mines. This resulted in the publication of [24][2] and [25]:

  > Ricardo Cannizzaro, Rhys Peter Matthew Howard, Paulina Lewinska and Lars Kunze. 'Towards probabilistic causal discovery, inference & explanations for autonomous drones in mine surveying tasks'. In: *Proceedings of the 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) Workshop 'Causality for Robotics: Answering the Question of Why'*. 2023.

  and

  > C Imrie, R P M Howard, D Thuremella, N M Proma, T Pandey, P Lewinska, R Cannizzaro, R Hawkins, C Paterson, L Kunze and V Hodge. 'Aloft: Self-Adaptive Drone Controller Testbed'. In: *Proceedings of the 19th International Symposium on Software Engineering for Adaptive and Self-Managing Systems*. SEAMS '24. Lisbon, AA, Portugal: Association for Computing Machinery, 2024, pp. 70–76. ISBN: 9798400705854. DOI: 10.1145/3643915.3644107.

### 1.4.3 Tangential Publications

These publications do not correspond to the contents of this thesis, nor the general themes. Nonetheless they share the thesis' author as a first author, and were developed and published during the years of work that culminated in this manuscript. To fully document the academic pursuits of the author, primarily for the benefit of this thesis' assessors, they are as follows:

[2] Cannizzaro and Howard are joint first authors on this work

- The author began work on tracking deformable objects within the battery recycling domain while at the University of Birmingham Extreme Robotics Lab. These contributions were refined and built upon throughout the initial years of the doctorate, resulting in the publication of [26][3]:

  > Alireza Rastegarpanah, Rhys Peter Matthew Howard and Rustam Stolkin. 'Tracking linear deformable objects using slicing method'. In: *Robotica* 40.4 (2022), pp. 1188–1206. DOI: 10.1017/S0263574721001065.

- The author worked as part of a multi-institute team in the lead up to the start of the doctorate course and into the first few terms of said course. This work was centred around identifying novel visual features of geological or industrial interest for rover applications. This culminated in the publication of [27] and [28]:

  > Rhys Peter Matthew Howard, Sam Barrett and Lars Kunze. 'Don't Blindly Trust Your CNN: Towards Competency-Aware Object Detection by Evaluating Novelty in Open-Ended Environments'. In: *2021 IEEE International Conference on Robotics and Automation (ICRA)*. 2021, pp. 13286–13292. DOI: 10.1109/ICRA48506.2021.9562116.

  and

  > J Ocón, I Dragomir, F Cordes, R Dominguez, R Marc, V Bissonnette, R Viards, A C Berthet, G Reina, A Ugenti, A Coles, A Coles, A Green, R P M Howard and L Kunze. 'ADE: Enhancing Autonomy for Future Planetary Robotic Exploration'. In: *IAF Space Exploration Symposium 2021 at the 72nd International Astronautical Congress (IAC) 2021*. International Astronautical Federation (IAF). 2021.

[3] Rastegarpanah and Howard are joint first authors on this work

# Chapter 2

# Literature Review

The first step in laying the groundwork for this thesis is in comprehensively exploring and discussing the history and current state of the art when it comes to related work. The following sections aim to do exactly that, beginning with a general coverage of causality literature, and then moving on to see how this is actually applied to benefit autonomous agents. Lastly this review examines contemporary methods attempting to model behavioural interactions for AVs — i.e. the primary domain considered by this work.

## 2.1 Introduction to Causality

Before one can consider how causality might be applied in domains such as the AV and service-robotics domains, it is necessary to understand what is meant when one mentions causal techniques and how these methodologies came to be. This section aims to explore this, first in a general capacity, and then moving on to consider domains within which temporal information is of importance, as is typically the case in domains concerning autonomous agents.

### 2.1.1 Beginnings & Overview

Causal reasoning as a formal area of study began to take hold through the work of Pearl [4] based upon his earlier work upon Bayesian networks [29] and expanding upon previously established ideas such as Granger causality [30]. Generally speaking causal models provide information regarding the generative processes involved in the creation of data, rather than just the associations between variables / events.

Indeed causality as a field of study emerged in order to account for the shortcomings of traditional correlation-based techniques in statistics. This is primarily due to the potential for correlation to present itself between variables that have no effect on each other at all. Smith and Cordes [31] present a particularly extreme example in which they show a correlation between ice-cream sales and the number of murders across a few years. While there is nothing that would indicate that ice-cream consumption causes murder or visa versa, both variables happen to be affected by the climate. Hence the climate here is a confounding variable that has a causal effect upon both ice-cream sales and murder rates, and only causal models that incorporate this type of information can help autonomous systems account for such misleading associations. An example closer to the domains considered within this work can be seen in Figure 2.1. In this example, both the sensor readings (e.g. Light Detection and Ranging (LiDAR), cameras) and odometry of an autonomous vehicle are confounded by the weather affecting the vehicle and its operating environment. Failure to properly account for this while training a model with data featuring a variety of weather conditions could lead the model to build a false association between the sensor readings and odometry. Conversely if one wishes to build a system better able to handle domain shift, explicitly capturing the causal impact of the weather can more effectively enable this. However, it should be noted that within the context of this thesis the main draw of causality is from its links to cognition [5] and its ability to aid in formulating explanations. While aspects of the field such as handling confounding variables and identifying causal direction are important they are not the focus of this work. This is because in many cases the properties of the overarching task (e.g. temporal precedence) or individual scenarios (e.g. sparse causal relationships) minimise the issues associated with these traditional concerns.

In terms of how the field of causality has come to formally represent causal relationships, the two most popular formalisms are causal Bayesian networks and SCMs, both of which rely upon Directed Acyclical Graphs (DAGs). As mentioned previously, causal Bayesian networks evolved out of Pearl's earlier work with Bayesian networks [29]. Structurally speaking, a causal Bayesian network and a non-causal Bayesian network are identical, both relying upon a DAG formulation. The difference lies in that the DAG edges in a causal Bayesian network must be reflective of causal relationships, and not just conditional-independence, although in some cases

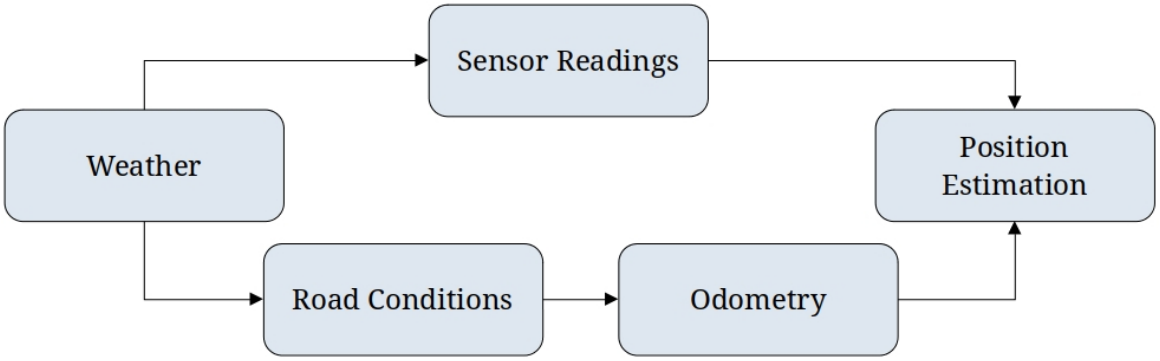



Figure 2.1: An example of the importance of causal modelling based upon the AV domain.

conditional independence can be used to infer such causal relationships. SCMs by contrast split the variables of their DAGs into exogenous and endogenous variables, in doing so creating a clear divide between the probabilistic and deterministic variables of the modelled system respectively. Typically causal Bayesian networks are preferred when performing causal reasoning over discrete events, while SCMs are used to causally reason about the interaction of numeric variables.

With the primary motivations and history behind the development of causal models established, these formulations of causality have been leveraged within big-data applications to better identify confounding factors, design experiments / studies, and ultimately discern useful causal relationships between the elements under consideration. Hence causality has found adoption within sociology [32], economics [33], medicine [34], and more besides. As it happens, approaches such as structural equation models and path analysis [35, 36, 37] were in use well before the beginnings of formal causality research and bear much resemblance to the causality formalisms later developed. Thus it can be argued that causal modelling as a statistical technique has a long history, even if its formal study only began in the 80s and 90s.

### 2.1.2 Causal Discovery

As interest in the use of causal models has grown approaches have been developed to discover causal models, with some of the first being the constraint-based Peter-Clark algorithm (PC) approach developed by Spirtes et al. [38] and the noise-based Linear Non-Gaussian Acyclical Model (LiNGAM) approach presented by Shimizu et al. [39]. A comprehensive review of such structural causal-discovery methods utilising observational data is provided by Glymour et al. [40]. Furthermore, while interventional approaches such as those proposed by Kocaoglu

et al. [41] exist, they are not always suitable for certain tasks or domains given they require experimentation. For example, the AV domain considered throughout much of this thesis is largely unsuitable for interventional techniques given a mixture of safety-concerns in online instances, and the inability to retrospectively intervene in offline post-hoc instances. With this in mind, interventional approaches will inevitably play a lesser role in this work, with preference given to techniques utilising observational data or counterfactual inference.

Note that some works that will be touched upon later utilise forms of simulation to gather causal information [16, 42]. While interventional techniques may be too risky to carry out in the real world for domains such as AVs, simulation provides a safe environment to do so. However, assuming that such intervention is done in order to better explain an observed scene post-hoc, such techniques can be considered counterfactual as opposed to purely interventional methods. This is because in reproducing the conditions of the observed scene to provide the simulation environment to make inferences, one has effectively abducted upon a number of exogenous variables. Thus such methods aim to examine the impact of the interventions 'all other things being equal' from the original scene, in effect matching the concept of counterfactual inference.

### 2.1.3 Causal Modelling & Discovery for Time-Series Data

Initial research utilising causality did not make attempts to explicitly encode or investigate the presence of time steps in data. In other words, while a specific event or variable may be captured at several points in time, these were effectively treated as distinct events / variables. However, given the frequency of time-series data in the domains within which causality has become popularised, there have naturally been extensions made to causality formalisms. A comprehensive documentation of these formulations is provided by Peters et al. [43].

Arguably the most notable shift comes in the approach one must take to carrying out causal discovery with such data though. Assaad et al. [44] provide an extensive survey of observational approaches applied to time-series data in addition to a quantitative evaluation of the approaches on synthetic and real data. The evaluated methods in question are the Granger-causality-

based PWGC[4][30], MVGC [45], and TCDF [46] approaches; the constraint-based PCMCI [47], oCSE [48], and tsFCI [49] approaches; the noise-based VAR-LiNGAM [50] and TiMINo [51] approaches; and the score-based DYNOTEARS [52] approach.

Beyond these more traditional categories NNs are increasingly being deployed to tackle temporal causal discovery. An immediate example of this is the TCDF method mentioned above, but other approaches such as the NAVAR method presented by Bussmann et al. [53] are also being explored. Lastly, an alternative avenue is being explored by the likes of Zhang et al. [54] and Huang et al. [55] that specifically aims to exploit causal non-stationarity — i.e. inconsistency in causal relationships across time — in order to carry out causal discovery.

## 2.2 Causality for Autonomous Agents

Having now given an overview of causality literature it is now possible to discuss how causal techniques can be applied to augment or extend the capabilities of autonomous agents. This discussion will begin with coverage of generic techniques that can be applied across autonomous-agent domains before shifting the focus to techniques developed around the AV domain, as these relate more closely to the contents of this thesis.

### 2.2.1 Non-Domain-Specific Applications

Multiple works have described the increasing need to incorporate causal models as a component of autonomous agents [12, 13] and particular attention has been applied to its use in XAI and robotics [13, 56]. Of relevance to this work are the sub-fields of algorithmic recourse [57] and temporal causal discovery [44]. Both of these sub-fields offer some level of XAI capabilities in the form of contrastive explanations and causal-structure identification respectively.

A popular use of causality is to describe the behaviour of autonomous systems at various levels of abstraction. At the highest level, Smith and Ramamoorthy [58] describe the application of overhead-image saliency analysis to allow causal reasoning on robot behaviour. At a more discrete level, Baumann et al. [59] extend upon existing work [60] by identifying the causal structures inherent to dynamic system using the controllability of certain variables to carry out interventions. Indeed there has been several works exploring the application of causality

[4]For the purposes of succinctness, several methods that will be discussed in greater depth later in the thesis are only introduced by their acronym here.

to ordinary differential equations, although these works do not necessarily relate directly to autonomous agents [60, 61, 62]. Lastly, there is other work which aims to give robots an understanding of the behaviour of other agents (e.g. humans) through causal modelling [63].

Some approaches aim to utilise causality to augment their methodology, either using predefined models or by simultaneously learning models alongside their use. Here causality has attracted attention from the Reinforcement Learning (RL) community, who aim to develop methods that can balance exploring to learn a causal model and exploiting it [14, 64]. While this makes these techniques applicable across a number of domains, it may not be suitable for high-risk applications such as autonomous driving. Within planning there exists works that fall outside the remit of RL. For example, work by Cannizzaro and Kunze [11] explores combining SCMs and POMDPs for the purposes of probabilistic robot planning under confounded decision-making. Other works attempt to actively learn causal models as part of their overall pipeline [65, 66] with goals of avoiding confounding factors, augmenting performance, and reducing the impact of domain shift. However, both of the cited methods rely upon black-box architectures that lack transparency and require large amounts of data, which may make them unsuitable for some autonomous-agent domains.

Another field of research that has been touched upon is utilising existing causal models to aid in discovering root causes or providing explanations. To this end Ibrahim et al. [67] present a semi-autonomous method for combining a mixture of fault / attack models and domain knowledge in order to provide a holistic causal model. This in turn supports the derivation of explanations for the cyber-physical systems upon which they focus. In the same field, Diehl and Ramirez-Amaro [10] provide a method entirely focused upon using an existing causal model to produce explanations for failures by comparing the observed variable values with the closest variable values that produce a non-failure outcome. They have since built upon their previous work on post-hoc explanations by presenting a system that anticipates failures and takes corrective actions [68]. Other works take a slightly different approach, focusing on responsibility attribution for agents producing a given outcome [69, 70]. In a similar thread Foerster et al. [71] describe a counterfactual multi-agent RL method that approximates 'difference rewards' by determining how a particular course of action for a given agent contributes to the overall

scene reward versus a default course of action. While their work certainly draws inspiration from causal conceptions of counterfactual inference, the work does not directly utilise any causal formalisms nor does the work contextualise itself within causality literature. Finally, Nashed et al. [72] integrate MDPs and SCMs as a means of providing causal explanations for the actions of agents operating in a grid-world environment.

The final notable area causality has found use in autonomous agents is in augmenting their observation and state-representation capabilities. In the realm of computer vision, Zhang et al. [73] provide an overview of computer-vision tasks and the utilisation of causality within them. A primary reason for applying causality here is to minimise the risk of spurious correlations between variables [74]. Outside of computer vision, causality has also been discussed or applied to problems involving the epistemic [75] and aleatoric [76] uncertainty associated with observations.

### 2.2.2 Autonomous-Vehicle Applications

Before discussing contemporary works integrating causality into AVs it should be noted that this section will focus more generally upon the intersections of these research fields rather than on the behavioural-interaction modelling that is the focus of this thesis. Related work more specifically concerned with the topic of behavioural interactions will be discussed below in Section 2.3 such that the work in question can be discussed in greater detail.

One area of research bridging the AV domain and causality is the creation of datasets for benchmarking. McDuff et al. [77] present 'CausalCity', a technique for generating semi-realistic data exhibiting vehicle-following behaviour. This work is useful in that it evaluates three causal-discovery techniques: the Neural Relational Inference (NRI) method [78] which operates on Two-Dimensional (2D) particle systems and motion-capture data; the Neuro-Symbolic Dynamic Reasoning (NS-DR) method [79] that presents itself by testing on Three-Dimensional (3D) particle systems; and the Visual Causal Discovery Network (V-CDN) method [80] which consists of an entire pipeline dedicated to detecting key points on clothing and determining the physical relationships between said points with a causal model. All three of these methods can be considered visual causal discovery in that they are restricted to working upon visual input. Furthermore, due to all three of them utilising neural networks it decreases the transparency of how their output is derived and increases the risk of performance degradation from domain

shift. In terms of the generated data itself, the causal relationships between the agents depicted in the data reflect coordinated following behaviour between two AVs rather than the type of spontaneous interactions considered throughout this thesis. Sun et al. [81] utilises causality to help refine data used in benchmarking motion-prediction techniques by attempting to identify 'causal agents'. However, the work does rely upon human labelling and thus the criteria for which agents are considered causal is not entirely clear. This is not necessarily a problem in itself, but does make it hard to differentiate poor method performance from what might just be a difference in criteria of what constitutes a causal link between agents.

In contrast to attempting to create AV datasets in order to evaluate causal methodologies, some works aim to exploit causal models in order to better perform AV-domain data generation tasks. For example, Ding et al. [82] utilise causal models for the purposes of better generating data for driving scenarios, given that causal models better represent inter-variable relationships than associative probabilistic models alone.

Other work has attempted to utilise causal discovery as part of a wider process within AVs, for example de Haan et al. [83] integrate causal discovery as part of imitation learning operating on human-driving behaviours. Meanwhile the work of Lin et al. [84] integrates causal representations into an RL pipeline for AV decision making in an attempt to reduce the impact of domain shift. Additionally, some work considers the influence image data has on decision making in vehicles [85, 86], utilising concepts such as visual saliency. However, this is only relevant to camera-based systems whereas this work primarily focuses upon abstract state data devoid of sensor-specific information. Despite this, there is importance in developing explainability techniques for camera-based sensors [87], particularly given the rise of deep NNs in AV systems which effectively amount to black boxes.

Lastly, while the type of causality they explore does not align with the established practices of the field, Thomas and Groth [88] nonetheless present an interesting discussion of causality in AVs. They propose a framework that satisfies the requirements presented by several existing standards. This makes the work valuable in terms of building a practical system, even if the framework in its current form does not use the kind of causal reasoning popularised by Pearl [4].

## 2.3 Modelling Behavioural Interactions for Autonomous Agents

With the general discussion of causality applied to autonomous agents given above there remains one aspect of literature which requires yet further discussion, namely those existing methodologies that aim to model behavioural interactions between agents — with an emphasis placed on AVs. Given the proximity and overlap between such literature and the work presented here, a deeper examination of these contributions to the field is required for this review to be comprehensive.

### 2.3.1 General Methodologies

While certainly not a requirement for modelling agent behavioural interactions, a frequent approach throughout literature in this area is to integrate the modelling of other agents into their frameworks. The rationale follows that by being able to model each agent within a given scene one can build an understanding of how the agents influence each other. Indeed this logic is applied throughout the latter portions of this thesis, whereby the planning and control systems of multiple vehicles are emulated at once in order to try and infer causal relationships. Given that the modelling of other agents has a long history dating back to conceptions of intentionalism [89], not to mention applying to a myriad of applications and domains, the reader is referred to the thorough survey of such literature provided by Albrecht and Stone [90].

In contrast to techniques utilising agent modelling others describe agent interactions purely in terms of spatio-temporal relationships. Often utilising measures such as distance, relative speed, or Time To Collision (TTC), these inter-agent metrics have been used to estimate risk associated with interactions [22, 91] and predict future agent behaviour [23, 92, 93]. In particular Helbing contributed towards the development of both Intelligent Driver Models (IDMs) [93] and social-forces models [92], which laid the groundwork for much research into agent interactions within the vehicular- and pedestrian-traffic domains respectively. However, such approaches typically do not claim to identify causal relationships between interaction participants as this is not always necessary to achieve their goals.

Some works considering spatial-temporal relationships employ similar ideas to causal methods, such as the work of Bellotto et al. [94], Dondrup et al. [95], and Mghames et al. [96], all of which base their work upon qualitative trajectory calculus. This method relies upon identifying

inter-agent-metric trends that can then be used for prediction and planning. The work also shares a similar goal to the previously-described service-robotics scenario of improving the social awareness of agent planning. Where this thesis differs from this work is the assumption of action-based interactions later on, in addition to placing greater emphasis upon the internal cognition of agents.

Many works that aim to approach agent interaction by modelling the agents themselves utilise techniques from IRL, either in order to better understand the motivations of agents or to go further and attempt to imitate them. For a survey of the extensive field of IRL the reader is referred to the work of Arora and Doshi [97], as the contents of this literature review focus upon the application of IRL to agent interaction modelling specifically. For example, Wen et al. [98] utilise IRL to determine how the behaviour of human-driven vehicles differs in the presence of AVs versus other human-driven vehicles. By contrast, Kuderer et al. [99] do not consider agent behavioural interactions but do propose a means of capturing the variations in behaviour across various agents in the context of AVs. Their work is mainly concerned with individual preferences in driving style, however one can extend this concept to how drivers might alter their behaviour depending upon context. Indeed a combination of these concepts inspire some of the work in the latter part of this thesis. These sections of the thesis are concerned with identifying reward parameters that reflect case-by-case observed agent behaviour. This is because relying upon approximated or trained parameters may overestimate the caution with which agents act which may not match reality, particularly in edge cases [100].

Other works apply methods more rooted in game theory and cognitive science, which mirrors the type of approaches considered later in this thesis. For example, Ma et al. [101] describe a pedestrian model in which each pedestrian makes movement decisions based upon their prediction of how other pedestrians will act. In a similar thread Rahimi et al. [102] describe a pedestrian trajectory prediction system that captures the causal influence of each agent upon other agents. This allows one to make counterfactual queries regarding how each agent within an observed scene may have acted had a given agent not been present. Lastly Camara et al. [103] utilise game theory to describe two pedestrians approaching an intersection approach. Each agent wishes to move as fast as possible while avoiding collision with the other agent,

thus the situation resembles the well established prisoner's dilemma [104] problem. The work tackles this problem with a Gaussian regression approach that calibrates the reward weighting for each agent that dictates their action selection. This facilitates behaviour prediction for the considered agents in counterfactual scenarios, however it should be noted that the methodology in question does not align itself with the body of established causality literature.

A key facet of this thesis is that the behavioural interactions it concerns itself with occur spontaneously, and that the agents involved may possess different internal models and motivations. This somewhat mirrors the concept of 'zero-shot coordination' as described by Hu et al. [105] whereby 'agents are placed into a cooperative situation with a novel partner and must quickly coordinate if they wish to earn high payoffs'. The key difference between this work and the aforementioned literature is that here there is no assumption of coordination, with this thesis often exploring cases where agents act with malicious intent or careless disregard for others.

In terms of approaches that do attempt to deal with these circumstances there are of course the inter-agent metric based approaches cited above [22, 91], however more complex approaches exist. For example, the proposed methods of Debbarma et al. [106] and Song et al. [107] both utilise information regarding the behaviour of a human driver within their vehicle to identify risk. Although, in order to integrate this information in behavioural-interaction modelling vehicle connectivity and specific sensors are required, raising privacy and hardware consistency concerns. Other approaches integrate external factors instead of [108, 109, 110] or in addition to [111, 112] vehicle-interior information. Such factors include driving manoeuvrers and lane deviations, yet these are typically only considered from an ego-agent perspective.

### 2.3.2 Causality-Based Techniques

One of the closer works in terms of content to the early portions of this thesis is presented by Maier et al. [42]. They use SCMs to describe a two-vehicle braking scenario and carry out causal inferences using it, specifically in the context of Advanced Driver-Assistance Systems (ADASs). The SCM in question is highly tailored to the scenario in question, and incorporates factors such as fog, rain, time, and traction. From here the structural equations associated with the variables are either learnt via Bayesian networks, or rely upon the Car Learning to Act (CARLA) simulator [113] to provide information on such variables. Overall this represents a promising

application of causality to an AV scenario, however it is specific to this scenario and is egocentric. That is to say that the model performs well for an ADAS system, yet it is specific to two-vehicle braking scenarios and cannot make inferences from the perspectives of other agents.

Alternatively Gyevnar et al. [16] offers some promising results in terms of offering natural language explanations for action selection based upon the counterfactual-effect-size model [114]. They split causation into teleological and mechanistic perspectives that aim to describe the actions of an ego-agent in terms of its goals and the actions of surrounding agents respectively. The framework takes a query action and counterfactually simulates new outcomes by starting from an earlier time step and utilising Monte Carlo tree search to probabilistically sample different behaviour. These counterfactual outcomes are then labelled based upon whether the queried outcome occurred in the sampled timeline. From here the teleological causation is established by comparing the ego-agent's rewards between counterfactual timelines that did and did not exhibit the queried action. Meanwhile the mechanistic causation is established by fitting a classifier that takes the features of other agents as input, and outputs whether or not the queried action should be expected. In theory one can then interpret the importance attributes of the classifier as indicating a potential mechanistic cause. Lastly, the discovered causes are converted into simple English linguistic terms for conveyance to a user. Similar to the previous approach this work can be considered largely egocentric in nature, although in this case the focus is more upon providing natural language explanations, a task which it appears to handle effectively. Importantly, their work demonstrates the capacity to scale to a larger number of AVs without issue, however the simulations in question are 'only a few seconds long'. This raises questions as to how well the approach can deal with notable time-lags between cause and effect.

In terms of multi-AV interactions a promising approach is given by Tang et al. [115] that extends upon the already discussed NRI method [78]. The proposed approach — named Grounded Relational Inference (GRI) — builds upon NRI by training two decoders for the interaction graphs represented in latent space. The first decoder outputs a policy from which vehicle trajectories can be synthesised given an initial state, meanwhile the second decoder takes a trajectory and outputs a real number reflecting the reward associated with it. The overall framework is trained in an adversarial manner. This means the policy decoder attempts

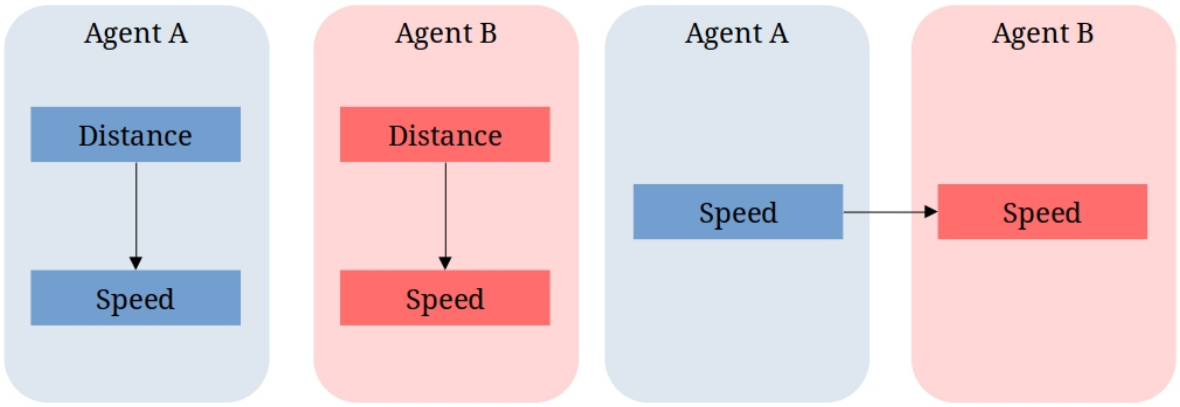
to synthesise trajectories indistinguishable from observed trajectories. Meanwhile the reward decoder attempts to approximate reward such that synthesised trajectories can be separated from observed trajectories by said reward. Overall the framework appears to perform well, achieving high accuracy rates across the synthetic and real-world data considered. However, it is worth noting that NRI actually outperforms GRI in terms of minimising the error associated with synthesising trajectories. One aspect in which this approach could be limited is that since this is ultimately a data-driven method, it may struggle in cases where atypical behaviour or interactions are exhibited. Indeed this is a concern for subsequently discussed data-driven methods too, however the GRI method may be more susceptible here, given that the design of the reward decoder makes the assumption that agents work cooperatively to maximise the overall reward. While such an assumption of cooperation might work for the most part in typical circumstances, often the instances where a post-hoc analysis is of greatest interest are those with atypical behaviour where cooperation is absent (e.g. road rage, insurance fraud, drunk driving). Nevertheless, the promising results of GRI mean it may provide a foundation for further work, with said work ideally alleviating the risks regarding atypical interactions in the process.

In a similar thread, Xiao et al. [116] provide a review of graph-based techniques for identifying hazardous events in the context of AVs. The majority of such works consider agent interactions and spatio-temporal relationships of importance yet their modelling is typically not the focus of the work. Instead they aim to anticipate accidents [117, 118, 119] or predict agent motion [120, 121]. However, the aforementioned review does highlight the work of Li et al. [122] that utilises graph convolutional NNs in order to predict goal-oriented agent behaviour and identify potential causes of deviations from this behaviour. In order to do so the scene is described via generated 'Ego-Thing' and 'Ego-Stuff' graphs that reflect the interactions of other agents and environmental elements with the ego-agent respectively. While the quantitative results of the work indicate an improvement over its contemporaries, it is debatable as to the extent to which causal agent interactions are exploited. Removal of spatial-relationship information only resulted in a 0.2 % decrease in performance, while the other type of Ego-Thing modelling just concerns correlation between object appearances rather than behaviour. This seems to indicate that the approach did not learn significantly useful behavioural-interaction information.

Additionally, this work is set apart from the methodology this thesis focuses upon in that: a) it lacks transparency due to heavy NN usage; b) it is vision-based; and c) it is once again egocentric. As such, while the work does indeed bear merit, it is ultimately applicable in different circumstances and for different tasks than the contributions of this thesis.

In terms of literature on pedestrian-agents Castri et al. has made several contributions beginning with the application of PCMCI to model causal relationships between egocentric agent variables [123]. The main motivation behind this was to model agent-object and agent-environment interactions, with the goal of better understanding interactions between pedestrian agents in domains such as warehouse automation. This was subsequently built upon by utilising transfer entropy to help filter out variables deemed unnecessary to describe agent behaviour [124]. They also proposed a method that would use an intervention-based approach to refine a discovered causal model in the event said model became inaccurate due to causal non-stationarity [15]. More recent work of theirs has focused upon the proposal and subsequent evaluation of their 'ROS-Causal' framework [125, 126]. The contributions of this work being to provide a set of causal-discovery tools integrated with the Robot Operating System (ROS) framework and a simulation environment with which to test them. Generally speaking, this collection of work has done a great deal to explore the application of causal-discovery methods to autonomous-agent data, and indeed there is some overlap between this and the area of work described in the earlier chapters of this thesis. However, the general focus of the previously cited works is upon intra-agent causal models, whereas this thesis focuses primarily upon inter-agent causal models, even if both offer a means of capturing behavioural interactions between agents.

Before continuing, to clarify the distinction between intra-agent and inter-agent causal models consider Figure 2.2. Both subfigures depict agents A and B where it is assumed that agent B is following agent A at some distance and thus the speed of agent A causally affects the speed of agent B. Each subfigure represents either an intra-agent or inter-agent causal representation of this scenario. In Figure 2.2a the causal relationships are strictly between variables specific to a given agent, whereas in Figure 2.2b there are causal relationships present between variables in different agents. The former can tell one more information regarding the inner workings of individual agents and are generally easier to transfer between scenarios, however the latter




(a) Intra-Agent Causal Models (b) Inter-Agent Causal Models

Figure 2.2: Illustration of how intra- and inter-agent causal models differ.

is needed to describe specific agent interactions. This is primarily because in order to utilise intra-agent causal models to explain specific scenarios one would still need to identify how the behaviour of one agent influences the root cause variables of other agents. For example, in the aforementioned scenario one would need to discover how the speed of agent A influences the distance variable of agent B, in essence forming an inter-agent causal model.

Other methods that integrate causality yet focus less specifically upon AVs include the contributions of Fujii et al. [127]. They aim to model and discover the causal attributes of movement interactions between animals. To do this they describe a framework that combines multi-layered perceptrons in a manner guided by existing theoretical models within biology — specifically movement ecology. These are then trained upon data, while utilising a loss function that incorporates a regularisation term that once again leverages existing biology-domain knowledge. In order to allow the extraction of causality information, the paper builds upon previous work that attempts to extract Granger causality from NNs [128] by apply the underlying principles of self-explaining NNs [129]. Such NNs are structured with the aim of maximising the explicitness, stability, and faithfulness of the networks. As such, after applying the learning process to animal-trajectory scene data it then becomes possible to attempt discovery of Granger-causality relationships by applying statistics metrics to the learnt network parameters. The work stands apart from this thesis in that it considers continuous relationships between the animals, and only considers simple interactions such as attraction or repulsion in movement. Nonetheless it demonstrates the wide-spread applications of and interest in such modelling techniques.

Some methods do not directly attempt to model agent behavioural interactions but instead consider adjacent tasks. For example, Tang et al. [130] attempt to identify instances where models trained on driving-trajectory data over-estimate the ability of non-ego-agents to predict an ego-agent's behaviour. They identify the source of this being that the overall predictive model has access to the future trajectory of the ego-vehicle, while realistically the non-ego-vehicles do not. Thus in order to identify instances of this occurring they measure the extent to which the future ego-agent trajectory influences the model's prediction of non-ego-agent trajectories. If this measure is particularly high it is indicative of a case in which the prediction of non-ego-agent trajectories is reliant upon future ego-agent trajectory data, and thus potentially unreliable.

Lastly, there has been some work that examines the theoretical underpinnings of intention recognition as a component of agent interaction [131]. Given this work has focused upon these interactions from a human-computer-interaction perspective, and a mostly theoretical one at that, the overlap with this thesis is mostly in relation to its motivations and philosophy.

## 2.4 Discussion

Before proceeding to the main-content chapters of this thesis, it is first worth briefly revisiting those works closest to the avenues of research considered here in order to contextualise the contributions that will be presented in the following chapters. The works identified as falling closest to the contents of this thesis are those of Maier et al. [42], Tang et al. [115], and Gyevnar et al. [16]. First and foremost this is because all these works are causal in nature and focus on the AV domain and thus attempt to model types of behavioural interaction that are mostly similar to this thesis.

For Maier et al. [42] the biggest limitation is that their approach assumes a fixed ADAS scenario which their entire causal model is built around. In the context of their work this is a strength, as they are not interested in modelling other scenarios and this well defined causal model instead allows them to focus on additional details such as weather conditions. By contrast the work of this thesis aims to present methods that can be applied in a variety of scenarios.

Tang et al. [115] on the other hand do support a wide range of scenarios, but their approach relies upon the use of graphical NNs. While this is not necessarily an issue in itself, it does mean that their framework is a black box in places that offers less transparency than the methods

presented in the latter parts of this thesis. Ultimately this is a reflection of different philosophies, as it is possible that their method can more accurately identify behavioural interactions but in doing so it cannot offer insight regarding the process used to make these identifications.

Lastly Gyevnar et al. [16] offer some of the most promising work of all compared with the contributions of this thesis. The approach offers greater transparency than the work of Tang et al. [115] yet is not fixed to a single scenario like Maier et al. [42]. Currently the main limitations associated with their work is the previously mentioned small time horizon of simulations along with the predominant focus of their work being upon the success of their textual explanations as evaluated by two user studies. By contrast this work examines interactions occurring over a longer timespan (e.g. $10\,s$) and its evaluation focuses upon the accuracy of output causal-summary graphs. Importantly the work of this thesis also primarily evaluates itself against real-world data, while Gyevnar et al. utilise synthetic data generated based upon scenarios described by Albrecht et al. [132].

Taking all this into consideration this work represents one of the first of its kind to merge the fields of causal reasoning and autonomous agents for the purposes of describing behavioural interactions. This is particularly true if one narrows the scope to AVs. One can observe this in the fact that all the publications discussed above are contemporary with this work, with all having been published after the start of research relating to this thesis. That is to say that the contributions of this work constitute some of the first to a particularly novel area of study and indeed there is still much work yet to be done within this niche.

# Chapter 3

# Adapting Causal Methods to Model Agent Behavioural Interactions

With the motivations and goals of this thesis laid out in Chapter 1, and the literature of the field discussed in Chapter 2, one can begin to formally define the research problem associated with this work. In this chapter existing formalisms for capturing time-series causal relations are introduced (Section 3.1) and the efficacy of contemporary causal-discovery approaches applied to agent data are evaluated (Section 3.2).

## 3.1 Causal Relationships Between Agent Proxy Variables

In this section the basic formalisms associated with causal modelling, and time series in particular, are introduced. This is followed by a discussion of how these might be utilised to capture behavioural interactions within an autonomous-agent domain.

### 3.1.1 Structural Causal Models

The basic structure utilised by many works within causality literature is the DAG.

**Definition 1** (Directed Acyclic Graph)**.** A directed acyclic graph $G$, as defined by Thulasiraman and Swamy [133], is a tuple:

$$G = \langle N,\ E \rangle \tag{3.1}$$

with its elements defined as follows:

- $N = \{N_0, N_1, ...\}$ is a set of nodes / vertices which within the context of causality could refer to variables or events.

- $E \subset N \times N$ is a set of directed edges defined as the edge going from the first element of the tuple, to the second element of the tuple. In the context of causality these can represent the causal relationships themselves.

The directed graph structure allows one to define a function to derive the set of the direct parents of a node $Pa(\cdot) : N \to 2^N$ and a function to derive the set of all the ancestors of a node $An(\cdot) : N \to 2^N$:

$$Pa(N_i) = \{N_j | \langle N_j, N_i \rangle \in E, N_j \in N\} \tag{3.2}$$

$$An(N_i) = Pa(N_i) \cup \left( \bigcup_{N_j \in Pa(N_i)} An(N_j) \right) \tag{3.3}$$

The final quality of a DAG is acyclic nature, which requires that the directed graph be devoid of cycles, which is formally defined here as follows:

$$\bigvee_{N_i \in N} N_i \notin An(N_i) \tag{3.4}$$

Note that in the event of a cycle being present this will not be possible to algorithmically compute via the aforementioned equations, as it will recurse infinitely. However, given that sets do not contain duplicate elements, an analytical solution is available, which is sufficient to demonstrate the acyclic quality of the graph.

While DAGs are useful as high level tools for describing causal systems, a more detailed causal-model structure is provided by SCMs.

**Definition 2** (Structural Causal Model)**.** Here a slightly modified conception of structural causal models is proposed. This is primarily based upon the SCMs described by Bareinboim and Pearl [134] which in turn merge the original SCM and probabilistic causal model formalisms proposed by Pearl [4]. The primary change made here is to integrate the underlying DAG associated with the SCM as part of the SCM definition, such that the causal model is fully described by said definition. Thus, a structural causal model $M$ is a tuple:

$$M = \langle U,\ V,\ E,\ F,\ P(U) \rangle \tag{3.5}$$

with its elements defined as follows:

- $U = \{U_0, U_1, ...\}$ is a set of exogenous variables that capture factors external to the model.
- $V = \{V_0, V_1, ...\}$ is a set of endogenous variables that derive their values as a function of other variables.
- $E \subset (U \cup V) \times V$ is a set of directed edges that indicates the first variable of the tuple is an input for the second variable of the tuple — which is endogenous by definition.
- $F = \{f_0, f_1, ...\}$ is a set of structural equations that derive the values of endogenous variables. Assuming $f_i$ is the structural equation associated with endogenous variable $V_i$, then $f_i$ will map from the domain of $Pa(V_i)$ to the domain of $V_i$. Note, that $Pa(\cdot)$ and $An(\cdot)$ are redefined for the SCM formalism, and as a result of the definition of $E$ it is impossible for an exogenous variable to have parents.
- $P(U)$ is the joint probability distribution over the domains of the exogenous variables $U$. It is required that all exogenous variables be independent of each other, or in other words:

$$\bigvee_{U_i \in U} \bigvee_{U_j \in U \setminus \{U_i\}} P(U_i, U_j) = P(U_i)P(U_j) \tag{3.6}$$

With this definition of an SCM one now has a means of representing causal relationships between variables that is the standard among most literature regarding causality. However, autonomous agents by their inherent nature should consider time as a factor, as they exist in a continuously evolving environment, in contrast to the types of experiment favoured by big-data applications (e.g. sociology, economics, medicine, etc.).

### 3.1.2 Temporal Extension of SCMs

While the exact means of representing temporal SCMs varies, the consensus within causality literature is to operate under the assumption that associated data is in time-series form, i.e. a time-indexed series of data points taken at equal intervals. Peters et al. [43] provide a comprehensive overview of this topic, from which the following formalisms are presented.

An important property which is almost universally assumed when considering time in causality is temporal precedence.

**Definition 3** (Temporal Precedence)**.** Temporal precedence requires that cause always come before effect. In other words, given time steps $T \subseteq \mathbb{N}$, if an endogenous variable $V_j$ at time $t' \in T$ is affected by the value of variable $N_i$ at time $t \in T$ then one can guarantee $t \leq t'$. This assumption is sometimes taken further to mean $t < t'$, however this assumption should only be made provided causal effects cannot occur in the time between time steps. Thus the strictness of temporal precedence usually depends upon the variables and temporal resolution in question.

With temporal precedence established as a concept, the most basic means of capturing time-series causal relations is the full-time-graph SCM, which captures all potential causal links between variables for every time step.

**Definition 4** (Full-Time-Graph SCM)**.** For time steps $T$ a full-time-graph SCM, based upon Definition 2 and Peters et al. [43], is a tuple:

$$M^T = \langle U^T,\ V^T,\ E^T,\ F^T,\ P(U^T) \rangle \tag{3.7}$$

with its elements defined as follows:

- $U^T \subseteq U \times T$ is a set of exogenous variables indexed to each time step in $T$. $U^T_{i,t} \equiv \langle U_i, t \rangle \in U^T$ refers to exogenous variable $U_i$ indexed to time $t$.
- $V^T \subseteq V \times T$ is a set of endogenous variables indexed to each time step in $T$. $V^T_{i,t} \equiv \langle V_i, t \rangle \in V^T$ refers to endogenous variable $V_i$ indexed to time $t$.
- $E^T \subset (U^T \cup V^T) \times V^T$ is a set of directed edges that indicate the first time-indexed variable of the tuple is an input for the second time-indexed variable of the tuple — which is endogenous by definition. Importantly the temporal-precedence assumption from Definition 3 means that the following condition must hold:

$$\bigvee_{\langle N_{i,t}, V_{j,t'} \rangle \in E^T} t \leq t' \tag{3.8}$$

- $F^T = \{f^T_{0,0}, f^T_{0,1}, ..., f^T_{1,0}, f^T_{1,1}, ...\}$ is a set of time-indexed structural equations that derive the values of time-indexed endogenous variables. Assuming $f_{i,t}$ is the structural equation

associated with endogenous variable $V_i$ at time $t$, then $f^T_{i,t}$ will map from the domain of $Pa^T(V^T_{i,t})$ to the domain of $V^T_{i,t}$. The full-time-graph variants of $Pa^T(\cdot) : N \times T \to 2^N$ and $An^T(\cdot) : N \times T \to 2^N$ are defined as follows:

$$Pa^T(N^T_{i,t}) = \{N^T_{j,t'} \mid \langle N^T_{j,t'}, N^T_{i,t} \rangle \in E^T,\ N^T_{j,t'} \in N^T\} \tag{3.9}$$

$$An^T(N^T_{i,t}) = Pa^T(N^T_{i,t}) \cup \left( \bigcup_{N^T_{j,t'} \in Pa^T(N^T_{i,t})} An^T(N^T_{j,t'}) \right) \tag{3.10}$$

where $N^T = U^T \cup V^T$, and $N^T_{i,t} \equiv \langle N_i, t \rangle \in N^T$ refers to variable $N_i$ indexed to time $t$.

- $P(U^T)$ is the joint probability distribution over the domains of the time-indexed exogenous variables $U^T$. Note that there is not a requirement for the distribution to remain the same across time under this temporal SCM definition, so long as the exogenous variables of $U$ are independent of one another.

It is worth noting that the underlying data type used for time indexing data need not be based upon natural numbers, so long as a bijective mapping between the data type and $T$ is possible. Additionally, in a non-temporal SCM a variable cannot typically have itself as a parent, as this would create a graphical cycle, and many causal techniques rely upon the underlying graph of an SCM being acyclic. However, within temporal SCMs so long as $t < t'$ one may have $N_{i,t}$ as a parent of $V_{i,t'}$, as this just implies that a variable causally affects its own future state.

A full-time-graph SCM allows for the maximum expressiveness in terms of being able to indicate one variable affecting another over a specific time lag. However, such a representation is not spatially efficient, with the maximum number of causal links for a full-time-graph SCM having complexity $\mathcal{O}(\,|T|^2\,|V|\,(|V| + |U|)\,)$. Typically full-time-graph SCMs will have a sparse population of causal links, but they nonetheless remain inefficient for representing consistent time-lagged causal relations, as the causal links must be replicated across time.

Thus there are alternate means of representing causal relations that are more spatially efficient, at the cost of expressiveness. The most simple example of this is the causal-summary graph and its variant the causal-adjacency graph.

**Definition 5** (Causal-Summary Graph)**.** A causal-summary graph, based upon Definitions 2 and Peters et al. [43], can be constructed from a full-time-graph SCM $M^T = \langle U^T,\ V^T,\ E^T,\ F^T,\ P(U^T)\rangle$, and is a tuple:

$$\mathcal{G}^S = \langle N^S,\ E^S\rangle \tag{3.11}$$

with its elements defined as follows:

- $N^S = \{U_i \mid U_{i,t}^T \in U^T\} \cup \{V_i \mid V_{i,t}^T \in V^T\}$ is a set of the combined exogenous and endogenous variables with any time-indexing information removed.
- $E^S = \{\langle N_i, V_j\rangle \mid \langle N_{i,t}^T, V_{j,t'}^T\rangle \in E^T\}$ is a set of directed edges which indicate the first variable of the tuple has a causal effect upon the second variable of the tuple — which is endogenous by definition — at some point in time.

One could hypothetically construct a causal-summary graph without a source SCM, however the resulting causal-summary graph would have little in the way of applications and would not summarise a temporal SCM as the name suggests. Importantly while temporal SCMs can be converted into causal-summary graphs, causal-summary graphs are not SCMs as they lack any information regarding structural equations or probability distributions. Thus they are effectively just DAGs and are typically used for high-level illustration, inspection, and evaluation.

**Definition 6** (Causal-Adjacency Graph)**.** A causal-adjacency graph consists of the undirected graphical version of a causal-summary graph. While somewhat counterintuitive, this graphical form can be useful in scenarios where the ground truth only specifies that there is a causal relationship between two variables, but not the direction of causation.

The third commonly used temporal causal representation is the window-time-graph SCM, which aims to describe relative causal relationships up to a Maximum Time Lag (MTL) $\tau$. Originating out of a desire to describe time-lagged Granger causality [135, 136], window time graphs have become a popular output format within causal discovery [44, 137]. Causal discovery pertains to methods aiming to identify a mixture of graphical structure and structural equations for an underlying SCM, and thus offers a good starting point for identifying causal relationships indicative of agent behavioural interactions.

**Definition 7** (Window-Time-Graph SCM)**.** For a set of relative time lags $L^\tau = \{l \mid l \leq \tau,\ l \in \mathbb{N}\}$ a window-time-graph SCM, based upon Definitions 2 and Runge [137], is a tuple:

$$M^\tau = \langle U^\tau,\ V^\tau,\ E^\tau,\ F^\tau,\ P(U) \rangle \tag{3.12}$$

with its elements defined as follows:

- $U^\tau \subseteq U \times L^\tau$ is a set of exogenous variables indexed to each relative time lag in $L^\tau$. $U^\tau_{i,l} \equiv \langle U_i, l \rangle \in U^\tau$ refers to exogenous variable $U_i$ indexed to relative time lag $l$.

- $V^\tau \subseteq V \times L^\tau$ is a set of endogenous variables indexed to each relative time lag in $L^\tau$. $V^\tau_{i,l} \equiv \langle V_i, l \rangle \in V^\tau$ refers to endogenous variable $V_i$ indexed to relative time lag $l$.

- $E^\tau \subset (U^\tau \cup V^\tau) \times V^\tau$ is a set of directed edges that indicate the first time-lag-indexed variable of the tuple is an input for the second time-lag-indexed variable of the tuple — which is endogenous by definition. Similar to the full-time-graph SCM the temporal-precedence assumption from Definition 3 means that the following condition must hold:

$$\bigvee_{\langle N_{i,l}, V_{j,l'} \rangle \in E^T} l \leq l' \tag{3.13}$$

- $F^\tau = \{f^\tau_0, f^\tau_1, ...\}$ is a set of structural equations that derive the values of time-lag-indexed endogenous variables. Assuming $f^\tau_i$ is the structural equation associated with endogenous variable $V_i$ across time, then $f^\tau_i$ will map from the domain of $Pa^\tau(V^T_{i,t})$ to the domain of $V^T_{i,t}$ for time $t$. The full-time-graph variants of $Pa^\tau(\cdot) : N \times T \to 2^N$ and $An^\tau(\cdot) : N \times T \to 2^N$ are defined as follows:

$$Pa^\tau(N^T_{i,t}) = \{N^T_{j,(t+l'-l)} \mid \langle N^\tau_{j,l'}, N^\tau_{i,l} \rangle \in E^\tau,\ N^\tau_{i,l}, N^\tau_{j,l'} \in N^\tau\} \tag{3.14}$$

$$An^\tau(N^T_{i,t}) = Pa^\tau(N^T_{i,t}) \cup \left( \bigcup_{N^T_{j,t'} \in Pa^\tau(N^T_{i,t})} An^\tau(N^T_{j,t'}) \right) \tag{3.15}$$

where $N^\tau = U^\tau \cup V^\tau$, $N^T_{i,t} \equiv \langle N_i, t \rangle \in N^T$ refers to variable $N_i$ indexed to time $t$, and $N^\tau_{i,l} \equiv \langle N_i, l \rangle \in N^\tau$ refers to variable $N_i$ indexed to relative time lag $l$.

- $P(U)$ is the joint probability distribution over the domains of the exogenous variables $U$. Again, the exogenous variables must be independent of one another. However, in contrast to a full-time-graph SCM, the distributions of each exogenous variable must be independent and identically distributed across time, as the SCM represents a sliding window rather than an absolute range of time steps.

Window-time-graph SCMs have a similar restriction to full-time-graph SCMs in that a variable can only have itself as a parent provided $l > 0$ in order to avoid cycles.

While this is otherwise close to the full-time-graph SCM in format, the crucial difference is that typically $\tau \lll |T|$, and thus window-time-graph SCMs are much more compact. Furthermore, while full-time-graph SCMs provide absolute causal relations between every time-indexed variable, window-time-graph SCMs specify relative time-lagged causal relations which are consistent across time. This is reflected in that while a full-time-graph SCM theoretically has a structural equation for each variable for every time step, i.e. $|F^T| = |V \times T|$, the window-time-graph SCM only has one for each variable, i.e. $|F^\tau| = |V|$. This relates to another property required for window-time-graph SCMs to effectively capture causal relations, causal stationarity [137].

**Definition 8** (Causal Stationarity)**.** Causal stationarity requires that the causal relationship between two variables for a given time lag remains constant throughout time. In Definition 7 this is achieved by defining the structural-equation set $F^\tau$ and causal-edge set $E^\tau$ such that regardless of where the sliding time window of size $\tau$ is placed across time steps $T$ the structural equations and causal edges remain accurate.

A good example of systems which exhibit strong causal stationarity are control processes, for example those in chemical or nuclear plants. In such systems, a change in control values will typically result in some causal effect upon other monitored variables albeit with a time lag. Since the causal relationships of the variables involved are purely dictated via physics they are likely to be consistent across time. By contrast, consider how the time of a person going to sleep causally affects the time that same person gets up. Intuitively it makes sense that when one sleeps strongly influences the time they will get up, but the time lag between these events is far from consistent. Furthermore, there may be instances where the time one gets up impacts when

one sleeps, thus reversing the direction of causality. As such in this example causal stationarity is likely either very weak, or absent entirely.

Overall this is a substantially more compact representation for cases where causal relationships remain consistent through time. The main reason causal discovery favours such a representation is due to observation-based methods being inherently data-driven. As such they rely upon establishing patterns seen consistently throughout the input time-series data.

### 3.1.3 Agent-Behavioural-Interaction Representation via Temporal SCMs

Much of what has been described so far relates to generic causal modelling techniques, not specific to autonomous agents. Thus the last step before one can begin examining how to identify behavioural interactions between agents is to describe how it may be possible to formulate such interactions in terms of the variables and causal relationships.

In order to contextualise the discussion of how to represent agent behavioural interactions it is helpful to select a suitable problem domain for analysis. This thesis primary focuses upon the AV domain, this is motivated by several points:

- Despite being a domain in which autonomous agents act, there is a general expectation that road vehicles follow a certain order dictated via a highway code, the road layout, and the presence of other vehicles. This in turn makes it easier to reason about behavioural interactions, in addition to simplifying the extraction of ground-truth causal relationships in some scenarios.

- There is a large number of datasets available for the AV domain given the amount of industry interest in the research area. The ability to evaluate proposed methodology on real-world data is valuable and not easily available in many other domains.

- There is immediate benefit to the domain in integrating identification of behavioural interactions via causal modelling. The AV domain presents a high-risk to human life, and thus being able to understand how road agents influence each other's behaviour is critical to furthering the development of safe AV technologies. Furthermore, in failure cases it is beneficial to be able to analyse scenes post-hoc in order to diagnose system issues and establish accountability [17].

Having established the motivations behind selecting the AV domain as the primary domain through which to contextualise behavioural interactions, the previously discussed scenario depicted in Figure 1.4 can be revisited:

> Two vehicles occupy the same lane in close proximity to one another. No other vehicle resides between the fore vehicle **c0** and aft vehicle **c1**. Furthermore, sufficient clearance is given ahead of **c0** and behind **c1** in order to consider the behaviour of the two vehicles in isolation. Henceforth, this setup will be referred to as a two-vehicle convoy. Because of their close proximity, should **c0** slow down or **c1** speed up, the other agent will have to respond in kind or risk collision. Similarly if **c0** speeds up or **c1** slows down, while it does not directly force the other agent to take action in the same way, it does enable them to act in kind without the risk of collision. Figures 1.4a and 1.4b depict the before and after of such a behavioural interaction, where **c0** braking caused **c1** to subsequently brake to avoid collision. One expects each vehicle to affect the other to varying degrees across time. In other words, one can consider **c0** and **c1** causally adjacent based upon Definition 6. Lastly, so long as they each occupy separate lanes from each other and **c0** / **c1**, a number of other independent vehicles may be present {**i0**, **i1**, ...} that lack causal adjacency with any other vehicle. Figure 1.4c depicts the causal-adjacency graph for the scenario.

While this scenario provides an example of how causal relationships might relate to behavioural interactions between vehicles, it is still necessary to determine what variables would best serve to capture this relationship. This work proposes two perspectives which indicate which variables might be used:

1. **Actuation-Oriented Variables:** This perspective argues that the behaviour of agents is best captured via the variables the agent actuates itself with as this offers a direct view of the instantaneous control output of the agent. Example Variables: Acceleration, Steering.

2. **Goal-Oriented Variables:** This perspective argues that the behaviour of agents can be best captured through the variables they aim to control since this could offer a greater insight into the agent's intentions. Example Variables: Speed, Lane.

This choice of the primary variable of interest offers some flexibility in the causal representation of AV scenarios. Despite this, if one assumes that there is a single variable per agent then the variable selection for a given scenario is quite restrictive. Although in the case of more complex scenarios there is greater flexibility in how one might represent behavioural interactions as causal links between variables. For example, in a three-vehicle convoy does the front vehicle affect the middle and rear vehicle directly, or does the front vehicle affect the rear vehicle via the middle vehicle acting as a mediator? The ambiguity over such representations is part of why this work focuses upon a simpler two-vehicle-convoy scenario, as otherwise the formulation of a ground truth for quantitative evaluation would inherently rely upon some strong assumptions. Of course, if one drops the assumption of single variable per agent one can then capture intra-agent causal relationships, however this was opted against for three reasons: Firstly, the focus of this thesis is upon inter-agent causal relationships, thus the introduction of intra-agent causal relationships could distract from this. Secondly, the introduction of several variables for each agent once again introduces ambiguity regarding the correct assumptions to make when formulating a ground truth. Lastly, some preliminary experimentation was carried out which did incorporate this, but these experiments indicate that if anything the methods performed worse when introducing more variables per agent. This is likely because the intra-agent causal relationships are stronger than inter-agent causal relationships, and thus these intra-agent relationships can 'overshadow' the inter-agent ones.

With causal relationships and variables contextualised under the behavioural interactions of an AV scenario, one can now begin to consider the potential means by which relevant causal relationships might be discovered.

## 3.2 Benchmarking Temporal Causal Discovery on Agent Data

The following section aims to examine the readiness of existing causality literature in being able to discover and represent behavioural interactions between autonomous agents. This is explored within the AV scenario described in previous section, utilising experiment data from two real-world datasets, and one synthetic dataset.

### 3.2.1 Concept of Observation-Based Causal Discovery

In order to utilise causal modelling to make inferences or otherwise, one must first produce a causal model. One approach to this is to construct the model manually. However, this relies upon the causal relationship between variables being known, which is not always possible. Occasionally the goal of research is to actually deduce the causal relationships between variables. One approach to this is to design and carry out experiments such as randomised controlled trials to gather causal information, however this can be expensive and in some cases unethical.

In the AV domain the risk to human life is high and in many cases one wishes to analyse scenes post-hoc, thus experimentation or intervention-based approaches are unsuitable. However, there is a variety of approaches that have been proposed to identify causal relationships via observation alone. Such observation-based causal-discovery approaches typically aim to identify one or more of the tuple elements of an SCM:

- $U$: The method may try to identify external factors that may not be included as part of an initial model. Most commonly approaches try to identify hidden confounders that affect several variables.

- $V$: Although little literature exists on the concept, in line with the previous item a causal-discovery approach could attempt to identify more complex elements external to an initial model.

- $E$: Arguably the most common objective of causal discovery is to determine the presence and direction of causal relationships between variables. Note that because these methods are observation-based there is no guarantee of being able to correctly identify the direction of all causal relationships.

- $F$: Some methods aim to approximate the structural equations or parametrisation of said structural equations rather than the overall structure.

- $P(U)$: Lastly there are cases where data derived from observation is used to approximate the probability distributions of external factors. This is most frequently done as part of the abduction step of counterfactual inference, although in this context it is not typically referred to as a form of causal discovery.

This work mainly considers the causal discovery of causal links $E$. Since one knows the agents present within a given scene, the focus is on determining the presence of behavioural interactions by discovering causal relationships between their variables. It is worth considering that some causal-discovery methods aim to identify hidden confounders as well. However, since most of the datasets that will be considered do not contain sufficient ground-truth information on potential confounders, it would be impossible to evaluate effectively.

In order to effectively carry out causal discovery some methods rely upon a condition known as causal faithfulness in order to function.

**Definition 9** (Causal Faithfulness)**.** A data distribution is faithful to a ground-truth SCM provided that the conditional-independence relations present in the data are reflective of parent-child relationships in the SCM. For example, while rare, there are cases where an underlying system produces data such that causally-related variables appear conditionally independent. This can occur when variables that share a common ancestor are both parents of a collider and the structural equations happen to function in such a way that the parents cancel out. Should this occur, a method that relies upon conditional independence to identify SCM structure would be unable to do so properly. The reader is referred to Figure 3.1 for an example instance of where causal faithfulness does not hold.

Lastly, methods which aim to carry out causal discovery on time-series data typically aim to produce a window-time-graph SCM, described in terms of relative time-lagged causal relations. This is primarily because observation-based causal discovery is data driven and thus requires causal relations to be consistently exhibited throughout time in order to identify them effectively. In the following section the different categories of temporal causal-discovery methods that aim to address this task will be explored.

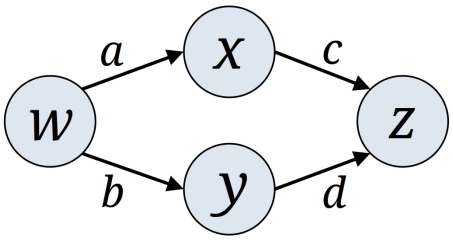



Figure 3.1: The above SCM offers an example of where causal faithfulness may fail to hold. Assume $a$, $b$, $c$, and $d$ represent linear coefficients of the causal relationships between the variables $w$, $x$, $y$, $z$. If the linear coefficients are parametrised in such a way that $ac + bd = 0$ then the causal effects of $x$ and $y$ on $z$ will cancel out as a result of their mutual ancestor $w$. As a result of this cancellation the conditional-independence relations that should hold for the SCM would not be reflected in the resulting data distributions. Therefore the data distribution would not be faithful to the underlying SCM under such circumstances. This example also demonstrates why causal unfaithfulness can be quite rare, as for the most part it relies upon very specific interactions between structural equations and their parametrisations.

### 3.2.2 Contemporary Temporal Causal-Discovery Methods

The following sections will introduce a variety of temporal causal-discovery approaches and the categories they fall under.

#### 3.2.2.1 Granger Causality

Granger causality is built upon the concept that causes should provide unique information that enables the prediction of their effects [30]. The first two Granger-causality-based approaches assume a linear relationship between variables. This should indeed be true of the variables considered here, as the tail convoy vehicle should roughly mirror the speed and acceleration of the head convoy vehicle, albeit with a time lag. The Pair-Wise Granger Causality (PWGC) [30] approach determines the likelihood of a causal link from variable $N_j \subset \mathbb{R}$ to endogenous variable $V_i \subset \mathbb{R}$ by considering two auto-regressive models:

$$V_{i,t} = \sum_{l=1}^{\tau} (w_{i,l}\, V_{i,(t-l)}) + \varepsilon_{i,t} \tag{3.16}$$

$$V_{i,t} = \sum_{l=1}^{\tau} (w_{i,l}\, V_{i,(t-l)} + w_{j,l}\, N_{j,(t-l)}) + \varepsilon_{i,t} \tag{3.17}$$

where $w_{i,l}, w_{j,l} \in \mathbb{R}$ represent coefficients specific to a time lag $l$ for variables $V_i$ and $N_j$ respectively, and $\tau$ is the maximum time lag considered. Meanwhile $\varepsilon_i$ and $\varepsilon_j$ are white-noise time

series which represent the influence of factors external to the auto-regressive models. Practically speaking, the coefficients described by $w_{i,l}$ and $w_{j,l}$ correspond to elements in a multidimensional array $\mathbf{W}$, which provides a complete description of the linear causal relationships between the variables it considers. If $N_j$ does indeed have a causal effect upon $V_i$ one would expect $N_j$ to possess unique information allowing one to better predict $V_i$ with (3.17) than with (3.16). An F-test can be applied to an accuracy metric (e.g. residual sum of squares) for each of the models to determine whether the difference is significant enough to consider a causal link to have been discovered. Furthermore, following the determination of all causal links, one can formulate approximations of the structural equations by merging the learnt auto-regressive models. Note that this particular variant of PWGC, along with the methods that follow, does not identify instantaneous causal relations. However, this is not typically an issue for the considered use case, as most data pertaining to autonomous agents is captured at smaller intervals than the average human reaction time.

The PWGC approach is limited in that it only ever considers pairs of variables and can therefore struggle with mediators relationships present in the underlying causal graph. Multi-Variate Granger Causality (MVGC) [45] builds upon PWGC by proposing the following auto-regressive models at the cost of higher computational overhead:

$$V_{i,t} = \sum_{N_k \in N_{-j}} \sum_{l=1}^{\tau} (w_{k,l}\, N_{k,(t-l)}) + \varepsilon_{i,t} \tag{3.18}$$

$$V_{i,t} = \sum_{N_k \in N} \sum_{l=1}^{\tau} (w_{k,l}\, N_{k,(t-l)}) + \varepsilon_{i,t} \tag{3.19}$$

where $N_{-j} = N \setminus \{N_j\}$. Unlike PWGC, which only considers self causation in the restricted case and the additional information provided by a single variable in the full case, MVGC considers all bar variable $N_j$ in (3.18) and all variables in (3.19). Because this approach considers all information available while only excluding the information of variable $N_j$ that is being examined as a cause, it is not only able to tackle mediator causal relationships, but is also robust against against confounding variables. This is because any confounding case other than an exogenous variable that only affects the potential cause and effect will have its information

already captured by (3.18) and therefore cannot lead to a misleading increase in accuracy in (3.19). MVGC otherwise operates similar to PWGC, comparing accuracy metrics associated with each auto-regressive model via an F-test, determining if a causal link is present.

The Temporal Causal Discovery Framework (TCDF) [46] offers a non-linear approach that is based upon Attention-based Dilated Depth-wise Separable Temporal Convolutional Networks (AD-DSTCNs) which are themselves based upon Attention-based Convolutional Neural Networks (ABCNNs) [138]. Rather than train pairs of auto-regressive models an AD-DSTCN is trained for each variable whereby a learnable attention-value limits the participation of each variable in predicting the future values of the network associated variable. The idea being that as the network is trained the attention values will come to indicate which of the time series are most likely to have a causal effect over the time series being predicted for. The most significant difference between AD-DSTCNs and ABCNNs is the use of a dilation mechanic in AD-DSTCNs which allows the use of smaller One-Dimensional (1D) kernels while maintaining a wide range of possible time lags between causes and effects.

The Neural Additive Vector Auto-Regression (NAVAR) [53] approach is another NN based technique which can model non-linear relationships between variables. In contrast to TCDF, NAVAR trains an NN — a Multi-Layer Perceptron (MLP) here — for each variable to predict for all other variables. The final predicted value for each variable is formulated from the sum of the contributions from all variables, plus an additional bias value. NAVAR then assumes that if a variable is a causal parent of another variable, it will hold useful predictive information and thus its contributions will vary more. Hence the variance of the contributions from variables can be used as a metric to detect causal parentage.

#### 3.2.2.2 Noise-Based Approaches

Noise-based approaches share a theoretical similarity with Granger-causality-based approaches in that these approaches also rely upon examining the flow of information between variables. However, in contrast to Granger causality, noise-based methods do not operate upon identifying information possessed by variables useful for predicting other variables. Instead they attempt to identify the direction of causal relationships by identifying information variables hold about the noise of other variables.

The first noise-based method is built upon the LiNGAM approach [39], which the Vector Auto-Regressive Linear Non-Gaussian Acyclical Model (VAR-LiNGAM) [139] approach extends to operate upon time-series data. To explain the premise of VAR-LiNGAM, first consider the following bivariate example of LiNGAM [39]. One assumes that if variable $N_j$ has a causal effect upon endogenous variable $V_i$ that the distribution of each variable is generated as follows:

$$N_j = \varepsilon_j \tag{3.20}$$

$$V_i = w_{i,j} N_j + \varepsilon_i \tag{3.21}$$

where $\varepsilon_i$ and $\varepsilon_j$ represent noise from the influence of factors external to the above models. Under this distribution, $V_i$ captures information on the noise provided by $\varepsilon_j$ because its value is derived from $N_j$. However, $N_j$ does not capture any information on $\varepsilon_i$, establishing an asymmetry that LiNGAM aims to exploit. It is important to note that $\varepsilon_i$ and $\varepsilon_j$ cannot both be Gaussian, as under these conditions it is impossible to determine the causal direction.

In order to solve the directions of causal links the model reframes the underlying model responsible for the data as $V = \mathbf{W}V + \varepsilon$ where $V$ is the set of variables represented as a vector, $\varepsilon$ is a vector of noise values, and $\mathbf{W}$ is a strictly lower triangular matrix which describes the causal links of $E$. If the model is further refined to $V = \mathbf{W}'\varepsilon$ where $\mathbf{W}' = (\mathbf{I} - \mathbf{W})^{-1}$ one can attempt to solve for $\mathbf{W}$. The initial work on LiNGAM [39] applied independent component analysis [140] in order to achieve this. However, VAR-LiNGAM utilises an alternate version known as the Direct Linear Non-Gaussian Acyclical Model (DirectLiNGAM) [141] approach. This method constructs an auto-regressive model and recursively checks independence between each variable acting as a predictor and the residuals given by applying that predictor to other variables. The most independent predictor is placed highest in the causal hierarchy. In each subsequent step the remaining variables are substituted for their residuals obtained during the previous step to remove the influence of variables with established positions. Since the above process only establishes the direction of causation, the strength of causal effects can be determined by conventional covariance-based regression, before pruning is carried out by applying the adaptive-lasso method [142]. From here the time-series extension provided by

VAR-LiNGAM [139] is facilitated by expanding the previously-described model over a time window defined by the MTL $\tau$:

$$V_t = \sum_{l=1}^{\tau} (\mathbf{W}_l \, V_{(t-l)}) + \varepsilon_t \tag{3.22}$$

Provided one can approximate $(\mathbf{W}_l)_{1 \geq l \geq \tau}$ the steps of DirectLiNGAM can be followed in similar manner albeit working with time-series variables. The use of adaptive lasso is once again important in order to reduce the number of false-positive causal edges. Note that because $(\mathbf{W}_l)_{1 \geq l \geq \tau}$ is formulated in terms of endogenous variables, this method cannot be used to identify causal edges from specific exogenous variables. Note, while $\mathbf{W}$ is a matrix for LiNGAM, it is a 3D array for VAR-LiNGAM in order to have weightings across several time lags.

The other noise-based approach, Time-series Models with Independent Noise (TiMINo) [51] consists not so much of a single model as much as it describes a class of models. The models described by TiMINo are assumed to adhere to the form:

$$V_{i,t} = f_i^{\varepsilon}(Pa^{\tau}(V_{i,t}), \varepsilon_{i,t}) \tag{3.23}$$

where $f_i^{\varepsilon} \in F^{\varepsilon}$ is a variation on structural equations that adds a noise term $\varepsilon_{i,t}$. In addition to the usual assumptions of the ground-truth graph being acyclic and causal faithfulness, the underlying model must adhere to one of the following:

- Instances of $f_i^{\varepsilon}$ are linear, and instances of $\varepsilon_{i,t}$ are non-Gaussian.
- Instances of $f_i^{\varepsilon}$ are non-linear, and instances of $\varepsilon_{i,t}$ are Gaussian.

Alternatively these requirements can be relaxed if the data follows a time structure — as opposed to independent and identically distributed time-indexed variables. Due to the nature of the data being worked with, this latter case captures the task more closely. In terms of how TiMINo operates, it utilises a supplied regression method and independence test. The implementation used here is comprised of a linear-regression model and a cross-covariance-based independence test with Bonferroni correction. TiMINo proceeds by learning a predictor for each time series and then determining how independent each time series involved in the predictor is from the

residuals produced from applying the predictor. The time series with the predictor that produces the greatest level of independence is deemed to be at the bottom of the causal hierarchy as little or no information regarding its own noise is present in other time series. This process is repeated until a full causal hierarchy is established. At this point the causal parents of each time series are refined by removing those time series unnecessary to produce independent residuals.

#### 3.2.2.3 Constraint-Based Approaches

Constraint-based approaches rely upon conditional-independence tests, assuming causal faithfulness as described in Definition 9. Both of the constraint-based methods considered here share a similar initial process of first identifying which which nodes are unconditionally independent. They then progressively check independence upon adjacent node pairs conditional upon their neighbours. Once adjacencies remain stable one can identify unshielded triples and convert these into collider structures as these are the only structures that can be directly identified under the assumptions of the methods considered. For example if $N_i$ and $N_j$ are both adjacent to $N_k$ but not to each other, and $N_i$ and $N_j$ are not independent conditional upon $N_k$, the only possible faithful graph structure would be for $N_k$ to be a collider of $N_i$ and $N_j$. At this point the behaviour of each of the considered approaches diverges.

One of the oldest constraint-based methods is the PC approach [38]. Following the steps outlined above PC proceeds to iteratively apply 3 rules for refining the causal direction of adjacent nodes. Depending upon the data available it will not always be possible to orient every edge. The resulting graph describes an Markov Equivalence Class (MEC) which contains all the possible fully-directed causal graphs based upon the remaining undirected edges, should any be present. The Peter-Clark algorithm with Momentary Conditional Independence (PCMCI) approach [47] extends PC to better work with time series. While the initial steps are similar to PC, the momentary-conditional-independence test is designed to avoid the influence of auto-correlations by evaluating the level of dependence between nodes while conditioning upon the parents of both nodes. It is important to note that PCMCI can be used with any conditional-independence test, though for this work only the partial-correlation [143] approach is considered.

The other constraint-based method builds upon Fast Causal Inference (FCI) algorithm [144] and is aptly named Time-Series Fast Causal Inference (tsFCI) [49] to reflect its modifications

specific to time series. FCI follows the same initial process as PC, however when it comes to the iterative application of rules there are 10 rules as opposed to 3. The purpose of these additional rules to allow FCI to accommodate the presence of exogenous hidden confounder ancestors and hidden descendants that have been inadvertently conditioned upon (e.g. selection bias). To achieve this FCI works upon the concept of a maximal ancestral graph rather than an MEC. This allows for bi-directional edges representing hidden confounder ancestors, and non-directional edges representing where hidden descendants have been conditioned upon. With tsFCI the conversion of the time series into a window-time-graph skeleton facilitates most of the work required to allow FCI to operate effectively on time-series data. Two restrictions of tsFCI are that it can no longer capture instantaneous causal links nor hidden descendants that have been conditioned upon, however neither of these are problematic for the use case considered.

#### 3.2.2.4 Non-Stationarity-Based Approaches

While most approaches struggle with non-stationarity, approaches such as Constraint-based causal Discovery from Nonstationary / heterogeneous Data (CD-NOD) proposed by Zhang et al. [54] actively aim to exploit non-stationarity. For the most part CD-NOD resembles a constraint-based approach except with the addition of a surrogate variable designed to capture either domain-based or time-based distribution shifts. The discovery of the causal-graph skeleton is carried out in a similar manner to a method such as PC, relying upon conditional-independence tests and applying orientation rules. In determining causal direction, CD-NOD bears some similarities to a noise-based approach. However, rather than examining the propagation of noise CD-NOD examines the influence of variable specific parameters calculated as a function of the surrogate variable. Since the actual parameters are unknown CD-NOD tackles the problem through the application of kernel density estimation. One limitation of this approach is that if there is a particularly influential hidden confounder present CD-NOD may struggle to correctly identify the correct causal direction between affected variables.

In the aforementioned work, CD-NOD was purely intended for application on instantaneous causal relationships, however it was discussed that it was natural to extend the method to include time-lagged relationships. This was subsequently accomplished by Huang et al. [145]. However, a shortcoming of CD-NOD is that while it can capture shifts in causal-model parametrisation, it

cannot capture changes to the causal skeleton that occur with time, nor can it capture shifts in causal direction with time. This is also true of the time-lagged extension, meaning that a causal relationship that varies in time lag over time but not in causal effect cannot be adequately captured. Nevertheless of the approaches evaluated in this work, CD-NOD should in theory offer the best framework to provide robustness against non-stationarity.

#### 3.2.2.5 Score-Based Approaches

Score-based approaches view the causal-discovery process as the task of finding a causal graph which maximises a scoring metric of how well the data fits the supplied graph. Since an exhaustive search would be computationally intractable this is typically tackled as an optimisation problem. In terms of the score metrics used for this type of approach, the Bayesian information criterion [146], the Bayesian Dirichlet equivalence [147] score, and a cross-validation-based [148] technique have all been proposed as options.

However, the approach considered for evaluation here is an evolution of Non-combinatorial Optimisation via Trace Exponential and Augment Lagrangian for Structure learning (NOTEARS). This method, Dynamic Non-combinatorial Optimisation via Trace Exponential and Augment Lagrangian for Structure learning (DYNOTEARS) [52], defines a new metric with which to assess how well the data fits the current model. The following consists of the metric they define without any of the components relating to instantaneous causal relationships:

$$\ell(\mathbf{W}) = \frac{1}{2(t_{max}+1-\tau)} \sum_{t=\tau}^{t_{max}} \sum_{V_i \in V} \left( V_{i,t} - \sum_{l=1}^{\tau} \sum_{N^j \in N} \left( w_{i,j,l}\, N_{j,(t-l)} \right) \right)^2 + \beta_{\mathbf{W}} \|\mathbf{W}\|_1 \tag{3.24}$$

where $\beta_{\mathbf{W}}$ is a regularisation constant and $\|\cdot\|_1$ is the Manhattan distance. Meanwhile $w_{i,j,l}$ corresponds to the $i$-th row and $j$-th column of a slice of multidimensional array $\mathbf{W}$ defined by the time lag $l$. The first half of the metric aims to minimise the error of approximation via multidimensional weight array $\mathbf{W}$, while the second half aims to simplify this array with the addition of a regularisation term. After optimising $\mathbf{W}$ one can then construct a causal graph by applying a threshold to each element $w_{i,j,l}$ within $\mathbf{W}$, and adding a graph edge should the value prove great enough. As a result $\mathbf{W}$ is both a weight array updated while learning a linear predictive model with DYNOTEARS, as well as a description of causal adjacency.

### 3.2.3 Benchmark Experiments

In this section the experiments carried out to benchmark the previously-described methodologies are detailed. This is followed by coverage of the subsequent results and the conclusions to be drawn from them.

#### 3.2.3.1 Data, Code, & Parameters

The experiments were carried out on three datasets independently, the 'Woven Prediction' dataset [149], the 'highD' dataset [150], and a novel synthetic dataset. For all datasets the variables utilised to represent the agents were acceleration and speed, representing the actuation-oriented and goal-oriented perspectives on capturing behaviour presented in Section 3.1.3.

The code used in the experiments builds upon an earlier framework provided by Assaad et al. [44]. This code along with any publicly-shareable data is made available in a Git repository[5].

The key parameters universal to all the methods were the significance alpha $\lambda_\alpha$ and MTL $\tau$. In order to avoid assessing the efficacy of the methods with a poor parameter selection these values were individually varied. The values used for the significance alpha were 0.001, 0.005, 0.01, 0.03, 0.05 and 0.1, while using an MTL of $2.5\,s$ based upon a conservative estimate of reaction and actuation times of human drivers. Meanwhile the values used for the MTL were $2.5\,s$, $3.6\,s$ and $4.9\,s$, while using a significance alpha of 0.05. Any remaining parameters involved in these experiments will be documented in Appendix A, but may be mentioned briefly in the following text. Note that this includes parameters relating to the generation of the synthetic dataset described below.

**highD Dataset** The highD dataset focuses upon stretches of highway at six locations across Germany [150]. It offers extensive coverage of the roads it considers by capturing overhead footage using a drone. This is automatically annotated, resulting in a positional error of labelled vehicles typically under $10\,cm$. The dataset recordings were captured at $25\,Hz$ over the course of several minutes. Rather than a semantic map, the dataset offers the y-position of the lane markings in pixel space, as all data is aligned to have the lanes run parallel to the x-axis.

[5] https://github.com/cognitive-robots/temporal-cd-evaluation-paper-resources

Due to additional details in terms of lane occupancy and inter-vehicular adjacency provided by the dataset, it was possible to automate the process of creating scenes. To do this one can scan the agent meta-data and look for the following conditions:

- All agents present in the extracted scene must not change lanes at any point in the scene. This is to simplify the process of identifying vehicles with behaviour independent of one-another.

- The convoy head **c0** and convoy tail **c1** should be vehicles in the same lane, with **c0** the next vehicle directly ahead of **c1**.

- The independent vehicles {**i0**, **i1**, ...} should be vehicles occupying different lanes to the convoy vehicles. No two independent vehicles should occupy the same lane. There must be at least one independent vehicle for a scene to be valid for extraction.

- Both **c0** and **c1** should have an absolute difference between their maximum and minimum speeds greater than 20 % of their maximum speed.

- **c0** should have a distance headway greater than $20\,m$ and **c1** should have a distance headway less than $10\,m$. This is both to avoid extracted scenes with convoy sizes above two and to ensure **c0** and **c1** are close enough for a causal interaction to actually occur.

- There must be a span of at least $10\,s$ for which all relevant vehicles have data captured. This is generally to ensure the scene is long enough for analysis. More importantly some of the baseline methods expect the scene to be provided in a tabular format, which requires data for every relevant agent at each of the captured time steps.

From this process a total of 115 scenes were extracted. Unfortunately, unlike the other two datasets it was not possible to automate the process of determining the direction of causal relationships, and thus the evaluation of this dataset was based on causal adjacency.

**Woven Prediction Dataset** This dataset is comprised of over a thousand hours of driving data collected over 20 AVs operating in Palo Alto, California [149]. In contrast to the highD dataset, this dataset is based around a much greater diversity of road structures and driving

styles. The data itself is comprised of egocentric data (e.g. position and orientation of capturing AV), details of nearby vehicles and pedestrians, and the status of any perceived traffic lights. This time-series data is structured into a series of scenes, each $\sim 30\,s$ in length and captured at $10\,Hz$. In addition to these scenes, there is a static semantic map describing the road network the AVs were operating on.

In order to transform the data available into two-agent convoy scenes upon which causal discovery could be carried out, a manual inspection was carried out upon each scene in turn. This was used to determine which scenes exhibited the scenario laid out in Section 3.1.3, and within these scenes which agents — besides the ego vehicle — would form part of the scene. In total 50 scenes were extracted from the dataset in this manner, all of which were at least $10\,s$ in length following any trimming of the scene that needed to be carried out. Each of these scenarios featured at least one acceleration or braking action from **c0** and a noticeable response from **c1**. As manual inspection offered greater control over selection criteria, it was possible to check the direction of the causal relationship shown in Figure 1.4c, with it being from **c0** to **c1**.

There exists a possibility of confounders in this dataset due to the greater variety of road elements and increased difficulty of isolating causal relationships for quantitative evaluation. In terms of the aforementioned road elements, the most significant impact comes from traffic lights which have the capacity to jointly affect several agents outside of any causal influence they may have on one another. However, other elements such as road signs or road markings can also act as confounders, e.g. a speed sign may cause multiple vehicles to change speed one after another despite no causal link existing between them. As for the increased difficulty in isolating causal relationships for evaluation, since agents are no longer occupying straight sections of parallel lanes it is harder to discern which agents are actually affecting one another in order to derive a ground truth graph. While one could develop a strategy for doing so and analyse each scene in turn, the sheer variety in road layouts makes it difficult to account for all cases that might be encountered. Nonetheless, while the possibility of confounders does exist it is argued here that this constitutes a fair challenge to the methods as the highD dataset lacks confounders — thus allowing this work to evaluate the methods in their absence — and several methods claim to be able to cope with confounders anyway.

Finally, the speeds and accelerations of agents were estimated based upon the change in agent positions between frames. Here in order to account for positional jitter, the speed and acceleration were smoothed via a 15-frame moving average.

**Synthetic Dataset** In addition to running experiments upon the two real-world datasets, experiments were also carried out upon a novel synthetic dataset. Doing so facilitates differentiating between performance loss due to the nature of the scenario from performance loss due to real-world complications (e.g. sensor noise, variations in human behaviour, etc.). In order to generate this dataset, a series of speed-goal objectives were assigned for the lead convoy agent and independent agent. A proportional-feedback controller was then used to actuate the acceleration of the aforementioned agents while adhering to a set of linear kinematic constraints. Meanwhile the tail convoy agent was actuated by a proportional-feedback controller that aims to maintain a convoy distance based upon a braking time of $2.24\,s$, with a $0.5\,s$ time lag to mimic a reasonable human reaction time. In both of these cases the proportional-feedback controller in question calculated the vehicle acceleration by multiplying the speed / distance error by a predefined gain parameter. For the experiments it was found a proportional gain of 1.0 sufficed to produce satisfactory causal scenes. The scenes were made to last for 50–70 $s$ and were generated at $10\,Hz$. Scenes were generated in such a way as to have 12 causal interactions within the convoy and 12 actions carried out by the independent agent. Similar to the Woven dataset, the greater level of control offered by synthesising the scenarios allowed the evaluation of a directed variant of the causal relationship in Figure 1.4c, with **c0** affecting **c1**. In total 100 synthetic scenes were generated before being parametrised in terms of speed and acceleration respectively.

#### 3.2.3.2 Evaluation Metric

To determine the efficacy of each evaluated method the True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) counts for each causal scene were calculated. For a discovered causal-summary graph $\hat{\mathcal{G}}^S = \langle \hat{N}^S,\ \hat{E}^S \rangle$ compared against a ground-truth causal-summary graph $\mathcal{G}^S = \langle N^S,\ E^S \rangle$, the calculations are as follows:

$$|TP|^O = |\hat{E}^S \cap E^S| \tag{3.25}$$

$$|FP|^O = |\hat{E}^S \setminus E^S| \tag{3.26}$$

$$|FN|^O = |E^S \setminus \hat{E}^S| \tag{3.27}$$

$$|TN|^O = |V^S|(|V^S| - 1) - (|TP|^O + |FP|^O + |FN|^O) \tag{3.28}$$

Similarly, these calculations can be made while treating the causal-summary graphs as causal-adjacency graphs — as is done with the highD dataset:

$$|TP|^A = 0.5|(\hat{E}^S \cup \{\langle \hat{N}_j^S, \hat{N}_i^S \rangle \mid \langle \hat{N}_i^S, \hat{N}_j^S \rangle \in \hat{E}^S\}) \cap (E^S \cup \{\langle N_j^S, N_i^S \rangle \mid \langle N_i^S, N_j^S \rangle \in E^S\})| \tag{3.29}$$

$$|FP|^A = 0.5|(\hat{E}^S \cup \{\langle \hat{N}_j^S, \hat{N}_i^S \rangle \mid \langle \hat{N}_i^S, \hat{N}_j^S \rangle \in \hat{E}^S\}) \setminus (E^S \cup \{\langle N_j^S, N_i^S \rangle \mid \langle N_i^S, N_j^S \rangle \in E^S\})| \tag{3.30}$$

$$|FN|^A = 0.5|(E^S \cup \{\langle N_j^S, N_i^S \rangle \mid \langle N_i^S, N_j^S \rangle \in E^S\}) \setminus (\hat{E}^S \cup \{\langle \hat{N}_j^S, \hat{N}_i^S \rangle \mid \langle \hat{N}_i^S, \hat{N}_j^S \rangle \in \hat{E}^S\})| \tag{3.31}$$

$$|TN|^A = 0.5|V^S|(|V^S| - 1) - (|TP|^A + |FP|^A + |FN|^A) \tag{3.32}$$

From these counts one can derive various evaluation metrics with which one can determine the efficacy of the temporal causal-discovery approaches:

$$\text{precision} = \begin{cases} \frac{|TP|}{|TP|+|FP|} & |TP| + |FP| > 0 \\ 0 & |TP| + |FP| = 0 \end{cases} \tag{3.33}$$

$$\text{recall} = \frac{|TP|}{|TP| + |FN|} \tag{3.34}$$

$$\text{fall-out} = \frac{|FP|}{|FP| + |TN|} \tag{3.35}$$

$$\text{F}_1 \text{ score} = \frac{2\,|TP|}{2\,|TP| + |FP| + |FN|} \tag{3.36}$$

where precision is the proportion of identified causal links that are true causal links, recall is the proportion of true causal links that are identified, and fall-out is the proportion of non-causal-links that are falsely identified as causal links. Lastly $\text{F}_1$ score offers a balanced

measure between precision and recall, and thus a general measure of overall performance. In order to calculate these metrics across the datasets for each method-parameter combination, the precision, recall, and $F_1$ score are calculated for each scene and then the mean and standard deviation are calculated across these.

Before proceeding, the method of matching scenes to the predefined ground-truth graph for each dataset is summarised here for the benefit of the reader. For the synthetic dataset the data-generation method was simply built around the two-vehicle-convoy scenario described in Section 3.1.3 and thus the resulting data was guaranteed to adhere to the ground truth. For the highD dataset it was possible to filter scenes by whether or not they adhered to the two-vehicle-convoy scenario — and thus the predefined ground-truth graph — by imposing a strict set of restrictions utilising the dataset metadata. Lastly the Woven dataset relied upon manual selection of scenes and agents by the author that adhered to the two-vehicle-convoy scenario. The act of manually selecting or manually labelling scenes is arguably the most practical method if one desired to carry out similar experiments with new data. However, the main caveats that come with such an approach are the additional labour that is required and the inherent subjectivity that comes with humans processing the data.

#### 3.2.3.3 Results

The results are displayed in Table 3.1 and illustrated in Figure 3.2. Overall the best mean $F_1$ scores are provided by DYNOTEARS, MVGC, and TiMINo. PCMCI and NAVAR also provide competitive results, but are outperformed by at least one other method in every case. PWGC, tsFCI, and CD-NOD generally under-perform compared with other methods, while TCDF and VAR-LiNGAM completely fail in almost every case.

While it is possible that these failures are due to an issue of implementation, TCDF directly calls the same code utilised by the original work [46] and VAR-LiNGAM directly calls code from the Python 'lingam' package [141, 151], making this unlikely. Theoretically speaking it is possible that TCDF is under-performing due to lack of training data, but this is made less likely by the fact NAVAR is competitive also training upon the same amount of data. Likewise with VAR-LiNGAM, it is possible that there is too weak of a direct coupling between variables for the noise of one variable to affect another, but considering TiMINo is competitive, this is

| | | | | highD | | | Woven | | | Synthetic | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Method | Var. | $\lambda_\alpha$ | $\tau$ (s) | Precision | Recall | $F_1$ Score | Precision | Recall | $F_1$ Score | Precision | Recall | $F_1$ Score |
| CD-NOD | acc. | 0.1 | N/A | 0.05±0.12 | 0.15±0.35 | 0.07±0.17 | 0.23±0.37 | 0.32±0.47 | 0.26±0.39 | 0.30±0.17 | 0.81±0.39 | 0.43±0.22 |
| CD-NOD | sp. | 0.1 | N/A | 0.02±0.11 | 0.05±0.22 | 0.03±0.13 | 0.34±0.45 | 0.40±0.49 | 0.36±0.45 | 0.32±0.33 | 0.57±0.50 | 0.39±0.37 |
| DYNOTEARS | sp. | N/A | 2.5 | 0.11±0.09 | 0.76±0.43 | 0.19±0.14 | **0.58±0.37** | 0.88±0.32 | **0.66±0.33** | 0.70±0.35 | 0.91±0.29 | 0.76±0.31 |
| DYNOTEARS | sp. | N/A | 3.6 | 0.12±0.12 | 0.80±0.40 | **0.20±0.15** | 0.57±0.34 | 0.90±0.30 | **0.66±0.30** | 0.69±0.35 | 0.90±0.30 | 0.76±0.31 |
| MVGC | acc. | 0.001 | 2.5 | 0.03±0.04 | 0.29±0.45 | 0.05±0.08 | 0.34±0.13 | 0.94±0.24 | 0.49±0.15 | 0.70±0.31 | 0.95±0.22 | 0.78±0.25 |
| MVGC | acc. | 0.03 | 2.5 | 0.02±0.04 | 0.33±0.47 | 0.05±0.07 | 0.32±0.08 | 0.94±0.24 | 0.47±0.12 | 0.60±0.26 | 0.98±0.14 | 0.72±0.20 |
| MVGC | sp. | 0.001 | 2.5 | 0.00±0.01 | 0.06±0.24 | 0.01±0.03 | 0.32±0.22 | 0.76±0.43 | 0.44±0.27 | 0.94±0.16 | **1.00±0.00** | **0.96±0.11** |
| PWGC | acc. | 0.001 | 2.5 | 0.07±0.03 | **0.97±0.16** | 0.12±0.05 | 0.38±0.21 | 0.90±0.30 | 0.52±0.22 | 0.70±0.29 | 0.98±0.14 | 0.78±0.22 |
| NAVAR | sp. | 0.1 | 2.5 | 0.00±0.00 | 0.00±0.00 | 0.00±0.00 | 0.52±0.30 | **0.96±0.20** | 0.63±0.25 | 0.77±0.29 | 0.99±0.10 | 0.83±0.21 |
| PCMCI | acc. | 0.05 | 3.6 | 0.08±0.15 | 0.28±0.45 | 0.12±0.21 | 0.27±0.41 | 0.34±0.47 | 0.29±0.42 | 0.91±0.23 | 0.97±0.17 | 0.93±0.20 |
| PCMCI | acc. | 0.05 | 4.9 | 0.07±0.12 | 0.33±0.47 | 0.12±0.18 | 0.35±0.40 | 0.50±0.50 | 0.39±0.42 | 0.78±0.33 | 0.91±0.29 | 0.82±0.30 |
| PCMCI | sp. | 0.1 | 2.5 | **0.13±0.26** | 0.23±0.42 | 0.16±0.30 | 0.27±0.35 | 0.42±0.49 | 0.32±0.39 | 0.63±0.44 | 0.70±0.46 | 0.65±0.44 |
| Random | N/A | N/A | N/A | 0.06±0.04 | 0.80±0.40 | 0.11±0.07 | 0.33±0.24 | 0.75±0.44 | 0.44±0.28 | 0.33±0.24 | 0.75±0.44 | 0.44±0.28 |
| TCDF | acc. | 0.05 | 3.6 | 0.01±0.09 | 0.01±0.09 | 0.01±0.09 | 0.09±0.28 | 0.10±0.30 | 0.09±0.28 | 0.19±0.39 | 0.19±0.39 | 0.19±0.39 |
| TCDF | sp. | 0.05 | 3.6 | 0.00±0.00 | 0.00±0.00 | 0.00±0.00 | 0.10±0.26 | 0.14±0.35 | 0.11±0.29 | 0.00±0.00 | 0.00±0.00 | 0.00±0.00 |
| TiMINo | acc. | 0.05 | 3.6 | 0.07±0.06 | 0.69±0.46 | 0.12±0.10 | 0.46±0.37 | 0.74±0.44 | 0.54±0.37 | 0.92±0.21 | 0.98±0.14 | 0.94±0.17 |
| TiMINo | sp. | 0.05 | 3.6 | 0.06±0.03 | 0.91±0.28 | 0.11±0.06 | 0.53±0.31 | 0.92±0.27 | 0.64±0.27 | 0.51±0.25 | 0.98±0.14 | 0.64±0.20 |
| TiMINo | acc. | 0.03 | 2.5 | 0.06±0.06 | 0.63±0.48 | 0.12±0.11 | 0.40±0.37 | 0.64±0.48 | 0.47±0.39 | **0.95±0.18** | 0.98±0.14 | 0.96±0.16 |
| tsFCI | acc. | 0.001 | 2.5 | 0.00±0.00 | 0.00±0.00 | 0.00±0.00 | 0.34±0.39 | 0.54±0.50 | 0.40±0.40 | 0.39±0.39 | 0.60±0.49 | 0.45±0.40 |
| tsFCI | sp. | 0.001 | 2.5 | 0.00±0.00 | 0.00±0.00 | 0.00±0.00 | 0.48±0.38 | 0.76±0.43 | 0.55±0.37 | 0.05±0.17 | 0.09±0.29 | 0.06±0.19 |
| VAR-LiNGAM | acc. | 0.005 | 2.5 | 0.06±0.08 | 0.45±0.50 | 0.11±0.14 | 0.08±0.27 | 0.08±0.27 | 0.08±0.27 | 0.07±0.26 | 0.07±0.26 | 0.07±0.26 |
| VAR-LiNGAM | sp. | 0.05 | 3.6 | 0.02±0.08 | 0.07±0.25 | 0.03±0.12 | 0.17±0.37 | 0.18±0.38 | 0.17±0.37 | 0.00±0.00 | 0.00±0.00 | 0.00±0.00 |

Table 3.1: Full benchmark results given in terms of mean and standard deviation values across precision, recall, and $F_1$ score. Rows are separated by method-parameter combinations, and columns depict the aforementioned metrics associated with each combination. The parameter combinations shown represent those that provide the maximum $F_1$ score of the associated method for each dataset, for a maximum of three unique parameterisations. Lastly, the values in bold indicate the maximum mean value of a metric within a given column.

again unlikely. Another possibility is that all variable noise is Gaussian, and TiMINo is still able to perform causal discovery despite this fact, while VAR-LiNGAM cannot.

In terms of parameter selection, significance alpha has the most influence although also the least consistent pattern of the two key parameters, making it an important parameter to tune for effective application of many methods. Overall MTL has a minimal impact across many approaches, but in general over-estimating the MTL leads to degradation in performance, therefore a good estimate may suffice.

A clear issue illustrated by these results is the lack of readiness for these methods to be applied to real-world data in these types of scenarios. While some methods such as MVGC and TiMINo are able to get close to a $F_1$ score of 1.0 on synthetic data, the best performance that is seen on real-world data is from DYNOTEARS at $\sim$0.565 $F_1$ score. This is clearly an unacceptable level of performance within the AV domain, demonstrating that either significant

improvements need to be made to existing methodologies or alternative directions of research considered. This work does however highlight which methods could be candidates for further work, namely the aforementioned MVGC, TiMINo and DYNOTEARS methods.

In Figures 3.3 and 3.4 qualitative results are offered in the form of the discovered graphs for two scenes utilising acceleration and speed as the variable of interest respectively. While these visualisations do offer some information regarding the general over-sensitivity (PWGC, TiMINo, VAR-LiNGAM) or over-specificity (MVGC, NAVAR, TCDF, tsFCI) of methods, it does show any consistent pattern of failure across the methods. With this in mind a further examination of the input data itself is warranted.

Between real-world data and synthetic data the main differences that could explain the degradation of performance are: increased levels of noise, causal non-stationarity, and causal sparsity. Of these issues the latter two are of greater interest here, as issues resulting from noise do not directly relate to causality research and can potentially be resolved at a lower level of abstraction (e.g. more accurate localisation). In terms of causal non-stationarity, the synthetic dataset effectively has stationarity as a result of the agent reaction time always being $0.5\,s$, meanwhile real-world drivers feature inconsistent reaction times. As for causal sparsity, the real-world scenes that are considered are typically $30\,s$ or under and contain a few causal interactions within the convoy at most. Meanwhile the synthetic scenes are 50–$70\,s$ in length and contain 12 causal interactions within the convoy. This highlights a key issue with attempting causal discovery in a system where causal interactions may be infrequent and the observation window brief. The combination of these issues also likely explains the relatively poor performance of a method leveraging non-stationarity like CD-NOD. Here synthetic scenes provide ample stationary data, while real scenes provide relatively little data and may also be non-stationary in nature.

Diagnosing which of these three aforementioned issues may be present within a given scene can be done by a human provided they can visualise the scene in question. In terms of noise, if the agents under consideration are moving erratically as a result of poor localisation then it is unlikely the causal-discovery methods will be able to make causal inferences, as such sudden movements interfere with the patterns created in data by the real-world causal relationships.

Causal sparsity is also easy to check, as a causally sparse scene will show the agents of interest only observably interacting for a small part of the overall scene, e.g. $3\,s$ of interaction in a $30\,s$ scene. Lastly causal non-stationarity can be diagnosed if the manner in which two agents behaviourally interact is inconsistent across the scene. Examples include the tail agent in a two-vehicle convoy reacting to the head agent with differing time lags, or alternatively the head agent and tail agent alternating in who affects who throughout the scene.

One last area of discussion relating to the results is the potential reasons why the methods perform notably better on the Woven dataset versus the highD dataset. This is particularly surprising given the increased scene complexity, sensor noise, and presence of confounding factors present in the Woven dataset. Here it is possible that the manual selection of scenes has had an impact. Because scenes were selected based upon instances where there was notable behavioural interaction between agents (e.g. braking to a stop, pulling away at an intersection) it is likely that the causal relationship was depicted strongly in the data. By contrast, because the highD dataset utilised automated extraction based upon metadata, its extracted scenes may contain instances of more subtle or gradual behavioural interactions which in turn may be depicted more weakly in data. While it is difficult to verify if this is indeed the cause for this discrepancy the potential for human bias was always a risk when considering the manually selected scenes and agents from the Woven dataset. Thus it seems prudent to focus upon the highD dataset moving forwards in order to minimise any such risk of human bias.

#### 3.2.3.4 Discussion

The experiments establish that the evaluated approaches struggle with tackling even a simple causal scenario in the AV domain when working with real-world data. However, these methods constitute the frontiers of the temporal-causal-discovery field, indicating that further development of existing methodologies or new lines of thought entirely may be necessary to overcoming the new challenges presented within the AV domain.

The chief qualities of these new types of problems are causal non-stationarity and causal sparsity. The first of these has been identified in some recent literature as a matter of concern [40]. There has been some progress in this direction with one work applying a non-stationary

causal-discovery approach to medical data [152] and another very recent work attempting causal discovery on conditionally-stationary data [153]. This latter example is interesting because it explores a physical causal relationship between particles through a spring, which more closely mirrors the types of relationships present in the AV domain. The Optimal Causation Entropy (oCSE) method [48] also claims to avoid assuming stationarity, but has the limitation of only working with time lags of a single time step.

Causal sparsity as a quality is harder to overcome, as observational approaches inherently rely upon evidence that is hard to accrue from the brief interactions that occur within autonomous-agent domains. Here a counterfactual approach to causal discovery could potentially succeed by using approximations of agent behaviour to discover agent behavioural-interaction causal relationships, and thus is a promising direction for future work.

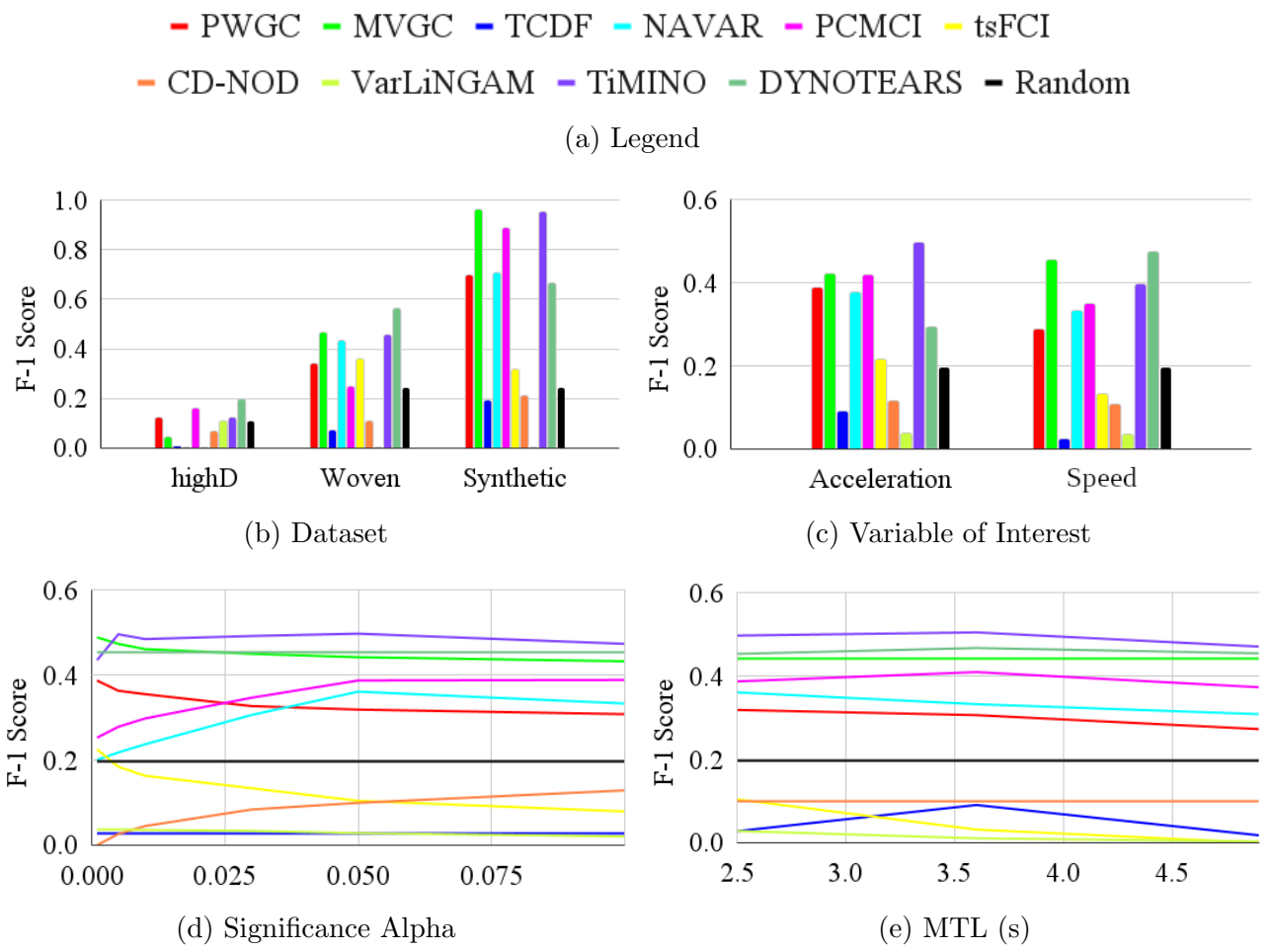
Before concluding this chapter it is necessary to tackle a potential criticism of the evaluated scenario, namely that it might fall victim to some of the challenges and limitations of causal discovery centred around two variables [154]. The main difficulty that comes from two-variable causal discovery is the difficulty in identifying the direction of causality. Here there are two matters to note regarding the evaluated scenario. The first is that given the temporal resolution instantaneous causal relationships between variables are of no relevance. This means that given sufficient correlation between two variables with a time lag, one can typically rely upon temporal precedence to identify the direction of causality. Secondly, and more crucially, the highD dataset experiments only evaluated the methods based upon the causal adjacency of discovered graphs. Thus, the inability to determine the direction of causality due to the focus upon two variables cannot possibly be the reason for the poor performance of the methods in this particular case. With these two matters taken into consideration it seems improbable that the results of these experiments are due to the simplicity of the scenario in question, but instead reflect the challenges of causal non-stationarity and sparsity discussed above.

Having discussed the benchmark results and their implications one can now summarise the contributions of this chapter as follows:

- An analysis of the benefits and challenges associated with describing AV scenarios via SCMs or more generally causal graphs. Causal representations offer high-level models that capture how the behaviour of autonomous agents affect one another. Such high-level models allow for increased interpretability and potentially the ability to generalise causal knowledge gained from a particular scene. However, this has the caveat of making it harder to define ground-truth graphs for complex scenarios without making strong assumptions regarding the underlying behavioural interactions.

- A set of benchmark experiments applying contemporary temporal causal-discovery methods to AV datasets. This work represents one of the first in its field to carry out such a benchmark and is a foundational step in establishing the readiness of causality to be applied to autonomous-agent domains. Critically it was identified that existing methods fall short of the necessary level of performance required for such a domain, along with reasons some of these methods may have performed poorly. More generally however it was found that the difficulties encountered by these methods is likely the result of qualities associated with the autonomous-agent data itself.

- An analysis of the challenges posed by autonomous-agent data when identifying causal relationships. While sensor noise was likely partially to blame for some failure cases, the predominant culprits for the poor causal-discovery performance were causal non-stationarity and causal sparsity. This could be identified by comparing how the qualities of the data differed between the synthetic and real-world datasets as the causal-discovery methods performed comparatively well on the former. Crucially this means that autonomous-agent data, in which interactions can be infrequent and not always consistent, cannot be effectively interpreted by traditional causal-discovery methods due to their data-driven nature, thus prompting the consideration of alternative approaches.




(a) Legend

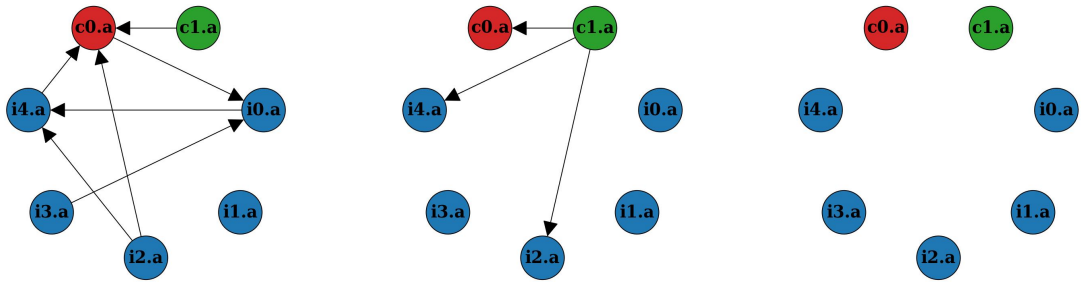
(b) Dataset

(c) Variable of Interest

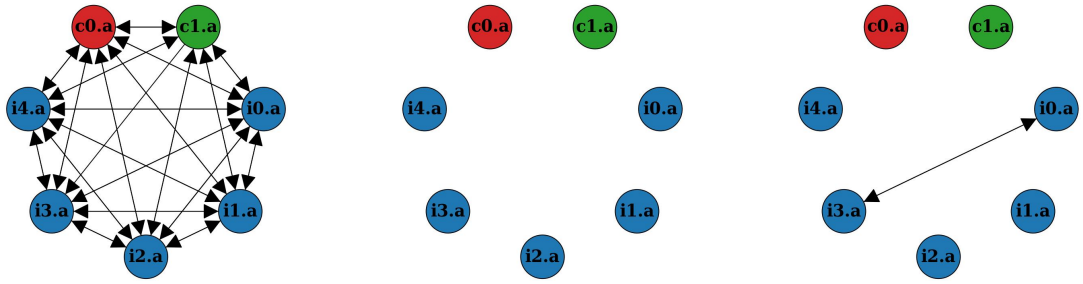
(d) Significance Alpha

(e) MTL (s)

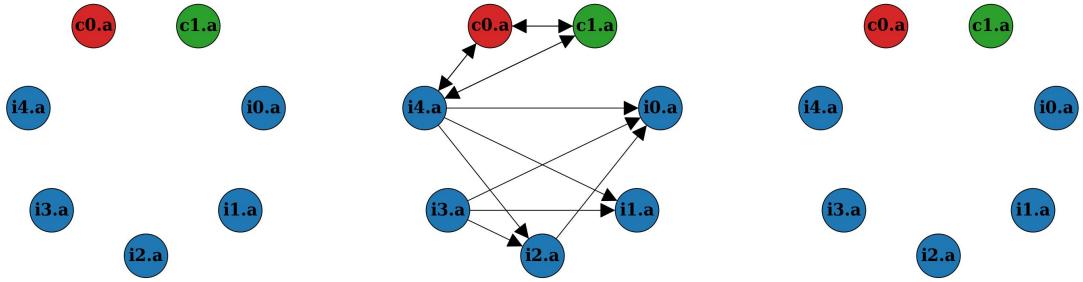
Figure 3.2: $F_1$ score from applying each method in the benchmark to the highD, Woven, and synthetic datasets. During the evaluation experiments the variable of interest (i.e. speed or acceleration), significant alpha, and MTL were varied independently in order to minimise the risk of poor performance from parameter selection. Figure 3.2a is the legend for all the other subfigures, matching line / bar colour to each of the evaluated methods. 3.2b illustrates the peak $F_1$ score achieved across all parameter combinations for each dataset. Figure 3.2b illustrates the mean taken across the peak $F_1$ scores for each dataset with parameter combinations determined by fixing the variable of interest and varying the other parameters. Figures 3.2d and 3.2e similarly illustrate the mean taken across the peak $F_1$ scores for each dataset except fixing the significance alpha and MTL respectively. The main take away points are that significance alpha and MTL have a minimal impact on the performance of the methods, but the variable of interest selection can have significant impact. Unfortunately there does not seem to be a consistent pattern across all methods, although there is some consistency within method types — i.e. Granger causality, noise-based approaches, and constraint-based approaches all seem to generally prefer acceleration. However, it appears the input data's nature has a far greater impact than the method or parameter selection as indicated by the synthetic dataset outperforming the Woven dataset, which in turn outperforms the highD dataset, each by notable margins.



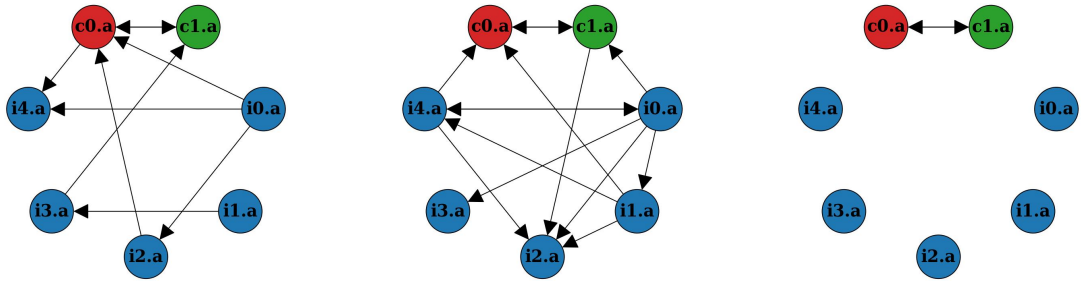

(a) CD-NOD (b) DYNOTEARS (c) MVGC




(d) PWGC (e) NAVAR (f) PCMCI

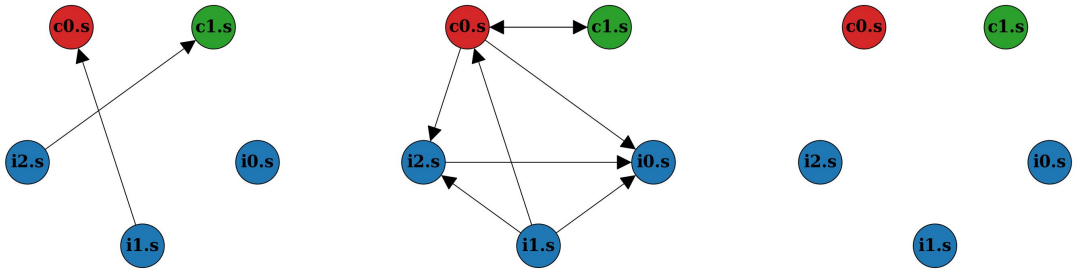



(g) TCDF (h) TiMINo (i) tsFCI




(j) VAR-LiNGAM (k) Random (l) Ground Truth

Figure 3.3: Depicts the causal-summary graphs from applying each of the evaluated methods to a scene from the highD dataset with acceleration as the variable of interest ($\lambda_\alpha = 0.05$, $\tau = 2.5\,s$), along with the ground truth (Figure 3.3l). Most methods appear to either exhibit too high a sensitivity (Figures 3.3a, 3.3d, 3.3h, and 3.3j) or too high a specificity (Figures 3.3c, 3.3e, 3.3g, and 3.3i), with only DYNOTEARS offering a competitive result (Figure 3.3b).




(a) CD-NOD (b) DYNOTEARS (c) MVGC

(d) PWGC (e) NAVAR (f) PCMCI

(g) TCDF (h) TiMINo (i) tsFCI

(j) VAR-LiNGAM (k) Random (l) Ground Truth

Figure 3.4: Depicts the causal-summary graphs from applying each of the evaluated methods to a scene from the highD dataset with speed as the variable of interest ($\lambda_\alpha = 0.05$, $\tau = 2.5\,s$), along with the ground truth (Figure 3.4l). Most methods appear to either exhibit too high a sensitivity (Figures 3.4b, 3.4d, 3.4h, and 3.4j) or too high a specificity (Figures 3.4c, 3.4e, 3.4g, and 3.4i), with only PCMCI offering a competitive result (Figure 3.4f).

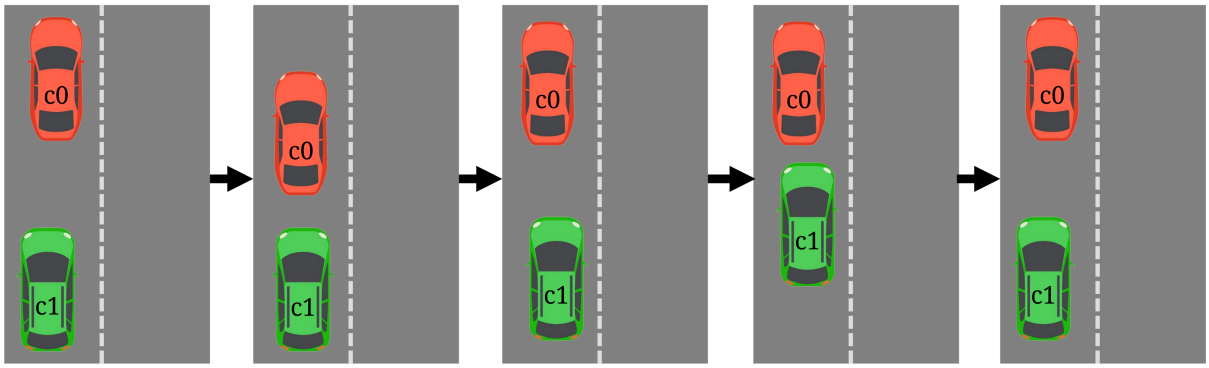
# Chapter 4

# Identifying Agent Behavioural Interactions via Theory of Mind

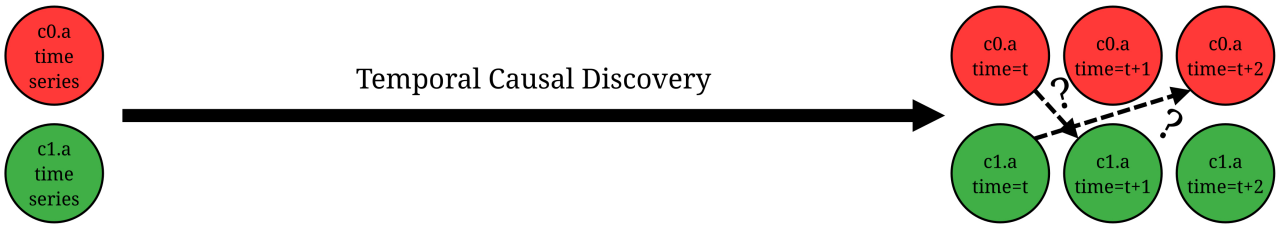
In Chapter 3 the idea of utilising causality for capturing behavioural interactions was introduced, alongside an evaluation of existing causal-discovery approaches for identifying such interactions. However the benchmark of temporal causal-discovery techniques demonstrated the unreadiness of such methods to be applied to autonomous-agent data. In this chapter an action-based perspective on the causal representation of behavioural interactions is introduced (Section 4.1) and a novel method of identifying behavioural interactions via theory of mind is proposed and evaluated (Section 4.2). Figure 4.1 offers some preliminary insight as to how the approach presented in this chapter differs from those considered in the previous chapter.

## 4.1 Casual Relationships Between Agent Actions

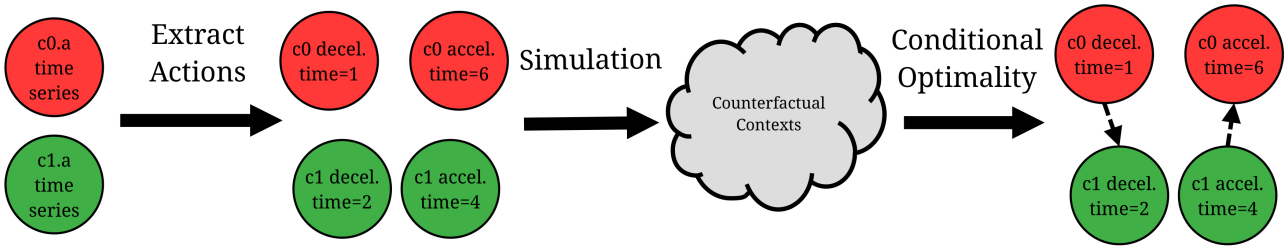
Section 3.2 identified two issues that traditional temporal causal-discovery approaches encounter in trying to identify behavioural interactions between autonomous agents. While causal sparsity is indeed a property that greatly increases the difficulty of such identifications, causal non-stationarity presents a more fundamental issue for the actual representation of behavioural interactions. If causal non-stationarity is present then a window-time-graph representation will not suffice, yet a full time graph would require one to identify links from variables across any number of previous time steps to the variables being analysed in the current time step. What is required then is a representation that provides a similar expressiveness to a full time graph, yet retains the size, simplicity, and interpretability of a window time graph.




(a) A scenario illustrative of causal non-stationarity in the behavioural interactions of convoy agents **c0** and **c1**. In this scenario **c0** and **c1** are initially travelling at similar speeds, however **c0** begins to decelerate, causing **c1** to have to decelerate in response. Shortly after however **c1** begins to tailgate **c0** by accelerating, causing **c0** to accelerate in response. Thus because of the switch in causal direction as well as the inconsistent causal time lags the scene lacks causal stationarity.




(b) Depiction of traditional temporal causal discovery applied to this scenario. Typically such methods work directly upon the input time-series data and then output a window-time graph, although in some cases only a causal-summary graph is given as output. In any case, as a result of the causal non-stationarity of the scenario most temporal causal-discovery methods would struggle to identify and subsequently represent the behavioural interactions taking place.




(c) Depiction of the theory of mind approach presented here. The first step in this process is extracting behaviour of each agent in terms of discrete actions rather than continuous variables. Since these actions now represent time-specific instances of behaviour rather than variables persisting throughout the scene, the issue of causal stationarity is removed, as the causal relationships are specific to each action pair. Once actions have been extracted, once can utilise simulation to envisage counterfactual outcomes. By considering how the presence of absence of actions affect the optimality of the outcomes for agents, one can then identify behavioural interactions as causal relationships between actions.

Figure 4.1: Visualisation of the difference between the presented approach and traditional temporal causal-discovery methods.

### 4.1.1 Viewing Agents as Decision Makers

In order to develop a solution that can tackle causal non-stationarity one must first identify the reasons for non-stationarity being present to begin with. One potential explanation for this is that autonomous agents cannot be modelled as simple control systems. This is backed up by the fact that the synthetic dataset in Section 3.2 performed quite well compared to real-data, and was generated with proportional-feedback control. Within the AV domain this runs counter to some existing research that makes use of frameworks such as IDMs [93]. Such models treat drivers as control systems which attempt to achieve a desired speed while maintaining a safe distance ahead. This is a valid assumption in many cases. However, even in these cases stationarity is not guaranteed due to inconsistencies such as reaction times. These models are also limited as they only consider target speed and not actions such as lane changing. This suggests that such models are a simplification of the true process behind agents, and if one wishes to gain insights into the behavioural interactions of agents, one must consider the world from their perspective.

While autonomous agents may exhibit some behaviour that resembles a control system such as an IDM, this work argues that such control is ultimately motivated by decisions made by the agents in the form of actions. Thus, a causal model that wishes to capture behavioural interactions between agents should reason in terms of said actions. Actions as a means of expressing the intentions and behaviour of agents are by no means novel, with commonly used frameworks such as MDPs [155] and action languages (e.g. Stanford research institute problem solver [156], planning domain definition language [157], action description language [158]) making them a foundational part of their functioning. This work does not attempt to advance planning nor action description, instead it focuses on how and why agent behavioural interactions should be formulated in terms of their discrete actions. In doing so there is a departure from the vast majority of work within causal discovery, which typically attempts to find generalisable causal relationships between continuous variables — versus individual relationships between discrete actions. In addition to the practical reason of overcoming non-stationarity as an obstacle, it is also argued that reasoning in terms of actions more closely matches human cognition. Consider the following statements, in the context of the scenario described in Section 3.1.3, both reflective of the same observation:

1. 'The speed of agent **c1** is proportional to the speed of agent **c0**, with a time lag of $1\,s$'

2. 'Agent **c1** braked at $5\,s$ in response to agent **c0** braking at $4\,s$'

Even assuming the stationarity assumed by the first statement holds, the latter statement is debatably much more in line with the manner which humans think and talk about events. It is with all these factors in mind, that a causal action-based representation of agent behavioural interactions is proposed.

### 4.1.2 Autonomous-Agent SCMs

Having discussed the motivations behind basing behavioural-interaction representation upon actions, the formalisms that permit this representation are now introduced. First and foremost, a novel temporal-SCM definition is proposed based upon a theory of mind for autonomous agents. While this is built as an extension of window-time-graph SCMs it is assumed the window time graph exhibits the Markov property.

**Definition 10** (Markov Property)**.** The Markov property assumes that within a stochastic process the probability distribution over the next state depends only upon the current state, i.e. the process is memoryless.

In the context of causality this effectively only permits causal links with a time lag of 1. It is still possible to practically have causal relationships with a time lag greater than one, however this must be achieved via proxy variables to store values across multiple time steps.

Within this work this term will be used to refer to any memoryless process represented via SCMs. This is despite the fact there are instances — due to interventions or exogenous variables with degenerate distributions — where SCMs may represent deterministic memoryless processes. Nevertheless, SCMs generally are stochastic as a result of their exogenous variables capturing factors outside of those modelled via structural equations.

Based upon these concepts a novel temporal-SCM definition can be made, designed to represent autonomous agents via an SCM-based framework.

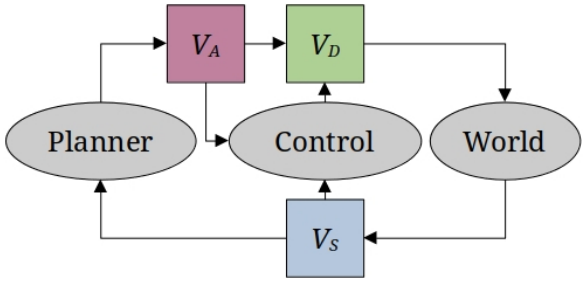



Figure 4.2: Overview of the autonomous-agent SCM structure.

**Definition 11** (Autonomous-Agent SCM)**.** For a set of relative time lags $L^A = \{0, 1\}$ an autonomous-agent SCM, based upon Definitions 7 and 10, is a tuple:

$$M^A = \langle U^A,\ V^A,\ E^A,\ F^A,\ P(U) \rangle \tag{4.1}$$

with its elements defined as follows:

- $U^A \subseteq U \times L^A$ is a set of exogenous variables indexed to each relative time lag in $L^A$. $U^A_{i,l} \equiv \langle U_i, l \rangle \in U^A$ refers to exogenous variable $U_i$ indexed to relative time lag $l$.
- $V^A \subseteq (\{V_S, V_A, V_D\} \cup V^O) \times L^A$ is a set of endogenous variables indexed to each relative time lag in $L^A$. Here $V_S$ is the world state captured as a variable, $V_A$ is the composite action of the agent stored as a variable, $V_D$ are the driving-actuation values of the agent in a variable, and $V^O$ are all other variables besides. $V^A_{i,l} \equiv \langle V_i, l \rangle \in V^A$ refers to endogenous variable $V_i$ indexed to relative time lag $l$.
- $E^A \subset (U^A \cup V^A) \times V^A$ is a set of directed edges that indicate the first time-lag-indexed variable of the tuple is an input for the second time-lag-indexed variable of the tuple — which is endogenous by definition. Otherwise matches Definition 7.
- $F^A = \{f^A_0, f^A_1, ...\}$ is a set of structural equations that derive the values of time-lag-indexed endogenous variables. Otherwise matches Definition 7.
- $P(U)$ is the joint probability distribution over the domains of the exogenous variables $U$. Otherwise matches Definition 7.

In addition to the above definition, an overview of the general structure of autonomous-agent SCMs is provided in Figure 4.2. The main aspects that differentiate this specific SCM definition

from a typical window-time-graph SCM is the integration of the Markov property along with the expectation of a particular structure based around the $V_S$, $V_A$, and $V_D$ variables. The operation of agents based upon this structure is expected to follow the ensuing steps:

1. $V_{S,t}$ is updated based upon the world state at the time $t$.
2. A planner takes the state variable $V_{S,t}$ and outputs the composite-action variable $V_{A,t}$ for time $t$.
3. The driving-actuation values for time $t$, $V_{D,t}$ are derived from a combination of action values that are directly controllable from $V_{A,t}$ in addition to a controller's outputs based upon inputs from goal-state values from $V_{A,t}$ and current state values in $V_{S,t}$.
4. Driving-actuation values in $V_{D,t}$ are used to interact with the world, the current time advances to $t+1$ before the loop returns to step 1.

Based upon this structure and series of operations one can identify variable $V_A$ as the primary variable of interest for the task of representing behavioural interactions between agents. With this in mind one can define a novel temporal graph that captures behavioural interactions based upon the full-time-graph and autonomous-agent SCMs.

**Definition 12** (Behavioural-Interaction Graph)**.** For time steps $T \subseteq \mathbb{N}$ a behavioural-interaction graph, based upon Definitions 4 and 11, is a tuple:

$$\mathcal{G}^{BI} = \langle N^{BI},\ E^{BI} \rangle \tag{4.2}$$

with its elements defined as follows:

- $N^{BI} = \{V_{0,A,0}, V_{0,A,1}, ..., V_{1,A,0}, V_{1,A,1}, ...\}$ is a set of the composite-action variables for multiple agents across the time steps described by $T$. Here $V_i$ denotes all variables for an agent $i$, $V_{i,A}$ the composite-action variable across time for agent $i$, and $V_{i,A,t}$ the composite-action variable for agent $i$ at time $t$. While the autonomous-agent SCM described by Definition 11 specifies causal relationships in terms of relative time lags, it is assumed that in this instance one is analysing a period of time over which the autonomous-agent SCM has been rolled out post-hoc.

- $E^{BI} = \{\langle V_{i,A,t}, V_{i',A,t'}\rangle \mid t < t',\ i \neq j,\ V_{i,A,t} \in An^T(V_{i',A,t'}),\ V_{i',A,t'} \in V_0 \cup V_1 \cup ...\}$ is a set of directed edges which indicate the first composite-action variable of the tuple has a causal effect upon the second composite-action variable of the tuple based upon their values at times $t$ and $t'$ respectively. This can effectively be used to describe a causal relationship between two discrete agent actions, or what can otherwise be considered a representation of behavioural interactions between agents.

With the definition of this structure one can redefine the objective of this work as identifying the edges $E^{BI}$ that describe the behavioural interactions present within a given scene.

Lastly, in order to contextualise the actions and the previously-described structures a working format for actions is described within the AV domain. The value $a_{i,t}$ of a composite-action variable $V_{i,A,t}$ of an agent $i$ at time $t$ is defined as follows:

$$a_{i,t} = \langle g_{i,t,sp},\ g_{i,t,la}\rangle \tag{4.3}$$

where $g_{i,t,sp}$ and $g_{i,t,la}$ refer to speed and lane goal values respectively. Here a goal $g_{i,t,x}$ is defined for agent $i$ with state variable $x$ at time $t$ as:

$$g_{i,t,x} = \langle\langle s_{i,t',x},\ t'\rangle,\ t\rangle \approxeq \langle s_{i,t',x},\ t'\rangle \tag{4.4}$$

where $s_{i,t',x}$ and $t'$ refer to the desired state variable value and target time respectively for a goal of agent $i$ at time $t$ — i.e. achieve value $s_{i,t',x}$ by time $t'$. As the notation suggests, often when working directly with goals (e.g. for their extraction) they include their start time as part of the goal tuple structure. However, in other instances it makes little sense to include this information — i.e. because the start time for the current goal will always be the current time step. Thus the exact format used in any particular place will depend upon what is most convenient considering the operations taking place, however this should be clear via the context provided by any given discussion. Lastly, note that $g_{i,t,x,x} \equiv s_{i,t',x}$ and $g_{i,t,x,t'} \equiv t'$ may be used as an alternative form of notation in some instances throughout this work, again depending upon which form offers the most clarity for any given discussion.

### 4.1.3 Extracting Autonomous-Agent Actions

The motivations for using an action-based representation for behavioural interactions have been established and the structures that can be used for this purpose have been formalised. Now one can begin to consider how the relevant actions may be inferred from data gathered from a scene. While identifying behavioural interactions between agents is largely domain independent, this step does require a level of understanding of the domain. Here the current speed and lane of a vehicular agent are considered as the state variables, the desired speed and lane as the action variables, and the acceleration and steer as the driving-actuation variables. However, the method shown here for extracting actions could be applied to any domain in which variables possess a similar relationship.

The first step to identifying vehicular actions is to identify the discrete speed goals. A set of potential speed-goal start times $T_i^{g,sp,A}$ and target times $T_i^{g,sp,\Omega}$, are derived as follows:

$$\begin{aligned} T_i^{g,sp,A} = \{\ t \in T \mid\ & (d_{i,t,a} < \lambda_{d,a} \wedge d_{i,(t+1),a} \geq \lambda_{d,a}) \vee \\ & (d_{i,t,a} > -\lambda_{d,a} \wedge d_{i,(t+1),a} \leq -\lambda_{d,a}), \\ & d_{i,t,a},\ d_{i,(t+1),a} \in \mathbb{R}\ \}\ \cup\ \{\ \min\ T\ \} \end{aligned} \tag{4.5}$$

$$\begin{aligned} T_i^{g,sp,\Omega} = \{\ t \in T \mid\ & (d_{i,t,a} \geq \lambda_{d,a} \wedge d_{i,(t+1),a} < \lambda_{d,a}) \vee \\ & (d_{i,t,a} \leq -\lambda_{d,a} \wedge d_{i,(t+1),a} > -\lambda_{d,a}), \\ & d_{i,t,a},\ d_{i,(t+1),a} \in \mathbb{R}\ \}\ \cup\ \{\ \max\ T\ \} \end{aligned} \tag{4.6}$$

where $d_{i,t,a} \in \mathbb{R}$ is the value associated with the acceleration component of the driving-actuation variable $V_D$ of agent $i$ at time step $t$. Likewise $d_{i,(t+1),a} \in \mathbb{R}$ is the same except for time step $(t+1)$. Meanwhile $\lambda_{d,a}$ is a threshold used to determine there has been sufficient actuation in order to conclude a new speed goal was set.

Assuming that $T_i^{g,sp,A}$ and $T_i^{g,sp,\Omega}$ are sorted and can be indexed, Algorithm 1 describes the process by which the set of time-indexed goals $\boldsymbol{g}_{i,sp}$ can be extracted for agent $i$. Here $s_{i,t,sp} \in S_s$ is the value associated with the speed component of the state variable $V_S$ of agent $i$ at time $t$. Then $\lambda_{\delta a,sp,t}$ is a threshold used to ensure there is enough time between the action start time and the goal target time in order to reflect a protracted effort from the agent. Lastly, $\lambda_{\delta a,sp,sp}$ is

**Algorithm 1:** Autonomous-Agent Speed Goal Extraction

**Data:** $V_S$, $T$, $T_i^{g,sp,A}$, $T_i^{g,sp,\Omega}$
**Result:** $\boldsymbol{g}_{i,sp}$
$\boldsymbol{g}_{i,sp} := \{\}$, $j := 0$, $k := 0$;
**while** $j < |T_i^{g,sp,A}| \ \wedge \ k < |T_i^{g,sp,\Omega}|$ **do**
  $t = T_i^{g,sp,A}[j]$, $t' = T_i^{g,sp,\Omega}[k]$;
  **if** $t' - t \geq \lambda_{\delta a,sp,t} \ \wedge \ |s_{i,t,sp} - s_{i,t',sp}| \geq \lambda_{\delta a,sp,sp}$ **then**
    $g_{i,t,sp} := \langle\langle s_{i,t',sp},\ t'\rangle,\ t\rangle$;
    $\boldsymbol{g}_{i,sp} := \boldsymbol{g}_{i,sp} \ \cup \ \{\ g_{i,t,sp}\ \}$;
    **while** $T_i^{g,sp,A}[j] < T_i^{g,sp,\Omega}[k]$ **do**
      $j$++;
    **end**
  **else**
    $k$++;
  **end**
**end**
**if** $|\boldsymbol{g}_{i,sp}| = 0$ **then**
  $t := \min\ T$;
  $g_{i,t,sp} := \langle\langle s_{i,t,sp},\ t\rangle,\ t\rangle$;
  $\boldsymbol{g}_{i,sp} := \{\ g_{i,t,sp}\ \}$;
**end**

used to determine whether there is a sufficient difference between the speed at decision time and the goal target speed in order to be significant. Assuming both of these conditions are met, an additional goal is added to $\boldsymbol{g}_{i,sp}$, otherwise the next potential goal target time is considered. Whenever a goal is added to $\boldsymbol{g}_{i,sp}$, the next potential goal start time is selected from those greater than or equal to the target time of the recently added goal, as it is assumed only one goal can be pursued at a time.

With the set of speed goals established, one must now extract the time-indexed lane goals for vehicular agent $i$. Unlike the speed goals, these can be extracted by set comprehension alone:

$$\begin{aligned}\boldsymbol{g}_{i,la} = \{\ \langle\langle s_{i,t,la},\ t\rangle,\ t - \lambda_{\delta a,la,t}\rangle \mid \forall_{t'=t+1}^{t+\lambda'_{\delta a,la,t}} s_{i,t',la} = s_{i,t,la},\ s_{i,t,la} \neq s_{i,(t-1),la},\\ s_{i,t,la},\ s_{i,(t-1),la} \in \mathcal{T}^{V_{S,la}},\ t \in T\ \}\ \cup\ \{\ \langle\langle s_{i,(\min T),la},\ \min\ T\rangle,\ \min\ T\rangle\ \}\end{aligned} \tag{4.7}$$

where $s_{i,t,la} \in \mathcal{T}^{V_{S,la}}$ is the value associated with the lane component of the state variable $V_S$ of agent $i$ at time $t$. Similar to the speed-goal extraction $s_{i,(t-1),la} \in \mathcal{T}^{V_{S,la}}$ is the same as $s_{i,t,la}$ except it considers time step $(t-1)$. Meanwhile $\lambda_{\delta a,la,t}$ represents a threshold on the time prior

**Algorithm 2:** Autonomous-Agent Action Extraction via Goal Merging

**Data:** $\boldsymbol{g}_{i,sp}$, $\boldsymbol{g}_{i,la}$
**Result:** $\boldsymbol{a}_i$
$\boldsymbol{a}_i := \{\}$, $j := 0$, $k := 0$;
**do**
  $g_{i,sp} := \boldsymbol{g}_{i,sp}[j]$;
  $g_{i,la} := \boldsymbol{g}_{i,la}[k]$;
  $a_i := \langle\langle g_{i,sp}[0],\ g_{i,la}[0]\rangle,\ \max\{\ g_{i,sp}[1],\ g_{i,la}[1]\ \}\rangle$;
  $\boldsymbol{a}_i := \boldsymbol{a}_i\ \cup\ a_i$;
  **if** $j = |\boldsymbol{g}_{i,sp}| - 1$ **then**
    $k := k + 1$;
  **else if** $k = |\boldsymbol{g}_{i,la}| - 1$ **then**
    $j := j + 1$;
  **else**
    **if** $\boldsymbol{g}_{i,sp}[j][1] = \boldsymbol{g}_{i,la}[k][1]$ **then**
      $j := j + 1$;
      $k := k + 1$;
    **else if** $\boldsymbol{g}_{i,sp}[j][1] < \boldsymbol{g}_{i,la}[k][1]$ **then**
      $j := j + 1$;
    **else**
      $k := k + 1$;
    **end**
  **end**
**while** $j < |\boldsymbol{g}_{i,sp}| \wedge k < |\boldsymbol{g}_{i,la}|$;

to lane change the lane-goal must have begun and $\lambda'_{\delta a,la,t}$ represents how long an agent must have remained in a lane for a change of lane-goal to have occurred.

With sets of speed and lane goals now established, one can formulate a series of time-indexed actions by combining the goal sets, as per Algorithm 2. Similar to the previous algorithm, it is required that the speed-goal set $\boldsymbol{g}_{i,sp}$ and lane-goal set $\boldsymbol{g}_{i,la}$ can be indexed and are sorted based upon their time-indexing. Thus the time-indexed actions for the scene have been extracted, and one can begin to consider how causal relationships between agent actions can be identified.

Before proceeding to discuss the means by which one might identify said causal relationships it is important to note the computational overhead of this extraction process. At least for the domains considered throughout this work the process described above scales linearly with scene duration and the number of agents. Regardless the extraction process typically takes up minimal computation time in comparison to the subsequent inference of causal relationships.

## 4.2 Discovering Causal Links Between Actions via Theory of Mind

Having remedied the obstacle of causal non-stationarity by adopting an action-based perspective to behavioural interactions (see Figure 4.1), one can now begin to tackle the difficulties presented by causal sparsity. Since the goal is to now identify causal links between individual actions rather than variables, the problem of causal sparsity is exasperated, as one must identify the presence of a causal link based upon only the information captured by the two actions involved. The solution presented in this section utilises the theory-of-mind concept introduced in Section 4.1, a variety of reward metrics, and counterfactual simulation in order to achieve this.

### 4.2.1 Simulation-Based Counterfactual Action-Causal-Link Identification

The actual identification of causal links between agent actions requires the establishment of several definitions pertaining to causality. From here the process of simulating counterfactual timelines can be outlined, and the resulting notations described.

#### 4.2.1.1 Causal Necessity & Conditional Optimality

Before discussing the means by which causal relationships are identified in this work, it should be qualified that it is specifically causal relationships of necessity that are considered here. Below this property is described, in terms of two actions $a_{\mathcal{C},t}$ and $a_{\mathcal{A},t'}$ of agents $\mathcal{C}$ and $\mathcal{A}$ respectively.

**Definition 13** (Causal Necessity)**.** Action $a_{\mathcal{C},t}$ was necessary for action $a_{\mathcal{A},t'}$ to have occurred, if $a_{\mathcal{A},t'}$ would not have occurred had $a_{\mathcal{C},t}$ not occurred. This does not imply that $a_{\mathcal{C},t}$ was solely capable of causing $a_{\mathcal{A},t'}$ to occur — as is the case with causal sufficiency — just that its occurrence was part of a chain of required events that caused $a_{\mathcal{A},t'}$ to happen.

With this established, there is now the question of how one can determine that one action was necessary in bringing about another. For this, one can consider the theory of mind introduced in Section 4.1 and depicted in Figure 4.2. Here the actions of an agent are captured through the variable $V_A$ and are derived from a planner. While it is difficult to infer the exact cognitive process by which an agent plans, in the absence of an immediate goal state it is reasonable to assume that either a heuristic or reward mechanism is used to guide agent decisions. Furthermore,

one would assume that agents prefer actions / plans that are optimal or near-optimal based upon said mechanism. In the context of causal-relationship identification this leads this work to introduce the concept of conditional optimality.

**Definition 14** (Conditional Optimality)**.** Action $a_{\mathcal{A},t'}$ is conditionally optimal on action $a_{\mathcal{C},t}$ occurring, if $a_{\mathcal{A},t'}$ would not have been optimal — or near-optimal — had $a_{\mathcal{C},t}$ not occurred.

While quite similar to Definition 13, this definition incorporates intuitions regarding agent planners. Conditional optimality for two potentially causally-linked actions is significantly easier to test. One only need determine how the occurrence of $a_{\mathcal{C},t}$ affects the optimality of $a_{\mathcal{A},t'}$ based upon some metric. Before one can consider this metric though, it is first necessary to consider how data on an alternate occurrence of actions can be gathered. To this end it is proposed that simulations be used to counterfactually gather information about alternate worlds in which certain actions did not occur.

#### 4.2.1.2 Deriving Counterfactual Data via Simulation

Given that SCMs describe the data-generation process for a given system, it is relatively simple to simulate counterfactual timelines by rolling out an SCM after abducting exogenous-variable distributions and intervening upon the desired variables. If considering agents at a high level, one can refer to Figure 4.2. In this figure, the action variable $V_A$ is being intervened upon by altering the occurrence of the potential cause and effect actions, resulting in the planner component of the SCM loop effectively being removed. Thus, provided one has determined how to emulate the control and world components of the SCM, it should be possible to perform the counterfactual roll-outs.

**World** Within many agent-oriented domains the high-level world component is knowable and can be well defined. Within simplified problems this could be a grid-world representation, e.g. [11, 159]. However, more frequently a physics model — typically based upon rigid bodies [160] — is utilised to simulate how agents interact with each other and the world around them, based upon their driving-actuation values and the previous world state. For the AV domain considered here, a 2D rigid-body simulator is utilised with vehicles represented via their

rectangular bounding box. Into this simulator, each vehicle supplies its acceleration and steering as an output from the control component, based upon the input from action variable $V_A$.

**Control** In contrast to the world component of an agent SCM, the control component is harder to infer. This component is responsible for converting the value of the action variable $V_A$ into the value of the driving-actuation variable $V_D$ for a given time step. Thus the exact structure and cognition of this component is highly domain dependant. As such while a general solution is not available, a description of the approach taken here for the AV domain is provided.

Based upon the AV action defined in (4.3) and the driving-actuation values described in Section 4.1.3 it is necessary to convert speed and lane goals into acceleration and steering values. For a speed goal $g_{i,t,sp} = \langle s_{i,t',sp}, t' \rangle$ the required acceleration $d_{i,t,a}$ is calculated as follows:

$$d_{i,t,a} = \frac{s_{i,t',sp} - s_{i,t,sp}}{\delta_t \ \max\{\ t' - t,\ 1\ \}} \tag{4.8}$$

where $\delta_t$ is the fixed time-step spacing in terms of real-time. This effectively just calculates the average acceleration required to reach the goal speed, at the goal time. However, this does come with the caveat, that if the goal time is in the present or past, the control will try to have the agent reach the goal speed within a single time step.

The steering is of greater difficulty to derive and here is defined in terms of curvature, i.e. radians per metre. From here the lane goal $g_{i,t,la} = \langle s_{i,t',la}, t' \rangle$ is used to calculate the required steering $d_{i,t,st}$ as follows:

$$d_{i,t,st} = \frac{st_{la}(s_{i,t',la}) \ + \ st_{mid}(g_{i,t,la},\ s_{i,t})}{2} \tag{4.9}$$

where $st_{la}(\cdot) : \mathcal{T}^{V_{S,la}} \to \mathbb{R}$ is a function that returns the curvature of the input lane identifier. Meanwhile $st_{mid}(\cdot) : \mathcal{T}^{V_{S,la}} \to \mathbb{R}$ is a function that utilises the input lane identifier to determine the steering required to face towards the midpoint of the lane, and is defined as follows:

$$st_{mid}(g_{i,t,la},\ s_{i,t}) = \frac{\text{atan2}\left((l_{mid}(s_{i,t',la}) - s_{i,t,pos}) \times s_{i,t,dir},\ (l_{mid}(s_{i,t',la}) - s_{i,t,pos}) \cdot s_{i,t,dir}\right)}{s_{i,t,sp}\,\delta_t \ \max\{\ t' - t,\ 1\ \}} \tag{4.10}$$

where $s_{i,t,pos}$ and $s_{i,t,dir}$ are the state values associated with the position and direction components of agent $i$ at time $t$. As for the function $l_{mid}(\cdot) : \mathcal{T}^{V_{S,la}} \rightarrow \mathbb{R}^2$, it takes the lane component of a state value and returns the midpoint associated with that lane. Thus the steering control uses a mixture of lane-curvature information along with directional error based upon current agent positioning in order to derive its output.

By combining these two control mechanisms, one has sufficient parts for the control component of the AV-domain framework utilised here, as described by the SCM design depicted in Figure 4.2. From here the concept of counterfactual contexts can be introduced, in turn allowing one to begin considering how optimality metrics can identify causal relationships between actions.

#### 4.2.1.3 Counterfactual Contexts

Having described the general concept behind the proposed method of this work, as well as how the necessary counterfactual data might be gathered via simulation, new notation is introduced in order to formally define the optimality metrics considered here. Specifically, the idea of a counterfactual context denoted via $[\,\cdot\,]^X$ is introduced. This corresponds to evaluating the contents of the square brackets under a counterfactual simulation in which the intervention specified by $X$ is applied. In this work, only three counterfactual contexts need be considered:

- The counterfactual context in which the potentially causing action $a_{\mathcal{C},t}$ of agent $\mathcal{C}$ does not occur. This is formally denoted by $[\,\cdot\,]^{\{\neg a_{\mathcal{C},t}\}}$, but for succinctness $[\,\cdot\,]^{\neg\mathcal{C}}$ will be used.

- The counterfactual context in which the potentially affected action $a_{\mathcal{A},t'}$ of agent $\mathcal{A}$ does not occur. This is formally denoted by $[\,\cdot\,]^{\{\neg a_{\mathcal{A},t'}\}}$, but for succinctness $[\,\cdot\,]^{\neg\mathcal{E}}$ will be used.

- The counterfactual context in which neither the potentially causing action $a_{\mathcal{C},t}$ of agent $\mathcal{C}$ nor the potentially affected action $a_{\mathcal{A},t'}$ of agent $\mathcal{A}$ occurs. This is formally denoted by $[\,\cdot\,]^{\{\neg a_{\mathcal{C},t},\, \neg a_{\mathcal{A},t'}\}}$, but for succinctness $[\,\cdot\,]^{\neg\mathcal{C},\,\neg\mathcal{E}}$ will be used.

Lastly, in the absence of a counterfactual context indicated through the square bracket notation, it can be assumed that any values referenced correspond to those from the originally observed scene, devoid of any interventions.

### 4.2.2 Inferring Causal Relationships via Optimality Metrics

Having now introduced the concepts of conditional optimality and the means by which data can be gathered via counterfactual simulation, one remaining task is to determine how to measure optimality for an autonomous agent. Once again this is considered from the perspective of the AV domain in keeping with previous sections. To this end three variations of optimality metric are presented here: reward-based, agency-based, and hybrid.

#### 4.2.2.1 Reward-Based Variant

The reward-based variant makes the assumption there is some numerical metric to measure how desirable a particular world state is from the perspective of an agent. For the state variable $V_S$ the value for agent $i$ at time $t$ is denoted as $s_{i,t}$. For this state value the reward function $r(\cdot) : \mathcal{T}^{V_S} \to \mathbb{R}$ is defined as follows:

$$r(s_{i,t}) = r_{ttc}(s_{i,t,ttc}) \cdot r_{cct}(s_{i,t,cct}) \cdot r_s(s_{i,t,sp}) \tag{4.11}$$

$$r_{ttc}(s_{i,t,ttc}) = 1 - \exp(-s_{i,t,ttc}) \tag{4.12}$$

$$r_{cct}(s_{i,t,cct}) = \begin{cases} 1, & s_{i,t,cct} \leq 0 \\ 0, & s_{i,t,cct} > 0 \end{cases} \tag{4.13}$$

$$r_s(s_{i,t,sp}) = 1 - 0.5 \exp(-\max\{\ 0.1\, s_{i,t,sp},\ 0\ \}) \tag{4.14}$$

where $s_{i,t,ttc}$, $s_{i,t,cct}$, and $s_{i,t,sp}$ refer to the TTC, Cumulative Collision Time (CCT), and speed values of $s_{i,t}$ respectively. Here the overall reward given by (4.11) is calculated as a product of the other reward metrics provided.

For a pair of potential cause and effect actions $a_{\mathcal{C},t}$ and $a_{\mathcal{A},t'}$, one can begin to derive a metric for the likelihood of a causal link being present between the actions by considering the minimum reward under different counterfactual contexts. To do this a measure of the extent to which the rewards indicate conditional optimality is defined:

$$\delta_{r,\mathcal{A}} = \delta^{+}_{r,\mathcal{A}} + \delta^{-}_{r,\mathcal{A}} \tag{4.15}$$

$$\delta^{+}_{r,\mathcal{A}} = \left(\min_{t'' \in (t', t'+\beta_h]} r(s_{\mathcal{A},t''})\right) - \left[\min_{t'' \in (t', t'+\beta_h]} r(s_{\mathcal{A},t''})\right]^{\neg\mathcal{E}} \tag{4.16}$$

$$\delta^{-}_{r,\mathcal{A}} = \left[\min_{t'' \in (t', t'+\beta_h]} r(s_{\mathcal{A},t''})\right]^{\neg\mathcal{C},\neg\mathcal{E}} - \left[\min_{t'' \in (t', t'+\beta_h]} r(s_{\mathcal{A},t''})\right]^{\neg\mathcal{C}} \tag{4.17}$$

where $\delta_{r,\mathcal{A}}$ is formed from a combination of $\delta^{+}_{r,\mathcal{A}}$ and $\delta^{-}_{r,\mathcal{A}}$. Here $\delta^{+}_{r,\mathcal{A}}$ measures the extent to which agent $\mathcal{A}$ taking $a_{\mathcal{A},t'}$ increases reward given that $a_{\mathcal{C},t}$ occurs. Meanwhile $\delta^{-}_{r,\mathcal{A}}$ measures the extent to which agent $\mathcal{A}$ taking $a_{\mathcal{A},t'}$ decreases reward given that $a_{\mathcal{C},t}$ does not occur. These are calculated based upon a simulation horizon $\beta_h$ which defines how far into the future reward is considered. Once $\delta_{r,\mathcal{A}}$ is derived, provided it is above a threshold $\lambda_{\delta r}$ a causal link is identified between $a_{\mathcal{C},t}$ and $a_{\mathcal{A},t'}$, as the following describes:

$$\xi^{\text{reward}}_{a_{\mathcal{A},t'}} = \begin{cases} 1, & \delta_{r,\mathcal{A}} \geq \lambda_{\delta r} \\ 0, & \delta_{r,\mathcal{A}} < \lambda_{\delta r} \end{cases} \tag{4.18}$$

#### 4.2.2.2 Agency-Based Variant

Agency in the context of this work refers to an agent's ability to carry out its actions as planned. The most common example of this property being violated is as a result of a collision between agents. Environmental factors such as wind and ice could also influence this, however such factors are harder to capture. With this in mind the agency function $ag(\cdot) : 2^{\mathcal{T}^{V_S} \times T} \rightarrow 2^{\mathcal{T}^{V_S} \times T} \rightarrow \mathbb{B}$ is defined given two state variable time series $s_{\mathcal{C}}, s_{\mathcal{A}} \in 2^{\mathcal{T}^{V_S} \times T}$ as follows:

$$ag(s_{\mathcal{C}},\ s_{\mathcal{A}}) = \bigwedge_{t'' \in (t', t'+\beta_h]} \begin{cases} 0, & s_{\mathcal{C},t'',cct} > 0 \wedge s_{\mathcal{A},t'',cct} > 0 \wedge s_{\mathcal{C},(t''-1),cct} \leq 0 \wedge s_{\mathcal{A},(t''-1),cct} \leq 0 \\ 1, & \text{otherwise} \end{cases} \tag{4.19}$$

The function relies upon states from both agents $\mathcal{C}$ and $\mathcal{A}$ in order to ensure that the loss of agency is due to a collision between the agents of interest, and not from a third entity. It is assumed that an agent will act in such a way as to maintain their ability to enact their own

actions. As such, now there is a way of calculating agency, one can describe four agency patterns which pertain to causal relationships or the lack thereof:

$$\xi^{\text{active}}_{a_{\mathcal{A},t'}} = (ag(s_{\mathcal{C}},\ s_{\mathcal{A}})) \wedge \neg[ag(s_{\mathcal{C}},\ s_{\mathcal{A}})]^{\neg\mathcal{E}} \wedge [ag(s_{\mathcal{C}},\ s_{\mathcal{A}})]^{\neg\mathcal{C},\neg\mathcal{E}} \tag{4.20}$$

$$\xi^{\text{passi.}}_{a_{\mathcal{A},t'}} = (ag(s_{\mathcal{C}},\ s_{\mathcal{A}})) \wedge \neg[ag(s_{\mathcal{C}},\ s_{\mathcal{A}})]^{\neg\mathcal{C}} \wedge [ag(s_{\mathcal{C}},\ s_{\mathcal{A}})]^{\neg\mathcal{C},\neg\mathcal{E}} \tag{4.21}$$

$$\xi^{\text{facil.}}_{a_{\mathcal{A},t'}} = (ag(s_{\mathcal{C}},\ s_{\mathcal{A}})) \wedge \neg[ag(s_{\mathcal{C}},\ s_{\mathcal{A}})]^{\neg\mathcal{C}} \wedge \neg[ag(s_{\mathcal{C}},\ s_{\mathcal{A}})]^{\neg\mathcal{C},\neg\mathcal{E}} \tag{4.22}$$

$$\xi^{\text{m.e.m.}}_{a_{\mathcal{A},t'}} = (ag(s_{\mathcal{C}},\ s_{\mathcal{A}})) \wedge \neg[ag(s_{\mathcal{C}},\ s_{\mathcal{A}})]^{\neg\mathcal{E}} \wedge \neg[ag(s_{\mathcal{C}},\ s_{\mathcal{A}})]^{\neg\mathcal{C},\neg\mathcal{E}} \tag{4.23}$$

The active and passive cases describe situations where a causal link should be present. The active case describes a situation where the action $a_{\mathcal{A},t'}$ must be taken given that $a_{\mathcal{C},t}$ has occurred in order to avoid a loss of agency, or in other words the agent is actively forced to take the affected action. The passive case in contrast describes a situation where taking action $a_{\mathcal{A},t'}$ in the circumstance where $a_{\mathcal{C},t}$ has not occurred would lead to a loss of agency, and thus the occurrence of $a_{\mathcal{C},t}$ passively allows the action $a_{\mathcal{A},t'}$ to be made. In contrast the facilitation and mutual-effect-motive cases describe cases where a causal link is unlikely to be present or impossible to test for. In the facilitation case, because agency is lost regardless of whether $a_{\mathcal{A},t'}$ is taken assuming $a_{\mathcal{C},t}$ did not occur, it is impossible to evaluate the impact of the presence of $a_{\mathcal{C},t}$ upon the choice to take action $a_{\mathcal{A},t'}$. Meanwhile the mutual-effect motive describes a case where failing to take $a_{\mathcal{A},t'}$ regardless of whether $a_{\mathcal{C},t}$ has occurred will lead to a loss of agency, and therefore it is in the agent's interest to take action $a_{\mathcal{A},t'}$ in either case.

With these metrics defined, one can now determine whether a causal link is present:

$$\xi^{\text{agency}}_{a_{\mathcal{A},t'}} = (\xi^{\text{active}}_{a_{\mathcal{A},t'}} \vee \xi^{\text{passi.}}_{a_{\mathcal{A},t'}}) \wedge \neg(\xi^{\text{facil.}}_{a_{\mathcal{A},t'}} \vee \xi^{\text{m.e.m.}}_{a_{\mathcal{A},t'}}) \tag{4.24}$$

#### 4.2.2.3 Hybrid Variant

Here an approach that combines elements from the reward-based and agency-based variants is described. Given the previously defined causal-link tests and metrics, one can define the hybrid test as follows:

$$\xi^{\text{hybrid}}_{a_{\mathcal{A},t'}} = (\xi^{\text{active}}_{a_{\mathcal{A},t'}} \vee \xi^{\text{passi.}}_{a_{\mathcal{A},t'}} \vee \xi^{\text{reward}}_{a_{\mathcal{A},t'}}) \wedge \neg(\xi^{\text{facil.}}_{a_{\mathcal{A},t'}} \vee \xi^{\text{m.e.m.}}_{a_{\mathcal{A},t'}}) \tag{4.25}$$

Here the left side of the conjunction states that the counterfactual contexts must either indicate active or passive causal links — based upon the agency-based variant — or demonstrate sufficient conditional optimality as defined by the reward-based variant. Meanwhile the right side matches the agency-based variant completely in specifying that the counterfactual contexts must not indicate that a facilitation or mutual-effect-motive case is present. This effectively amounts to the agency-based causal-link test with the additional opportunity to add causal links where the reward-based metric is above the predefined threshold.

### 4.2.3 Evaluation Experiments

With the optimality-based causal-link testing methods defined, one can now evaluate the efficacy of these variants, comparing against the methods benchmarked in Section 3.2.3.

#### 4.2.3.1 Data, Code, & Parameters

The experiments were carried out upon the same highD dataset [150] scenes as those documented in Section 3.2.3.1, foregoing the Woven, and synthetic datasets. The synthetic dataset lacks sufficient information regarding bounding boxes and lateral movement. Meanwhile the Woven dataset has lane structures that are heavily segmented. This is fine for AV-navigation purposes, but it raises questions such as how lane transitions can be effectively represented, and what the default behaviour should be when a simulated agent reaches a branching point. Since the highD dataset offers a high scene count and was the most challenging dataset from Chapter 3 it made sense to focus solely on this dataset.

The code relating to this work along with any data that can be publicly shared is made available in a Git repository[6]. The method described in this chapter will henceforth be referred to as the Simulation-based Causal Analysis and Reasoning System Version 1 (SimCARSv1) framework. The code relating directly to this method's implementation can be found in a different yet associated Git repository[7].

For action-extraction parameters, the acceleration actuation threshold is $\lambda_{d,a} = 0.2\,m/s^2$; the speed-goal speed-difference threshold is $\lambda_{\delta a,sp,sp} = 1\,m/s$; the speed-goal duration threshold is $\lambda_{\delta a,sp,t} = 1\,s$; and the pre- and post-lane-change lane-goal start-time thresholds are $\lambda_{\delta a,la,t} = 1\,s$

[6]https://github.com/cognitive-robots/counterfactual-cd-paper-resources
[7]https://github.com/cognitive-robots/SimCARSv1

and $\lambda'_{\delta a,la,t} = 2.5\,s$ respectively. For the causal-link identification, the reward-based conditional-optimality threshold $\lambda_{\delta r}$ is varied across the set of values specified by $\{\ 0.1x \mid x \in [1, 10] \subset \mathbb{N}\ \}$. As before, see Appendix A for additional parameters.

#### 4.2.3.2 Evaluation Metric

For this set of experiments the same evaluation metrics are used as those from the benchmark experiments, as documented in Section 3.2.3.2. The one departure from the evaluation metrics described there is that the precision for each scene is calculated as follows:

$$\text{precision} = \begin{cases} \frac{|TP|}{|TP|+|FP|} & |TP| + |FP| > 0 \\ -1 & |TP| + |FP| = 0 \end{cases} \tag{4.26}$$

where precision values of $-1$ are then subsequently disregarded during the calculation of the precision mean and standard deviation. This offers a better reflection of what precision actually represents, as otherwise scenes that identify false-positive causal links are treated the same as those which identify none — potentially due to overly high specificity. Importantly, this still allows the metrics to be calculated on a scene-by-scene basis.

#### 4.2.3.3 Results

Here the results of the evaluation experiments are documented. These are primarily quantitative results, however a qualitative example was captured in order to better illustrate the process of testing causal links based upon conditional optimality.

**Quantitative** Results regarding the efficacy of the proposed-approach variants are presented in Table 4.1 and illustrated in Figure 4.3. In general all of the variants outperform DYNOTEARS, MVGC, and TiMINo by a notable margin, with the hybrid and agency-based variants outperforming all three baselines regardless of the $\lambda_{\delta r}$ parameter. The peak mean $F_1$ scores for the reward-based, agency-based, and hybrid approaches are 0.349, 0.851, and 0.725 respectively. These all outperform the top baseline, DYNOTEARS, by at least 0.223, which is a notable margin given there is a smaller difference in $F_1$ score between DYNOTEARS and the random baseline.

The differences in performance between the reward-based and agency-based variants indicate that the simplistic collision-driven nature of the agency calculations provide a strong starting

| Method | Var. | $\lambda_\alpha$ / $\lambda_{\delta r}$ | $\tau$ (s) | Precision | Recall | $F_1$ Score |
|---|---|---|---|---|---|---|
| DYNOTEARS | sp. | N/A | 3.6 | 0.118±0.116 | 0.800±0.400 | 0.197±0.147 |
| MVGC | acc. | 0.03 | 2.5 | 0.025±0.039 | 0.330±0.470 | 0.046±0.071 |
| Random | N/A | N/A | N/A | 0.060±0.039 | 0.800±0.400 | 0.110±0.069 |
| SimCARSv1 (Reward-Based) | N/A | 0.5 | N/A | 0.333±0.300 | **0.904±0.295** | 0.420±0.283 |
| SimCARSv1 (Agency-Based) | N/A | N/A | N/A | **0.887±0.238** | 0.896±0.307 | **0.837±0.317** |
| SimCARSv1 (Hybrid) | N/A | 1 | N/A | 0.790±0.311 | 0.896±0.307 | 0.773±0.335 |
| TiMINo | acc. | 0.05 | 3.6 | 0.070±0.062 | 0.687±0.464 | 0.125±0.104 |

Table 4.1: The precision, recall, and $F_1$ score of applying each of the SimCARSv1 variants to the highD dataset. Additionally a subset of Table 3.1 is included, depicting the highD-dataset results for the DYNOTEARS, MVGC, and TiMINo methods, to act as baselines. Note that for the baseline methods, specificity increases as $\lambda_\alpha$ decreases, whereas for the reward-based and hybrid SimCARSv1 variants, specificity increases as $\lambda_{\delta r}$ increases.

point for this type of causal discovery. It was expected that the reward calculations based upon several factors would provide a more nuanced test which could capture edge cases that the agency-based variant could not. As such the hybrid variant was intended to combine the strengths of these two variants to compliment one another, however the evaluation metrics show that the agency-based variant significantly outperforms the reward-based variant. The result is that the hybrid variant resembles a version of the agency-based variant which is only hindered by the inclusion of elements relating to the reward-based variant. This is reflected in the fact that the hybrid variant's peak $F_1$ score is found when $\lambda_{\delta r} = 1.0$, or in other words when the barrier for a causal link to be accepted by the reward-based variant is at its highest. This also explains why the hybrid variant increases in performance as the reward-based variant decreases in performance. While this does mean that for the purposes of this work the agency-based variant provides the best performance, the concept of a reward-based variant still holds merit in that it can better capture a variety of agent motivations across a continuous set of values. But with this greater complexity comes a need for additional tuning, thus further research is needed.

**Qualitative** Figure 4.4 illustrates a two-agent-convoy scenario based off the one described in Section 3.1.3, and the process of applying SimCARSv1 to it. Figure 4.4a depicts both the original scene as well as counterfactual outcomes associated with the scenario. In this scenario the front vehicle **c0** — indicated in red — brakes in order to avoid collision with the vehicle in front of it; and in response the rear vehicle **c1** — indicated in green — brakes in order to

avoid collision with **c0**. Within this scenario **c0** can be considered as the causing agent $\mathcal{C}$ and its braking as the causing action $a_{\mathcal{C},t}$. Meanwhile **c1** can be considered as the affected agent $\mathcal{A}$ and its braking in response to the behaviour of **c0** as the affected action $a_{\mathcal{A},t'}$.

The counterfactual outcomes show that the absence of $a_{\mathcal{A},t'}$ with $a_{\mathcal{C},t}$ still occurring would cause **c1** to collide with **c0**. Meanwhile the absence of $a_{\mathcal{C},t}$ will cause **c0** to collide with the vehicle in front of it regardless of whether $a_{\mathcal{A},t'}$ is taken, leaving **c1** unscathed in both cases.

In terms of the optimality metrics this leads to a large $\delta^{+}_{r,\mathcal{A}}$ from the perspective of **c1**, as the outcomes represent the difference between normal driving conditions and **c1** experiencing a collision. Meanwhile $\delta^{-}_{r,\mathcal{A}}$ is negative given that it is still beneficial for **c1** to brake in either case, even if it is less critical assuming the absence of $a_{\mathcal{C},t}$. Still the combination of $\delta^{+}_{r,\mathcal{A}}$ and $\delta^{-}_{r,\mathcal{A}}$ is still sufficient to indicate that a causal link is present between **c0** and **c1** provided $\lambda_{\delta r} \leq 0.69$.

Based upon the agency conditional-optimality metrics defined in (4.20), (4.21), (4.22), and (4.23), the type of relationship indicated by the metrics reflects $a_{\mathcal{C},t}$ actively causing $a_{\mathcal{A},t'}$. This matches with the intuition that **c0** forces **c1** to brake with its actions, as to do otherwise would cause **c1** to have a collision. This also indicates that both the agency-based and hybrid variants will positively identify the causal link between **c0** and **c1**.

Lastly the causal-summary graph outputs of applying SimCARSv1 to two highD scenes are shown in Figures 4.5 and 4.6. These graphs mostly reflect the behaviour implied by the quantitative results, with the agency-based variant performing well, the reward-based variant being overly sensitive, and the hybrid variant being hampered by its reward-based component. Nonetheless these graphs offer greater clarity regarding the exact outputs given by these variants. Importantly this also facilitate direct comparison with Figures 3.3 and 3.4 from Chapter 3 as they utilise the same highD scenes.

#### 4.2.3.4 Discussion

The above results along with the inherent rationale of the theory-of-mind approach demonstrate that there is merit to this type of methodology. However, a few concerns ought to be addressed and as such these direct the next steps of the overarching work documented in this thesis.

First among these is the use of a discrete-action representation over continuous variables. While the reasons for doing this have been well established it still raises questions. One of these

is how patterns regarding the causal relationships between actions can be determined such that identified causal relationships can be utilised to aid in agent decision making. Similarly it is debatable whether the current manner of extracting actions is optimal, not to mention the uncertainty over how such an extraction method could be deployed in an online setting. There is no immediately obvious solution here and further work is required to identify the best path forwards for research in this area.

The second issue that needs to be diagnosed and remedied is the difference in performance between the reward-based and agency-based variants of SimCARSv1. In theory the more diverse criteria considered by the reward-based variant should lead to a more expressive and adaptable metric of optimality. In reality the reward elements considered might not accurately reflect the variations of motivations between drivers, nor does it account for irrational behaviour from humans (e.g. reckless driving, road rage). By contrast the agency-based metric just carries the expectation that drivers try to avoid collisions, which can generally be assumed to be true. If one wishes to continue building upon the concept of reward-based conditional-optimality metrics it seems prudent to incorporate some mechanism for capturing these variations from agent-to-agent and thus more closely match the cognition of these agents.

The last issue is that the relationship identified between actions is not strictly causal. The planning process of the potentially affected agent is somewhat emulated by exploring the conditional optimality of the potentially affected action. However, intervening upon the action variable of the affected agent in theory severs any causal path between the agents. The most obvious means of remedying this would be to try and directly emulate agents and their planning processes as part of an SCM. However, there are several obstacles to directly modelling autonomous agents via SCMs, a matter which motivates the work of the next chapter.

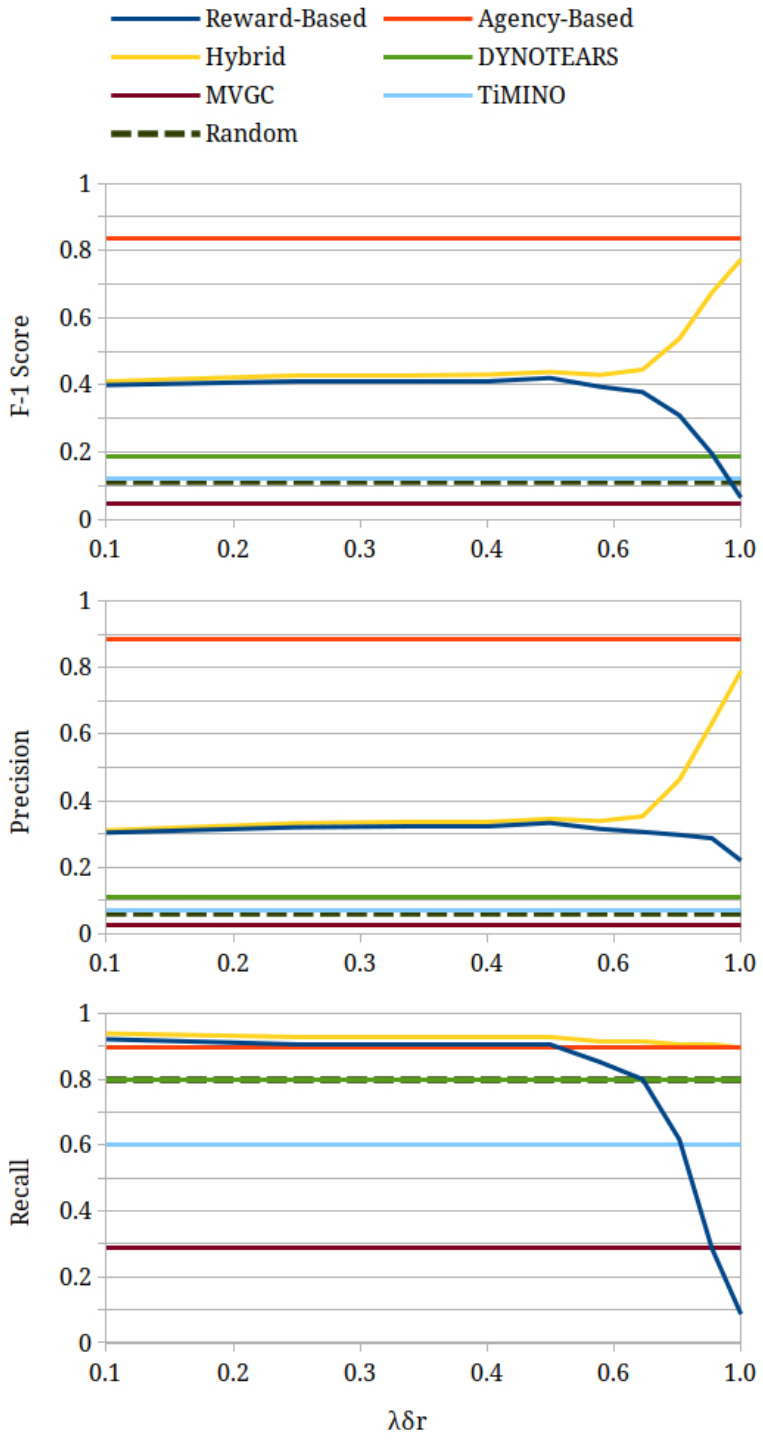
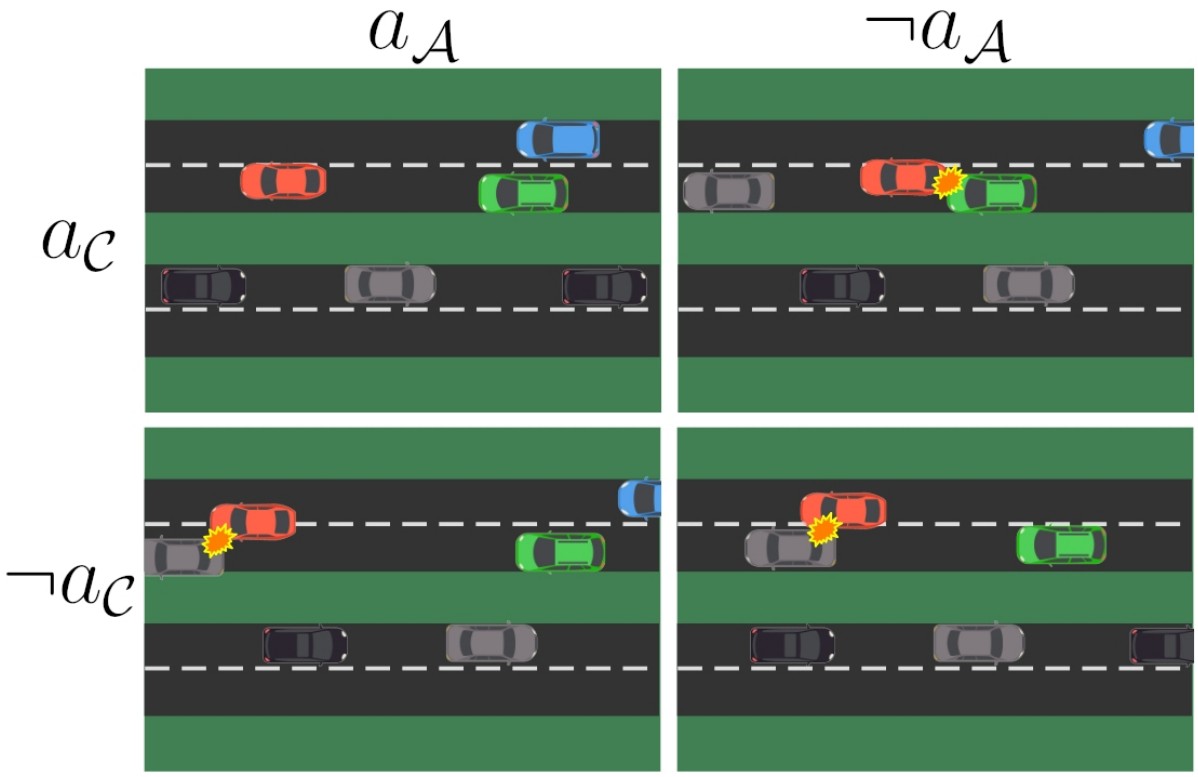



Figure 4.3: $F_1$ score, precision, and recall from applying each variant of SimCARSv1 to the highD dataset. The x-axis varies the reward-based conditional-optimality threshold $\lambda_{\delta r}$, which influences just the reward-based and hybrid variants of SimCARSv1. The DYNOTEARS, MVGC, and TiMINo results plotted here reflect their performance metrics from the benchmark in Section 3.2.3, evaluated upon the same highD dataset. (© 2023 IEEE)




(a) Counterfactual-Outcome Illustrations

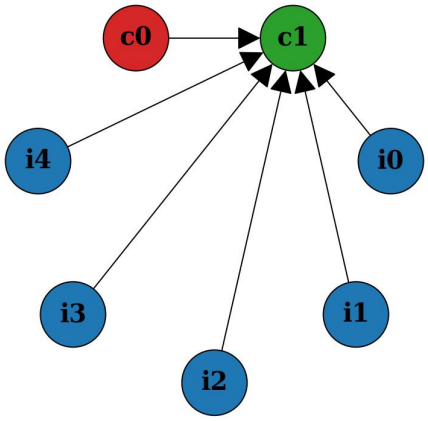
| | $a_{\mathcal{A}}$ | $\neg a_{\mathcal{A}}$ |
|---|---|---|
| $a_{\mathcal{C}}$ | 0.97 | 0.00 |
| $\neg a_{\mathcal{C}}$ | 0.97 | 0.69 |

(b) Reward-Based Optimality Metrics

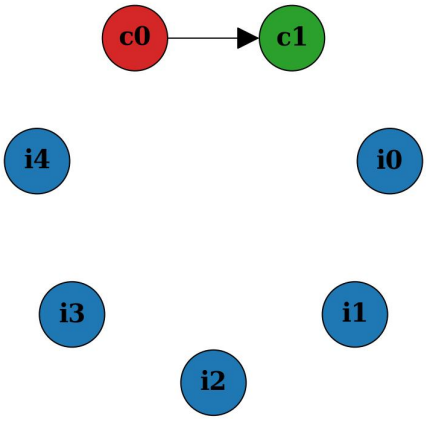
| | $a_{\mathcal{A}}$ | $\neg a_{\mathcal{A}}$ |
|---|---|---|
| $a_{\mathcal{C}}$ | 1 | 0 |
| $\neg a_{\mathcal{C}}$ | 1 | 1 |

(c) Agency-Based Optimality Metrics

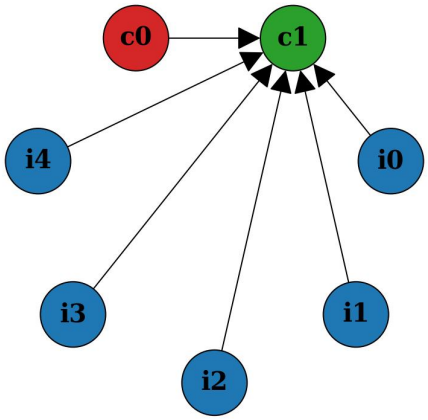
Figure 4.4: Qualitative example of applying SimCARSv1 to a two-agent-convoy scenario. The rows annotated with $a_{\mathcal{C},t}$ and $\neg a_{\mathcal{C},t}$ indicate the outcomes associated with the counterfactuals in which action $a_{\mathcal{C},t}$ did and did not occur respectively. Similarly the columns annotated with $a_{\mathcal{A},t'}$ and $\neg a_{\mathcal{A},t'}$ indicate the outcomes associated with the counterfactuals in which action $a_{\mathcal{A},t'}$ did and did not occur respectively. All of the metrics depicted here are done so from the perspective of the affected agent, i.e. agent $\mathcal{A}$. Lastly, while the illustrative depictions of the outcomes are purely artistic, they are based upon a direct superimposition of art assets onto a user-interface visualisation from SimCARSv1. (© 2023 IEEE)



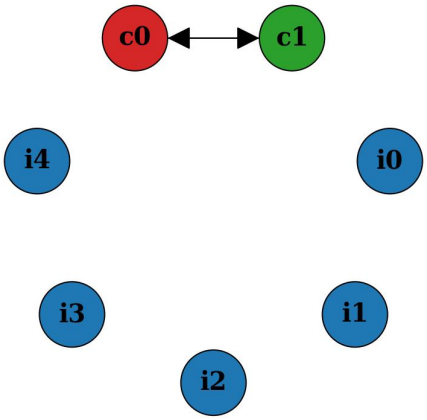

(a) Reward-Based Variant




(b) Agency-Based Variant

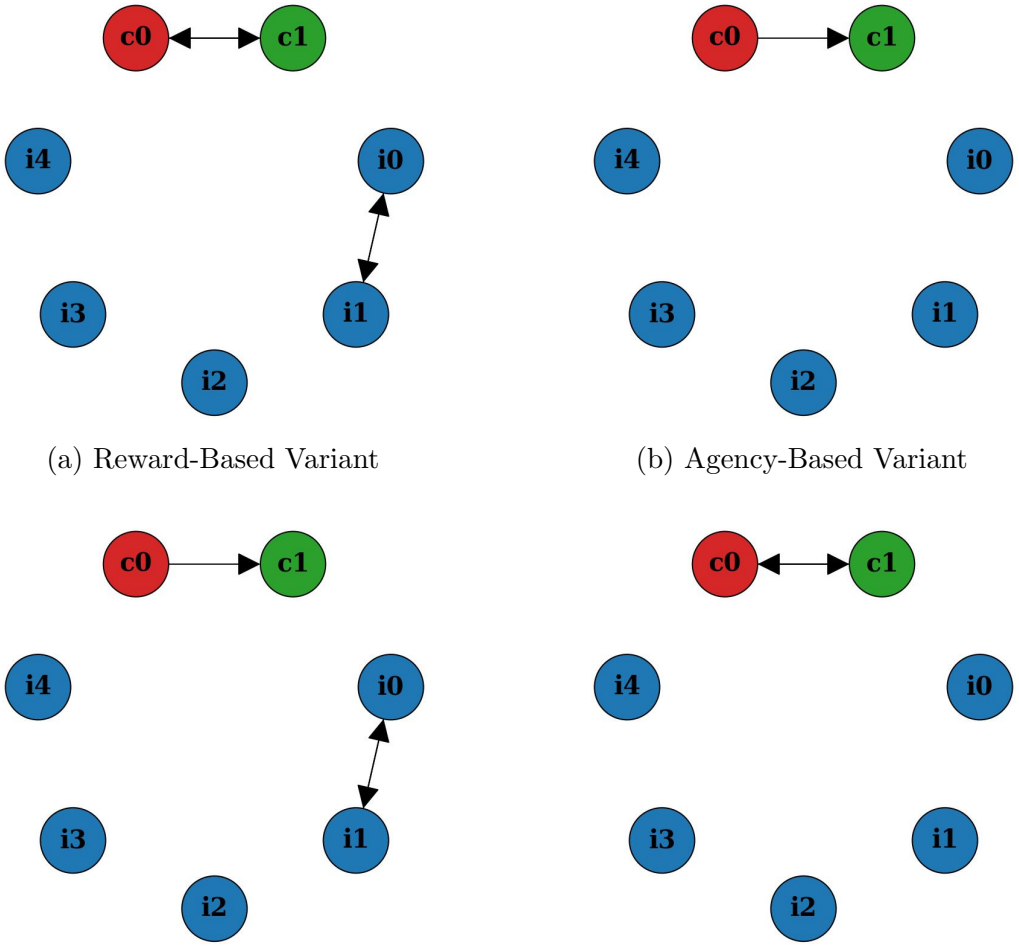



(c) Hybrid Variant




(d) Ground Truth

Figure 4.5: Depicts the causal-summary graphs from applying SimCARSv1 to a scene from the highD dataset ($\lambda_{\delta r} = 0.5$), along with the ground truth (Figure 4.5d). The agency-based variant performs well (Figure 4.5b) identifying a causal relationship between the convoy vehicles. Meanwhile the reward-based variant is overly sensitive (Figure 4.5a) and while it does does correctly identify a causal relationship between the convoy vehicles, it also identifies every independent vehicle as affecting the tail convoy vehicle. The hybrid variant is seemingly hampered here by the reward-based component (Figure 4.5c) as it exhibits the exact same issue of over-sensitivity. Lastly it should be noted that the scene utilised for this figure is the same scene utilised for Figure 3.3.




(a) Reward-Based Variant

(b) Agency-Based Variant

(c) Hybrid Variant

(d) Ground Truth

Figure 4.6: Depicts the causal-summary graphs from applying SimCARSv1 to a scene from the highD dataset ($\lambda_{\delta r} = 0.5$), along with the ground truth (Figure 4.6d). Similar to Figure 4.5 the agency-based variant performs well (Figure 4.6b). However, the reward-based variant does perform better here (Figure 4.6a) only identifying a single spurious causal relationship. Interestingly the reward-based variant is also the only variant to identify the potentially bi-directional nature of the causal relationship between the convoy agents. In contrast, the hybrid variant does not capture the bi-directionality of this relationship (Figure 4.6c) which could imply the presence of facilitation or mutual-effect-motive cases between the actions of the convoy vehicles. Similar to the previous figure the input scene here is the same used for Figure 3.4.

# Chapter 5

# Extending Structural Causal Models for Autonomous Agents

Chapter 4 introduced the concept of utilising theory of mind for identifying behavioural interactions between autonomous agents. However, the proposed method which utilised this concept — i.e. SimCARSv1 — ultimately was not grounded in true causality due to the planner cognition not being emulated during counterfactual simulation. To elaborate, there were two key flaws with this system. The first is that much of the system acts as a black-box, or at least does not make an attempt to expose the flow of data through the system, limiting the system's transparency. Secondly, and more seriously however, the system intervenes upon both the potentially causing and potentially affected action when attempting to identify behavioural interactions. This breaks any possible link of causal ancestry as the intervention on the potentially affected action causes its relationship with its parents to be severed. While the system remains effective despite these flaws, they should nonetheless be addressed, both to increase the system transparency and increase the causal grounding of any identified behavioural interactions. Despite this the implementation of Autonomous Embodied Systems (AESs) via SCMs is by no means a trivial feat, and thus this chapter aims to address the challenges facing one wishing to carry out such an implementation. This is done in the hopes of alleviating inconvenient limiting factors in the development of causality within this field. As such in this chapter these challenges are explored (Section 5.1), a series of extensions to SCMs addressing them are proposed (Section 5.2), and novel case studies on the integration of SCMs across two autonomous-agent domains are discussed (Section 5.3).

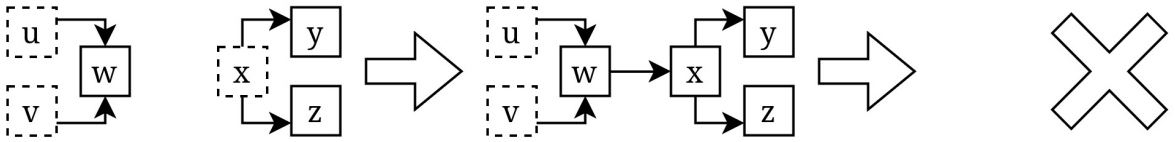



Figure 5.1: SCM without Socket Variables

Figure 5.2: Illustration of merging SCMs and the lack of ability to subsequently un-merge them. Here solid borders denote endogenous variables, dashed borders exogenous variables, and dotted borders socket variables.

## 5.1 Challenges to SCM Utilisation in Autonomous Agents

While it has been stated here that there are several challenges presented to implementing AESs via SCMs these challenges have yet to be detailed. Thus an explanation of the exact nature of these challenges is provided here, in order to motivate and guide the extensions of Section 5.2.

### 5.1.1 Modularisation & Encapsulation

While SCM use in AESs remains rare, those works that do utilise SCMs typically either utilise a single monolithic SCM, or several entirely independent SCMs. Furthermore, it is rare to store the data relating to an SCM within the structure of an SCM itself. This is a reasonable approach for experimental settings or when working in data science. However, effective system design would suggest one should aim to modularise and encapsulate SCMs in order to minimise repetitions in the graphical structure, restrict data access based upon system design, and cater each module to a particular aspect of the system [161]. While these are desirable qualities in system design across domains, AESs have in particular shown a preference for modular libraries which are easily ported and integrated across systems. The best showcase of this preference is provided by the ROS framework [162], which features a decentralised modular design. This is especially important when parts of the system may need to be replicated an arbitrary number of times. For example, assuming a given SCM represents a single vehicle, this will need to be replicated for each vehicle present in a scene.

SCMs are to some degree already modular by nature [4], as a causal graph may be cut and endogenous variables replaced with exogenous. However this leaves system designers few options for hiding or exposing exogenous variables. While it is true that exogenous variables

describe the full breadth of the external flow of data into an SCM, it should be possible as system designers to specify which of these variables are exposed and how they are exposed.

Furthermore, there is no inherent mechanism by which the merging of two SCMs can be reversed. For example if one considers Figure 5.2 it can be seen that after merging the two SCMs there remains no record of $x$ originally being an exogenous variable. Having lost this information, along with the distribution associated with $x$ when it was an exogenous variable, un-merging the SCMs becomes difficult in the absence of external information. As such, extending SCMs to inherently support dynamic merging and un-merging would significantly increase their inherent modularity in addition to that of any architecture built upon them.

### 5.1.2 Constant-Space Representation of Time Series

Window time graphs offer a compact representation of temporally-lagged causal relationships, particularly for AESs, which often assume the Markov property (see Definition 10). However, even the window-time-graph representation still requires one to roll-out the graph during inference to cover the relevant time period. This can lead to the rolled-out graph taking up a great deal more space than the initial window time graph, not to mention that inferences over different time periods typically each require their own independent roll-out. This becomes especially complicated when one wishes to use the generative properties of SCMs to counterfactually simulate different outcomes, creating separate branches for each inferred history.

Such roll-outs also represent a problem for SCM-wrapper design. If the window time graph for an SCM captures the underlying system then it is unclear whether performing a roll-out should mutate the original SCM within the wrapper, or whether the roll-out operation should return a new rolled-out SCM. The former loses the original SCM representation, while the latter could mean managing a large number of SCMs, neither option is ideal.

### 5.1.3 Arbitrary Number of Agent Interactions

When operating within their environments AESs belonging to certain domains — such as those considered in this chapter's case studies — will interact with various agents. The number and nature of these agents will often vary, thus an SCM capturing an AES should allow the handling of a fluid set of values. While the structural equations of an SCM can derive an output from

an input set, the greater issue is being able to fluidly formulate such a set within an SCM. In instances where this set is derived directly from observation, this can simply be captured via an exogenous variable. However, given that such a fluid set would be comprised of agents, one may also wish to explicitly model these agents within the SCM. In particular, causal methods grounded in intentionalism [89] may wish to utilise a theory of mind for other agents when making counterfactual inferences. Furthermore, assuming the conditions under which agents are liable to influence one another depends upon the actions of said agents, abducting an exogenous set of inputs may not suffice for the task in question.

## 5.2 Building Upon the Functionality of SCMs

This section details several extensions to the SCM formalism to address the challenges associated with incorporating SCMs into AESs introduced in Section 5.1. This is done with the goal of facilitating further work within causality that concerns agent behavioural interactions.

### 5.2.1 Structural Equations with Side-Effects

Before one can properly introduce any extensions to the SCM formalism, it is first necessary to describe the process of integrating side-effects into the structural equations of an SCM. Side-effects in this context of this work refer to computations that occur beyond the purely mathematical calculations of structural equations (e.g. printing, communications, memory). In order to incorporate this into the SCM formalism, a similar approach is taken to that typically observed in functional-programming paradigms by utilising monadic actions [163].

**Definition 15** (Monad)**.** A monad $\mathfrak{M}$ based upon the definition given by Wadler [163] is comprised of the following elements:

- $\mathfrak{M}(\cdot)$, a type constructor.
- $u_{\mathfrak{M}}(\cdot) : \mathcal{T} \to \mathfrak{M}(\mathcal{T})$, a unit / return function that wraps a value of type $\mathcal{T}$ with the monad of type $\mathfrak{M}$, giving the monadic value of type $\mathfrak{M}(\mathcal{T})$.

- $b_{\mathfrak{M}}(\cdot) : \mathfrak{M}(\mathcal{T}) \to (\mathcal{T} \to \mathfrak{M}(\mathcal{T}')) \to \mathfrak{M}(\mathcal{T}')$, a bind function that applies a function to the wrapped value of the input monadic value while propagating the monad forwards, wrapping the output of the aforementioned function. The bind function will also occasionally be denoted via a binary operator $\triangleright_{\mathfrak{M}}$ for a cleaner presentation.

Additionally a monadic action may be associated with data constructors that describe how additional data pertaining to the monadic action can be passed when instantiating a monad of the corresponding type. This can be utilised to embed additional computations within otherwise mathematically pure functions. Effectively, the monadic action applied to a value provides a means of keeping track of which computations need to be done and where. The primary utilisation of monads within this work is to keep track of meta-variable information while conducting inference on an SCM.

### 5.2.2 Variable Context

A variable context $C$ is a set of meta-variables that are globally accessible by the structural equations of the SCM to which it belongs. Utilisation of such a concept must be handled with care, as one could inadvertently incorporate causal relationships via $C$ that would not be represented within the causal graph of the SCM. However, $C$ only captures the current time $C_T \in C$ and the time-step size $C_{\delta t} \in C$, and these are only used to emulate a temporal roll-out and carry out other time-related computations. Thus $C$ does not invalidate SCMs that make use of it provided they do so via the structural equations described in Section 5.2.3.

From a theoretical point of view one can utilise the well documented state monad[8] [163] to implement such functionality, given its similar features.

**Definition 16** (State Monad)**.** A state monad $\mathfrak{S}$ based upon the definition given by Wadler [163] is a monad with elements defined as follows:

$$\mathfrak{S}(\mathcal{T}) = (\ \mathcal{S} \to \mathcal{T} \times \mathcal{S}\ ) \tag{5.1}$$

[8]Despite the name this monad and its state domain do not directly relate to the state variables as described previously in this thesis. These are just already well established terms within functional-programming literature.

$$u_{\mathfrak{S}}(v) = (\ \lambda x.\ \langle v,\ x \rangle\ ) \tag{5.2}$$

$$b_{\mathfrak{S}}(\mathfrak{S}(v), f) = (\ \lambda x.\ (f\ (\mathfrak{S}(v)\ x)[0])\ (\mathfrak{S}(v)\ x)[1]\ ) \tag{5.3}$$

where $x \in \mathcal{S}$ is a state value drawn from the state domain of choosing, and $v \in \mathcal{T}$ is the wrapped value of a predefined type. Furthermore, $f(\cdot) : \mathcal{T} \to \mathfrak{S}(\mathcal{T}')$ is a non-specific function that takes an input value of a given type and outputs a state monadic value of another type. Note that this wrapped output type may be the same as the input type. Lastly, (5.3) makes use of a lambda-calculus-style application of functions for the sake of clarity, and square brackets are used to denote tuple indexing.

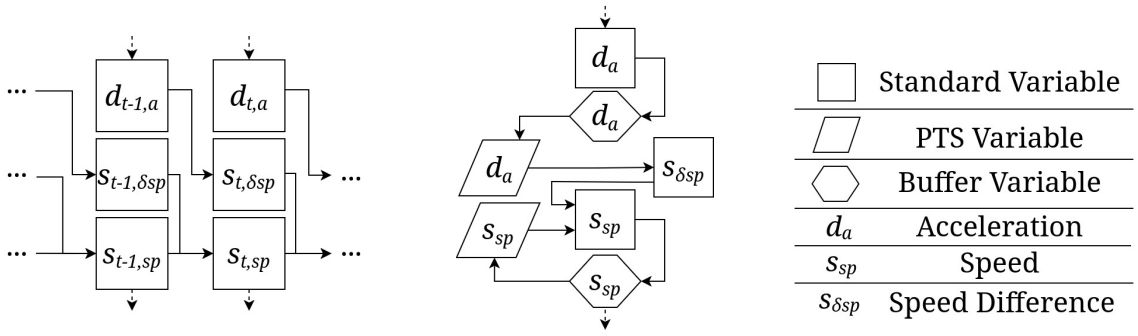
One can utilise this monad with the variable-context definition providing the data-type for the state domain $\mathcal{S}$. This effectively permits passing a variable context $C$ into a state monadic value returned by an SCM in order to carry out computations involving its meta-variables. However, for the sake of simplicity in notation, and in order to focus upon the core contributions of this work, it will be assumed that $C$ is globally accessible via this monadic representation even if not explicitly stated. Furthermore, this should not prove to be an issue with the inclusion of other monads as the practice of composing monads together is well established [164].

### 5.2.3 Temporal Variables

Temporal variables do not refer to a specific variable or structural equation but instead a class of endogenous variables that operate on, or with, time meta-variables. These can be further separated into variables whose computations rely upon meta-variables and Previous Time Step (PTS) variables which establish the temporal structure of the SCM. The latter of these will be discussed first, as they are essential to understanding temporal representation within the SCM-formalism extensions.

#### 5.2.3.1 Previous-Time-Step Variables

PTS variables are used within the context of this work to address the challenge introduced in Section 5.1.2. Namely they enable the representation of temporal SCMs utilising a recursive




(a) Window-Time-Graph SCM (b) Extended SCM (c) Legend

Figure 5.3: SCMs depicting the calculation of speed $s_{sp}$ from acceleration $d_a$ by calculating the speed difference $s_{\delta sp}$ and adding it to the speed from the previous time step. A typical time-series SCM would be coupled along with time-series data, corresponding with time-indexed variables — as indicated via subscripts. In contrast the extended-SCM representation uses PTS variables and buffer variables. These allow the SCM a fixed-size graph while still capturing causal links between time steps, in addition to providing data encapsulation. Despite the similarities between these causal models and a traditional Kalman filter, the models shown here do not capture uncertainty. This is functionality that SCMs can undoubtedly provide, however as this is not the focus of this work, it has been omitted here for simplicity.[9]

structure that does not have to be rolled out during inference. A PTS variable $V^{\delta} \in V$ is defined based upon it having an associated structural equation $f_{V^{\delta}} \in F$ defined as follows:

$$f_{V^{\delta}}(V_{Pa}) = \delta_{-} \ \triangleright_{\mathfrak{S}} \ (\lambda\emptyset.\ V_{Pa}) \ \triangleright_{\mathfrak{S}} \ \delta_{+} \tag{5.4}$$

$$\delta_{-} = (\ \lambda C_T.\ \langle \emptyset,\ C_T - C_{\delta t} \rangle\ ) \tag{5.5}$$

$$\delta_{+}(v) = (\ \lambda C_T.\ \langle v,\ C_T + C_{\delta t} \rangle\ ) \tag{5.6}$$

where $\emptyset$ denotes the empty set / value, $\delta_{-} : \mathfrak{S}(\emptyset)$ is a monadic value, and $\delta_{+}(\cdot) : \mathcal{T}^{V_{Pa}} \rightarrow \mathfrak{S}(\mathcal{T}^{V_{Pa}})$ is a function. Importantly $\delta_{+}$ and $\delta_{-}$ utilise their monadic action to increment or decrement the current time $C_T$ by $C_{\delta t}$ respectively. This has the effect of shifting the current time to the previous time step, evaluating the parent variable $V_{Pa}$, before returning the current time to where it was initially. The function expects that the parent $V_{Pa}$ also possesses the monad $\mathfrak{S}$. If this is not already the case — i.e. from $V_{Pa}$ having PTS-variable ancestors of its own — this can easily be achieved via use of the unit function $u_{\mathfrak{S}}(\cdot)$.

[9]The use of $s_{sp}$, $s_{\delta sp}$, and $d_a$ to represent variables rather than the proper variable notation $V_{S,sp}$, $V_{S,\delta sp}$, and $V_{D,a}$ is an abuse of notation. This is primarily done to increase clarity within the figures, although this also avoids confusion over whether graph vertices represent abstractions or realised instances of variables when discussing roll-outs.

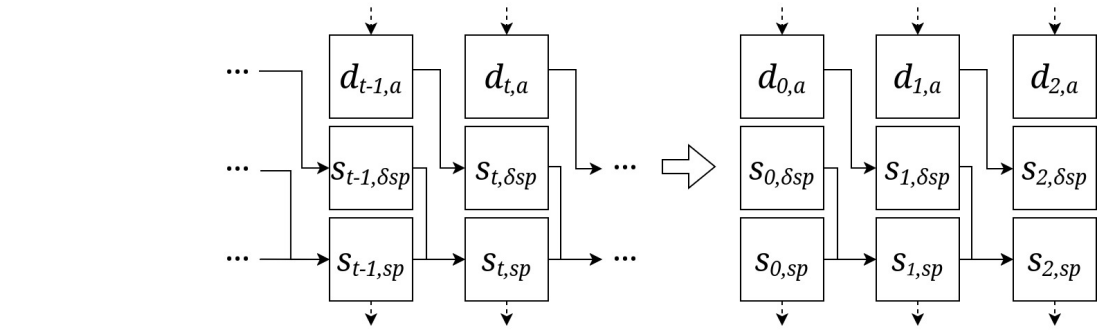
By utilising PTS variables one can represent temporal causal relationships in an SCM while maintaining the size and shape of the causal graph even during inference. Figure 5.3 illustrates this by showing how temporal causal relations in kinematics can be captured via the introduction of recursive cycles of the causal graph, as opposed to performing a roll-out of a window time graph during inference. Typically cycles would invalidate the SCM and break with assumptions made by many causal-reasoning techniques. However the combination of graphical structure, PTS variables, and variable context means that inference upon the SCM effectively emulates a roll-out while avoiding any actual modification to the SCMs graphical structure. With this in mind, this work proposes the following theorem:

**Theorem 1** (Window-Time-Graph Representation – PTS Representation Isomorphism)**.** *Any casual reasoning method which can operate upon a window-time-graph SCM should equally function upon an SCM utilising the proposed PTS-variable representation.*

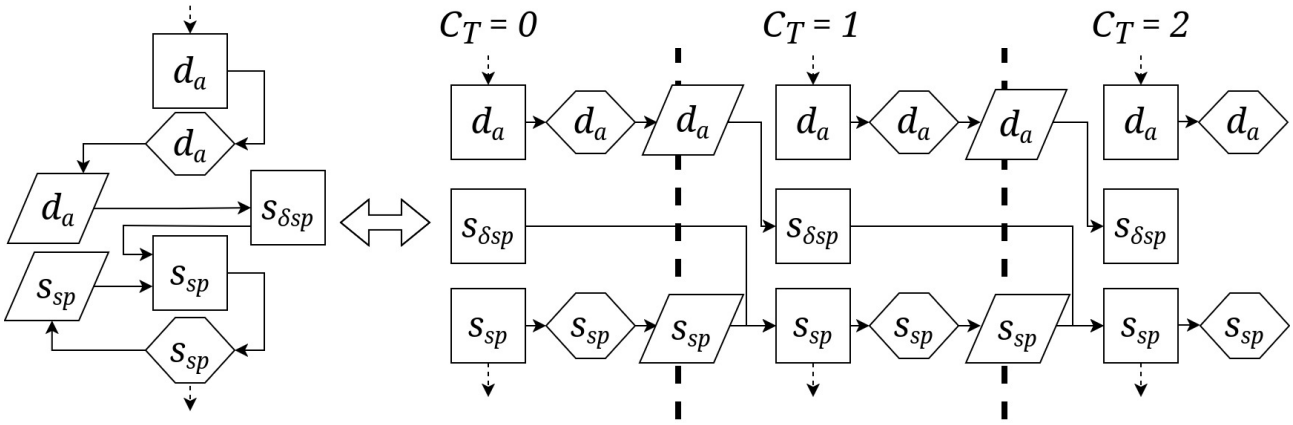
For which the following proof is offered:

*Proof of Theorem 1.* Consider Figure 5.4, in which the two SCMs from Figure 5.3 are rolled out for time steps 0–2 inclusive. Given that PTS variables are only responsible for managing $C_T$ and buffer variables just cache values indexed to $C_T$ — in other words fulfilling utility roles — the structure of the roll-outs are otherwise identical. Thus any series of operations one could perform on Figure 5.4a could be mapped directly to 5.4b without issue, and indeed theoretically this roll-out process could continue indefinitely. Ultimately the key difference between these roll-outs is that in Figure 5.4a each variable is replicated across time steps based upon the window-time-graph SCM, whereas in Figure 5.4b the variables are re-used. In essence, the utilisation of PTS variables virtually emulates the replication of variables by maintaining the current time step via meta-variable $C_T$, which can be used by buffer variables to time-index data. The key benefit this offers is not only in avoiding having to actually replicate the original SCM variables, but being able to maintain the same structure — as indicated by the double-ended arrow — which may be beneficial for certain system designs. □

Note that while the graphs presented here do bear some resemblance to the extended summary graphs described by Assaad et al. [165] they are arguably closer to a memoryless




(a) Window-Time-Graph SCM




(b) Extended SCM

Figure 5.4: Depicts the roll-out of the SCMs in Figure 5.3 for time steps 0–2 inclusive. Here it should be noted that while the window-time-graph SCM in Figure 5.4a is actually rolled out to cover the time steps in question, that the extended SCM in Figure 5.4b is virtually rolled out, utilising a variable context to do so. In the process of carrying out this virtual roll-out the cycles present in the left graph of Figure 5.4b are for all intents and purposes removed during roll-out as shown in the graph on the right. This is due to the current time meta-variable $C_T$ being decremented and incremented during the backwards and forwards passes of inference respectively, in doing so altering the context under which variables are accessed.

window time graph. Extended summary graphs split their variables into a 'present' set, and an all-encompassing 'past' set. Whereas the approach given here can be seen as splitting variables into a 'present' set, and a 'previous' set, which only concerns the previous time step and no further back. However due to this representation being defined recursively it can still encompass the whole history modelled by the SCM. The key difference here is that extended summary graphs eschew causal stationarity while this work relies upon its presence. Similarly the memoryless nature of the graphs shares similarities to past work modelling dynamic systems via local-independence graphs [166]. Yet techniques often associated with local independence — which typically treat each variable as a Markov process — are not necessarily suitable for the types

of architectures considered here. Within SCMs all structural equations are deterministic and exogenous variables are by definition independent of one another. As such, a local-independence-graph representation might work well for a smaller number of high-level endogenous variables, most of which have an exogenous-variable parent. However, this work consists of a large number of low-level endogenous variables with well defined structural equations. Meanwhile the presence of exogenous variables is rare by comparison, thus the use-case is entirely different.

#### 5.2.3.2 Time-Step-Size-Product & Time-Step-Size-Quotient Variables

A Time-Step-Size Product (TSSP) variable $V^{\cdot\delta t} \in V$ or Time-Step-Size Quotient (TSSQ) variable $V^{\div\delta t} \in V$ simply aims to multiply or divide their input by the time-step size $C_{\delta t}$ respectively. Formally, one can define their functions as follows:

$$f_{V^{\cdot\delta t}}(V_{Pa}) = V_{Pa} \cdot C_{\delta t} \tag{5.7}$$

$$f_{V^{\div\delta t}}(V_{Pa}) = V_{Pa} \div C_{\delta t} \tag{5.8}$$

These are primarily useful for carrying out of kinematics and dynamics calculations in rigid-body and point-mass representations. While linear scaling may not be suitable for modelling all embodied-system variables, it suffices in the context of this work for modelling AV dynamics in a similar manner to Kalman filtering. This is demonstrated within Figure 5.3 where $s_{\delta sp}$ — i.e. the speed difference between time steps — must be calculated by multiplying the acceleration $d_a$ by the time-step size as dictated by $C_{\delta t}$. Hypothetically one could hardcode such a value into structural equations in order to avoid this, however this definitions opens the option to have time-step size fluidly change across time while maintaining the same structural equations. This is important given that real-world autonomous systems cannot always guarantee a constant time-step size in-between sensor readings.

#### 5.2.3.3 Current-Time-Difference Variables

The slightly more complex Current Time Difference (CTD) variable $V^{\delta T} \in V$ calculates the duration between the input of the variable and the current time $C_T$ as follows:

$$f_{V^{\delta T}}(V_{Pa}) = V_{Pa} - C_T \tag{5.9}$$

CTD variables are primarily useful within controllers where they can be utilised to determine the duration of time between now and when a planned action goal should be accomplished by. Once again, this can be achieved by hardcoding the variables at each time step with the times associated with them. The main reason to avoid this, is firstly to allow structural equation reuse, and secondly to enable the recursive structure required for utilisation of PTS variables.

#### 5.2.3.4 Time-Conditional Variables

Lastly the time-conditional variable $V^{T?} \in V$ is introduced. This takes the current time $C_T$ and compares it against a predefined time in order to select which parent from which to derive its output. Their structural equations are defined as follows:

$$f_{V^{T?}}(V_{Pa,0}, V_{Pa,1}) = \begin{cases} V_{Pa,0} & C_T < \theta_T \\ V_{Pa,1} & C_T \geq \theta_T \end{cases} \tag{5.10}$$

where $\theta_T$ is a configurable time parameter specified during structural-equation construction. Time-conditional variables have multiple uses including the fluid introduction and removal of agents in environment interaction calculations. Furthermore, they can be utilised for counterfactual-simulation purposes, as one can effectively splice together a new counterfactual SCM together with an original SCM. From here $\theta_T$ can be used to dictate the divergence point between the original- and counterfactual-SCM histories. This is primarily useful within planners in order to envisage a series of alternate actions diverging at each action start time with the use of a time-conditional variable.

### 5.2.4 Buffer Variables

Having introduced the variable context and explored how one can utilise temporal variables to interact with the meta-variables of $C$ one can now utilise this to store time-indexed variable data encapsulated within the SCM. A buffer variables $V^B \in V$ has a single parent, and as the name suggests they act as a buffer for this parent, storing data relevant to it. In order to do so

a buffer variable $V^B$ maintains a time-indexed dictionary $\Phi_{V^B} = \langle T_{V^B}, \phi_{V^B} \rangle$ where $T_{V^B} \subset T$ is the set of time steps with indexed data, and $\phi_{V^B}(\cdot) : T \to \mathcal{T}^{V^B}$ is a function that maps a time step to a value or distribution for variable $V^B$ of type $\mathcal{T}^{V^B}$. With this established the structural equation for a buffer variable $f_{V^B} \in F$ is defined as follows:

$$f_{V^B}(V_{Pa}) = \begin{cases} \text{UP}_{\{\}}(\phi_{V^B}(C_T)) & C_T \in T_{V^B} \\ \text{UP}_{\{\langle \Phi_{V^B}, C_T, V_{Pa} \rangle\}}(V_{Pa}) & C_T \notin T_{V^B} \end{cases} \tag{5.11}$$

where $f_{V^B}(\cdot) : \mathcal{T}^{V_{Pa}} \to \mathfrak{B}(\mathcal{T}^{V_{Pa}})$ either returns the output of its parent variable, or retrieves stored data, depending upon whether the current time $C_T$ is found within the dictionary time steps $T_{V^B}$. Importantly the output type of $V^B$ matches that of its parent $V_{Pa}$, denoted as $\mathcal{T}^{V_{Pa}}$, albeit with the addition of wrapping the type in the buffer monad $\mathfrak{B}$. The buffer monad takes a single type and provides a data constructor $\text{UP}_X(\cdot)$ — i.e. update the dictionary based upon $X$. The data constructor takes a set $X$ of tuples structured $\langle \Phi, t, x \rangle$ and for each tuple stores the value of $x$ at the time of tuple creation within the dictionary $\Phi$ time-indexed to time step $t \in T$. To comply with monad requirements, the following unit and bind functions are provided:

$$u_{\mathfrak{B}}(v) = \text{UP}_{\{\}}(v) \tag{5.12}$$

$$b_{\mathfrak{B}}(\text{UP}_X(v), f) = (\lambda \text{UP}_Y(w).\, \text{UP}_{X \cup Y}(w))\ f(v) \tag{5.13}$$

This monad bears resemblance to the writer monad [164] albeit with the addition of time-indexed writing to the specified buffer-variable dictionaries.

### 5.2.5 Socket Variables

As mentioned in Section 5.1.1 SCMs are inherently well suited to modularisation [4]. After all, exogenous variables capture factors outside of the model that nonetheless influence the model. As such if one then produces a model for these factors they can easily combine the SCMs together to form a new greater whole. The main limitations of the naive combination of SCMs is the lack of ability for system designers to set limits on how SCMs are joined, and the difficulty in fluidly separating SCMs.

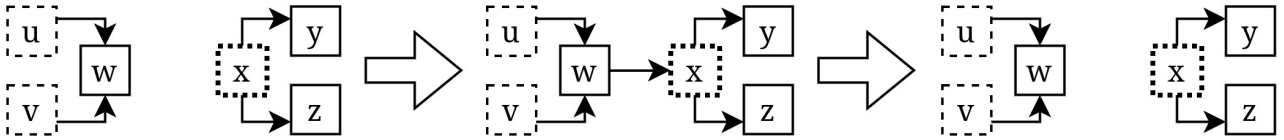



Figure 5.5: SCM with Socket Variables

Figure 5.6: Illustration of merging and un-merging SCMs with the aid of socket variables. Once again solid borders denote endogenous variables, dashed borders exogenous variables, and dotted borders socket variables.

Thus in this work the exogenous variables $U$ are decomposed into socket variables $U^S$ and hidden exogenous variables $U^H$. A socket variable $U_i^S \in U^S$ differs from a typical exogenous variable in that its value is conditionally derived as follows:

$$U_i^S \leftarrow \begin{cases} P(U_i^S) & |Pa(U_i^S)| = 0 \\ Pa(U_i^S) & |Pa(U_i^S)| = 1 \end{cases} \tag{5.14}$$

This allows one to join SCMs by assigning $U_i^S$ to take the output of a variable from another SCM. Importantly tracking socket variables separately ensures that the act of joining SCMs is reversible, allowing the fluid reconfiguration of SCM-based modules. Note that the above function only accounts for cases of zero or one socket-variable parents. This is because socket variables impose a maximum of one parent variable as a structural constraint on SCMs — similar to endogenous variables only allowing a maximum of one exogenous parent.

Figures 5.2 and 5.6 illustrate the process of merging SCMs, and in the latter case the un-merging of SCMs too. Here if one attempts to utilise SCMs without socket variables (see Figure 5.2), the SCMs $\langle U = \{u, v\}, V = \{w\}, E = \{\langle u, w\rangle, \langle v, w\rangle\}, F, P(U)\rangle$ and $\langle U' = \{x\}, V' = \{y, z\}, E' = \{\langle x, y\rangle, \langle x, z\rangle\}, F', P(U')\rangle$ are merged to get a new SCM $\langle U'' = \{u, v\}, V'' = \{w, x, y, z\}, E'' = \{\langle u, w\rangle, \langle v, w\rangle, \langle w, x\rangle, \langle x, y\rangle, \langle x, z\rangle\}, F'' = F \cup F', P(U'')\rangle$. However, there is no easy way to un-merge the SCMs into their constituent components, at least not without allowing one to arbitrarily split an SCM at any endogenous variable — which is hardly conducive to good system design. In contrast if one utilises SCMs with socket variables (see Figure 5.6), the SCMs $\langle U^H = \{u, v\}, U^S = \{\}, V = \{w\}, E = \{\langle u, w\rangle, \langle v, w\rangle\}, F, P(U^H), P(U^S)\rangle$ and $\langle U^{H\prime} = \{\}, U^{S\prime} = \{x\}, V' = \{y, z\}, E' = \{\langle x, y\rangle, \langle x, z\rangle\}, F', P(U^{H\prime}), P(U^{S\prime})\rangle$ are merged to form a new

SCM $\langle U^{H\prime\prime} = \{u, v\}, U^{S\prime\prime} = \{x\}, V'' = \{w, y, z\}, E'' = \{\langle u, w\rangle, \langle v, w\rangle, \langle w, x\rangle, \langle x, y\rangle, \langle x, z\rangle\}, F'' = F \cup F', P(U^{H\prime\prime}), P(U^{S\prime\prime}))\rangle$. Because this formulation specifically identifies the socket variables — and importantly maintains their probability distributions — one can split the new SCM back into its constituent parts if desired. Thus socket variables are not used within SCM modules so much as in-between them, functioning as an interface. One can substitute $w$ and $x$ in Figure 5.6 for say the output of a controller module and input of a physical-representation module respectively. Allowing their fluid reconfiguration at runtime in a way that otherwise would not be possible.

Having introduced the concept of socket variables, the previously-described decomposition leaves a set of hidden exogenous variables $U^H$, which effectively just consists of regular exogenous variables. However, as a matter of system design it is assumed that access to modify $P(U^H)$ is determined by a wrapper interface for the SCM following principles of encapsulation, and as such these variables are effectively 'hidden' by the wrapper.

As a final note, socket variables do bear some resemblance to the use of context variables in existing work [167, 168]. Within these works context variables are typically used to integrate data with varying levels of intervention across exogenous variables during causal discovery. This is done by having the context variables conditionally draw from either the original exogenous variable distribution or alternatively an input specified by an intervention. In the context of robotics this could allow a robot to learn a causal model from a combination of passive observation and active experimentation [168]. For this use case it is important to ensure that certain properties regarding the independence of exogenous variables are upheld [167], as this is are core assumption regarding SCMs. However this is where socket variables differ from context variables. Socket variables make no claims regarding their independence from one another once they are connected to parent variables. This is because the intended use of socket variables is to facilitate the merging and un-merging of SCM modules. Thus for all intents and purposes once two SCMs have been merged the socket variables situated between the SCMs act as endogenous variables, which need not be independent of one another. Therefore while context variables and socket variables utilise a similar mechanism, their goals and the guarantees associated with them are notably different.

### 5.2.6 Retrospective Causal Stationarity with Mutable Input Sets

In order to capture a system of multiple agents interacting within a shared environment, it is necessary to be able to support the collation of a fluid number of sources of input data into a set. It is trivial to construct a structural equation which takes input from a fixed number of known agents. Thus the challenge is extending SCMs to allow the fluid introduction and removal of SCM modules corresponding to the varying presence of agents in the environment.

The most simple and often practical approach to this is to simply reconfigure the structural equations and graphical structure as appropriate while the overall SCM is actively in use. This arguably makes the most sense for online deployments as one cannot pre-construct the SCM to account for agents one is not aware of yet. Additionally deployments taking place over a protracted length of time will have the overall SCM continue to grow unless there is a mechanism for pruning irrelevant agent SCM modules.

The issue with this is that altering the structural equations of SCMs amounts to a breach of causal stationarity [137], which poses an issue for window-time-graph roll-outs, the PTS-variable temporal representation introduced here, as well as many inference techniques applied to SCMs. To illustrate this, consider the scene depicted in Figure 5.7a. In this scene, initially a red and green vehicle are present, with a blue vehicle later joining the scene, and finally the red vehicle leaving the scene. A naive approach to represent this would just be to add or remove variables relating to these agents as they come or go, as shown in Figure 5.7b. The issue with such a representation is that the SCM given for each time step is only usable for inference in that time step, as other time steps might have data for variables that are not present, or be missing data for variables that are present.

Despite these aforementioned issues it is argued here that online mutation of the input set to a structural equation can be supported while maintaining a form of causal stationary:

**Definition 17** (Retrospective Causal Stationarity)**.** Let $T$ be defined in such a way that $\min T = 0$. A fixed SCM $M_t$ that accurately models causal relations for time $t \in T$ has Retrospective Causal Stationarity (RCS) if it accurately models all previous time steps $T_{<t} = \{\, t' \,|\, t' < t,\, t' \in T \,\}$. This effectively means that if $t$ is the most recently observed time step, $M_t$ can be utilised for inference without concern for a lack of causal stationarity.

Assuming the most up to date SCM in use has RCS one can safely utilise the SCM across all previous time steps while remaining faithful to the underlying data. Hypothetically it is also possible to use such an SCM for future time-step prediction, but such inferences suffer the same caveats that most extrapolations do.

**Theorem 2** (Mutable-Input-Set RCS)**.** *One can capture the fluid introduction and removal of sources of type* $\mathbb{N} \times \mathcal{T}$ *to and from the input set for a variable* $V_i$ *with structural equation* $f_{V_i}(\cdot) : 2^{\mathbb{N}\times\mathcal{T}} \to \mathcal{T}^{V_i}$ *via a series of SCMs* $\{M_0, M_{\delta t}, ..., M_t\}$ *that each provide RCS.*

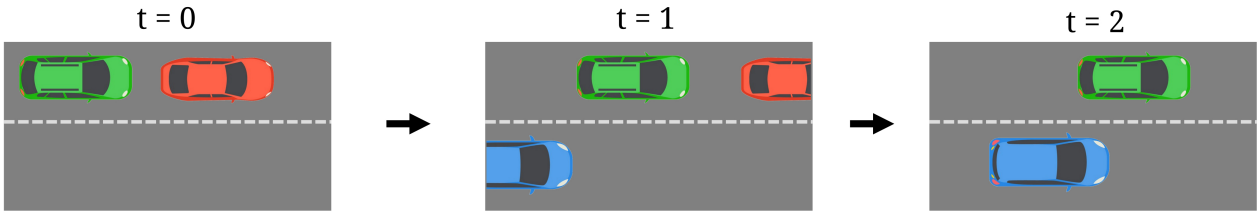
Here the set of natural numbers $\mathbb{N}$ is used to uniquely identify contributions from different sources (e.g. agents) and importantly allows the same value of data type $\mathcal{T}$ to be captured several times within the input set.

Figure 5.7c offers a visual overview of the approach this work takes to handle mutable input sets and thus model scenes such as the one depicted in Figure 5.7a. The following text provides the theoretical underpinning for said approach by providing a proof of RCS by induction.

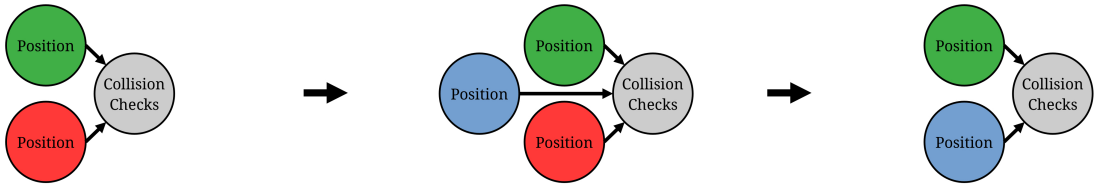
The structure relied upon by this approach to construct sets from a variable number of inputs within SCMs must first be described. The input set that is passed to $V_i$ is incrementally built up from the combination of:

- A socket variable $U^{S,\emptyset}$ where the probability distribution $P(U^{S,\emptyset})$ is a degenerate distribution that always returns the empty set $\emptyset$.
- A union variable $V^{\cup}$ that has two parents providing input sets of type $2^{\mathbb{N}\times\mathcal{T}}$ and outputs the union of said sets. The first of these parents provides a singleton set from a given source (e.g. an agent), while the second parent is a socket variable as described above.

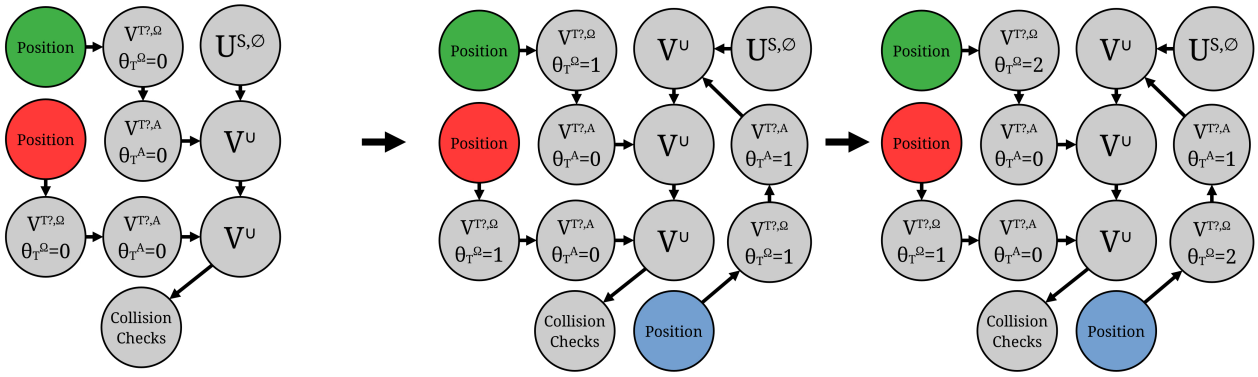
This effectively allows the system to iteratively construct a set beginning with $U^{S,\emptyset}$ providing an empty set. From here one can connect a union variable $V^{\cup\prime}$ in order to incorporate a new input into the set, while providing its own socket variable $U^{S\prime}_{\emptyset}$ from which another additional union variable could be connected in future. This structure allows for the fluid introduction and removal of input-set sources, as it effectively emulates a linked list.




(a) A scene illustrating a varying number of agents. At $t = 0$ just two vehicles, red and green respectively, are present. At $t = 1$ a blue vehicle enters the scene. Lastly at $t = 2$ the red vehicle leaves the scene. Despite the varying presence of agents across time it is desirable to have a stationary SCM that remains valid across time. Within the context of this work it has been identified that although a truly stationary SCM that can support this may not be possible, an SCM supporting a weaker form of stationarity — i.e. RCS — may be possible.




(b) A naive approach at causally modelling collision checking for agents across time. Because the SCM simply has variables added or removed based upon which agents are present for a given time step it is not possible to use a given SCM structure consistently across the scene. For example, attempting to utilise the SCM from $t = 2$ at $t = 0$ would result in one having no variable to pass the position data for the red vehicle to, and a spare variable for the blue vehicle's position with no input data.




(c) The approach given here based upon the SCM extensions provided by this work, in contrast to the above naive approach. While new variables are added across time, access to each of them is gated behind two time-conditional variables. These ensure that the agent variables are only referenced in time steps for which there is data relating to said agents. As such, the most up to date SCM under this approach can always be used across all previous time steps, thus affording it RCS.

Figure 5.7: A depiction of a scene with a varying number of agents, and a comparison of a naive approach against the extensions presented here in terms of representing the scene in a stationary manner.

One can now consider the following basic lemma:

**Lemma 2.1** (Base Case)**.** *Provided the SCM $M_0$ at the first time step is accurate for said time step, it has RCS.*

*Proof of Lemma 2.1.* Given that $M_0$ accurately models causal relations for $t = 0$, and $T_{<0} = \emptyset$, it automatically follows that $M_0$ accurately models for all time steps in $T_{<0}$. □

So that one may capture the fluid nature of the input-set sources, rather than having an input source feed directly into a union variable $V^{\cup}$, it should instead feed through a pair of time-conditional variables $V^{T?,A}$ and $V^{T?,\Omega}$. These variables should only return the input-set source for $\theta_T^A \leq t \leq \theta_T^\Omega$ where $\theta_T^A$ and $\theta_T^\Omega$ are configurable time parameters for each of the aforementioned time-conditional variables respectively. Should $t$ fall outside of these bounds, the time-conditional variables should return the output of a fixed exogenous variable $U^{S,\emptyset}$ where $P(U^{S,\emptyset})$ is a degenerate distribution that always returns the empty set $\emptyset$.

Finally, in order to demonstrate that all subsequent SCMs can also have RCS, one must show that the two means of modifying the input set — i.e. the introduction of an input-set source and the removal of an input-set source — can be carried out while maintaining RCS for the overall SCM.

**Lemma 2.2** (Introduction Induction Step)**.** *If $M_{t-\delta t}$ is an SCM with RCS and at least one new input-set source is introduced at the next time step $t$ then one can construct a new SCM $M_t$ that accurately captures causal relationships at time step $t$ and has RCS.*

*Proof of Lemma 2.2.* Construct $M_t$ by extending the SCM to incorporate the new input-set sources as described above and assign $\theta_T^A = t$ and $\theta_T^\Omega = t$ for their time-conditional variable parameters. Based upon the configuration described above, $\theta_T^A \leq t \leq \theta_T^\Omega$ and thus $M_t$ will include the new input-set source, accurately capturing the causal links at time $t$.

Meanwhile, given that $T_{<t}$ is defined as all time steps less than $t$, and for this source $\theta_T^A = t$, one can conclude that for all previous time steps captured by $T_{<t}$, the new SCM extension will only contribute $\emptyset$ to the input set. This emulates the input-set source not being present in the past, accurately reflecting the causal links present during those time steps. Thus, one can conclude that $M_t$ also has RCS after the new input-set-source introduction. □

**Lemma 2.3** (Status-Quo / Removal Induction Step)**.** *If $M_{t-\delta t}$ is an SCM with RCS and zero or more input-set sources are removed at the next time step $t$ then one can construct a new SCM $M_t$ that accurately captures causal relationships at time step $t$ and has RCS.*

*Proof of Lemma 2.3.* At each time step $t$, for each input-set source that is present in the modelled system, assign $\theta_T^{\Omega} = t$ for the configurable parameter of the associated $V^{T?,\Omega}$ in $M_t$ respectively. SCM input-set sources for which $\theta_T^A \leq t \leq \theta_T^{\Omega}$ have their contributions included in the input set. Thus, this update combined with Lemma 2.2 ensures that all input-set sources that are present in the modelled system will have their contributions captured within the input set, as is accurate for time step $t$.

All the while, $\theta_T^{\Omega}$ is always first initialised to the time step in which the associated input-set source was introduced — as per Proof of Lemma 2.2 — and is only updated as described above. As such, if an update is missed for a given input-set source — i.e. due to it no longer being present in the modelled system — then it follows that $t > \theta_T^{\Omega}$, and thus the SCM input-set source will return $\emptyset$, no longer contributing to the input set. In doing so the SCM remains accurate for time step $t$, and additionally retains the same behaviour it possessed for previous time steps as $\theta_T^A$ and $\theta_T^{\Omega}$ remain the same as in $M_{t-\delta t}$. Thus, one can conclude that $M_t$ also has RCS for instances where the status quo is maintained, and for instances where one or more input-set sources are removed. □

With the base case and induction steps defined, it is now possible to give the following proof by induction for the theorem.

*Proof of Theorem 2.* It has been demonstrated through Lemma 2.1 that the accurate SCM associated with the first time step $M_t$ automatically has RCS, effectively establishing a base case. Additionally it has been shown that input-set-source introductions (Lemma 2.2), removals (Lemma 2.3), and maintaining the status quo of input-set sources between time steps (Lemma 2.3) can be done while maintaining RCS for $M_t$ given that $M_{t-\delta t}$ had RCS. Thus via induction one can infer that the sequence of SCMs $\{M_0, M_{\delta t}, ..., M_t\}$ accurately captures these fluid mutations and all possess RCS. □

This provides a well defined process by which one can model a varying number of agent interactions via SCMs as shown previously in Figure 5.7c. This is undoubtedly a valuable trait for AVs with causal models operating in a shared environment which will have agents coming and going that need to be captured properly. Additionally the utility of socket variables and time-conditional variables has been demonstrated in making such an approach possible.

## 5.3 Case Studies in Modelling Autonomous Agents via SCMs

In order to demonstrate the extensions that have been introduced, two AES architectures were developed for the AV and service-robotics domains as case studies. Given that that these architectures are comprised of several SCM modules and even a single of these modules requires substantial documentation only a high-level overview is given here, which does happen to include the graphical structure of the SCM modules.

The full details of the SCM modules both in terms of their structural equations and explicit implementation are provided in a Git repository[10]. Furthermore, data and utility scripts relating to any experiments described here are stored in a separate auxiliary Git repository[11].

Before proceeding further it should be noted that although the architecture code and SCM formulations represent novel work, the underlying models represented by these SCM modules (e.g. rigid bodies, dynamic-bicycle models, social-forces models) are largely based upon existing literature. As such throughout the following case studies citations will be provided to ensure proper attribution is provided where SCM modules have adapted existing models.

In terms of the overall design of the case-study architectures, both represent AES agents by implementing the following elements:

- **Physical Representation:** Captures the dynamics and control inputs of the AES in question. This representation may be composed of several SCM modules put together. For example, there might be a basic point-mass representation, a rigid-body representation that builds upon it, and finally an SCM that captures how the driving-actuation inputs influence the forces and torques of the rigid body. These combine to provide a comprehensive representation of the AES's physical presence in the world.

[10]https://github.com/cognitive-robots/SimCARSv2
[11]https://github.com/cognitive-robots/av-extended-scms-paper-resources

- **Controller:** Provides a means of translating an input action to driving-actuation outputs, namely those suitable to be passed to the physical representation element described above. There is not necessarily any one correct implementation for any given AES, however the domain and particular AES type does influence the sorts of data that need be considered and what potential peripheral components may be required (e.g. a map).
- **Planner:** Determines an optimal, or near-optimal, course of behaviour for the AES based upon the current world state. This is provided in the form of an action, and similar to the controller described above, there is not necessarily an objectively best planner for any given AES. However, most planners do typically rely upon the use of a reward model — to determine the optimality of an outcome — and either a forward or inverse model in order to associate actions with their corresponding outcomes.
- **Entity and Link Representation:** Entity representations capture all of the interactions within a shared environment that affect a given AES. Among other matters, entity representations collate the influences of link representations. These in turn represent the the interactions between two entities within the shared environment. Importantly while there need only be one entity-representation SCM for a given AES, there may need to be several differing link-representation SCMs in order to account for interactions between different types of environment entity.

These elements are then reproduced for each agent and connected via the link representations, thus enabling the inference of high-level causal relationships by exploiting low-level knowledge of the systems involved.

The architecture modules themselves are described graphically in terms of their SCMs with the legend for said graphical depictions given in Figure 5.8. Of particular note is that exogenous variables of either type — i.e. socket or hidden — are depicted as rounded rectangles, while endogenous variables are depicted as rectangles with sharp-corners. Other shapes are utilised to depict other specialised variable types introduced in this chapter, while the colour of variables indicates to which module they belong.

With the generic structure and manner of presentation for the case-study architectures established, one can begin to examine the case studies themselves.

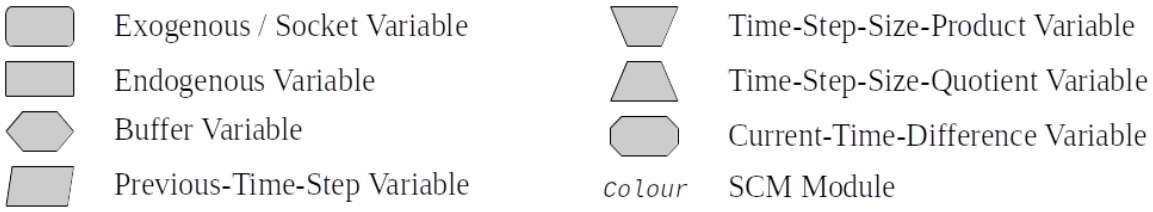



Figure 5.8: SCM Legend

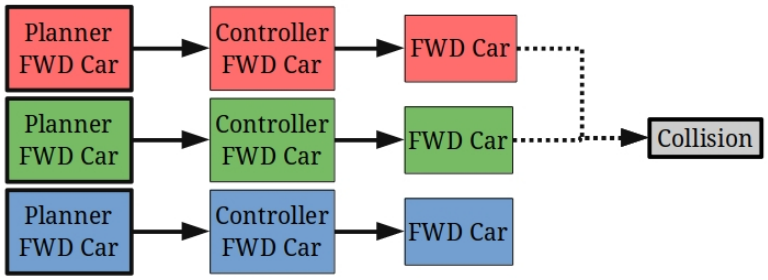



Figure 5.9: Depiction of the AV SCM architecture components causally interacting within the collision scenario depicted in Figure 1.1 and described in the text.

### 5.3.1 Autonomous Vehicles

A potential motivation behind developing an SCM architecture capturing the interactions of AESs in AV scenarios is for post-hoc analysis of scenarios. Critically, the AV domain is safety critical yet nonetheless presents a greater degree of order in terms of the expectation that road agents follow the laws of the road. The combination of these traits makes it both appealing and convenient to model causally.

#### 5.3.1.1 Example Scenario

An example AV scenario of the beneficial utilisation of SCM integration in post-hoc analysis was provided in Section 1.1.2 and illustrated in Figure 1.1. To reiterate, the scene depicts a red vehicle overtaking and a green vehicle accelerating from the opposite direction, resulting in a collision. The causal links related to this are depicted in Figure 5.9. This shows both the actions of the red agent and the green agent were necessary to cause the collision. While failures could in theory occur at the mechanical or controller level this work is interested in decisions made by the agent at the planner level. After all this is the level at which the agent has the most influence and thus carries greatest responsibility.

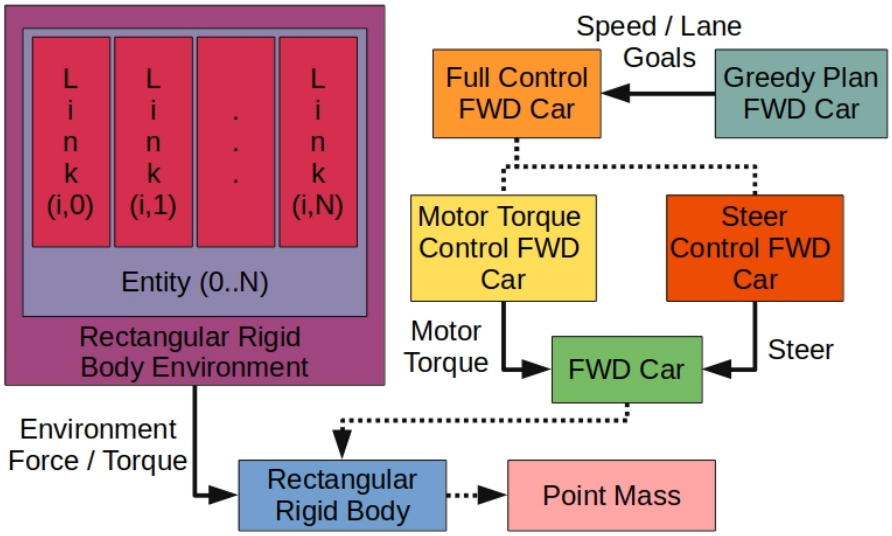



Figure 5.10: An overview of the AV case-study SCM architecture. Solid and dotted lines represent composition and inheritance relationships respectively. Although the individual components of the architecture are based upon well established models from literature the novelty of this architecture is in its SCM implementation, relying upon the extensions provided in this work. Specifically the architecture utilises the modularity and encapsulation provided by socket and buffer variables, the constant-space temporal representation avoiding the need for explicit roll-outs, and the support for a varying number of agents across time.

#### 5.3.1.2 Architecture

An overview of the AV SCM architecture is depicted in Figure 5.10. For this domain, here AES goals have been formulated as a combination of a target lane and a target speed with corresponding times by which to achieve these goals. These goals are converted into motor-torque and steering actuation values via a proportional-feedback controller. The motor-torque and steering values are then fed into a dynamic bicycle model [169]. Meanwhile the entity modules and their associated link modules are used to provide agent collision and drag force / torque information, treating vehicles as rectangular rigid bodies. The mechanics of the vehicles also assume they use Front-Wheel Drive (FWD).

The following text will examine the individual SCM modules of the architecture in detail, while also discussing where the extensions are used within the context of the example scenario.

**Point Mass** The SCM depicted in Figure 5.11 represents a 2D dynamic point mass [160] that does not explicitly occupy any portion of space yet provides the represented object with a mass, position, linear velocity, and linear acceleration, aswell as input forces which influence these.

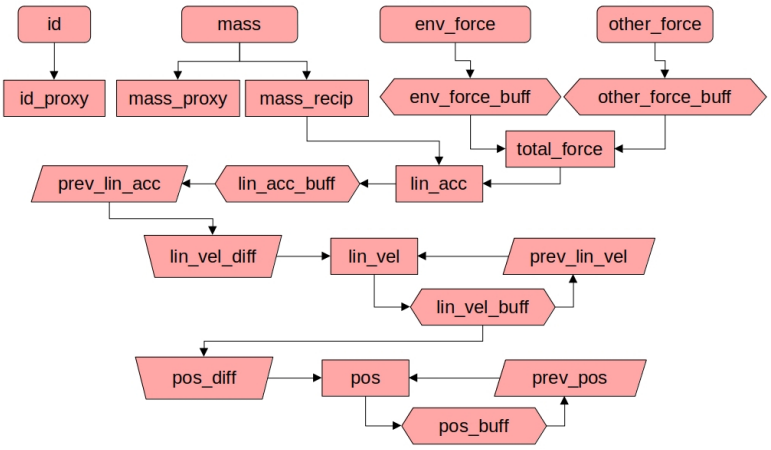



Figure 5.11: *Point Mass* SCM

Within the context of the example scenario, this is utilised to model the basic dynamics of the vehicles. While this is certainly not something that is only possible via the extensions provided here, the extensions do provide the following benefits:

- Succinct temporal representation through the *prev_pos*, *prev_lin_vel*, and *prev_lin_acc* PTS variables, preventing the need to duplicate variables across time steps.
- Encapsulation of SCM data within the *pos_buff*, *lin_vel_buff*, *lin_acc_buff*, *env_force_buff*, and *other_force_buff* buffer variables.
- Utilisation of the *env_force* and *other_force* socket variables, allowing the fluid reconfiguration of force inputs from the environment, and other forces (e.g. self-propelled motion).

**Rectangular Rigid Body** The SCM depicted in Figure 5.12 inherits from the *Point Mass* SCM, extending it to represent a 2D rectangular rigid body [160]. In doing so it allows the object to occupy a rectangular portion of space and provides it with a moment of inertia, rotation, angular velocity, and angular acceleration, aswell as input torques which influence these. This SCM additionally tracks the open-space in-front of the rectangular rigid body, as this information is calculated via the *Rectangular Rigid Body Entity* SCM via which a rectangular rigid body is represented in a shared environment.

Similar to the *Point Mass* SCM, this is just used to provide further dynamics to the vehicles within the example scenario. Again, utilisation of the extensions is not strictly necessary here, but provides the following benefits:

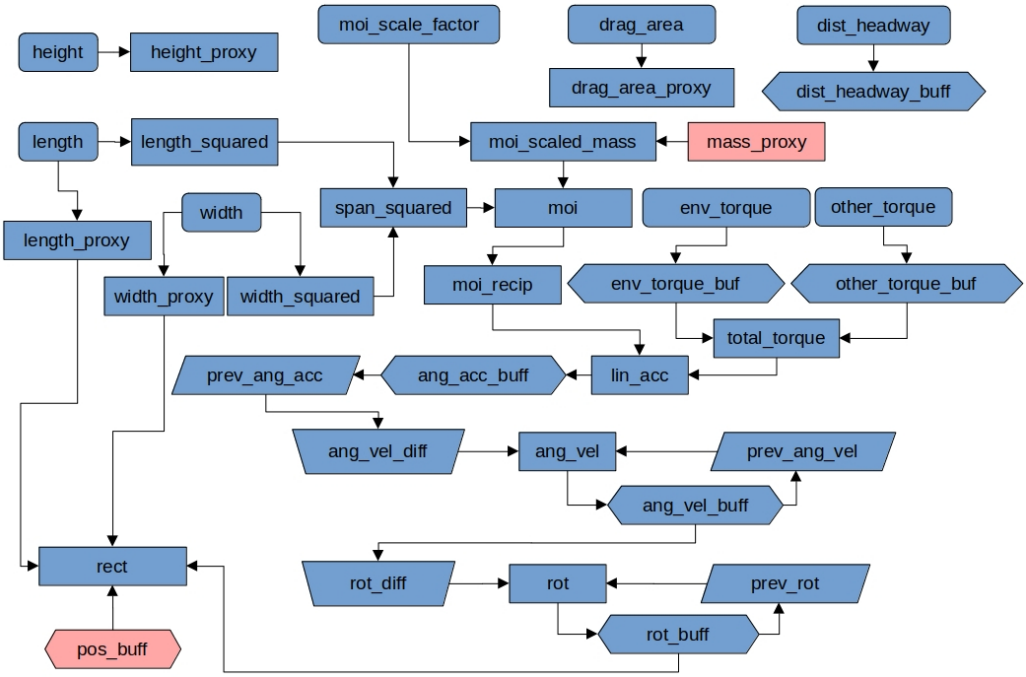



Figure 5.12: *Rectangular Rigid Body* SCM

- Succinct temporal representation through the *prev_rot*, *prev_ang_vel*, and *prev_ang_acc* PTS variables, preventing the need to duplicate variables across time steps.

- Encapsulation of SCM data within the *rot_buff*, *ang_vel_buff*, *ang_acc_buff*, *dist_headway_buff*, *env_torque_buff*, and *other_torque_buff* buffer variables.

- Utilisation of the *dist_headway*, *env_force* and *other_force* socket variables, allowing the fluid change of force / distance-headway inputs from the environment, and other forces.

**Rectangular Rigid Body Entity** The SCM depicted in Figure 5.13 represents rectangular rigid-body objects within a shared environment and performs two functions. The first is to calculate environmental forces. This involves combining drag force and the collision forces calculated by the *Rectangular Rigid Body Link* SCMs associated with a given *Rectangular Rigid Body Entity* SCM. The second is to calculate the minimum distance headway given by the *Rectangular Rigid Body Link* SCMs in order to determine the overall distance headway.

This SCM facilitates all interactions between a given vehicle and the shared environment within the example scenario. Unlike the previous two SCMs, this SCM inherently relies upon the formalisms introduced here:

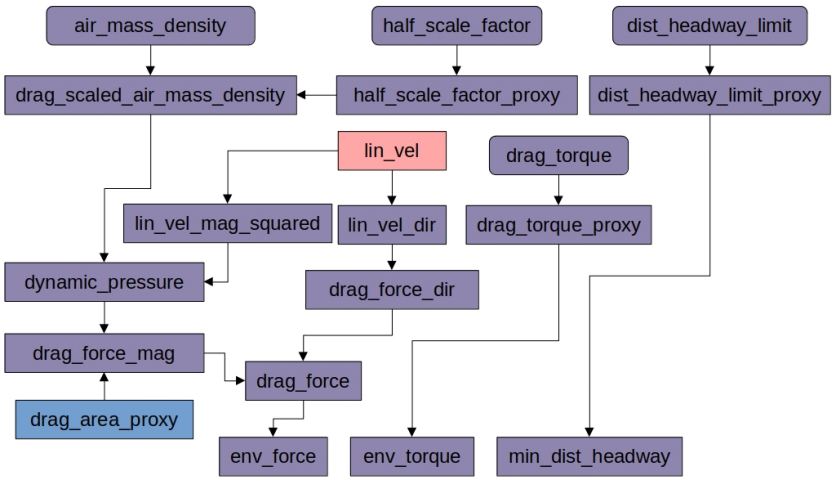



Figure 5.13: *Rectangular Rigid Body Entity* SCM

- The structure used to demonstrate that mutable input sets can be facilitated while maintaining RCS is used here in order to sum up the various environmental forces / torques that affect a given vehicle, as well as calculate the minimum distance headway. In particular this allows vehicles to come and go throughout the lifetime of a scenario such as the one described above. The implementation of this structure in turn requires the use of time-conditional and socket variables.

- Due to the SCM not utilising any buffer variables, this SCM does not store any data associated with its variables, thus helping to optimise memory usage.

**Rectangular Rigid Body Link** The SCM depicted in Figure 5.14 captures interactions between two rectangular rigid bodies in a shared environment. It most critically performs the collision computations for the two rigid bodies in question. It additionally computes distance headway between the two rigid bodies as this is frequently used within the driving domain as a safety metric [22]. Variables belonging to the primary and secondary *Rectangular Rigid Body* SCMs associated with the link are indicated via $A$ and $B$ subscripts respectively.

While the *Rectangular Rigid Body Entity* SCM represented the sum total of factors affecting a given vehicle, this SCM captures the interactions between two vehicles in particular, such that SCMs of this type can then have their values collated by a *Rectangular Rigid Body Entity* SCM. This SCM makes a small amount of usage of the extensions:

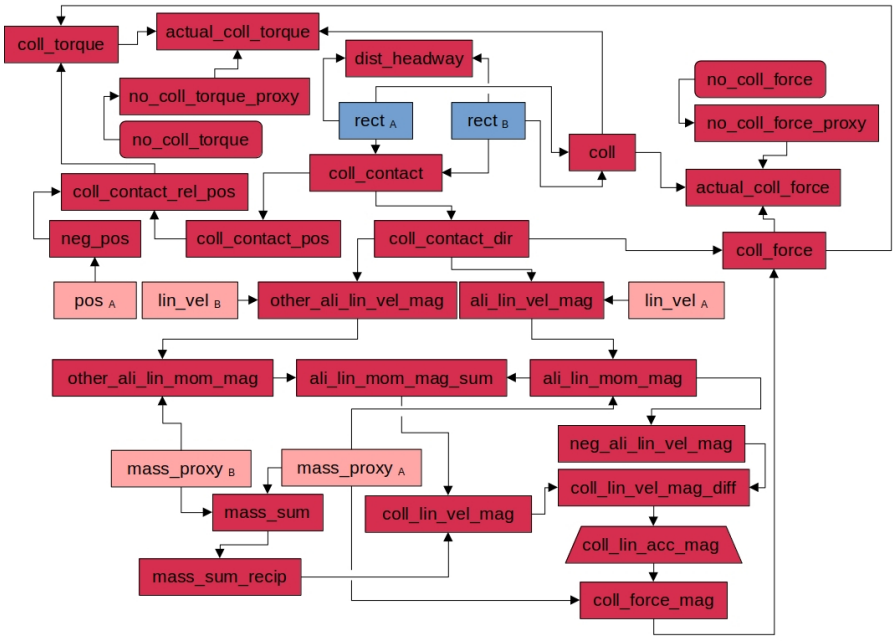



Figure 5.14: *Rectangular Rigid Body Link* SCM

- The *coll_lin_acc_mag* TSSQ variable allows the calculation of the required acceleration resulting from a collision given a required linear-velocity change and the time-step size.

- Due to the SCM not utilising any buffer variables, this SCM does not store any data associated with its variables, thus helping to optimise memory usage.

**FWD Car** The SCM depicted in Figure 5.15 inherits from the *Rectantular Rigid Body* SCM to represent a road vehicle, in particular one that has FWD. This SCM provides forces and torques as output to the underlying *Rectangular Rigid Body* SCM, instead providing variables for motor torque and steering as input. In order to calculate the forces and torques required, the SCM utilises these two inputs and treats the represented object as a dynamic bicycle [169], a simple means by which one can model a variety of vehicles.

In contrast to the *Point Mass* and *Rectangular Rigid Body* SCMs, this SCM provides the highest level of the physical representation of vehicles within the AV domain and is comprised of vehicle-specific variables. Once again, utilisation of the extensions is not completely necessary here, but is desirable:

- Encapsulation of SCM data in the *motor_torque_buff* and *steer_buff* buffer variables. This also saves on the memory consumption of the SCM given the large number of variables.

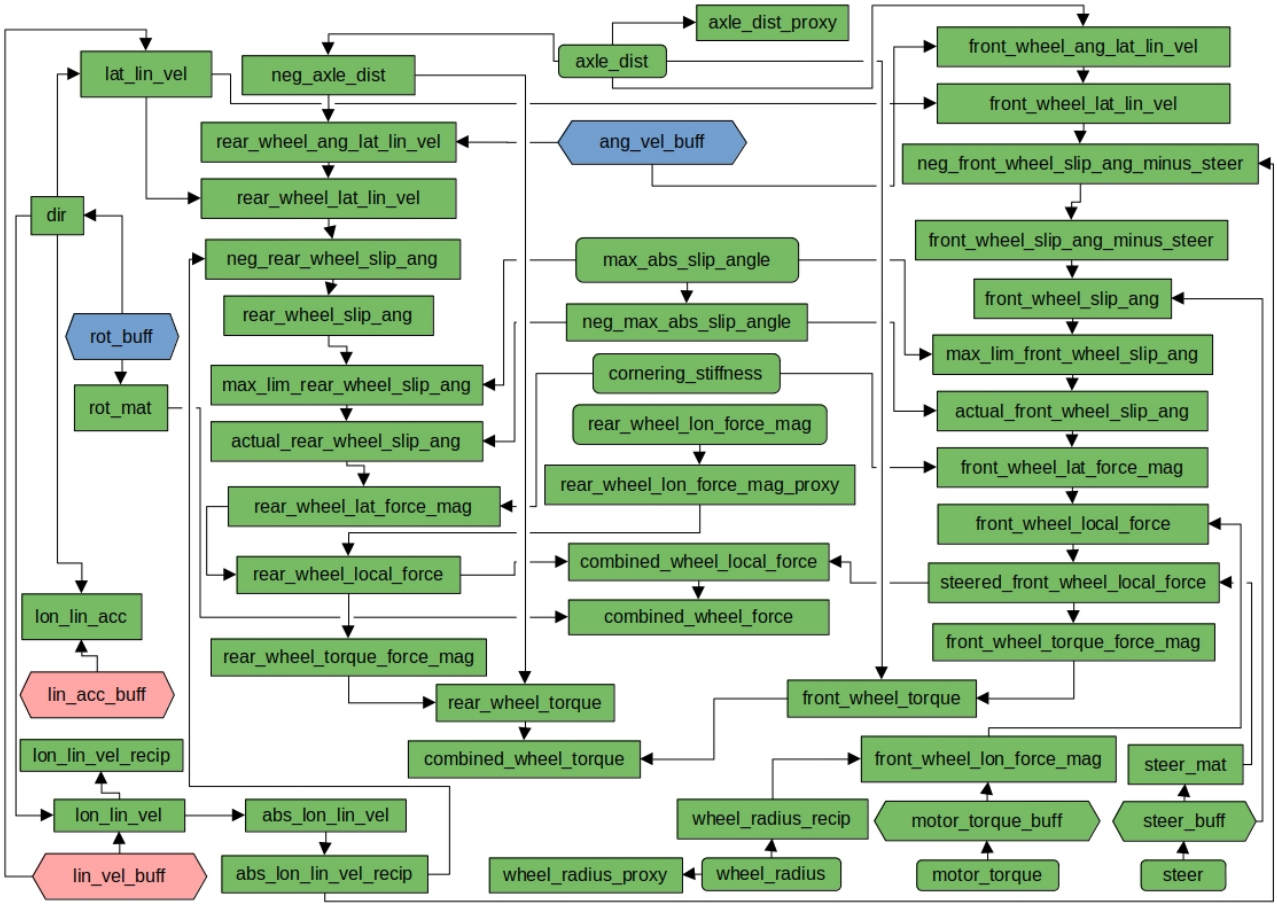



Figure 5.15: *FWD Car* SCM

- Utilisation of the *motor_torque* and *steer* socket variables, allowing the fluid reconfiguration of force / distance-headway inputs from the environment, and other forces. This is arguably of greater use here than in the *Point Mass* and *Rectangular Rigid Body* SCMs as this SCM interfaces directly with controller components within the architecture.

**Motor Torque Control FWD Car** The SCM depicted in Figure 5.16 provides a control mechanism which outputs motor torque in return for giving a goal speed and a goal time as input. From the difference between the longitudinal linear velocity given by the *FWD Car* SCM and the goal speed, and the difference between the current time and the goal time it is possible to calculate the necessary acceleration. Using the mass of the *FWD Car* SCM one can calculate the required force, before making adjustments to the force to account for environment forces also acting upon the *FWD Car* SCM. From the calculated force and wheel radius of the vehicle one can finally approximate the motor torque that can then be fed into the *FWD Car* SCM.

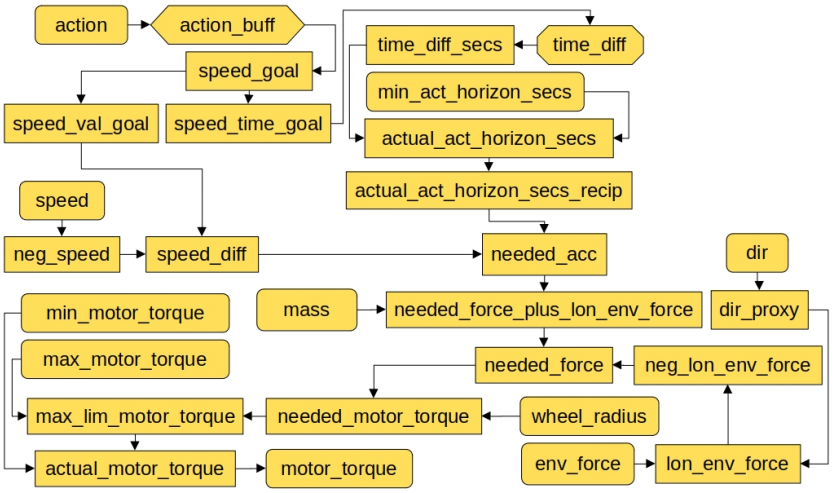



Figure 5.16: *Motor Torque Control FWD Car* SCM

Within the context of the example scenario this SCM, along with the *Steer Control FWD Car* SCM, is responsible for bridging the gap between the physical representation and action planner of the vehicles in question. Here, there is utilisation of several SCM-formalism extensions:

- The *time_diff* CTD variable allows the calculation of the time difference between the speed-goal target time and the current time.
- Encapsulation of SCM data within the *action_buff* buffer variable.
- Utilisation of the *action* socket variable, allowing the fluid reconfiguration of which planner ought to be used.
- Utilisation of the *speed*, *dir*, *mass*, *wheel_radius*, and *env_force* socket variables. Unlike the *action* socket variable, these are used as inputs from the physical representation, which are utilised to determine how the driving-actuation variables of the physical representation — i.e. the *motor_torque* variable — ought to be modulated.

**Steer Control FWD Car** The SCM depicted in Figure 5.17 provides a control mechanism which outputs steer in return for giving a lane-goal lane identifier and target time as input. In order to calculate the steering first the expected position of the rectangular rigid body is calculated for a predefined number of 'look-ahead' time steps. This is then projected onto the closest point of the central path of the lane associated with the lane-goal lane identifier. The

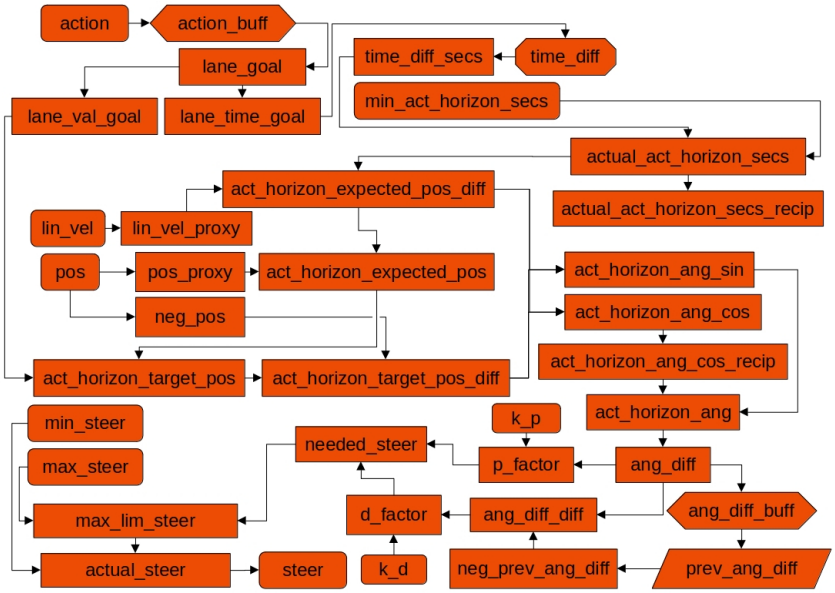



Figure 5.17: *Steer Control FWD Car* SCM

expected position combined with the projected position and current position given by the *FWD Car* SCM provide an angle error. This angle error, along with its time derivative are then used as inputs to a proportional-derivative-feedback controller [170], with the output steering then fed into the *FWD Car* SCM.

Within the context of the example scenario this SCM, along with the *Motor Torque Control FWD Car* SCM, is responsible for bridging the gap between the physical representation and action planner of the vehicles in question. Here, several SCM-formalism extensions are exploited:

- The *time_diff* CTD variable allows the calculation of the time difference between the lane-goal target time and the current time.

- Encapsulation of SCM data within the *action_buff* and *ang_diff_buff* buffer variables.

- Utilisation of the *action* socket variable, allowing the fluid reconfiguration of which planner ought to be used.

- Utilisation of the *pos* and *lin_vel* socket variables. Unlike the *action* socket variable, these are used as inputs from the physical representation, which are utilised to determine how the driving-actuation variables of the physical representation — i.e. the *steer* variable — ought to be modulated.

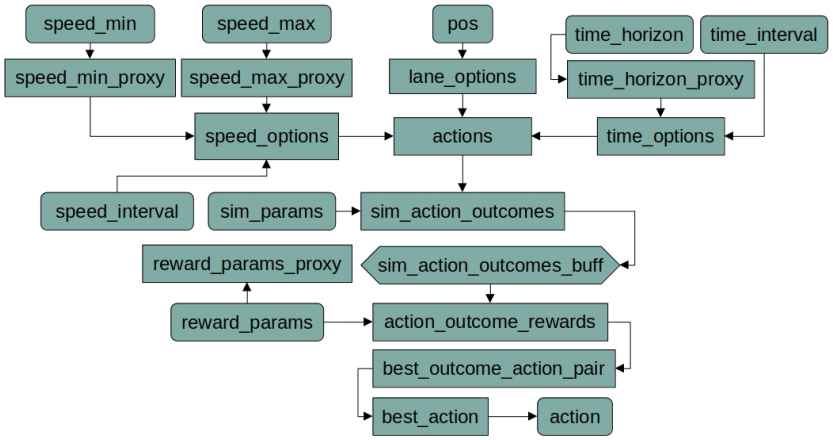



Figure 5.18: *Greedy Plan FWD Car* SCM

**Full Control FWD Car** Simply a wrapper SCM that inherits from both *Motor Torque Control FWD Car* and *Steer Control FWD Car* SCMs. These SCMs combined are able to take a speed and lane identifier associated with goals along with the target times by which to accomplish these goals. In turn the combined SCM can then output motor-torque and steer values for the associated *FWD Car* SCM.

**Greedy Plan FWD Car** The SCM depicted in Figure 5.18 plans the next action for a *Full Control FWD Car* SCM specified in terms of goals. In order to do so, a range of candidate speeds, lanes, and target times are calculated and combined to form a variety of possible actions. The outcomes associated with taking these possible actions are then simulated, and the rewards associated with the action-outcome pairs are calculated. This does rely upon a predefined simulator and predefined reward calculator being provided to the *Greedy Plan FWD Car* SCM. In any case, once the rewards have been calculated the action associated with the maximum reward is selected and fed into the corresponding *Full Control FWD Car* SCM.

Arguably the most important SCM besides the *FWD Car* SCM in the context of the example scenario, this SCM captures the decision-making capabilities of vehicular agents within the scene. In other words the decisions that are being identified as part of behavioural interactions are assumed to be made within a roughly comparable form of cognition. To achieve this the following extensions are utilised:

- Encapsulation of SCM data within the *sim_action_outcomes_buff* buffer variable.
- Utilisation of the *pos* socket variable. This is used as an input from the physical representation, which is utilised to determine how an action ought to be selected. In particular the *pos* variable is used to identify candidate lanes for the vehicular agent.

#### 5.3.1.3 Qualitative Experiment

Having described the architecture and how the formalisms have been used to facilitate and enhance aspects of it, one can now reconsider the scenario depicted in Figure 1.1 and Figure 5.9.

**Methodology** Similar to previous experiments within this work the tracks from the vehicular highD dataset [150] are considered. Here the aim is to demonstrate the expressiveness of the new architecture by examining a collision scenario, however such incidents do not typically feature in AV-domain datasets, and none are present throughout the highD dataset to the author's knowledge. Before proceeding, it should be noted that the format of such a scenario differs from the two-vehicle-convoy scenario discussed previously. As such use of the same identifiers would be illogical and thus the agents will be identified by colour henceforth. Since analysis of an observed collision scene from the highD dataset is not possible, instead a partially synthetic scene can be created. This can be done by taking an observed scene and intervening upon the actions of a vehicle — subsequently referred to as 'the cyan vehicle' or 'cyan' — to cause it to change lane, thus producing a collision with another vehicle — subsequently referred to as 'the magenta vehicle' or 'magenta'. Through this it is possible to produce a crash scene, which can then be analysed post-hoc. This can be done by modulating the actions of the cyan and magenta vehicles and simulating the outcome. Thus one can identify how the behaviour of each agent influences the occurrence of a crash, taking inspiration from the work of Chapter 4.

**Discussion** In Figure 5.19 one can see a scenario analogous to the one shown in Figure 1.1, only this one is generated from data drawn from the highD dataset and analysed counterfactually through the presented SCM architecture. As the figure shows, in this scenario the cyan vehicle attempts to merge into a lane after the magenta vehicle begins braking, resulting in a collision.

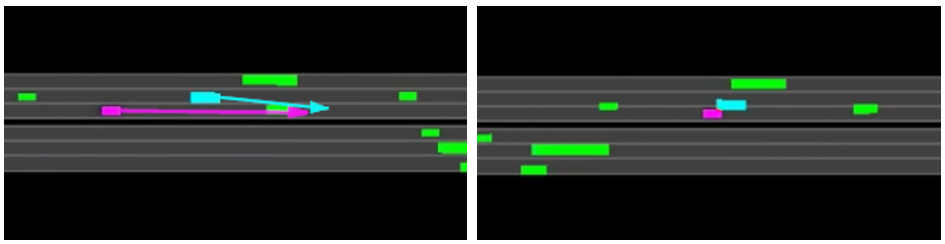
By modulating the actions of each agent via counterfactual interventions as described above one can attempt to identify the culpable agent. Conveniently, because the agent actions each relate to specific points in time one can rely upon the properties of temporal precedence and the existence of reaction times to avoid the formation of any cycles in the resulting causal graph. This does of course depend upon the length of time steps between data measurements being less than what one might expect in terms of a reaction time. For highD this is $40\,ms$, far below the average reaction time of a human — i.e. $\sim 220\,ms$ [171] — avoiding any concern.

Through these simulations one can infer that not only would the crash not have occurred had cyan not attempted to merge, but if magenta had not braked then it would have collided with another vehicle instead. Thus at least from a decision-making standpoint, cyan appears to carry greater culpability than magenta within the scenario. This assertion of culpability is based off the previously discussed 'but-for' test as laid out in the Model Penal Code [1] and contextualised within causality by Pearl and Mackenzie [2]. As the crash would not have occurred but for cyan attempting to change lanes, one can establish it as a culpable agent. Meanwhile a crash would have occurred regardless of whether magenta braked or not, which not only alleviates magenta of culpability, but also reflects sensible driving behaviour.

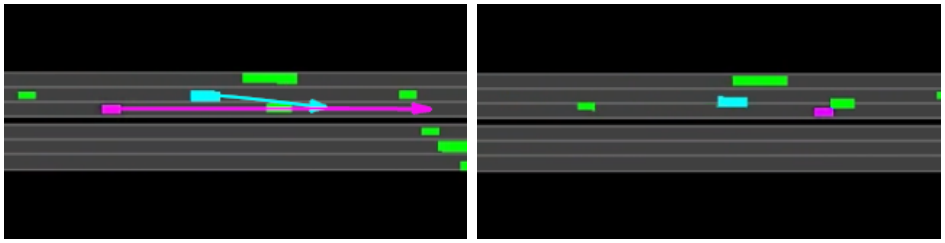
These outcomes make sense from the perspective of the experiment's design, as cyan's behaviour has been modified in a way that does not reflect typical human driving behaviour. Therefore the capability of the AV SCM architecture to make such inferences has been demonstrated to a sufficient degree for the purposes of this case study, although such a task will be revisited later in this thesis.

### 5.3.2 Service Robotics

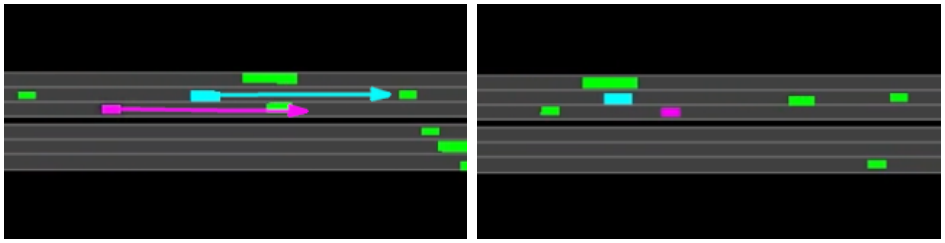
In contrast to the AV domain the service-robotics domain tends to carry substantially less risk. Yet this comes with a significantly greater level of chaos resulting from the free-form nature of action and interaction within pedestrian environments. While this arguably makes modelling such a system that much harder, it also gives one greater motivation to utilise causality in order to try and make sense of the chaos. Furthermore, because the domain carries less risk, one can utilise such information proactively.



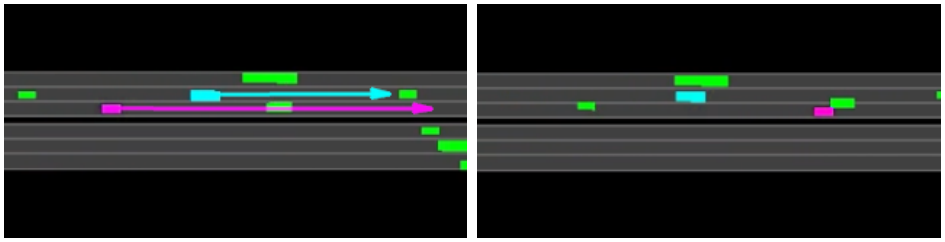
(a) Observed scene with collision



(b) Counterfactual without magenta braking



(c) Counterfactual without cyan meging



(d) Counterfactual without either action

Figure 5.19: Visualisation of a crash scene analysis by the AV SCM architecture. Subfigures show the initial scenes with agent plans on the left and final outcomes on the right. In the observed crash scene (Figure 5.19a) the cyan vehicle has just changed lanes, and the magenta vehicle has just braked. By modulating the agent actions one can see that if cyan had not attempted to merge (Figure 5.19c) a crash would not have occurred. Meanwhile in the other cases where magenta does not break (Figures 5.19b and 5.19d) a collision occurs regardless of what cyan does. As magenta had to brake to avoid collision, and braking in the absence of cyan changing lane does not cause a collision one can absolve magenta of blame here. This leaves cyan, whose decision to attempt a merge when it did can be considered a necessary cause of the collision, as the collision would not have occurred were it not for the lane change.

#### 5.3.2.1 Example Scenario

An example service-robotics scenario of the beneficial utilisation of SCM integration in online self-diagnosis was provided in Section 1.1.2 and illustrated in Figure 1.2. To reiterate, the scenario describes a green agent and a yellow agent both wishing to enter a narrow hallway,

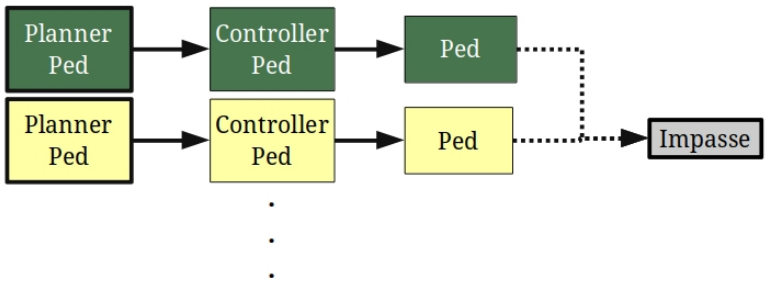



Figure 5.20: The components of the service-robotics SCM architecture and how they interact causally with the impasse scenario depicted in Figure 1.1 and described in the text.

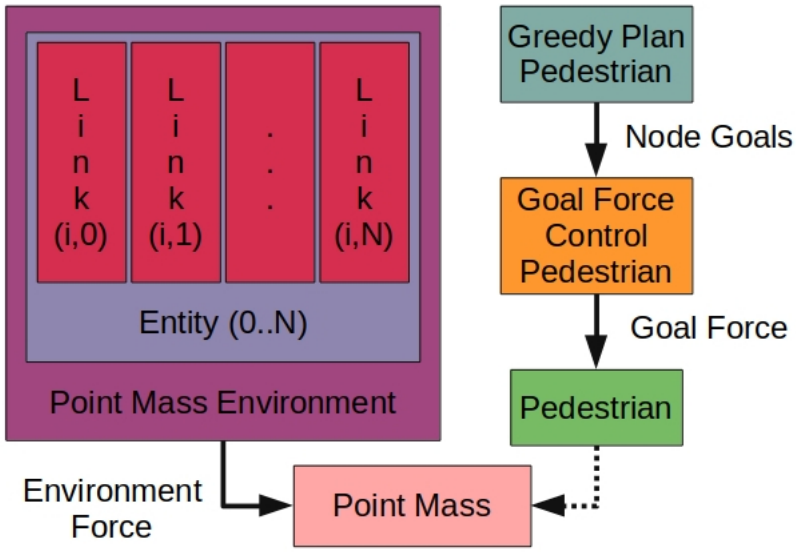



Figure 5.21: An overview of the service-robotics case-study SCM architecture. Solid and dotted lines represent composition and inheritance relationships respectively. Similar to Figure 5.10 the individual components of the architecture are based upon well established models from literature and the novelty of this architecture is in its SCM implementation, relying upon the extensions provided in this work. Once again, this architecture leverages the extensions made in relation to the modularisation and encapsulation of SCMs, constant-space temporal representation in SCMs, and the capacity to support a varying number of elements within SCMs.

and each attempting to give way to the other, resulting in an impasse. The causal links related to this are depicted in Figure 5.20. This shows that the actions — or lack thereof — of both the green agent and the yellow agent were necessary for the impasse to occur. Similar to the AV case, it is the planner level that is of interest. However, in this case it is because the green robot agent within the scenario may be able to learn in order to improve future decision-making rather than to ascribe some kind of blame.

#### 5.3.2.2 Architecture

An overview of the service-robotics SCM architecture is depicted in Figure 5.21. Since this domain does not have a direct analogue to lanes, the architecture specifies goals in terms of target nodes from a generated graphical road-map and the times by which these nodes should be reached. Given these nodes represent fixed points rather than channels of movement, and the maximum speed of pedestrians is comparatively low, it is also unnecessary to represent speed as a goal separate from the nodes. The controller once again utilises a simple proportional-feedback approach in order to provide a goal force which is applied to a point-mass representation of a human / robot. The environment entity and link modules are then used to represent the repulsive forces between agents that mimic the desire for personal space. The goal and repulsion forces combine to approximate the social-forces model developed by Helbing and Molnár [92].

The following text will examine the individual SCM modules of the architecture in detail, while also discussing where the extensions are used within the context of the example scenario.

**Point Mass** Effectively the same as the *Point Mass* SCM introduced in Section 5.3.1.2.

**Point Mass Entity** The SCM depicted in Figure 5.22 provides a representation of a point mass within a shared environment. Because this is tailored to the service-robotics domain this amounts to a sum of avoidance forces provided by the *Point Mass Link* SCMs associated with a given *Point Mass Entity* SCM.

This SCM facilitates all interactions between a given pedestrian and the shared environment within the example scenario. Unlike the *Point Mass* SCM, this SCM inherently relies upon the formalisms introduced here:

- The structure used to demonstrate that mutable input sets can be facilitated while maintaining RCS is used here in order to sum up the various environmental forces that affect a given pedestrian. In particular this allows pedestrians to come and go throughout the lifetime of a scenario such as the one described above. The implementation of this structure in turn requires the use of time-conditional and socket variables.

- Due to the SCM not utilising any buffer variables, this SCM does not store any data associated with its variables, thus helping to optimise memory usage.

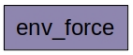



Figure 5.22: *Point Mass Entity* SCM

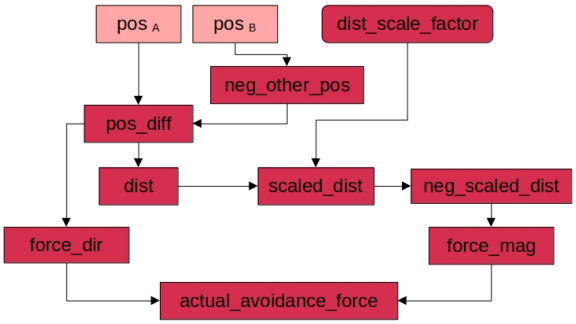



Figure 5.23: *Point Mass Link* SCM

**Point Mass Link** The SCM depicted in Figure 5.23 captures the interactions between two point masses in a shared environment. Again because this is tailored to a service-robotics scenario, these links capture avoidance forces in line with a social-forces model [92], with the repulsive forces scaling inversely to the exponent of the distance between the point masses. Variables belonging to the primary and secondary *Point Mass* SCMs associated with the link are indicated via $A$ and $B$ subscripts respectively.

While the *Point Mass Entity* SCM represented the sum total of factors affecting a given pedestrian, this SCM captures the interactions between two pedestrians in particular, such that SCMs of this type can then have their values collated by a *Point Mass Entity* SCM. This SCM only really benefits from one of the extensions:

- Due to the SCM not utilising any buffer variables, this SCM does not store any data associated with its variables, thus helping to optimise memory usage.

**Pedestrian** Inherits from the *Point Mass* SCM to represent a pedestrian agent, be it a human or robot. It is functionally speaking just a wrapper of the *Point Mass* SCM. Unlike the *FWD Car* SCM of the AV architecture, additional functionality was not required above the base *Point Mass* SCM at the time of the architecture's development.

**Goal Force Control Pedestrian** The SCM depicted in Figure 5.24 provides a control mechanism which outputs a force in return for providing a node-goal node identifier and target time as input. The node identifier corresponds to a vertex / position within a 2D spatial graph.

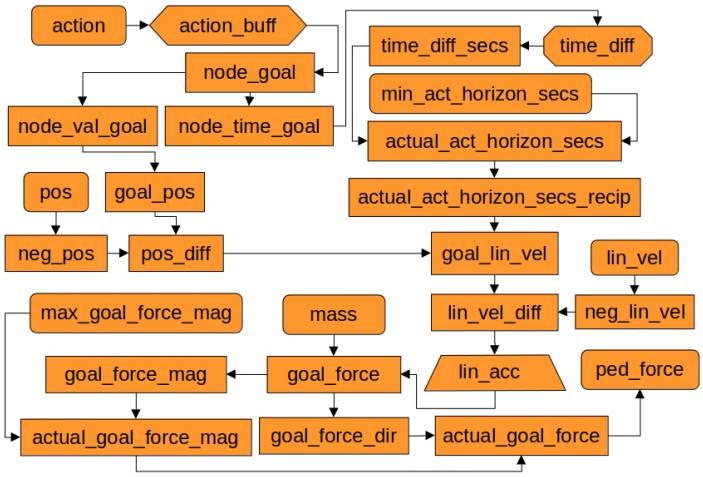



Figure 5.24: *Goal Force Control Pedestrian* SCM

The controller calculates a force magnitude proportional to the distance the node is from the position given by the *Pedestrian* SCM, and inversely proportional to the difference between the current time and the node-goal target time. This magnitude is limited in order to ensure the resulting force is not unrealistic for what a human / robot might be capable of. Meanwhile the force direction is calculated as the unit vector pointing from the current position to the node-goal node position. The combination of the force magnitude and direction provides the output for the SCM which can then be fed into the *Pedestrian* SCM.

Within the context of the example scenario this SCM is responsible for bridging the gap between the physical representation and action planner of the pedestrians in question. Here, there is utilisation of several SCM-formalism extensions:

- The *lin_acc* TSSQ variable allows the calculation of the required acceleration based upon a required linear-velocity difference and the size of a time step.

- The *time_diff* CTD variable allows the calculation of the time difference between the node-goal target time and the current time.

- Encapsulation of SCM data within the *action_buff* buffer variable.

- Utilisation of the *action* socket variable, allowing the fluid reconfiguration of which planner should be used at any given time.

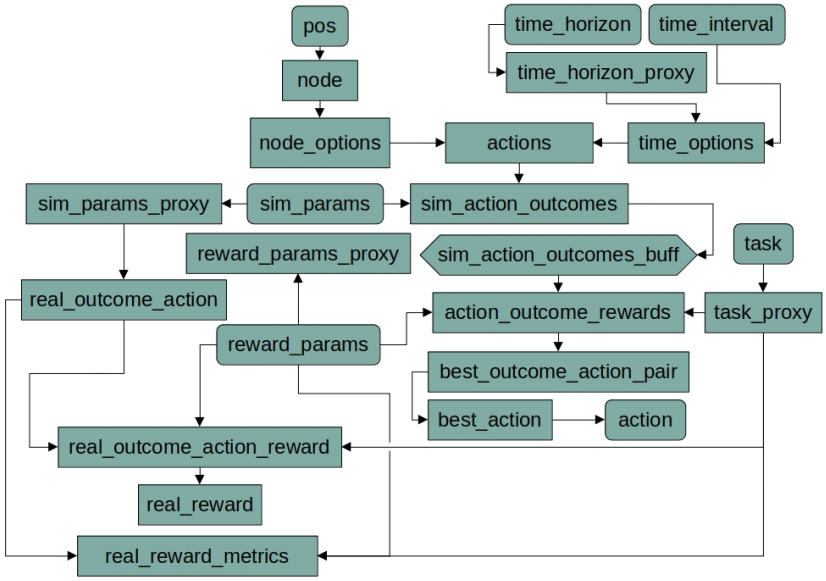



Figure 5.25: *Greedy Plan Pedestrian* SCM

- Utilisation of the *pos*, *lin_vel*, *mass*, and *env_force* socket variables. Unlike the *action* socket variable, these are used as inputs from the physical representation, which are utilised to determine how the driving-actuation variables of the physical representation — i.e. the *ped_force* variable — ought to be modulated.

**Greedy Plan Pedestrian** The SCM depicted in Figure 5.25 plans the next action for a *Goal Force Control Pedestrian* SCM specified in terms of goals. Mostly identical to the *Greedy Plan FWD Car* SCM from Section 5.3.1.2 except for the format of the actions / goals, and the nature of the predefined simulators and reward calculators used by the SCM. However, this SCM also relies upon a task socket variable that indicates a long-term task an agent is trying to achieve (e.g. reach a given node). Similar to the *Greedy Plan FWD Car* SCM once the maximum-reward action is selected it will be fed into an associated controller, in this case a *Goal Force Control Pedestrian* SCM.

Arguably the most important SCM besides the *Pedestrian* SCM in the context of the example scenario, this SCM captures the decision-making capabilities of pedestrian agents within the scene. In other words the decisions that are being identified as part of behavioural interactions are assumed to be made within a roughly comparable form of planner. In order to achieve this the following extensions are utilised:

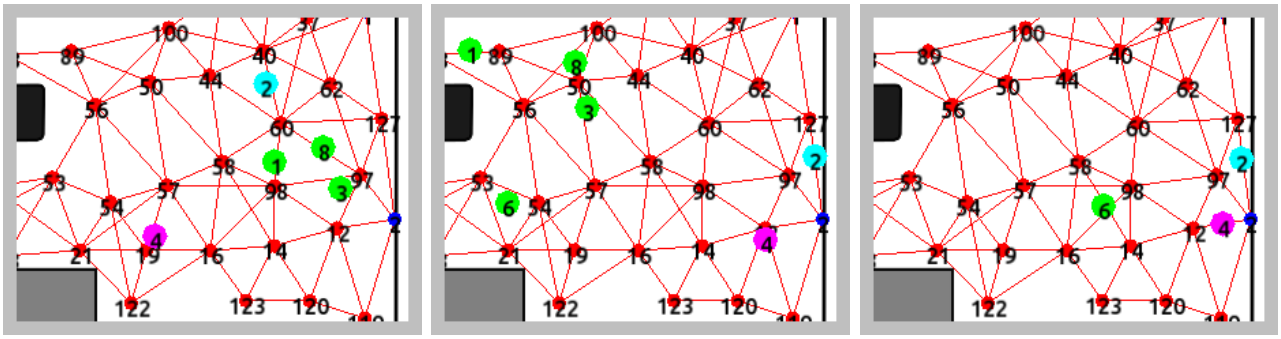
- Encapsulation of SCM data within the *sim_action_outcomes_buff* buffer variable.
- Utilisation of the *task* socket variable, allowing the fluid reconfiguration of the task which the pedestrian agent ought to try and achieve.
- Utilisation of the *pos* socket variable. This is used as an input from the physical representation, which is utilised to determine how an action ought to be selected. In particular the *pos* variable is used to identify candidate nodes for the pedestrian agent.

#### 5.3.2.3 Qualitative Experiment

The architecture and how the formalisms have been used to facilitate and enhance aspects of it has been described. Now one can reconsider the scenario depicted in Figure 1.2 and Figure 5.20.

**Methodology** The tracks from the pedestrian 'THÖR-MAGNI' dataset [172] are considered. Here a scene where two agents find themselves at an impasse is found, emulating the scenario described earlier in this section. This scene can then be analysed in the self-diagnostic manner mentioned before. Here this is approached by modulating alternative actions for each agent involved and inspecting how this affects the reward measures of each agent. These alternative actions are determined by replanning for the agents from the target time of their action prior to the impasse — i.e. when they reached the navigation node adjacent to the task-objective node. While a relatively simple method, it is sufficient to provide a demonstration of the service-robotics case-study SCM architecture.

**Discussion** The selected scene is depicted within Figure 5.26. Within this scene agent **p2** — depicted in cyan — and agent **p4** — depicted in magenta — are both attempting to reach node 2 as their task objective as part of a pick-and-place job (see Figure 5.26a). Here the numbers in these agent identifiers correspond those shown in the aforementioned figure. Initially **p2** approaches node 2 from the north via node 127 and **p4** approaches node 2 from the west via node 12. However, then they both hesitate for several seconds at an impasse as to which agent should proceed first (see Figure 5.26b). At last **p2** takes the initiative, moving to node 2 to complete its task and then moving away in order to allow **p4** to move to node 2 (see Figure 5.26c). Note that for the purposes of experiment reproduction, **p4** was originally accompanied by an agent **p5**, the two of them acting together to complete pick-and-place tasks. However, as



(a) Start of Scene (b) Beginning of Impasse (c) End of Scene

Figure 5.26: Visualisations of the service-robotics agent-impasse scene. Subfigures reflect the world state at various points in time throughout the scene. In the observed impasse scene the cyan and magenta agents are both attempting to reach node 2 — shown in dark blue — but as they both get close they stop and remain at an impasse for a number of seconds. By pre-empting the actions of each agent one can see how the subjective rewards of each agent would be affected, as discussed further in Figure 5.27.

this experiment was primarily to demonstrate the usage of the architecture, the scenario was simplified by omitting **p5**. Additionally, a brief tracking pause for **p2** around 206–207 *s* resulted in an abrupt change in its position, which is in turn reflected in its reward values.

The evolution of the reward values approximated for **p2** and **p4** is illustrated in Figure 5.27. As described in the methodology above, the scene analysis comprised of re-planning for each agent from an earlier time than their initial 'move to node 2' action. Given the task objective assigned to the agents of reaching node 2, this earlier re-planning just resulted in the agents moving to node 2 sooner, perhaps reflecting a more aggressive mode of behaviour. From there the overall rewards and individual reward metrics were approximated for each agent. As one might expect, in the cases where only a single agent pre-empts their original action to reach node 2, it results in a higher immediate reward for themselves while negatively impacting the reward of the other agent. This negative impact typically manifests through their space reward metric — i.e. the amount of personal space around each agent. On the other hand, both agents pre-empting their original action to reach node 2 results in a lower peak overall reward than most cases for all involved. Only an agent being the disadvantaged party in a single-agent-pre-emption case seems to result in a lower reward. Indeed this pattern of overall reward across each case reflects social norms that are intuitive to humans and also happens to match the pattern described in the prisoner's dilemma considered in game theory [104]. Of

course these findings are not unexpected, but nonetheless demonstrate a successful use of the architecture to make such inferences, which was the main purpose of this experiment in the context of this case study.

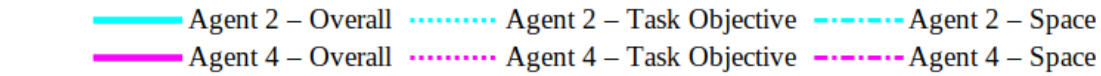
As a final note, another important factor which holds a large degree of sway over the results is the weighting attributed to the task-objective and space reward metrics. These weightings would not only influence the final overall-reward approximations for each agent, but also hold sway over which course of action they opted to pursue in cases where they were selected for pre-emption. For this experiment a weighting of 0.8 was given to the task-objective reward metric and a weighting of 0.2 was given to the space reward metric. This was done based upon the assumption that the agent's priority would still ultimately be upon getting their tasks done, while simply having a preference for extra personal space. In any case, this clearly indicates the importance of correctly approximating reward-metric weighting for agents, a topic that will be discussed in greater depth in Chapter 6.

### 5.3.3 Summary

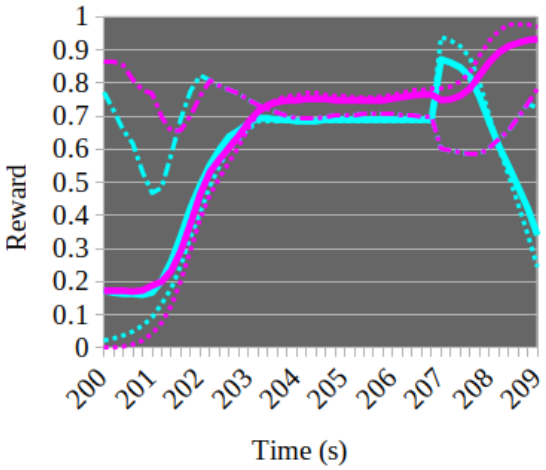
To summarise, the presented extensions represent a step forward for the expressive capabilities of SCMs within causality research as well as work integrating AESs with causality. In terms of how this ties in with the case studies discussed here, consider the following practical benefits:

- Utilisation of PTS-variable temporal representation combined with sparse data storage with buffer variables keeps memory storage for a large number of agents tractable.
- In a similar manner, when processing a large number of counterfactual queries using multiple different controllers and planners, socket variables make the reuse of SCM modules easier. Meanwhile time-conditional variables enable the reuse of data and easier management of multiple timelines.
- The RCS proof for mutable input sets, combined with socket variables and time-conditional variables, allows for a temporally-shifting number of elements to be represented in an SCM — in this case AES agents.

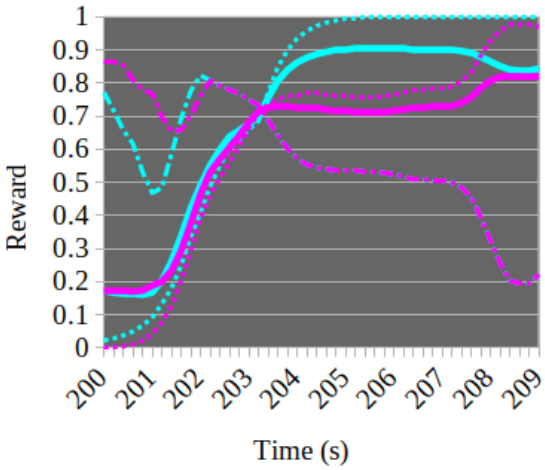
As such while each one of these contributions might offer a smaller contribution, their gestalt offers many new tools that can be utilised both within the AES space and beyond.



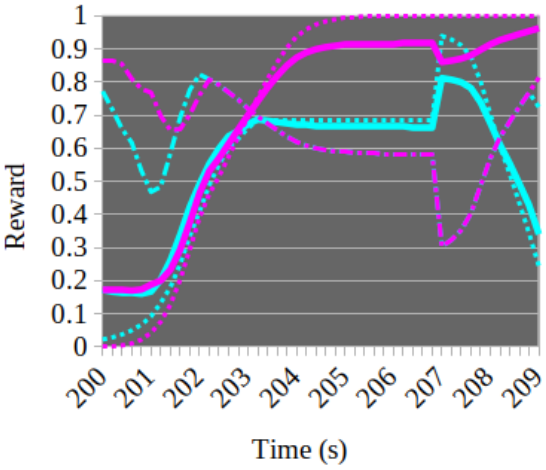

(a) Legend



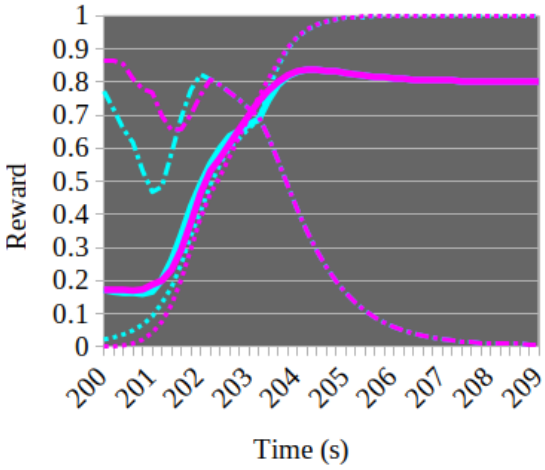

(b) Original




(c) Agent **p2** takes alternative action




(d) Agent **p4** takes alternative action




(e) Both agents take alternative actions

Figure 5.27: Illustrations of the overall rewards and reward metrics of **p2** and **p4** during the impasse scene depicted in Figure 5.26. Subfigures reflect how the evolution of said reward values differs depending upon whether or not the agents involved act pre-emptively. Based upon the data shown if an agent were to move to node 2 earlier it would result in a better subjective reward for that agent at the expense of the other agent. However if both agents attempt to move to node 2 earlier they both suffer in terms of their subjective reward. While this matches what one would reasonably expect given social norms, this nonetheless demonstrates the capacity of the case-study architecture to identify this autonomously.

# Chapter 6

# Accounting for Variations in Agent Motivating Factors

While Chapter 4 introduced the concept of utilising a theory of mind to identify agent behavioural interactions, Chapter 5 provided the tools required to unite this concept with SCM formalisms. In particular by integrating the planning process of agents within the causal model one can approximate the impact agent actions have upon the planning of other agents rather than relying upon a game-theoretic analysis of reward outcomes alone. Yet to effectively emulate the planning process of an autonomous agent one must estimate the motivations which direct their actions. In this chapter the approximation of the instantaneous motivations that govern the actions of autonomous agents is explored (Section 6.1). Furthermore, an integration of this motivation approximation as part of behavioural-interaction identification is proposed and then evaluated through a series of experiments (Section 6.2).

## 6.1 One-Shot Approximation of Instantaneous Agent Motivations

As alluded to above, in order to formulate an accurate theory of mind with which to reason about agent behavioural interactions one must understand the motivations that drive said agents to act. Furthermore, as such motivations are likely to vary across time one must approximate a parametrisation that describes these motivations for a specific point in time in order to accurately reflect the 'headspace' of the agent at that instant. In particular, whatever parametrisation is used to capture agent motivations must allow the facilitation of behaviours that may seem inoptimal or irrational, as the range of human actions an autonomous system is likely to encounter in the real-world will undoubtedly encompass such oddities.

### 6.1.1 Relation to Inverse Reinforcement Learning

Attempting to parametrise or learn a reward model which describes the motivations of autonomous agents is not a new concept and belongs to a set of techniques considered under the IRL umbrella. The general goal of such techniques is to take a series of observed states and actions and from this infer the persistent reward model that was being followed by the agents involved. This is typically associated with imitation-learning techniques that wish to utilise a reward model learnt from observing humans in order to imitate their behaviour. This work has similar aims, albeit mimicking the thought processes of observed agents rather than their behaviour.

However, rather than aiming to learn a persistent reward model, the goal here is to identify the instantaneous motivations behind singular actions. This follows previous work in IRL [97, 99] proposing a reward function varying across agents / situations. The idea behind this is two-fold: firstly it enables a more accurate imitation of an agent's planning process at a given instant, and secondly it facilitates the production of teleological explanations of observed behaviour. Because of the focus on instantaneous reward parametrisation the work done here is set apart from the majority of literature in the IRL field despite drawing inspiration from it.

### 6.1.2 Parametrising Reward for Autonomous-Vehicle Scenarios

The first step in formulating a reward parametrisation for a scenario in a given domain is deciding the overall structure of the reward model. For the types of AV scenarios that have been discussed throughout this thesis (see Section 3.1.3) it is argued here that a simple weight vector allows one to represent mixed motivations while remaining transparent and interpretable. This work is not alone in suggesting this with contemporary work applying causality to AVs by Gyevnar et al. [16] providing similar motivation behind their weight-vector representation.

With the overall structure of the reward model defined, one can formally establish its details. Calculations for reward are derived from an outcome $s_o = \langle s_{lt}, s_{sp}, s_{dh}, s_{ef}, s_{ad} \rangle$, which tracks the lane transitions ($s_{lt} \in \mathbb{Z}$), final speed ($s_{sp} \in \mathbb{R}$), distance headway ($s_{dh} \in \mathbb{R}$), and maximum environmental-force magnitude ($s_{ef} \in \mathbb{R}$) of a given vehicle. It also specifies whether the action in question had its goals accomplished ($s_{ad} \in \mathbb{B}$). From here reward is calculated as follows:

$$r^*(s_o) = \mathbf{r}(s_o)\boldsymbol{\gamma} \tag{6.1}$$

$$\mathbf{r}(s_o) = [\, r_0(s_o)\ r_1(s_o)\ r_2(s_o)\ r_3(s_o)\ r_4(s_o)\ 1\,] \tag{6.2}$$

where $\boldsymbol{\gamma} = [\, \gamma_0\ \gamma_1\ \gamma_2\ \gamma_3\ \gamma_4\ \gamma_5\,]^\intercal$ is a reward profile consisting of a weight vector. These weights are combined via dot product with a vector of reward metrics with a bias term $\mathbf{r}(\cdot)$ in order to calculate the overall reward associated with the outcome. The reward metrics capture a range of measures of utility that may be of varying levels of importance, and are defined as follows:

$$r_0(s_o) = \sigma(-s_{lt}) \tag{6.3}$$

$$r_1(s_o) = \min(\frac{s_{dh}}{\lambda_{dh} \cdot s_{sp}},\ 1) \tag{6.4}$$

$$r_2(s_o) = \exp(0.05(s_{sp} - \lambda_{sp})) \tag{6.5}$$

$$r_3(s_o) = \exp(-0.05 s_{sp}) \tag{6.6}$$

$$r_4(s_o) = \begin{cases} 1 & s_{ef} \leq \lambda_{ef} \\ 0 & \text{otherwise} \end{cases} \tag{6.7}$$

Here $r_0(\cdot)$ captures reward associated with lane transitions using the sigmoid function $\sigma(\cdot)$. Meanwhile $r_1(\cdot)$ represents the desire for a vehicle to maintain a safe distance with the vehicle in front. Here a parameter $\lambda_{dh} = 2\ s$ is used alongside the final speed to mirror the two-second rule [173] often recommended by driving authorities. Next $r_2(\cdot)$ and $r_3(\cdot)$ reflect the motivations of achieving a faster and slower final speed respectively, with $\lambda_{sp} = 31.3\ m/s$ acting as a speed-limit. Lastly, $r_4(\cdot)$ returns 1 if external forces imposed on the vehicle exceed $\lambda_{ef} = 1000\ N$ and 0 otherwise, effectively acting as a check for collisions.

### 6.1.3 Estimating Instantaneous Agent Reward Parametrisation

Since it is assumed here that the agents are effectively intentional systems that are acting to maximise some concept of reward, one can try to infer $\boldsymbol{\gamma}$ by considering the choice of action $a$ for a given agent. For a given action $a$, it is possible to consider a set of actions $\{\hat{a}_0, \hat{a}_1, ...\}$ that an agent could have alternatively taken at the time $a$ was executed. Using the generative

properties of the SCM architecture, one can simulate the outcomes associated with each action $\{\hat{s}_{o,0}, \hat{s}_{o,1}, ...\}$. From here these outcomes can be passed along with the observed outcome $s_o$ into the following distance function:

$$\zeta_o(s_o, {s_o}') = \alpha_o \Bigg( \alpha_{lt}(s_{lt} - {s_{lt}}')^2 + \alpha_{sp} \left( \frac{2(s_{sp} - {s_{sp}}')}{s_{sp} + {s_{sp}}'} \right)^2 + \alpha_{dh} \left( \frac{s_{dh}}{s_{sp}} - \frac{{s_{dh}}'}{{s_{sp}}'} \right)^2 + \alpha_{ef}(s_{ef} - {s_{ef}}')^2 + \alpha_{ad}(s_{ad} - {s_{ad}}')^2 \Bigg)^{\frac{1}{2}} \tag{6.8}$$

where $\alpha_o = 0.1$ is an overall distance scaling parameter, while $\alpha_{lt} = 100$, $\alpha_{sp} = 1$, $\alpha_{dh} = 0.1$, $\alpha_{ef} = 0.01$, and $\alpha_{ad} = 100$ are distance scaling parameters corresponding to the components of outcome tuples. This distance function was largely based off the Euclidean norm, albeit with the aforementioned scaling factors adjusting each of the constituent component distances. The distance of the $s_{lt}$, $s_{ef}$, and $s_{ad}$ components was just calculated as the difference between the components from the two outcomes. For the $s_{sp}$ component the difference of the speeds was divided by the mean of the speeds in order to make the resulting distance a relative distance. This was done so that similar speed differences are considered more significant when the speeds considered are smaller, e.g. $0\,m/s$ versus $10\,m/s$ should be considered more significant than $60\,m/s$ versus $70\,m/s$. For the $s_{dh}$ component, the distance headway of each outcome is divided by the associated speed in order to get the braking time for each outcome. From here the difference is taken between the braking times of each outcome before being fed into the Euclidean norm. This was done in order to capture that a greater distance headway is typically desired when travelling at greater speeds.

Through this comparison each hypothetical outcome can be assigned an overall reward based upon the negative exponent of the distance function. Along with (6.1), one can formulate this as a linear-regression task:

$$\begin{bmatrix} \mathbf{r}(\hat{s}_{o,0}) \\ \mathbf{r}(\hat{s}_{o,1}) \\ \vdots \end{bmatrix} \boldsymbol{\gamma} = \begin{bmatrix} \exp(-\zeta_o(s_o, \hat{s}_{o,0})) \\ \exp(-\zeta_o(s_o, \hat{s}_{o,1})) \\ \vdots \end{bmatrix} \tag{6.9}$$

From here the 'Householder rank-revealing QR[12] decomposition with column pivoting' approach implemented by Eigen [174] is utilised to provide a solution for $\boldsymbol{\gamma}$. The reward profile $\boldsymbol{\gamma}$ should offer insight into the motivations behind the agent choosing $a$, given that the hypothetical actions with outcomes closest to $s_o$ should have been assigned the highest rewards on the right-hand side of (6.9). Importantly, this should enable one to more accurately reason about how an agent's behaviour may have altered had events differed at the time of decision making.

## 6.2 Agent-Behavioural-Interaction Identification with Motivation Parametrisation

This chapter ultimately aims to leverage the learnt reward profiles along with the integration of planning within the SCM architecture in order to better explain agent behavioural interactions. Hence the objective is to establish explanations in the form of causal links between actions, where the occurrence of an action $a_{\mathcal{C},t}$ of agent $\mathcal{C}$ was necessary for agent $\mathcal{A}$ to select action $a_{\mathcal{A},t'}$.

### 6.2.1 Testing Causal Necessity Between Actions with Motivation Integration

Similar to the work in Chapter 4 the data associated with autonomous embodied systems — and in this case AVs — is by default predominantly continuous in nature. Thus similar to that chapter the approach detailed in Section 4.1.3 is utilised in order to convert this continuous data into discrete actions that describe the behaviour of the relevant agents. Once one has extracted the actions for vehicles using the aforementioned approach, one can iterate over pairs of actions and test for the presence of causal necessity. Here it is possible to utilise the property of temporal precedence (see Definition 3) to limit the number of pairs that need be considered.

To test for the presence of causal necessity the approach depicted in Figure 6.1 is followed in order to emulate the decision-making process of $\mathcal{A}$ via counterfactual inference. The first step in achieving this is to obtain a reward profile $\boldsymbol{\gamma}_{\mathcal{A}}$ for the time-indexed action $a_{\mathcal{A},t'}$ using the process detailed in Section 6.1.3. This allows one to emulate the planning process of $\mathcal{A}$ while utilising a similar conception of reward.

[12]QR is not an acronym here but refers to decomposing into matrices $Q$ and $R$ as part of the solving process.

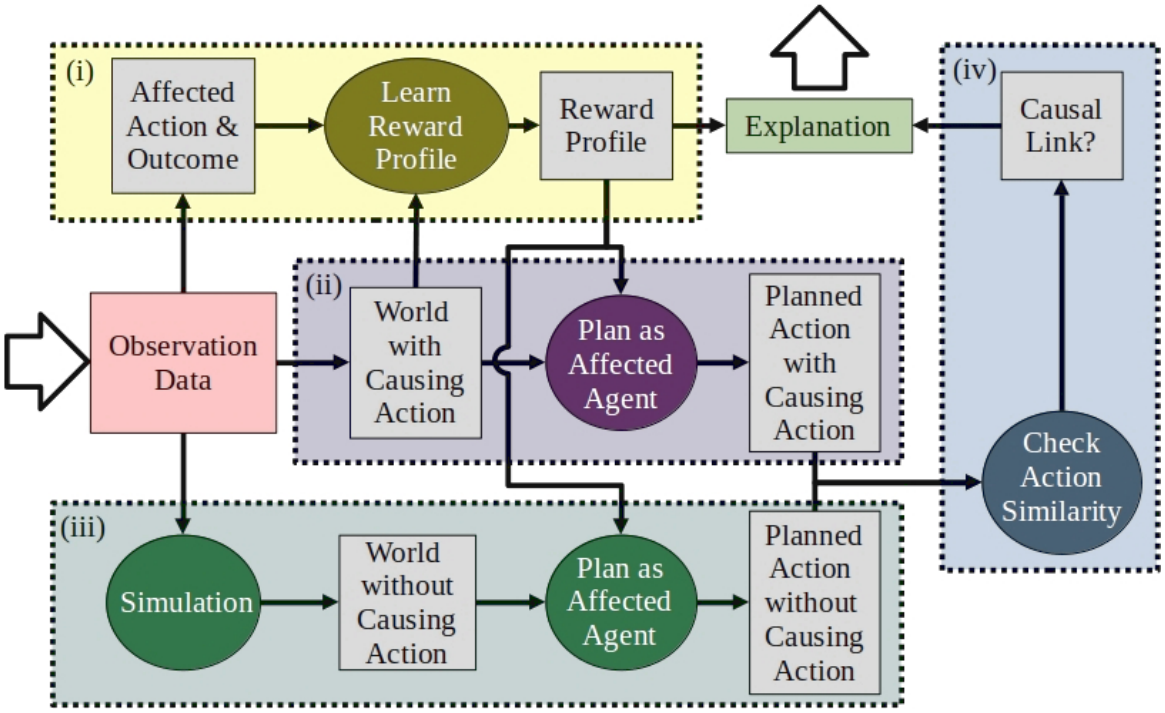



Figure 6.1: Depiction of the process by which a measure of causal necessity can be inferred between agent actions. Step (i) learns a reward profile for the action of a primary agent. Step (ii) plans for the primary agent under the observed world state. Meanwhile step (iii) simulates a world in which a secondary agent did not take a particular action, and plans under for the primary agent the resulting world state. Step (iv) compares these these two plans to determine if the secondary agent's action had a causal effect upon the primary agent's behaviour. The reward profiles and causal links combine to provide a causal explanation of a scene's behavioural interactions. (© 2025 IEEE)

From here, utilising the AV SCM architecture introduced in Section 5.3.1 one can plan for agent $\mathcal{A}$ at time $t'$ under two contrasting timelines. In the first case the approximated values of $\boldsymbol{\gamma}_{\mathcal{A}}$ are used to re-plan at time $t'$, providing a new action $\hat{a}_{\mathcal{A},t'}$. The second case also utilises $\boldsymbol{\gamma}_{\mathcal{A}}$ in re-planning yet intervenes upon the action variable $V_{\mathcal{C},A}$ in order to remove $a_{\mathcal{C},t}$ from the behaviour of agent $\mathcal{C}$. In terms of the counterfactual-context notation introduced in Section 4.2.1.3 one can thus define the resulting new output action as $\hat{a}^{\neg\mathcal{C}}_{\mathcal{A},t'} \equiv [\,\hat{a}_{\mathcal{A},t'}\,]^{\neg\mathcal{C}}$.

For each re-planning operation the planner SCM takes a range of possible actions $\{\hat{a}_0, \hat{a}_1, ...\}$ and simulates them by intervening upon the input action to the controller SCM specified by $V_{\mathcal{A},A}$. The SCM can then be used to forward propagate and generate outcomes $\{\hat{s}_{o,0}, \hat{s}_{o,1}, ...\}$ for a predefined simulation horizon $\beta_h$. From here one can apply (6.1) to the outcomes in order to derive the best actions $\hat{a}_{\mathcal{A},t'}$ and $\hat{a}^{\neg\mathcal{C}}_{\mathcal{A},t'}$ for the observed and counterfactual worlds respectively.

Finally in order to actually test for causal necessity between the actions Pearl and Mackenzie [2] discuss the concept of the 'but-for' test in relation to causal necessity, which in this case can be summarised as the following question:

Would $a_{\mathcal{A},t'}$ have not occurred, but for $a_{\mathcal{C},t}$ having occurred beforehand?

As discussed in Chapter 4 since it is impossible in many instances to know the exact planning process by which the action $a_{\mathcal{A},t'}$ was selected, answering this proves to be difficult. However, one can reformulate the question in order to approximate an answer:

Would $\hat{a}_{\mathcal{A},t'}$ have not occurred, but for $a_{\mathcal{C},t}$ having occurred beforehand?

Thus one can test for causal necessity against an approximated planning process for $a_{\mathcal{A},t'}$ represented through the planner SCM, even if replicating the exact planning process is not known. While there is indeed a risk of inaccurately reflecting the underlying planning process, the use of the motivation information captured in $\boldsymbol{\gamma}_{\mathcal{A}}$ helps to mitigate this.

In order to test for causal necessity here one can utilise the counterfactual action $\hat{a}^{\neg\mathcal{C}}_{\mathcal{A},t'}$ which reflects the planning of agent $\mathcal{A}$ in the absence of action $a_{\mathcal{C},t}$. Assuming $a_{\mathcal{C},t}$ was indeed necessary for $\hat{a}_{\mathcal{A},t'}$ to occur, one should expect to see a significant difference between $\hat{a}_{\mathcal{A},t'}$ and $\hat{a}^{\neg\mathcal{C}}_{\mathcal{A},t'}$. Hence, the two actions are passed into the following distance function:

$$\begin{aligned}\zeta_a(a, a') = \alpha_a \Bigg( \left( \frac{2(g_{sp,sp} - g_{sp,sp}{}')}{g_{sp,sp} + g_{sp,sp}{}'} \right)^2 + (g_{sp,t'} - g_{sp,t'}{}')^2 \\ + \alpha_{la} \left( 1 - \delta^K_{g_{la,la}, g_{la,la}{}'} \right) + (g_{la,t'} - g_{la,t'}{}')^2 \Bigg)^{\frac{1}{2}}\end{aligned} \tag{6.10}$$

where $\alpha_a = 0.1$ and $\alpha_{la} = 10$ are scaling parameters for the whole function and the lane-goal lane-identifier component distance respectively. Otherwise this function follows a similar pattern to (6.8) with the use of a Euclidean norm and scaling factors. Here the two target-time components each have their distance calculated by taking the difference in their value across actions. Meanwhile the distance between the speed-goal-value components of the actions is calculated similarly to the speed components in (6.8), i.e. based upon their relative difference. Lastly for the lane-goal lane-identifier components of the actions the distance is calculated by applying a Kronecker delta $\delta^K$ before scaling by $\alpha_{la}$.

Provided the output of (6.10) is greater than a predetermined threshold $\lambda_{\zeta a}$, one can say that the actions are sufficiently different. This indicates that $a_{\mathcal{C},t}$ was indeed necessary for the action $\hat{a}_{\mathcal{A},t'}$ to have occurred, as the action planned in the absence of $a_{\mathcal{C},t}$ — i.e. $\hat{a}^{\neg\mathcal{C}}_{\mathcal{A},t'}$ — is significantly different. Thus in terms of causal necessity, one can approximate $a_{\mathcal{C},t}$ as a cause of $a_{\mathcal{A},t'}$.

By determining that causality exists between two actions one can construct a behavioural-interaction graph with agent actions as the vertices and directional causal links between actions as edges that describe the causal influence that an agent taking one action had on an action taken by another agent. The resulting behavioural-interaction graph along with the reward profiles for the relevant agent actions describe how and why the scene unfolded as it did, both in terms of agent interactions and motivations.

### 6.2.2 Evaluation Experiments

Having now introduced a novel means of identifying behavioural interactions between agents while integrating motivation estimation it is now necessary to evaluate the method against the results of Sections 3.2.3 and 4.2.3.

Before proceeding it should be stated that these experiments do not attempt to compare themselves against the AV qualitative experiment discussed in Section 5.3.1.3. This aforementioned experiment was carried out to illustrate the functionality of the AV SCM architecture by considering a semi-synthetic collision scenario and attempting to identify levels of culpability for the agents involved. In contrast these experiments share more in common with the earlier main-content chapters in evaluating the performance of the proposed approach against real-world data. For this reason the topic of culpability identification is not revisited here.

#### 6.2.2.1 Data, Code, & Parameters

The experiments were carried out upon the same highD dataset [150] scenes as those documented in Section 3.2.3.1 for the same reasons as those given in Section 4.2.3.1. In addition to the highD dataset [150] utilised for the quantitative experiments, here the exiD [175] and inD [176] datasets are also used for qualitative experiments in order to capture a wider range of scenarios. These datasets are similarly formatted to the highD dataset, but instead of focusing on typical stretches of highway, they encompass highway on / off ramps and road intersections respectively.

This code along with any data that can be publicly shared is made available in a Git repository[13]. The method described in this chapter will henceforth be referred to as the Simulation-based Causal Analysis and Reasoning System Version 2 (SimCARSv2) framework to differentiate it from the SimCARSv1 approach proposed in Section 4.2.3.

For action-extraction parameters, the acceleration actuation threshold is $\lambda_{d,a} = 0.1\,m/s^2$; the speed-goal speed-difference threshold is $\lambda_{\delta a,sp,sp} = 1\,m/s$; the speed-goal duration threshold is $\lambda_{\delta a,sp,t} = 2.5\,s$; and the pre- and post-lane-change lane-goal start-time thresholds are $\lambda_{\delta a,la,t} = 2.5\,s$ and $\lambda'_{\delta a,la,t} = 0.5\,s$ respectively. For the identification of causal necessity, the action-distance conditional-optimality threshold $\lambda_{\zeta a}$ is varied across the set of values specified by $\{\ 0.1x \mid x \in [0, 10] \subset \mathbb{N}\ \}$. Once again, additional parameters are provided in Appendix A.

#### 6.2.2.2 Evaluation Metric

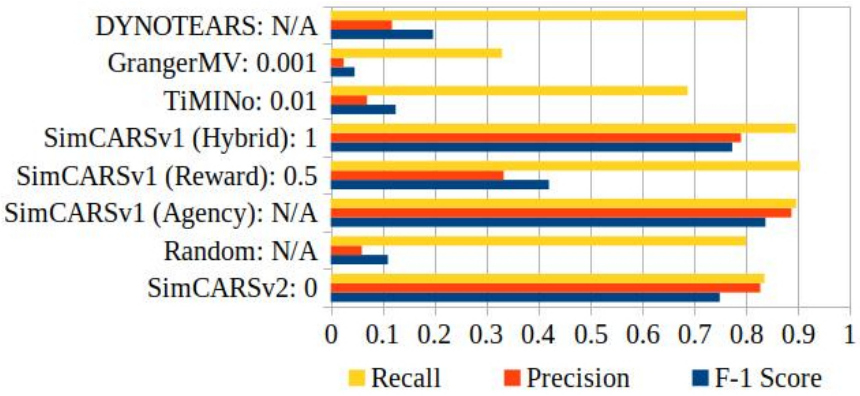
For this set of experiments the same evaluation metrics are used as those from the evaluation experiments of SimCARSv1. These metrics are initially documented in Section 3.2.3.2 and then altered slightly in Section 4.2.3.2.

#### 6.2.2.3 Results

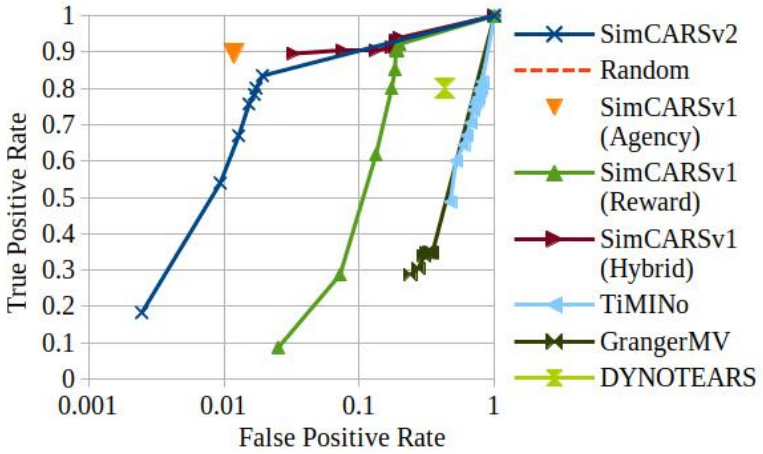
Here the results of the evaluation experiments are documented. The quantitative results are primarily focused upon comparison against SimCARSv1 and the benchmark experiments. Meanwhile the qualitative experiments offer greater insight than those conducted with SimCARSv1, largely due to the greater flexibility of the AV SCM architecture introduced in Chapter 5.

**Quantitative** In Table 6.1 the quantitative experiment results are shown, with Figure 6.2 providing an illustration of the results. Overall SimCARSv2 is highly competitive against existing methods, with only the agency-based variant of SimCARSv1 scoring higher than the proposed method in $F_1$ score. Importantly SimCARSv2 demonstrates a massive improvement over the SimCARSv1 reward-based variant in terms of precision, which is the most similarly designed method — both being reward based. This indicates that the linear-regression approach to learning reward-metric weightings is effective in capturing the priorities of autonomous agents.

[13] https://github.com/cognitive-robots/gce_vbai_lrp_paper_resources




(a) The precision, recall and $F_1$ score of evaluated methods. Numbers suffixed to the methods show the threshold associated with the maximum $F_1$ score depicted.




(b) Receiver Operating Characteristic (ROC) curve. Single points represent methods without sensitivity thresholds. Note the x-axis is logarithmic for the sake of clarity. Additionally, in contrast to the other metrics covered in this section, the TP-rate / recall and FP-rate / fall-out are calculated based upon the TP, TN, FP, and FN counts summed up across scenes, rather than the mean of the aforementioned metrics across scenes. The motivation is to illustrate the overall sensitivity versus specificity of the methods across threshold values, rather than the expected performance for each scene.

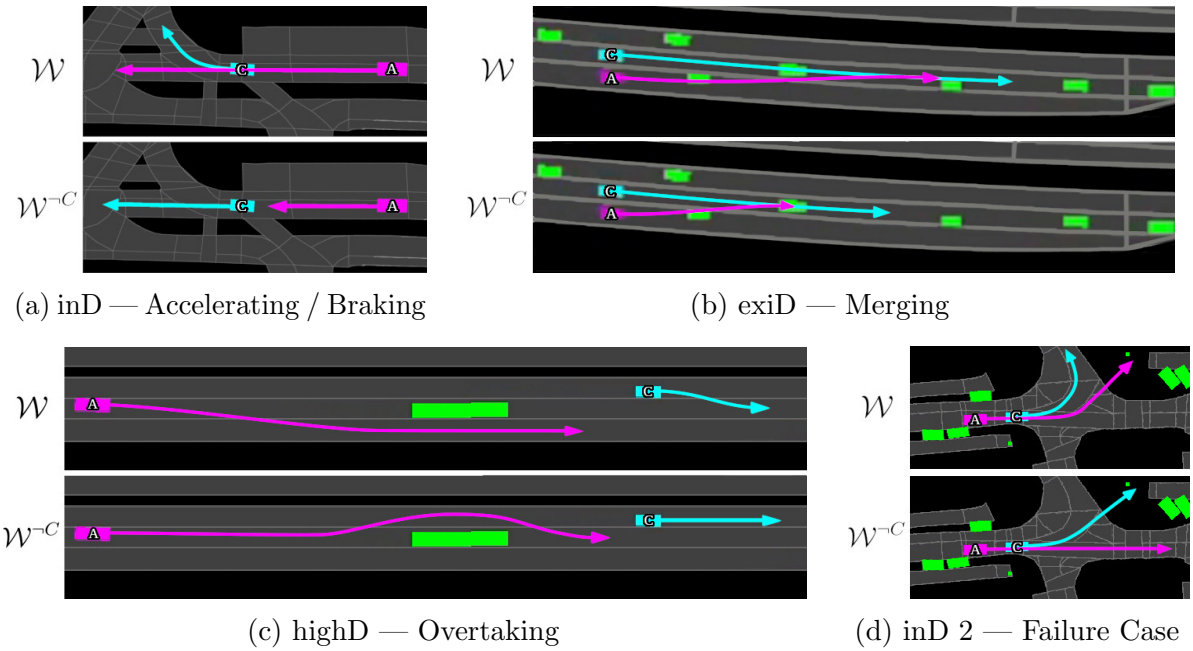
Figure 6.2: Illustrations of the quantitative results from the SimCARSv2 evaluation experiments. (© 2025 IEEE)

Overall the main factor that limits the performance of SimCARSv2 is its comparatively low sensitivity, given that it gives a maximal $F_1$ score and sensitivity of 0.749 and 0.835 respectively for a threshold $\lambda_{\zeta a} > 0$. Such a threshold effectively means any difference in action is sufficient for a causal link to test positive. A positive of this is that for higher thresholds SimCARSv2 is more precise than even the agency-based variant of SimCARSv1, which could make it useful in situations where precision is of greater importance. A potential cause of the comparatively lower sensitivity could be the reward metrics comprising the overall reward function. If the reward metrics present are not expressive enough to capture the motivations of a particular agent, then

| Method | Var. | $\lambda_\alpha$ / $\lambda_{\delta r}$ / $\lambda_{\zeta a}$ | $\tau$ (s) | Precision | Recall | $F_1$ Score |
|---|---|---|---|---|---|---|
| DYNOTEARS | sp. | N/A | 3.6 | 0.118±0.116 | 0.800±0.400 | 0.197±0.147 |
| MVGC | acc. | 0.03 | 2.5 | 0.025±0.039 | 0.330±0.470 | 0.046±0.071 |
| Random | N/A | N/A | N/A | 0.060±0.039 | 0.800±0.400 | 0.110±0.069 |
| SimCARSv1 (Reward-Based) | N/A | 0.5 | N/A | 0.333±0.300 | **0.904±0.295** | 0.420±0.283 |
| SimCARSv1 (Agency-Based) | N/A | N/A | N/A | **0.887±0.238** | 0.896±0.307 | **0.837±0.317** |
| SimCARSv1 (Hybrid) | N/A | 1 | N/A | 0.790±0.311 | 0.896±0.307 | 0.773±0.335 |
| SimCARSv2 | N/A | 0 | N/A | 0.827±0.284 | 0.835±0.373 | 0.749±0.371 |
| TiMINo | acc. | 0.05 | 3.6 | 0.070±0.062 | 0.687±0.464 | 0.125±0.104 |

Table 6.1: The precision, recall, and $F_1$ score of applying SimCARSv2 to the highD dataset. Additionally the contents of Table 4.1 are included, depicting the highD-dataset results for the DYNOTEARS, MVGC, and TiMINo methods as well as the SimCARSv1 variants, to act as baselines. Note that for methods besides SimCARSv1 and SimCARSv2, specificity increases as $\lambda_\alpha$ decreases. Meanwhile for the reward-based and hybrid SimCARSv1 variants, specificity increases as $\lambda_{\delta r}$ increases. Lastly, for SimCARSv2, specificity increases as $\lambda_{\zeta a}$ increases.

other metrics may end up being utilised as proxies when estimating the reward profile. If one then attempts to use this reward profile during counterfactual inference the agent cognition might deviate significantly from the original agent, in turn leading it to overlook certain causal relationships. Of course, the dependence of the method performance on a suitable set of reward metrics is indeed a limitation, as this may be hard to infer during system design. Another limiting factor is a lack of information with which to refine parameters. While SimCARSv2 offers an enhanced level of accuracy over SimCARSv1 in terms of its dynamics modelling, the ability to exploit this is limited by the type of meta-data available in the highD dataset. On a real-world deployment of SimCARSv2 one could configure the parameters of the vehicles based upon vehicle specifications, rather than relying upon rough approximations. Furthermore, the utilisation of an SCM architecture allows the use of distributions rather than fixed values for model inputs, something not possible with SimCARSv1. Lastly, since the agency-based variant of SimCARSv1 is based upon a subset of the metrics considered by SimCARSv2 it should hypothetically be possible to match or even surpass the performance of this baseline through further parameter tuning. This does however indicate that a potential weakness of the proposed approach is the number of parameters, and future work to simplify the means of outcome comparison could help alleviate these issues.




(a) inD — Accelerating / Braking

(b) exiD — Merging

(c) highD — Overtaking

(d) inD 2 — Failure Case

Figure 6.3: Illustration of twin-world analysis of driving scenes by SimCARSv2. Here $\mathcal{W}$ denotes the planned behaviour under the original world state at the time the affected action $a_{\mathcal{A},t'}$ was taken. Meanwhile $\mathcal{W}^{\neg\mathcal{C}}$ denotes the planned behaviour under the counterfactual world state in which the causing action $a_{\mathcal{C},t}$ was not taken, at the same time as before. Magenta indicates the affected agent, cyan the causing agent, and green any background agents. (© 2025 IEEE)

**Qualitative** Here four scenes in particular are selected for examination with the goal of exploring a variety of interaction types. These are depicted in Figure 6.3. Similar to the AV case-study experiments in Chapter 5 within these figures cyan is used to depict the causing agent $\mathcal{C}$, and magenta is used to depict the affected agent $\mathcal{A}$.

The first of these presents an intersection where the causing agent turns right while the affected agent continues on ahead, accelerating as it does (see Figure 6.3a). Whether the causing agent had turned right or not, the affected agent would have continued straight ahead. However, the proposed method suggests a behavioural interaction in that if the causing agent had continued straight ahead the affected agent would have instead slowed down — no longer having a clear path to accelerate ahead. The reward profile of this scenario (see Figure 6.4) demonstrates that agent $\mathcal{A}$ both wished to maintain a particular speed — determined via $r_2$ and $r_3$ — and maximise its distance headway — via $r_1$. This justifies the behaviour seen in both twin-worlds as agent $\mathcal{A}$ would have aimed to accelerate to its desired speed, but only if it did not decrease its distance headway.

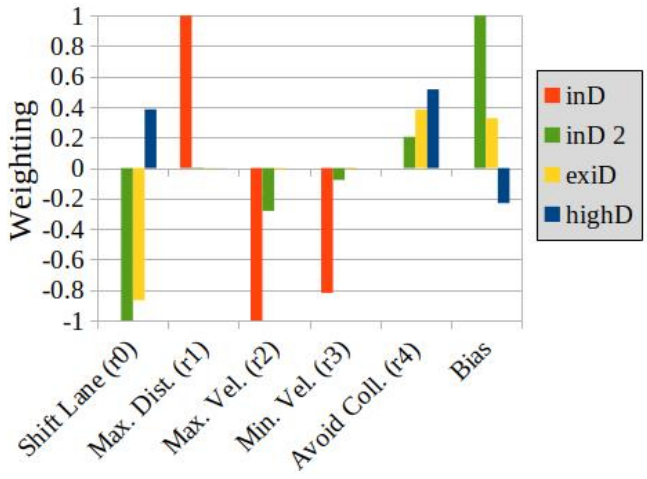



Figure 6.4: Scene reward profiles approximated by SimCARSv2. (© 2025 IEEE)

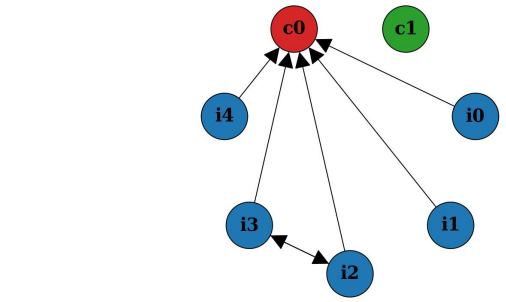
The second scene has the affected agent merging from an on-ramp into the lane occupied by the causing agent (see Figure 6.3b). However, in this scene the causing agent accelerates moments before the merge takes place. The proposed method is able to infer that without this acceleration taking place, agent $\mathcal{A}$ would have slowed down before merging, or else it would have risked a collision. Hence a behavioural interaction is present between the agents. The reward profile this time around (see Figure 6.4) indicates the affected agent wished to shift lane — via $r_0$ — while avoiding a collision — through $r_4$. Again this justifies the behaviour seen in each of the twin-worlds as agent $\mathcal{A}$ would have aimed to shift lane in either case, but must brake in the counterfactual case in order to not collide with agent $\mathcal{C}$.

The third scene depicts agent $\mathcal{A}$ overtaking another vehicle (see Figure 6.3c) shifting right to do so. The proposed methodology infers that this course of behaviour was the result of a lane change by agent $\mathcal{C}$. Otherwise the architecture reasons that agent $\mathcal{A}$ would have preferred to shift left before moving to the lane agent $\mathcal{C}$ would have originally moved to. The reward profile (see Figure 6.4) is similar to the previous except with the lane bias flipped, indicating that agent $\mathcal{A}$ deciding to shift right over left was primarily driven by a desire to avoid collision with agent $\mathcal{C}$ following its lane shift. This again makes sense, given that agent $\mathcal{A}$ in the observed case would want to avoid overtaking to prevent collision with agent $\mathcal{C}$, preferring to shift into a lane absent of vehicles. Meanwhile agent $\mathcal{A}$ in the counterfactual case could carry out an overtaking manoeuvre without risk of collision with agent $\mathcal{C}$, given that agent $\mathcal{C}$ remained in its original lane for this case. What is unusual is the decision of agent $\mathcal{A}$ to shift left rather

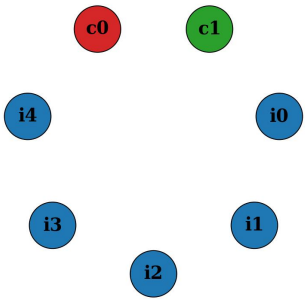
than right in the counterfactual case. This can perhaps be attributed to two underlying reasons, the first being that the reward model does not capture any legal considerations regarding left-hand-side or right-hand-side overtaking. The other reason could be linked to the fact that agent $\mathcal{A}$ still ultimately shifts right even after shifting left, which could throw reward metric $r_0$ off. Nonetheless this scene offers a good example of how critical a robust reward metric selection is in being able to properly ground any causal relationships that are identified via this process.

The final scene demonstrates a failure case for the proposed methodology (see Figure 6.3d). Here the two agents in question likely do have a behavioural interaction, which the system identifies. However, the planned actions via which it determines this are nonsensical, with one of the agents veering off the road in each case. This is a combination of two issues. The first being that the behaviour of agents when they reach the end of a lane segment is undefined. For example, if one intervenes to prevent an agent from turning left at an intersection, then does it turn right or go straight? The second issue is that whether or not the vehicle stays on the road is not part of the outcome or reward metrics. As such the agent will happily drive the vehicle off of the road provided it maximises its reward. Once again this highlights the importance of defining a diverse and comprehensive reward model that offers grounding to the identifications made by the proposed approach.

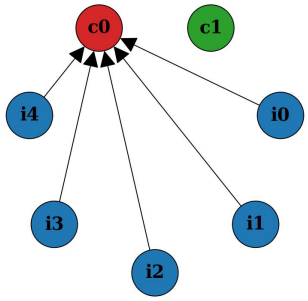
In addition to the aforementioned scenes selected for deeper examination, two sets of causal-summary graphs produced as part of the quantitative experiments are depicted in Figures 6.5 and 6.6. It should be mentioned that these cases were selected specifically because they demonstrate instances where SimCARSv2 outperforms SimCARSv1 as this can offer some interesting insights given that the agency-based variant of SimCARSv1 performed better than SimCARSv2 within the quantitative experiments. From these figures one can surmise that although SimCARSv2 may perform slightly worse than the agency-based variant of SimCARSv1 on average, in a situation where more subtle behavioural interactions occur SimCARSv2 may be able to offer greater insight.



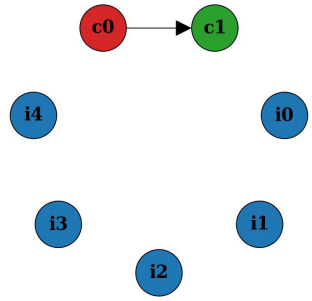

(a) SimCARSv1 (Reward-Based)



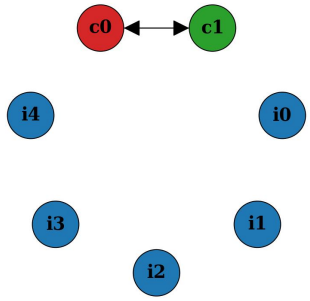

(b) SimCARSv1 (Agency-Based)




(c) SimCARSv1 (Hybrid)




(d) SimCARSv2

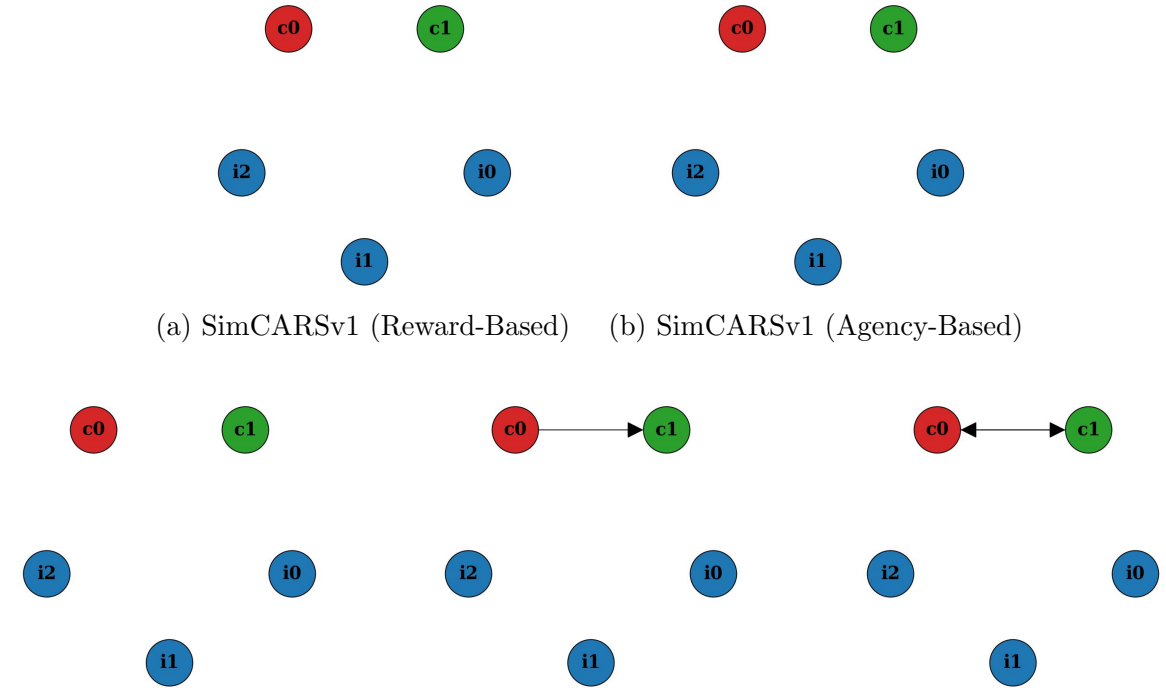



(e) Ground Truth

Figure 6.5: Depicts the causal-summary graphs from applying SimCARSv2 and SimCARSv1 to a scene from the highD dataset ($\lambda_{\zeta a} = 0$, $\lambda_{\delta r} = 0.5$), along with the ground truth (Figure 6.5e). SimCARSv2 performs well (Figure 6.5d), identifying only the true causal relationship between the convoy agents. By comparison the agency-based variant of SimCARSv1 (Figure 6.5b) did not identify any causal relationships, while the reward-based and hybrid variants (Figures 6.5a and 6.5c) identified several spurious causal relationships. The scene in question depicts a busy highway with the head convoy vehicle gradually speeding up throughout the scene with a similar response from the tail convoy vehicle. The subtle nature of the change in behaviour is perhaps the reason SimCARSv2 performed better than SimCARSv1 here. Meanwhile the presence of many vehicles increased the likelihood of discovering spurious causal relationships for the more sensitive variants of SimCARSv1.

#### 6.2.2.4 Discussion

Taking into account the results of the quantitative experiments it is necessary to discuss where SimCARSv2 stands in relation to SimCARSv1. When designing SimCARSv2 it was thought that the SCM architecture, along with the integration of planning and estimation of instantaneous agent motivations would lead to a quantitative increase in performance. A naturally added benefit would be that the estimated reward-metric weights could themselves provide some level of explanation regarding the potential thought processes of agents. However, SimCARSv2 failed to outperform the best performing variants of SimCARSv1, despite providing a significant improvement in performance over its reward-based variant. This is likely the result




(a) SimCARSv1 (Reward-Based) (b) SimCARSv1 (Agency-Based)

(c) SimCARSv1 (Hybrid) (d) SimCARSv2 (e) Ground Truth

Figure 6.6: Depicts the causal-summary graphs from applying SimCARSv2 and SimCARSv1 to a scene from the highD dataset ($\lambda_{\zeta a} = 0$, $\lambda_{\delta r} = 0.5$), along with the ground truth (Figure 6.5e). SimCARSv2 once again performs well (Figure 6.5d), identifying only the true causal relationship between the convoy agents. However, this time none of the SimCARSv1 variants (Figures 6.5a, 6.5b, and 6.5c) identified any causal relationships. Once again the scene is comprised of subtle changes to the speeds of the convoy vehicles across time, however in this instance the convoy vehicles initially decelerate before later speeding up again. Similar to Figure 6.5 the better performance of SimCARSv2 over SimCARSv1 can likely be attributed to the subtlety of the behavioural interactions. This scene was however slightly less busy and only consisted of two lanes per driving direction, which in turn might explain why the reward-based and hybrid variants of SimCARSv1 identified no spurious causal relationships for this case.

of the agency-based conditional-optimality metrics being of much greater simplicity than the composite reward models considered throughout SimCARSv1 and SimCARSv2. Given their increased complexity such composite reward models are that much harder to attune to the needs of the domain in question. Nonetheless, despite SimCARSv2 not outperforming SimCARSv1 it is not without its advancements, a summary of which is provided here:

- The reward-metric weights learnt by SimCARSv2 offer an insight into the motivations of agents. Such functionality was unavailable in SimCARSv1.
- At higher action-distance threshold values SimCARSv2 provides a lower FP-rate than even the agency-based variant of SimCARSv1.

- As indicated by Figures 6.5 and 6.6 there may be scenarios to which SimCARSv2 is better suited than SimCARSv1. At least in the cases analysed, this appears to be the result of SimCARSv2 being more sensitive to subtle behavioural interactions than SimCARSv1.

- The four scenes analysed in the qualitative experiments demonstrate that SimCARSv2 can use the planning component of the SCM architecture to analyse how the behaviour of agents might be altered in the absence of certain agent actions. Given that SimCARSv1 only functioned by toggling the presence of agent actions, it is more limited in its capacity to make counterfactual inferences. Put directly, SimCARSv2 intervenes upon agent actions and infers how other agents might react, while SimCARSv1 can only intervene upon agent actions and infer how this will affect the subjective rewards of agents. Thus it can be stated that SimCARSv2 offers a superset of the counterfactual-inference functionality offered by SimCARSv1.

Even having highlighted some of the benefits associated with SimCARSv2 above, it does have its limitations. While the issue of undefined behaviour following lane segments is raised, this could be solved on a real-world system where destination data is present. The bigger issue is the lack of expressiveness in the outcome and reward representations, which effectively permits any behaviour so long as it maximises reward. Assumptions were also made about control and planning, and where the dividing line ought to lie between them.

Thus the most sensible direction for future work would be to consider alternate means of modelling reward and instantaneously capturing this reward via IRL. A challenge is that any variable one wishes to consider as part of a reward model must be captured by the SCM architecture. For example, distance headway had to be explicitly implemented within the SCM to be used in the reward model. The task of deciding which information to include in a model is ultimately a design decision. However, one could potentially train some form of data-driven model (e.g. an NN) that can be used to sanity check the weight-vector-based reward model. Note that one likely ought to avoid replacing the weight-vector representation entirely as the primary reason for this reward model format was due to its inherent interpretability.

One could also explore integrating behavioural-interaction causal modelling into an RL loop, potentially allowing for greater efficiency and socially-awareness when operating around humans. This would differentiate itself from existing work merging causal reasoning with RL [14, 64] by its behavioural-interaction focus over typical egocentric perspectives. This is particularly relevant for domains involving a great deal of interaction with humans. Furthermore, given the safety-critical nature of the AV domain, it may be preferable to consider exploring this avenue of research in an alternate domain, such as service robotics. After all, long-term experimental deployments within service robotics [177] offer plenty of chances for opportunistic learning, not to mention a suitable environment for testing the socially-aware capabilities of methodologies.

In this work it has been demonstrated how information regarding the motivations of a vehicular agent can be incorporated into twin-world counterfactual inference in order to detect causal behavioural interactions. To do so it was also shown how a reward profile representing the instantaneous motivations of a vehicular agent can be approximated via simulation and linear regression. Quantitative experiments have verified that this approach is competitive against previous work, and significantly improves over the next-best reward-based model. Lastly several scenarios have been examined which illustrate the capabilities and limitations of the proposed approach via qualitative experiments.

# Chapter 7

# Conclusion

The work presented in this thesis comprises both a literature review concerning the utilisation of causality in modelling agent behavioural interactions as well as four chapters' worth of novel research and innovation within this field. Given the quantity of work and diverse contents of these aforementioned chapters this conclusion chapter serves to summarise the contributions of this thesis as a whole. Furthermore, the impact this work represents within literature will be discussed along with an overview of some of the limitations associated with the work as a whole. Finally, taking into consideration the status of the field as well as the contributions and limitations of the work a series of potential future avenues for research will be presented, providing a close to the contents of this thesis.

## 7.1 Contributions

This work began by providing a comprehensive literature review of the subject in question — modelling agent behavioural interactions via causal methodologies — which would in turn direct the contributions of the following four chapters of main content. Additionally, in the wake of carrying out several years worth of work, this literature review was revisited and revised to reflect recent developments in the field. However, the primary contributions of this work are undoubtedly contained within the aforementioned four main-content chapters, as described briefly in the following paragraphs.

**Chapter 3** This chapter introduced existing formalisms within causality literature and applied them to the task of modelling behavioural interactions between agents — AV agents in particular. Then in order to evaluate the efficacy of existing techniques in being able to discover or identify

these behavioural interactions a benchmark was carried out across ten contemporary temporal causal-discovery methods. The results indicated that these existing techniques were ill-suited to capturing these behavioural interactions given the causal sparsity and non-stationarity of inter-agent variable relationships.

**Chapter 4** Building upon the previous chapter, here an action-based causal-model format was proposed in order to facilitate the representation of behavioural interactions between agents despite their causal non-stationarity. Given the shift in focus from persistent variable-based causal relationships, to singular action-based causal relationships, an alternative to observation-based causal discovery was required. Therefore a novel game-theoretic approach was proposed that compares the counterfactual outcomes associated with simulations that toggle the presence of two potentially interacting observed actions. This follows from the idea that an agent's action is only caused by another action, if the first action were optimal for the agent only given that the second action occurred. The optimality of the outcomes were compared via three approaches: a product of reward metrics, a basic collision-based agency test, and a hybrid approach combining the two. The subsequent experimental results indicate all three approaches offered an improvement in performance over the contemporary temporal causal-discovery approaches, with the agency-based variant performing best, and the hybrid variant seemingly not benefiting from the combination of the other two variants. The reward-based variant, while not out-performing the agency-based variant nonetheless offered the potential for greater flexibility given it did not rely upon a single metric to measure optimality.

**Chapter 5** Overall the novel game-theoretic approach offered a notable improvement in the capacity to identify behavioural interactions between agents. However, a shortfall of the approach was the severing of true causal ancestry between agent actions during conditional optimality analysis. Additionally the approach utilised a black-box simulation, reducing the transparency of the overall pipeline. While the approach was certainly causally inspired, in order to truly model the chain of causal links that describe the causal relationship between two agent actions an SCM-based architecture would be preferable. In doing so one would both maximise transparency and allow for the utilisation of existing techniques within causality literature. However, given the focus of SCM applications is typically on big-data non-embodied

applications, there were several obstacles to engineering such a system. To remedy this, several novel theoretical formalisms were introduced to ease the integration of SCMs within autonomous embodied systems. This culminated in two case-study architectures, one addressing the AV domain which provides the real-world context for much of this thesis, and the other addressing the service-robotics domain in order to demonstrate the flexibility of the formalism extensions.

**Chapter 6** Here the idea of utilising an agent's conception of reward in identifying causality between agent actions was built upon, while utilising the SCM-based architectures developed as part of the previous chapter. In contrast to the earlier game-theoretic approach proposed, a new approach comprised of a comparison of twin-world counterfactuals with and without a potential cause is utilised. Due to the integration of planning within the SCM-based architecture it then became possible to compare the resulting planned actions to determine whether the potential cause did in fact influence the agent's behaviour at the time. However, in order to better capture the agent's motivations for the purposes of planning a form of one-shot IRL is utilised to approximate a reward-metric weight vector that reflects said motivations. In terms of quantitative results the new approach far outperformed the previously conceived reward-based methodology, making it competitive with the agency-based methodology overall, and slightly more precise. Furthermore, by integrating both planning and motivation estimation into the approach the expressiveness of the model was greatly improved as demonstrated via qualitative experiments. Overall this chapter represented the synthesis of all the lessons and contributions presented in the chapters that came before.

## 7.2 Impact

Having presented the contributions of this thesis it is now possible to discuss the ramifications of this work in the context of the research field and literature. The nature of the agent behavioural-interaction modelling tackled by this work means that the contributions made here are relevant to literature within autonomous systems as well as literature in causality. Hence the implications of the outcomes of this thesis are discussed in the context of each of these fields below.

**Autonomous Systems** At the beginning of this thesis the primary motivation behind this work was to enable autonomous agents — such as AVs — to better understand the type of behavioural interactions that occur between agents. In doing so, such systems can be developed with responsibility to better ensure safety and reliability in the presence of humans, animals, and other AI systems, not to mention potentially allowing the generation of explanations for observed behaviours / outcomes. The contributions documented above illustrate that this work goes some way to furthering progress towards these goals by enabling the identification of causal links between agent actions in an AV setting, as well as estimating the motivations behind those actions in terms of providing a reward profile. A positive side-effect of exploring this area of research within the AV domain is the raising of awareness of causality literature and its methodologies to researchers that may have otherwise not been exposed to the field previously.

**Causality** In contrast to the more general contributions to autonomous systems in terms of responsible development, the impact upon the field of causality is more subtle. Generally speaking, the work presented here offers significant opportunities for the field of causality to branch out beyond the typical focus on big-data applications. Specifically, the theoretical formalisms contributed in Chapter 5 present a bridge between the causal techniques rooted in data science and the type of object-oriented / modular programming seen in autonomous systems. However, the actual theoretical contributions themselves may have applications beyond just autonomous systems. For example, the proof regarding the fluid introduction and removal of set elements as part of an SCM is not limited to just autonomous systems. Therefore theoretically any modelled system which has elements coming and going throughout the modelled time span could benefit from the assurances provided by such a proof.

## 7.3 Limitations

Despite the utility provided by the contributions presented in this thesis, there are limitations that restrict the extent of their applicability for certain usages, and as a result identifying these limitations may help in guiding the next steps of research. Arguably the greatest limitation of these contributions is that while the identification of causal links between actions is well explored, the exploitation of this knowledge is not. Of course this information still provides

merit in the form of explanations for observed behaviour, but the lack of presented methodology with which to integrate the information into planning and control is something that future work should undoubtedly explore further.

Another limitation is that while the techniques developed here can identify behavioural interactions between agents via observation, it cannot do so via experimentation. This was partially by design, as the AV domain is a safety-critical field where an agent — i.e. a vehicle — cannot afford to experiment in order to identify causal links between agent actions. Nonetheless, other domains — as discussed below — might offer greater opportunity to carry out subtle experimentation as a secondary objective.

Next, while the modelling of agent motivations was a step in the right direction for understanding behavioural interactions, the mechanism for doing so remains quite simple. Part of this was intentional as the weight-vector format was selected for its inherent interpretability, however the approach of relying upon instantaneous motivations can prove unreliable, particularly if not all actual agent motivations are represented within the weight vector. This can lead to a false attribution towards a proxy reward metric or several completely random reward metrics that then lead to inaccurate simulations during counterfactual inference.

In terms of practical concerns rather than limitations of functionality, the proposed approaches do make assumptions regarding the planner and controller methodologies involved in the cognition of agents. This is an inevitability for any approach based upon a theory of mind that is working amidst humans. For the most part so long as the assumed methods are reasonably accurate and transparent — as the ones utilised here are — this is not of significant concern. This is because output data can be reliably used for the most part with the caveat that divergences in simulated agent behaviour may occur. However, in the case of the simplified planner — mainly the fact it only plans for a single action to be carried out immediately — this does limit the horizon which can be simulated while remaining accurate enough for counterfactual inference. Furthermore, in some instances it may be desirable to consider composite actions (e.g. over-taking within the AV domain) which would require the simulated planner to be capable of planning several low-level actions ahead of time.

Further to this previous point, the uncertainty inherent to some variables is not currently represented within the architectures introduced by this work. This was largely due to the focus of this thesis being upon high-level aspects such as theory of mind rather than low-level aspects such as sensing and actuation. However, previous work in robotics has been carried out on capturing and accounting for such uncertainties in variables within causality [11, 178]. Thus while integrating uncertainty into the methods and architectures presented here is definitely an important next step, this alone may not constitute significantly novel work as this area has already been well explored within causality.

Lastly, besides the service-robotics case study in Chapter 5 the proposed approaches have only been applied to the AV domain. While the approaches were developed in such a way that they ought to be applicable to any autonomous embodied system, it would be desirable to see this tested further in practice. Additionally, as mentioned above the AV domain has restrictions as a safety-critical domain which limit research regarding the active acquisition and exploitation of behavioural-interaction knowledge. Thus, application of the proposed methods in new domains is needed before one can make hard claims concerning their portability.

## 7.4 Revisiting the Research Questions

Before finally turning to the future of this research, a summary of the outcomes in relation to this thesis' RQs (see Section 1.2.3) is presented. This summary may overlap with some of the other discussions contained within this conclusion, however it is intended to provide a concise description of each RQ outcome for the reader's benefit. The RQ outcomes are as follows:

RQ-1. ***'Can contemporary causal-discovery approaches identify autonomous-agent behavioural interactions?'***

Generally speaking it is possible, but under the specific conditions presented by the AV scenario that was being considered, no. Arguably the biggest barrier presented by this scenario — and similar scenarios — is the relatively short window of time covered by the scene. This, combined with the fact that agent interactions are not continuously on-going and that the time-lags between cause and effect vary, makes it very difficult for data-driven causal-discovery methods to reliably identify behavioural-interactions.

RQ-2. ***'Can the nature of autonomous agents be exploited to overcome the challenges associated with identifying autonomous-agent behavioural interactions?'***

Certainly. Because one can make the assumption that autonomous agents typically act in their best interest — or at least to accomplish an objective of some kind — it is possible to exploit this to make counterfactual inferences regarding said agents. It was demonstrated that this could in turn be used to effectively identify behavioural interactions between agents. The main limitation here is the reliance upon building an expressive and accurate enough theory of mind for the agents under consideration.

RQ-3. ***'How can SCMs be extended in order to better represent a system of multiple interacting autonomous agents?'***

In a multitude of ways, however, this work focused upon three key areas. The first was to increase the modularity and encapsulation capabilities of SCMs such that they could be used closer to how classes are utilised within object-oriented programming. The second was to provide a new means of capturing temporal causal relations that does not rely upon roll-outs while remaining space efficient. Lastly a proof was given that demonstrated that it is possible to model a fluidly changing number of elements within an SCM while maintaining a form of causal stationarity. The biggest qualifier of these contributions is the similarity between the two case studies considered, as many of the extensions were proposed with these domains in mind. As such it would be valuable to see how these extensions fair within a completely different autonomous-agent domain with a different architecture design, in order to properly identify their limitations.

RQ-4. ***'How can autonomous-agent motivations be factored into the identification of behavioural interactions?'***

The straightforward answer would be via IRL, however when considering agent actions moment-by-moment one is faced with the ask of identifying an agent's motivations based upon a single instant. Within this thesis, this task was tackled by simulating a range of alternative actions for an agent at that instant, and then assigning an overall reward estimate based upon how similar the associated outcomes were to the observed agent outcome. From here it was possible to use linear regression to approximate a weight vector

over individual reward metrics (e.g. maximise distance-headway, maximise speed, etc.) which could then be used in counterfactual inference. This approach led to a significant increase in performance over the previously-proposed reward-based method, although it still faired similarly to other previously-proposed variants. The greatest limitation here shares similarities with RQ-2 in that the greatest weaknesses of the method were exposed by the limited expressiveness of the reward and outcome representations. This potentially does merit the use of a non-parametric alternative, however any such approach should be mindful of the need for transparency and interpretability within certain domains.

## 7.5 Future Work

While the previously-documented limitations undoubtedly offer some guidance as to the next steps one could take in researching the modelling of agent behavioural interactions via causality, this nonetheless leaves a lot of room for discussion of more specific avenues for inquiry. Below are suggestions of some such avenues.

In the limitations it was mentioned that although behavioural interactions are identified they are not necessarily exploited past being used as an explanatory tool. Given the emphasis in this work upon building systems that are better able to operate amongst other agents — particularly humans — it may seem wise to continue development here. This could be done by integrating a level of social awareness into agent planning based upon behavioural interactions and motivation approximation. Such a system would aim to learn social norms by observing humans and attempting to identify how agents select their actions, mindful of the subjective rewards of nearby agents. It could then try to apply this logic in similar circumstances, ideally providing a more comfortable experience for humans in the environment. Meanwhile it could also facilitate the opportunity for further learning via subtle experimentation in the course of applying the learnt social norms. In addition to overcoming the limitations of the techniques presented in this thesis, this would also offer the opportunity to explore other domains. While the AV domain could in theory be used, until the methodologies involved have sufficiently matured it may make more sense to trial the implementation of learnt social norms in a less safety-critical setting. Service robotics could potentially offer a good alternative here, as there

is a heavy emphasis upon the need for socially-acceptable behaviour while there is relatively little in the way of risk to humans in the surrounding environment.

Solving the issue of modelling agent reward, control, and planning provides a more difficult challenge. While it may be possible to mimic these processes for AVs, this does not necessarily reflect the cognition of a rational human driver. Although techniques such as IDMs exist, they remain fairly simplistic and are only useful in situations where pre-planned high-level routes are available. Furthermore IDMs also assume the agents one wishes to model do not attempt complex manoeuvrers (e.g. over-taking, merging, etc.) which limits the circumstances in which they can be applied. Given that relying upon fully-designed components alone can produce less natural behaviours and lead to unexpected edge cases it seems inevitable that some level of integrating human-driving data is required. While this may not be a new concept, a clear requirement if one is attempting to continue the course of responsible development of AI is ensuring transparency and the ability to ground the claims contained within system outputs. For example, if a planner built upon human-driving data suggests a particular course of action would be followed during counterfactual inference, it should be able to point to a similar case found in its training data in order to justify its output. This should allow one to provide more natural predictions of human behaviour while maintaining important system qualities for XAI.

Both of these proposed avenues if pursued will represent substantial contributions to the field in their own right, picking up from where this thesis leaves off and carrying the torch forwards. Through these efforts one can hope that ours is a future where AI and autonomous systems are not doomed to mindlessly navigate amongst humans while the latter suffers their social ineptitude. Instead, this work envisions a day whereby comprehending how AI and humans interact with each other, a better harmony and mutual understanding can be built.

# Appendix A

# Parameters

## A.1 Method Parameters for Experiments

### A.1.1 Random Baseline

- Likelihood of Assigning Causal Edge: 0.5

### A.1.2 CD-NOD

- Significance Alpha / $\lambda_\alpha$: { 0.001, 0.005, 0.01, 0.03, 0.05, 0.1 }
- Conditional-Independence Test: Kernel-Based Conditional Independence [179]
- Surrogate Variable: Time Index
- Maximum Number of Conditioning Variables: 1
- Type: 0 (A value of zero means all phases of CD-NOD will be ran)
- Pairwise: False
- Bonferroni Correction: False
- Conditional-Independence-Test Gaussian-Process Optimisation: True
- Direction-Determination Gaussian-Process Optimisation: True
- Observational Variable Kernel Width: Automatically calculated via Gaussian-process optimisation
- Time-Index Kernel Width: 0.1

### A.1.3 DYNOTEARS

- Maximum Time Lag / $\tau$: { 25, 36, 49 }
- Absolute-Edge-Weight Threshold: 0.01
- Regularisation Constant for $\mathbf{W}$ / $\beta_{\mathbf{W}}$: 0.05
  - This value is also used as the regularisation constant for a second linear coefficient array regarding variable auto-causation. However, since this work is interested in interactions between vehicles this second array is of little relevance. Nonetheless, it is important to make note of this for the purposes of experiment reproduction.
- Maximum Number of Iterations: 100
- Acyclicity Tolerance: $1.0 \cdot 10^{-8}$

### A.1.4 MVGC

- Significance Alpha / $\lambda_\alpha$: { 0.001, 0.005, 0.01, 0.03, 0.05, 0.1 }
- Maximum Time Lag / $\tau$: { 25, 36, 49 }
- Statistical Test: Chi-Squared
- Multiple-Hypothesis-Test Correction: Benjamini-Hochberg [180]
- Vector Auto-Regressive Model-Estimation Regression Mode: Ordinary Least Squares
- Information-Criteria Regression Mode: Locally-Weighted Regression [181]
- Model Order: Akaike Information Criterion [182]
- Maximum Auto-Covariance Lags: 1000
- Random Seed: Unseeded

### A.1.5 PWGC

- Significance Alpha / $\lambda_\alpha$: { 0.001, 0.005, 0.01, 0.03, 0.05, 0.1 }
- Maximum Time Lag / $\tau$: { 25, 36, 49 }
- Statistical Test: Sum of Squares Regression F-Test

### A.1.6 NAVAR

- Significance Alpha / $\lambda_\alpha$: { 0.001, 0.005, 0.01, 0.03, 0.05, 0.1 }
- Maximum Time Lag / $\tau$: { 25, 36, 49 }
- Hidden Nodes: 10
- Hidden Layers: 1
- Epochs: 2000
- Batch Size: 32
- Sparsity Penalty: 0.1
- Weight Decay: 0.001
- Dropout: 0.5
- Learning Rate: $3.0 \cdot 10^{-4}$
- Validation Proportion: 0.0
- Network Type: Multi-Layer Perceptron
- Normalize: True
- Split Time Series: False

### A.1.7 PCMCI

- Significance Alpha / $\lambda_\alpha$: { 0.001, 0.005, 0.01, 0.03, 0.05, 0.1 }
- Maximum Time Lag / $\tau$: { 25, 36, 49 }
- Minimum Time Lag: 0
- Conditional-Independence Test: Partial Correlation [143]
- Multiple-Hypothesis-Test Correction: None
- Maximum Number of Combinations of Conditions: 1
- Maximum Number of Conditions to Test: Unrestricted
- Maximum Number of Conditions of Y to Use: Unrestricted
- Maximum Number of Conditions of X to Use: Unrestricted

### A.1.8 TCDF

- Significance Alpha / $\lambda_\alpha$: { 0.001, 0.005, 0.01, 0.03, 0.05, 0.1 }
- Kernel Size & Dilation Coefficient: { 5, 6, 7 }
  - Although these are two separate parameters, while utilising a single hidden layer and identical values for kernel size and dilation coefficient, the maximum time lag consists of the square of this shared value.
- Hidden Layers: 1
- Epochs: 1000
- Learning Rate: 0.01
- Optimizer: Adam [183]
- CUDA: False
- Random Seed: 1111

### A.1.9 TiMINo

- Significance Alpha / $\lambda_\alpha$: { 0.001, 0.005, 0.01, 0.03, 0.05, 0.1 }
- Maximum Time Lag / $\tau$: { 25, 36, 49 }
- Assumed Time-Series Model: Linear
- Independence Test: Cross Covariance
- Include Instant Effects: False
- Check for Confounders: False

### A.1.10 tsFCI

- Significance Alpha / $\lambda_\alpha$: { 0.001, 0.005, 0.01, 0.03, 0.05, 0.1 }
- Maximum Time Lag / $\tau$: { 25, 36, 49 }
- Include Instant Effects: False
- Data Type: Continuous
- Algorithm: Time-Series Conservative Fast Causal Inference [49, 184]
  - This is a variation of the tsFCI algorithm [49] that incorporates the relaxed faithfulness assumptions applied to the PC approach in the work of Ramsey et al. [184].

### A.1.11 VAR-LiNGAM

- Significance Alpha / $\lambda_{\alpha}$: { 0.001, 0.005, 0.01, 0.03, 0.05, 0.1 }
- Maximum Time Lag / $\tau$: { 25, 36, 49 }
- Vector Auto-Regressive Model-Estimation Regression Mode: Ordinary Least Squares
- Vector Auto-Regressive Model-Trend Assumption: No Trend
- Pruning: True
- Regularisation Criterion: Bayesian Information Criterion [146]
- Algorithm: DirectLiNGAM [141]
- Random Seed: Unseeded

### A.1.12 SimCARSv1

- Reward-Based Conditional-Optimality Threshold / $\lambda_{\delta r}$: { 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0 }
- Action-Extraction Parameters:
  - Speed-Goal Actuation Acceleration Threshold / $\lambda_{d,a}$: 0.2 $m/s^2$
  - Speed-Goal Speed-Difference Threshold / $\lambda_{\delta a,sp,sp}$: 1.0 $m/s$
  - Goal Duration Threshold: 1.0 $s$
  - Lane Minimum Duration: 2.5 $s$
- Controller Parameters:
  - Steering Look-Ahead Steps: 10
  - Maximum Actuation Acceleration: 3.5 $m/s^2$
  - Minimum Actuation Acceleration: $-6.56$ $m/s^2$

### A.1.13 SimCARSv2

#### A.1.13.1 Autonomous Vehicles

- Action-Distance Conditional-Optimality Threshold / $\lambda_{\zeta a}$: { 0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0 }
- Action- and Outcome-Distance Scale Factors:
  - Overall Outcome-Distance Scale Factor / $\alpha_o$: 0.1
  - Lane-Transition Outcome-Distance Scale Factor / $\alpha_{lt}$: 100
  - Speed Outcome-Distance Scale Factor / $\alpha_{sp}$: 1
  - Distance-Headway Outcome-Distance Scale Factor / $\alpha_{dh}$: 0.1
  - Maximum-Environment-Force-Magnitude Outcome-Distance Scale Factor / $\alpha_{ef}$: 0.01
  - Action-Done Outcome-Distance Scale Factor / $\alpha_{ad}$: 100
  - Overall Action-Distance Scale Factor / $\alpha_a$: 0.1
  - Lane-Goal-Value Action-Distance Scale Factor / $\alpha_{la}$: 10
- Action-Extraction Parameters:
  - Speed-Goal Actuation Acceleration Threshold / $\lambda_{d,a}$: 0.1 $m/s^2$
  - Speed-Goal Duration Threshold / $\lambda_{\delta a,sp,t}$: 2.5 $s$
  - Speed-Goal Speed-Difference Threshold / $\lambda_{\delta a,sp,sp}$: 1.0 $m/s$
  - Pre-Lane-Change Lane-Goal-Start-Time Threshold / $\lambda_{\delta a,la,t}$: 2.5 $s$
  - Post-Lane-Change Lane-Goal-Start-Time Threshold / $\lambda'_{\delta a,la,t}$: 0.5 $s$
  - Time-Step Size: 0.5 $s$

- Vehicle & Controller Parameters:
  - Cars / Trucks / Other (Based off 'Toyota Ascent Sport (Hybrid), 1.8L' Datasheet)
    - Width: Determined from Data
    - Length: Determined from Data
    - Approximate Height: $(0.806 \cdot \text{Width})\ m$
    - Approximate Mass: $(116 \cdot \text{Width} \cdot \text{Length} \cdot \text{Height})\ kg$
    - Approximate Axle Distance: $(0.292 \cdot \text{Length})\ m$
    - Approximate Wheel Radius: $0.19\ m$
  - Motorcycles / Bicycles (Based off 'Honda Super Cub C125 2022' Datasheet)
    - Width: $0.72\ m$
    - Length: $1.915\ m$
    - Approximate Height: $1\ m$
    - Approximate Mass: $110\ kg$
    - Approximate Axle Distance: $0.6225\ m$
    - Approximate Wheel Radius: $0.17\ m$
  - Approximate Drag Area: $0.631\ m^2$
  - Cornering Stiffness: 49675
  - Maximum Motor Torque: $3260\ Nm$
  - Minimum Motor Torque: $-3260\ Nm$
  - Maximum Absolute Steer Angle: $0.616\ rad$
  - Maximum Slip Angle: $0.5\pi\ rad$

- Planner Parameters:
  - Action-Done Speed Threshold: 1 $m/s$
  - Speed Limit / $\lambda_{sp}$: 31.3 $m/s$
  - Distance-Headway Braking-Time Limit / $\lambda_{dh}$: 2 $s$
  - Maximum-Environment-Force-Magnitude Threshold / $\lambda_{ef}$: 1000 $N$
  - Minimum Speed-Goal Value: 0 $m/s$
  - Maximum Speed-Goal Value: 45 $m/s$
  - Speed-Goal Value Interval: 2.5 $m/s$
  - Maximum Goal Time Horizon: 5 $s$
  - Goal Time Interval: 2.5 $s$

#### A.1.13.2 Service Robotics

- Action-Extraction Parameters:
  - Pre-Node-Change Node-Goal-Start-Time Threshold: 0.2 $s$
  - Post-Node-Change Node-Goal-Start-Time Threshold: 0.2 $s$
  - Time-Step Size: 0.2 $s$
- Pedestrian & Controller Parameters:
  - Human
    - Approximate Mass: 70 $kg$
  - Maximum Node-Goal Force Magnitude: 200 $N$
- Planner Parameters:
  - Action-Done Node-Distance Threshold: 0.375 $m$
  - Maximum Goal Time Horizon: 5 $s$
  - Goal Time Interval: 2.5 $s$
  - Task-Objective Reward-Metric Weight: 0.8
  - Space Reward-Metric Weight: 0.2

## A.2 Synthetic-Dataset-Generation Parameters

- Variable: { Acceleration, Speed }
  - Each of these parameter settings was used once independently to generate 100 acceleration-based causal scenes and 100 speed-based causal scenes.
- Frequency: 10.0 $Hz$
- Duration: 50.0–70.0 $s$
- Convoy Actions: 12
- Independent Actions: 12
- Minimum Convoy Distance: 10.0 $m$
- Maximum Convoy Distance: 100.0 $m$
- Proportional-Feedback Coefficient: 1.0
- Integral-Feedback Coefficient: 0.0
- Differential-Feedback Coefficient: 0.0
- Minimum Action Interval: 1.0 $s$
- Minimum Speed: 0.0 $m/s$
- Maximum Speed: 44.7 $m/s$
- Minimum Start Speed: 4.47 $m/s$
- Maximum Start Speed: 26.8 $m/s$
- Minimum Acceleration: $-6.56$ $m/s^2$
- Maximum Acceleration: 3.5 $m/s^2$
- Braking-Time Limit: 2.24 $s$

- Reaction Time 0.5 $s$
- Fixed Actuary Noise: 0.1–1.6 $m/s^2$
- Proportional Actuary Noise: 0.1–1.6
- Fixed Sensory Noise: 0.01–0.16 $m$
- Proportional Sensory Noise: 0.005–0.08

# Appendix B

# Experimental Setup

## B.1 Hardware

- Motherboard: X570 AORUS PRO
- CPU: AMD Ryzen 9 3950X 16-Core Processor (3500 $MHz$)
- RAM: 32 $Mb$
- GPU: Nvidia GeForce RTX 2070 SUPER
- Storage: Samsung Electronics 970 EVO Plus NVMe M.2 Internal SSD

## B.2 Software

- Kernel: Linux Version 5.15.0-136-generic
- Operating System: Ubuntu 22.04.5 LTS
- C++ Standard: 20
- CMake: 3.22.1
- GCC/G++: 11.4.0
- Eigen: 3.4.0
- RapidJSON: 1.1.0
- Qt: 5.15.3

- SFML: 2.5.1
- Lanelet (ROS): 1.2.2
- Python: 3.12.4
  - numpy: 2.1.2
  - pandas: 2.2.3
  - scikit-learn: 1.6.1
  - scipy: 1.15.2
  - statsmodels: 0.14.4
  - joblib: 1.4.2
  - graphviz: 0.20.3
  - networkx: 3.4.2
  - matplotlib: 3.9.2
  - torch: 2.5.0
  - l5kit: 1.5.0 (Project Discontinued)
  - lingam: 1.9.0
  - numba: 0.61.2
  - causalnex: 0.12.1 (Not Updated for Python 3.11 and Above)
- R: 4.2.1
- Matlab: R2022a Update 4, 9.12.0.2009381

# Appendix C

# Additional Acknowledgements

The top-down car images utilised in Figures 1.4, 4.1, 4.4, and 5.7 were designed by Freepik. The low-polygon car meshes used in Figure 1.1 were made by Raphael Gonçalves.

# Bibliography


[1] American Law Institute. 'Model Penal Code'. In: Philadelphia, PA, USA: ALI, 1962. Chap. 2.03, p. 35. URL: https://archive.org/details/ModelPenalCode_ALI.

[2] Judea Pearl and Dana Mackenzie. *The Book of Why: The New Science of Cause and Effect*. 1st. USA: Basic Books, Inc., 2018. ISBN: 046509760X.

[3] Joseph Y Halpern. *Actual Causality*. The MIT Press, 2016. ISBN: 9780262336611. DOI: 10.7551/mitpress/10809.001.0001.

[4] Judea Pearl. 'Causality'. In: Cambridge university press, 2009. Chap. 7.

[5] Tobias Gerstenberg. 'What would have happened? Counterfactuals, hypotheticals and causal judgements'. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 377.1866 (2022), p. 20210339. DOI: 10.1098/rstb.2021.0339.

[6] Wesley C Salmon. *Causality and explanation*. Oxford University Press, 1998.

[7] Severin Kacianka, Amjad Ibrahim, Alexander Pretschner, Alexander Trende and Andreas Lüdtke. 'Extending Causal Models from Machines into Humans'. In: *Electronic Proceedings in Theoretical Computer Science* 308 (Oct. 2019), pp. 17–31. DOI: 10.4204/eptcs.308.2.

[8] David Fernández Llorca, Vicky Charisi, Ronan Hamon, Ignacio Sánchez and Emilia Gómez. 'Liability regimes in the age of AI: a use-case driven analysis of the burden of proof'. In: *Journal of Artificial Intelligence Research* 76 (2023), pp. 613–644.

[9] Amina Adadi and Mohammed Berrada. 'Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)'. In: *IEEE Access* 6 (2018), pp. 52138–52160. DOI: 10.1109/ACCESS.2018.2870052.

[10] Maximilian Diehl and Karinne Ramirez-Amaro. 'Why Did I Fail? A Causal-Based Method to Find Explanations for Robot Failures'. In: *IEEE Robotics and Automation Letters* 7.4 (2022), pp. 8925–8932. DOI: 10.1109/LRA.2022.3188889.

[11] Ricardo Cannizzaro and Lars Kunze. 'CAR-DESPOT: Causally-Informed Online POMDP Planning for Robots in Confounded Environments'. In: *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2023, pp. 2018–2025. DOI: 10.1109/IROS55552.2023.10342223.

[12] Thomas Hellström. 'The relevance of causation in robotics: A review, categorization, and analysis'. In: *Paladyn, Journal of Behavioral Robotics* 12.1 (2021), pp. 238–255. DOI: 10.1515/pjbr-2021-0017.

[13] Matthew Gadd, Daniele de Martini, Letizia Marchegiani, Paul Newman and Lars Kunze. 'Sense–Assess–eXplain (SAX): Building Trust in Autonomous Vehicles in Challenging Real-World Driving Scenarios'. In: *2020 IEEE Intelligent Vehicles Symposium (IV)*. 2020, pp. 150–155. DOI: 10.1109/IV47402.2020.9304819.

[14] Elias Bareinboim, Andrew Forney and Judea Pearl. 'Bandits with Unobserved Confounders: A Causal Approach'. In: *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2015.

[15] Luca Castri, Sariah Mghames and Nicola Bellotto. 'From Continual Learning to Causal Discovery in Robotics'. In: *Proceedings of the First AAAI Bridge Program on Continual Causality*. Ed. by Martin Mundt, Keiland W Cooper, Devendra Singh Dhami, Adéle Ribeiro, James Seale Smith, Alexis Bellot and Tyler Hayes. Vol. 208. Proceedings of Machine Learning Research. PMLR, Feb. 2023, pp. 85–91. URL: https://proceedings.mlr.press/v208/castri23a.html.

[16] Balint Gyevnar, Cheng Wang, Christopher G Lucas, Shay B Cohen and Stefano V Albrecht. 'Causal Explanations for Sequential Decision-Making in Multi-Agent Systems'. In: *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems*. AAMAS '24. Auckland, New Zealand: International Foundation for Autonomous Agents and Multiagent Systems, 2024, pp. 771–779. ISBN: 9798400704864.

[17] Alan F T Winfield, Anouk van Maris, Pericle Salvini and Marina Jirotka. 'An Ethical Black Box for Social Robots: a draft Open Standard'. In: *Proceedings of the 7th International Conference on Robot Ethics and Standard (ICRES)*. 2022, pp. 99–110.

[18] Rhys Peter Matthew Howard and Lars Kunze. 'Evaluating Temporal Observation-Based Causal Discovery Techniques Applied to Road Driver Behaviour'. In: *Proceedings of the Second Conference on Causal Learning and Reasoning*. Ed. by Mihaela van der Schaar, Cheng Zhang and Dominik Janzing. Vol. 213. Proceedings of Machine Learning Research. PMLR, Apr. 2023, pp. 473–498. URL: https://proceedings.mlr.press/v213/howard23a.html.

[19] Rhys Peter Matthew Howard and Lars Kunze. 'Simulation-Based Counterfactual Causal Discovery on Real World Driver Behaviour'. In: *2023 IEEE Intelligent Vehicles Symposium (IV)*. 2023. DOI: 10.1109/IV55152.2023.10186705.

[20] Rhys Peter Matthew Howard and Lars Kunze. 'Extending Structural Causal Models for Autonomous Vehicles to Simplify Temporal System Construction & Enable Dynamic Interactions Between Agents'. In: *Proceedings of the Fourth Conference on Causal Learning and Reasoning*. Ed. by Biwei Huang and Mathias Drton. Vol. 275. Proceedings of Machine Learning Research. PMLR, May 2025, pp. 1477–1505. URL: https://proceedings.mlr.press/v275/howard25a.html.

[21] Rhys Peter Matthew Howard, Nick Hawes and Lars Kunze. 'Generating Causal Explanations of Vehicular Agent Behavioural Interactions with Learnt Reward Profiles'. In: *2025 IEEE International Conference on Robotics and Automation (ICRA)*. 2025, pp. 10416–10423. DOI: 10.1109/ICRA55743.2025.11127745.

[22] Richa Nahata, Daniel Omeiza, Rhys P M Howard and Lars Kunze. 'Assessing and Explaining Collision Risk in Dynamic Environments for Autonomous Driving Safety'. In: *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*. 2021, pp. 223–230. DOI: 10.1109/ITSC48978.2021.9564966.

[23] Enrik Maci, Rhys P M Howard and Lars Kunze. 'Generating and Explaining Corner Cases Using Learnt Probabilistic Lane Graphs'. In: *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*. 2023, pp. 4201–4208. DOI: 10.1109/ITSC57777.2023.10422229.

[24] Ricardo Cannizzaro, Rhys Peter Matthew Howard, Paulina Lewinska and Lars Kunze. 'Towards probabilistic causal discovery, inference & explanations for autonomous drones in mine surveying tasks'. In: *Proceedings of the 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) Workshop 'Causality for Robotics: Answering the Question of Why'*. 2023.

[25] Calum Imrie, Rhys P M Howard, Divya Thuremella, Nawshin Mannan Proma, Tejas Pandey, Paulina Lewinska, Ricardo Cannizzaro, Richard Hawkins, Colin Paterson, Lars Kunze and Victoria Hodge. 'Aloft: Self-Adaptive Drone Controller Testbed'. In: *Proceedings of the 19th International Symposium on Software Engineering for Adaptive and Self-Managing Systems*. SEAMS '24. Lisbon, AA, Portugal: Association for Computing Machinery, 2024, pp. 70–76. ISBN: 9798400705854. DOI: 10.1145/3643915.3644107.

[26] Alireza Rastegarpanah, Rhys Peter Matthew Howard and Rustam Stolkin. 'Tracking linear deformable objects using slicing method'. In: *Robotica* 40.4 (2022), pp. 1188–1206. DOI: 10.1017/S0263574721001065.

[27] Rhys Peter Matthew Howard, Sam Barrett and Lars Kunze. 'Don't Blindly Trust Your CNN: Towards Competency-Aware Object Detection by Evaluating Novelty in Open-Ended Environments'. In: *2021 IEEE International Conference on Robotics and Automation (ICRA)*. 2021, pp. 13286–13292. DOI: 10.1109/ICRA48506.2021.9562116.

[28] Jorge Ocón, Iulia Dragomir, Florian Cordes, Raúl Dominguez, Robert Marc, Vincent Bissonnette, Raphael Viards, Anne Claire Berthet, Giulio Reina, Angelo Ugenti, Andrew Coles, Amanda Coles, Adam Green, Rhys P M Howard and Lars Kunze. 'ADE: Enhancing Autonomy for Future Planetary Robotic Exploration'. In: *IAF Space Exploration Symposium 2021 at the 72nd International Astronautical Congress (IAC) 2021*. International Astronautical Federation (IAF). 2021.

[29] Judea Pearl. 'Bayesian networks: A model of self-activated memory for evidential reasoning'. In: *Proceedings of the 7th conference of the Cognitive Science Society, University of California, Irvine, CA, USA*. 1985, pp. 15–17.

[30] C W J Granger. 'Investigating Causal Relations by Econometric Models and Cross-spectral Methods'. In: *Econometrica* 37.3 (1969), pp. 424–438. ISSN: 00129682, 14680262. URL: http://www.jstor.org/stable/1912791.

[31] Gary Smith and Jay Cordes. 'Introduction: Surely You Jest'. In: *The Phantom Pattern Problem: The Mirage of Big Data*. Oxford University Press, Sept. 2020. ISBN: 9780198864165. DOI: 10.1093/oso/9780198864165.003.0001.

[32] Jirakom Sirisrisakulchai and Supanika Leurcharusmee. 'Revisiting Returns to Education in Thailand: Structural Causal Model Framework'. In: *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 31 (2023), pp. 223–241. DOI: 10.1142/S0218488523400135.

[33] Guido W Imbens. 'Potential Outcome and Directed Acyclic Graph Approaches to Causality: Relevance for Empirical Practice in Economics'. In: *Journal of Economic Literature* 58.4 (Dec. 2020), pp. 1129–1179. DOI: 10.1257/jel.20191597.

[34] Jacob C Reinhold, Aaron Carass and Jerry L Prince. 'A Structural Causal Model for MR Images of Multiple Sclerosis'. In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*. Ed. by Marleen de Bruijne, Philippe C Cattin, Stéphane Cotin, Nicolas Padoy, Stefanie Speidel, Yefeng Zheng and Caroline Essert. Springer International Publishing, 2021, pp. 782–792. ISBN: 978-3-030-87240-3.

[35] Karl G Jöreskog. 'A General Method for Estimating a Linear Structural Equation System'. In: *ETS Research Bulletin Series* 1970.2 (1970). DOI: 10.1002/j.2333-8504.1970.tb00783.x.

[36] Arthur S Goldberger. 'Econometric theory'. In: Wiley, 1964. Chap. 7.

[37] Hubert M Blalock. 'Controlling for background factors: spuriousness versus developmental sequences.' In: *Sociological Inquiry* 34.1 (1964).

[38] Peter Spirtes, Clark Glymour and Richard Scheines. *Causation, Prediction, and Search*. 2nd. The MIT Press, Dec. 2001. ISBN: 978-1-4612-7650-0. DOI: 10.1007/978-1-4612-2748-9.

[39] Shohei Shimizu, Patrik O Hoyer, Aapo Hyvärinen and Antti Kerminen. 'A Linear Non-Gaussian Acyclic Model for Causal Discovery'. In: *Journal of Machine Learning Research* 7.72 (2006), pp. 2003–2030. URL: http://jmlr.org/papers/v7/shimizu06a.html.

[40] Clark Glymour, Kun Zhang and Peter Spirtes. 'Review of Causal Discovery Methods Based on Graphical Models'. In: *Frontiers in Genetics* 10 (2019), p. 524. ISSN: 1664-8021. DOI: 10.3389/fgene.2019.00524.

[41] Murat Kocaoglu, Karthikeyan Shanmugam and Elias Bareinboim. 'Experimental Design for Learning Causal Graphs with Latent Variables'. In: *Proceedings of NeurIPS 2017*. Ed. by I Guyon, U V Luxburg, S Bengio, H Wallach, R Fergus, S Vishwanathan and R Garnett. Vol. 30. Long Beach, CA, USA: Curran Associates, Inc., 2017. URL: https://proceedings.neurips.cc/paper/2017/hash/291d43c696d8c3704cdbe0a72ade5f6c-Abstract.html.

[42] Robert Maier, Lisa Grabinger, David Urlhart and Jürgen Mottok. 'Causal Models to Support Scenario-Based Testing of ADAS'. In: *IEEE Transactions on Intelligent Transportation Systems* 25.2 (2024), pp. 1815–1831. DOI: 10.1109/TITS.2023.3317475.

[43] Jonas Peters, Dominik Janzing and Bernhard Schölkopf. 'Elements of causal inference: foundations and learning algorithms'. In: The MIT Press, 2017. Chap. 10.

[44] Charles K Assaad, Emilie Devijver and Eric Gaussier. 'Survey and Evaluation of Causal Discovery Methods for Time Series'. In: *Journal of Artificial Intelligence Research* 73 (2022), pp. 767–819.

[45] John Geweke. 'Measurement of Linear Dependence and Feedback between Multiple Time Series'. In: *Journal of the American Statistical Association* 77.378 (1982), pp. 304–313. DOI: 10.1080/01621459.1982.10477803.

[46] Meike Nauta, Doina Bucur and Christin Seifert. 'Causal Discovery with Attention-Based Convolutional Neural Networks'. In: *Machine Learning and Knowledge Extraction* 1.1 (2019), pp. 312–340. ISSN: 2504-4990. DOI: 10.3390/make1010019.

[47] Jakob Runge, Peer Nowack, Marlene Kretschmer, Seth Flaxman and Dino Sejdinovic. 'Detecting and quantifying causal associations in large nonlinear time series datasets'. In: *Science Advances* 5.11 (2019). DOI: 10.1126/sciadv.aau4996.

[48] Jie Sun, Dane Taylor and Erik M Bollt. 'Causal Network Inference by Optimal Causation Entropy'. In: *SIAM Journal on Applied Dynamical Systems* 14.1 (2015), pp. 73–106. DOI: 10.1137/140956166.

[49] Doris Entner and Patrik O Hoyer. 'On causal discovery from time series data using FCI'. In: *5th European Workshop on Probabilistic Graphical Models*. Helsinki Institute for Information Technology HIIT. 2010, pp. 121–128.

[50] Aapo Hyvärinen, Kun Zhang, Shohei Shimizu and Patrik O Hoyer. 'Estimation of a Structural Vector Autoregression Model Using Non-Gaussianity'. In: *Journal of Machine Learning Research* 11.56 (2010), pp. 1709–1731. URL: http://jmlr.org/papers/v11/hyvarinen10a.html.

[51] Jonas Peters, Dominik Janzing and Bernhard Schölkopf. 'Causal Inference on Time Series using Restricted Structural Equation Models'. In: *Advances in Neural Information Processing Systems*. Ed. by C J C Burges, L Bottou, M Welling, Z Ghahramani and K Q Weinberger. Vol. 26. Curran Associates, Inc., 2013. URL: https://proceedings.neurips.cc/paper/2013/hash/47d1e990583c9c67424d369f3414728e-Abstract.html.

[52] Roxana Pamfil, Nisara Sriwattanaworachai, Shaan Desai, Philip Pilgerstorfer, Konstantinos Georgatzis, Paul Beaumont and Bryon Aragam. 'DYNOTEARS: Structure Learning from Time-Series Data'. In: *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*. Ed. by Silvia Chiappa and Roberto Calandra. Vol. 108. Proceedings of Machine Learning Research. PMLR, Aug. 2020, pp. 1595–1605. URL: https://proceedings.mlr.press/v108/pamfil20a.html.

[53] Bart Bussmann, Jannes Nys and Steven Latré. 'Neural Additive Vector Autoregression Models for Causal Discovery in Time Series'. In: *Discovery Science*. Ed. by Carlos Soares and Luis Torgo. Springer International Publishing, 2021, pp. 446–460. ISBN: 978-3-030-88942-5.

[54] Kun Zhang, Biwei Huang, Jiji Zhang, Clark Glymour and Bernhard Schölkopf. 'Causal discovery from nonstationary/heterogeneous data: Skeleton estimation and orientation determination'. In: *IJCAI: Proceedings of the Conference*. Vol. 2017. NIH Public Access. 2017, p. 1347.

[55] Biwei Huang, Kun Zhang, Mingming Gong and Clark Glymour. 'Causal Discovery and Forecasting in Nonstationary Environments with State-Space Models'. In: *Proceedings of the 36th International Conference on Machine Learning*. Ed. by Kamalika Chaudhuri and Ruslan Salakhutdinov. Vol. 97. Proceedings of Machine Learning Research. PMLR, June 2019, pp. 2901–2910. URL: https://proceedings.mlr.press/v97/huang19g.html.

[56] David Gunning and David Aha. 'DARPA's Explainable Artificial Intelligence (XAI) Program'. In: *AI Magazine* 40.2 (June 2019), pp. 44–58. DOI: 10.1609/aimag.v40i2.2850.

[57] Amir-Hossein Karimi, Gilles Barthe, Bernhard Schölkopf and Isabel Valera. 'A Survey of Algorithmic Recourse: Contrastive Explanations and Consequential Recommendations'. In: *ACM Comput. Surv.* 55.5 (Dec. 2022). ISSN: 0360-0300. DOI: 10.1145/3527848.

[58] Simón C Smith and Subramanian Ramamoorthy. 'Semi-supervised Learning From Demonstration Through Program Synthesis: An Inspection Robot Case Study'. In: *Electronic Proceedings in Theoretical Computer Science* 319 (July 2020), pp. 81–101. DOI: 10.4204/eptcs.319.7.

[59] Dominik Baumann, Friedrich Solowjow, Karl Henrik Johansson and Sebastian Trimpe. 'Identifying Causal Structure in Dynamical Systems'. In: *Transactions on Machine Learning Research* 2022.7 (2022). ISSN: 2835-8856. URL: https://openreview.net/forum?id=X2BodlyLvT.

[60] Paul K Rubenstein, Stephan Bongers, Bernhard Schoelkopf and Joris M Mooij. 'From Deterministic ODEs to Dynamic Structural Causal Models'. In: *Proceedings of the 34th Conference on Uncertainty in Artificial Intelligence (UAI)*. 2018, pp. 114–123.

[61] Joris M Mooij, Dominik Janzing and Bernhard Schölkopf. 'From ordinary differential equations to structural causal models: the deterministic case'. In: *Proceedings of the Twenty-Ninth Conference on Uncertainty in Artificial Intelligence*. UAI'13. Bellevue, WA: AUAI Press, 2013, pp. 440–448.

[62] Stephan Bongers, Patrick Forré, Jonas Peters and Joris M Mooij. 'Foundations of structural causal models with cycles and latent variables'. In: *The Annals of Statistics* 49.5 (2021), pp. 2885–2915. DOI: 10.1214/21-AOS2064.

[63] Alexander Wich, Holger Schultheis and Michael Beetz. 'Empirical Estimates on Hand Manipulation Are Recoverable: A Step Towards Individualized and Explainable Robotic Support in Everyday Activities'. In: *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems*. AAMAS '22. Virtual Event, New Zealand: International Foundation for Autonomous Agents and Multiagent Systems, 2022, pp. 1382–1390. ISBN: 9781450392136.

[64] Tom He, Jasmina Gajcin and Ivana Dusparic. 'Causal Counterfactuals for Improving the Robustness of Reinforcement Learning'. In: *AAMAS 2023 Workshop - Autonomous Robots and Multirobot Systems*. 2023.

[65] Zhongwei Yu, Jingqing Ruan and Dengpeng Xing. 'Explainable reinforcement learning via a causal world model'. In: *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*. Macao, P.R.China, 2023. ISBN: 978-1-956792-03-4. DOI: 10.24963/ijcai.2023/505.

[66] Tabitha E Lee, Jialiang Alan Zhao, Amrita S Sawhney, Siddharth Girdhar and Oliver Kroemer. 'Causal Reasoning in Simulation for Structure and Transfer Learning of Robot Manipulation Policies'. In: *2021 IEEE International Conference on Robotics and Automation (ICRA)*. 2021, pp. 4776–4782. DOI: 10.1109/ICRA48506.2021.9561439.

[67] Amjad Ibrahim, Severin Kacianka, Alexander Pretschner, Charles Hartsell and Gabor Karsai. 'Practical Causal Models for Cyber-Physical Systems'. In: *NASA Formal Methods*. Ed. by Julia M Badger and Kristin Yvonne Rozier. Springer International Publishing, 2019, pp. 211–227. ISBN: 978-3-030-20652-9.

[68] Maximilian Diehl and Karinne Ramirez-Amaro. 'A causal-based approach to explain, predict and prevent failures in robotic tasks'. In: *Robotics and Autonomous Systems* 162 (2023), p. 104376. ISSN: 0921-8890. DOI: 10.1016/j.robot.2023.104376.

[69] Matija Franklin, Hal Ashton, Edmond Awad and David Lagnado. 'Causal Framework of Artificial Autonomous Agent Responsibility'. In: *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*. AIES '22. Oxford, United Kingdom: Association for Computing Machinery, 2022, pp. 276–284. ISBN: 9781450392471. DOI: 10.1145/3514094.3534140.

[70] Stelios Triantafyllou and Goran Radanovic. 'Towards Computationally Efficient Responsibility Attribution in Decentralized Partially Observable MDPs'. In: *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems*. AAMAS '23. London, United Kingdom: International Foundation for Autonomous Agents and Multiagent Systems, 2023, pp. 131–139. ISBN: 9781450394321.

[71] Jakob Foerster, Gregory Farquhar, Triantafyllos Afouras, Nantas Nardelli and Shimon Whiteson. 'Counterfactual Multi-Agent Policy Gradients'. In: *Proceedings of the AAAI Conference on Artificial Intelligence* 32.1 (Apr. 2018). DOI: 10.1609/aaai.v32i1.11794.

[72] Samer B Nashed, Saaduddin Mahmud, Claudia V Goldman and Shlomo Zilberstein. 'Causal Explanations for Sequential Decision Making Under Uncertainty'. In: *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems*. AAMAS '23. London, United Kingdom: International Foundation for Autonomous Agents and Multiagent Systems, 2023, pp. 2307–2309. ISBN: 9781450394321.

[73] Kexuan Zhang, Qiyu Sun, Chaoqiang Zhao and Yang Tang. 'Causal reasoning in typical computer vision tasks'. In: *Science China Technological Sciences* 67.1 (2024), pp. 105–120.

[74] Hao Ding, Jintan Zhang, Peter Kazanzides, Jie Ying Wu and Mathias Unberath. 'CaRTS: Causality-Driven Robot Tool Segmentation from Vision and Kinematics Data'. In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022*. Ed. by Linwei Wang, Qi Dou, P Thomas Fletcher, Stefanie Speidel and Shuo Li. Springer Nature Switzerland, 2022, pp. 387–398. ISBN: 978-3-031-16449-1.

[75] Kaylene Caswell Stocking, Alison Gopnik and Claire Tomlin. 'From Robot Learning To Robot Understanding: Leveraging Causal Graphical Models For Robotics'. In: *Proceedings of the 5th Conference on Robot Learning*. Ed. by Aleksandra Faust, David Hsu and Gerhard Neumann. Vol. 164. Proceedings of Machine Learning Research. PMLR, Nov. 2022, pp. 1776–1781. URL: https://proceedings.mlr.press/v164/stocking22a.html.

[76] Ricardo Cannizzaro, Jonathan Routley and Lars Kunze. 'Towards a Causal Probabilistic Framework for Prediction, Action-Selection & Explanations for Robot Block-Stacking Tasks'. In: *IROS 2023 Workshop - Causality for Robotics*. Oct. 2023.

[77] Daniel McDuff, Yale Song, Jiyoung Lee, Vibhav Vineet, Sai Vemprala, Nicholas Alexander Gyde, Hadi Salman, Shuang Ma, Kwanghoon Sohn and Ashish Kapoor. 'CausalCity: Complex Simulations with Agency for Causal Discovery and Reasoning'. In: *Proceedings of the First Conference on Causal Learning and Reasoning*. Ed. by Bernhard Schölkopf, Caroline Uhler and Kun Zhang. Vol. 177. Proceedings of Machine Learning Research. PMLR, Apr. 2022, pp. 559–575. URL: https://proceedings.mlr.press/v177/mcduff22a.html.

[78] Thomas Kipf, Ethan Fetaya, Kuan-Chieh Wang, Max Welling and Richard Zemel. 'Neural Relational Inference for Interacting Systems'. In: *Proceedings of the 35th International Conference on Machine Learning*. Ed. by Jennifer Dy and Andreas Krause. Vol. 80. Proceedings of Machine Learning Research. PMLR, July 2018, pp. 2688–2697. URL: http://proceedings.mlr.press/v80/kipf18a.html.

[79] Kexin Yi, Chuang Gan, Yunzhu Li, Pushmeet Kohli, Jiajun Wu, Antonio Torralba and Joshua B Tenenbaum. 'CLEVRER: Collision Events for Video Representation and Reasoning'. In: *International Conference on Learning Representations*. 2020. URL: https://openreview.net/forum?id=HkxYzANYDB.

[80] Yunzhu Li, Antonio Torralba, Anima Anandkumar, Dieter Fox and Animesh Garg. 'Causal Discovery in Physical Systems from Videos'. In: *Advances in Neural Information Processing Systems*. Ed. by H Larochelle, M Ranzato, R Hadsell, M F Balcan and H Lin. Vol. 33. Curran Associates, Inc., 2020, pp. 9180–9192. URL: https://proceedings.neurips.cc/paper/2020/hash/6822951732be44edf818dc5a97d32ca6-Abstract.html.

[81] Liting Sun, Rebecca Roelofs, Ben Caine, Khaled S Refaat, Ben Sapp, Scott Ettinger and Wei Chai. 'CausalAgents: A Robustness Benchmark for Motion Forecasting'. In: *2024 IEEE International Conference on Robotics and Automation (ICRA)*. 2024, pp. 6820–6827. DOI: 10.1109/ICRA57147.2024.10610186.

[82] Wenhao Ding, Haohong Lin, Bo Li and Ding Zhao. 'CausalAF: Causal Autoregressive Flow for Safety-Critical Driving Scenario Generation'. In: *Proceedings of the 6th Conference on Robot Learning*. Ed. by Karen Liu, Dana Kulic and Jeff Ichnowski. Vol. 205. Proceedings of Machine Learning Research. PMLR, Dec. 2023, pp. 812–823. URL: https://proceedings.mlr.press/v205/ding23a.html.

[83] Pim de Haan, Dinesh Jayaraman and Sergey Levine. 'Causal Confusion in Imitation Learning'. In: *Advances in Neural Information Processing Systems*. Ed. by H Wallach, H Larochelle, A Beygelzimer, F d'Alché-Buc, E Fox and R Garnett. Vol. 32. Curran Associates, Inc., 2019. URL: https://proceedings.neurips.cc/paper/2019/file/947018640bf36a2bb609d3557a285329-Paper.pdf.

[84] Haohong Lin, Wenhao Ding, Zuxin Liu, Yaru Niu, Jiacheng Zhu, Yuming Niu and Ding Zhao. 'Safety-Aware Causal Representation for Trustworthy Offline Reinforcement Learning in Autonomous Driving'. In: *IEEE Robotics and Automation Letters* 9.5 (2024), pp. 4639–4646. DOI: 10.1109/LRA.2024.3379805.

[85] Jinkyu Kim and John Canny. 'Interpretable Learning for Self-Driving Cars by Visualizing Causal Attention'. In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. Oct. 2017.

[86] Chengxi Li, Stanley H Chan and Yi-Ting Chen. 'Who Make Drivers Stop? Towards Driver-centric Risk Assessment: Risk Object Identification via Causal Inference'. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2020, pp. 10711–10718. DOI: 10.1109/IROS45743.2020.9341072.

[87] Éloi Zablocki, Hédi Ben-Younes, Patrick Pérez and Matthieu Cord. 'Explainability of Deep Vision-Based Autonomous Driving Systems: Review and Challenges'. In: *Int. J. Comput. Vision* 130.10 (Oct. 2022), pp. 2425–2452. ISSN: 0920-5691. DOI: 10.1007/s11263-022-01657-x.

[88] Stephen Thomas and Katrina M Groth. 'Toward a hybrid causal framework for autonomous vehicle safety analysis'. In: *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability* 237.2 (2023), pp. 367–388.

[89] Daniel C Dennett. 'Intentional Systems'. In: *The Journal of Philosophy* 68.4 (1971), pp. 87–106. ISSN: 0022362X. URL: http://www.jstor.org/stable/2025382.

[90] Stefano V Albrecht and Peter Stone. 'Autonomous agents modelling other agents: A comprehensive survey and open problems'. In: *Artificial Intelligence* 258 (2018), pp. 66–95. ISSN: 0004-3702. DOI: https://doi.org/10.1016/j.artint.2018.01.002.

[91] Jieyu Zhu, Yanli Ma and Yining Lou. 'Multi-vehicle interaction safety of connected automated vehicles in merging area: A real-time risk assessment approach'. In: *Accident Analysis & Prevention* 166 (2022), p. 106546. ISSN: 0001-4575. DOI: https://doi.org/10.1016/j.aap.2021.106546.

[92] Dirk Helbing and Péter Molnár. 'Social force model for pedestrian dynamics'. In: *Phys. Rev. E* 51 (5 May 1995), pp. 4282–4286. DOI: 10.1103/PhysRevE.51.4282.

[93] Martin Treiber, Ansgar Hennecke and Dirk Helbing. 'Congested traffic states in empirical observations and microscopic simulations'. In: *Phys. Rev. E* 62 (2 Aug. 2000), pp. 1805–1824. DOI: 10.1103/PhysRevE.62.1805.

[94] Nicola Bellotto, Marc Hanheide and Nico Van de Weghe. 'Qualitative Design and Implementation of Human-Robot Spatial Interactions'. In: *Social Robotics*. Ed. by Guido Herrmann, Martin J Pearson, Alexander Lenz, Paul Bremner, Adam Spiers and Ute Leonards. Springer International Publishing, 2013, pp. 331–340. ISBN: 978-3-319-02675-6.

[95] Christian Dondrup, Nicola Bellotto, Marc Hanheide, Kerstin Eder and Ute Leonards. 'A Computational Model of Human-Robot Spatial Interactions Based on a Qualitative Trajectory Calculus'. In: *Robotics* 4.1 (2015), pp. 63–102. ISSN: 2218-6581. DOI: 10.3390/robotics4010063.

[96] Sariah Mghames, Luca Castri, Marc Hanheide and Nicola Bellotto. 'Qualitative Prediction of Multi-Agent Spatial Interactions'. In: *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. 2023, pp. 1170–1175. DOI: 10.1109/RO-MAN57019.2023.10309584.

[97] Saurabh Arora and Prashant Doshi. 'A survey of inverse reinforcement learning: Challenges, methods and progress'. In: *Artificial Intelligence* 297 (2021), p. 103500. ISSN: 0004-3702. DOI: 10.1016/j.artint.2021.103500.

[98] Xiao Wen, Sisi Jian and Dengbo He. 'Modeling the Effects of Autonomous Vehicles on Human Driver Car-Following Behaviors Using Inverse Reinforcement Learning'. In: *IEEE Transactions on Intelligent Transportation Systems* 24.12 (2023), pp. 13903–13915. DOI: 10.1109/TITS.2023.3298150.

[99] Markus Kuderer, Shilpa Gulati and Wolfram Burgard. 'Learning driving styles for autonomous vehicles from demonstration'. In: *2015 IEEE International Conference on Robotics and Automation (ICRA)*. 2015, pp. 2641–2646. DOI: 10.1109/ICRA.2015.7139555.

[100] Peter Bonsall, Ronghui Liu and William Young. 'Modelling safety-related driving behaviour—impact of parameter values'. In: *Transportation Research Part A: Policy and Practice* 39.5 (2005), pp. 425–444. ISSN: 0965-8564. DOI: https://doi.org/10.1016/j.tra.2005.02.002.

[101] Wei-Chiu Ma, De-An Huang, Namhoon Lee and Kris M Kitani. 'Forecasting Interactive Dynamics of Pedestrians With Fictitious Play'. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. July 2017.

[102] Ahmad Rahimi, Po-Chien Luan, Yuejiang Liu, Frano Rajič and Alexandre Alahi. 'Sim-to-Real Causal Transfer: A Metric Learning Approach to Causally-Aware Interaction Representations'. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2025, pp. 17271–17281.

[103] F Camara, S Cosar, N Bellotto, N Merat and C Fox. 'Continuous Game Theory Pedestrian Modelling Method for Autonomous Vehicles'. In: *Human Factors in Intelligent Vehicles*. Ed. by C Olaverri-Monreal, F García-Fernández and R J F Rossetti. River Publishers, Aug. 2020. ISBN: 978-8770222044. URL: https://eprints.whiterose.ac.uk/id/eprint/162717/.

[104] Steven Kuhn. 'Prisoner's Dilemma'. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N Zalta and Uri Nodelman. Winter 2024. Metaphysics Research Lab, Stanford University, 2024. URL: https://plato.stanford.edu/archives/win2024/entries/prisoner-dilemma/.

[105] Hengyuan Hu, Adam Lerer, Alex Peysakhovich and Jakob Foerster. '"Other-Play" for Zero-Shot Coordination'. In: *Proceedings of the 37th International Conference on Machine Learning*. Ed. by Hal Daumé III and Aarti Singh. Vol. 119. Proceedings of Machine Learning Research. PMLR, July 2020, pp. 4399–4410. URL: https://proceedings.mlr.press/v119/hu20a.html.

[106] Tina Debbarma, Tannistha Pal, Ashim Saha and Nikhil Debbarma. 'HCNNet: a hybrid convolutional neural network for abnormal human driver behaviour detection'. In: *Sādhanā* 50.1 (2025), p. 9.

[107] Wei Song, Guangde Zhang and Yicheng Long. 'Identification of dangerous driving state based on lightweight deep learning model'. In: *Computers and Electrical Engineering* 105 (2023), p. 108509. ISSN: 0045-7906. DOI: https://doi.org/10.1016/j.compeleceng.2022.108509.

[108] Zhijun Chen, Chaozhong Wu, Zhen Huang, Nengchao Lyu, Zhaozheng Hu, Ming Zhong, Yang Cheng and Bin Ran. 'Dangerous driving behavior detection using video-extracted vehicle trajectory histograms'. In: *Journal of Intelligent Transportation Systems* 21.5 (2017), pp. 409–421. DOI: 10.1080/15472450.2017.1305271.

[109] Shengxue Zhu, Chongyi Li, Kexin Fang, Yichuan Peng, Yuming Jiang and Yajie Zou. 'An Optimized Algorithm for Dangerous Driving Behavior Identification Based on Unbalanced Data'. In: *Electronics* 11.10 (2022). ISSN: 2079-9292. DOI: 10.3390/electronics11101557.

[110] Sarvesh Kolekar, Joost de Winter and David Abbink. 'Human-like driving behaviour emerges from a risk-based driver model'. In: *Nat Commun* 11.1 (Sept. 2020), p. 4850.

[111] Lien-Wu Chen and Hsien-Min Chen. 'Driver Behavior Monitoring and Warning With Dangerous Driving Detection Based on the Internet of Vehicles'. In: *IEEE Transactions on Intelligent Transportation Systems* 22.11 (2021), pp. 7232–7241. DOI: 10.1109/TITS.2020.3004655.

[112] Jia-Li Yin, Bo-Hao Chen, Kuo-Hua Robert Lai and Ying Li. 'Automatic Dangerous Driving Intensity Analysis for Advanced Driver Assistance Systems From Multimodal Driving Signals'. In: *IEEE Sensors Journal* 18.12 (2018), pp. 4785–4794. DOI: 10.1109/JSEN.2017.2765315.

[113] Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez and Vladlen Koltun. 'CARLA: An Open Urban Driving Simulator'. In: *Proceedings of the 1st Annual Conference on Robot Learning*. 2017, pp. 1–16.

[114] Tadeg Quillien and Christopher G Lucas. 'Counterfactuals and the logic of causal selection.' In: *Psychological Review* (2023).

[115] Chen Tang, Nishan Srishankar, Sujitha Martin and Masayoshi Tomizuka. 'Grounded Relational Inference: Domain Knowledge Driven Explainable Autonomous Driving'. In: *IEEE Transactions on Intelligent Transportation Systems* 25.9 (2024), pp. 10617–10635. DOI: 10.1109/TITS.2024.3424667.

[116] Dannier Xiao, Mehrdad Dianati, William Gonçalves Geiger and Roger Woodman. 'Review of Graph-Based Hazardous Event Detection Methods for Autonomous Driving Systems'. In: *IEEE Transactions on Intelligent Transportation Systems* 24.5 (2023), pp. 4697–4715. DOI: 10.1109/TITS.2023.3240104.

[117] Wentao Bao, Qi Yu and Yu Kong. 'Uncertainty-based Traffic Accident Anticipation with Spatio-Temporal Relational Learning'. In: *Proceedings of the 28th ACM International Conference on Multimedia*. Seattle, WA, USA: Association for Computing Machinery, 2020, pp. 2682–2690. ISBN: 9781450379885. DOI: 10.1145/3394171.3413827.

[118] Ryo Takahashi, Naoya Inoue, Yasutaka Kuriya, Sosuke Kobayashi and Kentaro Inui. 'Explaining potential risks in traffic scenes by combining logical inference and physical simulation'. In: *International Journal of Machine Learning and Computing* 6.5 (2016), pp. 248–255.

[119] Arnav Vaibhav Malawade, Shih-Yuan Yu, Brandon Hsu, Deepan Muthirayan, Pramod P Khargonekar and Mohammad Abdullah Al Faruque. 'Spatiotemporal Scene-Graph Embedding for Autonomous Vehicle Collision Prediction'. In: *IEEE Internet of Things Journal* 9.12 (2022), pp. 9379–9388. DOI: 10.1109/JIOT.2022.3141044.

[120] Sergio Casas, Cole Gulino, Renjie Liao and Raquel Urtasun. 'SpAGNN: Spatially-Aware Graph Neural Networks for Relational Behavior Forecasting from Sensor Data'. In: *2020 IEEE International Conference on Robotics and Automation (ICRA)*. 2020, pp. 9491–9497. DOI: 10.1109/ICRA40945.2020.9196697.

[121] Amir Sadeghian, Ferdinand Legros, Maxime Voisin, Ricky Vesel, Alexandre Alahi and Silvio Savarese. 'CAR-Net: Clairvoyant Attentive Recurrent Network'. In: *Proceedings of the European Conference on Computer Vision (ECCV)*. Sept. 2018.

[122] Chengxi Li, Yue Meng, Stanley H Chan and Yi-Ting Chen. 'Learning 3D-aware Egocentric Spatial-Temporal Interaction via Graph Convolutional Networks'. In: *2020 IEEE International Conference on Robotics and Automation (ICRA)*. 2020, pp. 8418–8424. DOI: 10.1109/ICRA40945.2020.9197057.

[123] Luca Castri, Sariah Mghames, Marc Hanheide and Nicola Bellotto. 'Causal discovery of dynamic models for predicting human spatial interactions'. In: *International Conference on Social Robotics*. Springer. 2022, pp. 154–164.

[124] Luca Castri, Sariah Mghames, Marc Hanheide and Nicola Bellotto. 'Enhancing Causal Discovery from Robot Sensor Data in Dynamic Scenarios'. In: *Proceedings of the Second Conference on Causal Learning and Reasoning*. Ed. by Mihaela van der Schaar, Cheng Zhang and Dominik Janzing. Vol. 213. Proceedings of Machine Learning Research. PMLR, Apr. 2023, pp. 243–258. URL: https://proceedings.mlr.press/v213/castri23a.html.

[125] Luca Castri, Gloria Beraldo, Sariah Mghames, Marc Hanheide and Nicola Bellotto. 'Ros-causal: A ros-based causal analysis framework for human-robot interaction applications'. In: *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction Workshop 'Causal-HRI: Causal Learning for Human-Robot Interaction'*. 2024.

[126] Luca Castri, Gloria Beraldo, Sariah Mghames, Marc Hanheide and Nicola Bellotto. 'Experimental Evaluation of ROS-Causal in Real-World Human-Robot Spatial Interaction Scenarios'. In: *2024 33rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN)*. 2024, pp. 1603–1609. DOI: 10.1109/RO-MAN60168.2024.10731290.

[127] Keisuke Fujii, Naoya Takeishi, Kazushi Tsutsui, Emyo Fujioka, Nozomi Nishiumi, Ryoya Tanaka, Mika Fukushiro, Kaoru Ide, Hiroyoshi Kohno, Ken Yoda, Susumu Takahashi, Shizuko Hiryu and Yoshinobu Kawahara. 'Learning interaction rules from multi-animal trajectories via augmented behavioral models'. In: *Advances in Neural Information Processing Systems*. Ed. by M Ranzato, A Beygelzimer, Y Dauphin, P S Liang and J Wortman Vaughan. Vol. 34. Curran Associates, Inc., 2021, pp. 11108–11122. URL: https://proceedings.neurips.cc/paper_files/paper/2021/file/5c572eca050594c7bc3c36e7e8ab9550-Paper.pdf.

[128] Ričards Marcinkevičs and Julia E Vogt. 'Interpretable Models for Granger Causality Using Self-explaining Neural Networks'. In: *International Conference on Learning Representations*. 2021. URL: https://openreview.net/forum?id=DEa4JdMWRHp.

[129] David Alvarez Melis and Tommi Jaakkola. 'Towards Robust Interpretability with Self-Explaining Neural Networks'. In: *Advances in Neural Information Processing Systems*. Ed. by S Bengio, H Wallach, H Larochelle, K Grauman, N Cesa-Bianchi and R Garnett. Vol. 31. Curran Associates, Inc., 2018. URL: https://proceedings.neurips.cc/paper_files/paper/2018/file/3e9f0fc9b2f89e043bc6233994dfcf76-Paper.pdf.

[130] Chen Tang, Wei Zhan and Masayoshi Tomizuka. 'Interventional Behavior Prediction: Avoiding Overly Confident Anticipation in Interactive Prediction'. In: *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2022, pp. 11409–11415. DOI: 10.1109/IROS47612.2022.9981524.

[131] Peter Goodison, Peter Johnson and Joanne Thoms. 'Establishing Causality in Complex Human Interactions: Identifying Breakdowns of Intentionality'. In: *Complex Sciences*. Ed. by Jie Zhou. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 1631–1641. ISBN: 978-3-642-02469-6. DOI: 10.1007/978-3-642-02469-6_42.

[132] Stefano V Albrecht, Cillian Brewitt, John Wilhelm, Balint Gyevnar, Francisco Eiras, Mihai Dobre and Subramanian Ramamoorthy. 'Interpretable Goal-based Prediction and Planning for Autonomous Driving'. In: *2021 IEEE International Conference on Robotics and Automation (ICRA)*. 2021, pp. 1043–1049. DOI: 10.1109/ICRA48506.2021.9560849.

[133] Krishnaiyan Thulasiraman and Madisetti N S Swamy. 'Graphs: theory and algorithms'. In: John Wiley & Sons, 2011. Chap. 5.7, p. 118.

[134] Elias Bareinboim and Judea Pearl. 'A General Algorithm for Deciding Transportability of Experimental Results'. In: *Journal of Causal Inference* 1.1 (2013), pp. 107–134. DOI: 10.1515/jci-2012-0004.

[135] Michael Eichler. 'Granger causality and path diagrams for multivariate time series'. In: *Journal of Econometrics* 137.2 (2007), pp. 334–353. ISSN: 0304-4076. DOI: 10.1016/j.jeconom.2005.06.032.

[136] Michael Eichler. 'Graphical modelling of multivariate time series'. In: *Probability Theory and Related Fields* 153 (2012), pp. 233–268.

[137] J Runge. 'Causal network reconstruction from time series: From theoretical assumptions to practical estimation'. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 28.7 (July 2018). ISSN: 1054-1500. DOI: 10.1063/1.5025050.

[138] Wenpeng Yin, Hinrich Schütze, Bing Xiang and Bowen Zhou. 'ABCNN: Attention-Based Convolutional Neural Network for Modeling Sentence Pairs'. In: *Transactions of the Association for Computational Linguistics* 4 (June 2016), pp. 259–272. ISSN: 2307-387X. DOI: 10.1162/tacl_a_00097.

[139] Aapo Hyvärinen, Shohei Shimizu and Patrik O Hoyer. 'Causal modelling combining instantaneous and lagged effects: an identifiable model based on non-Gaussianity'. In: *Proceedings of the 25th International Conference on Machine Learning*. ICML '08. Helsinki, Finland: Association for Computing Machinery, 2008, pp. 424–431. ISBN: 9781605582054. URL: 10.1145/1390156.1390210.

[140] Pierre Comon. 'Independent component analysis, A new concept?' In: *Signal Processing* 36.3 (1994). Higher Order Statistics, pp. 287–314. ISSN: 0165-1684. DOI: 10.1016/0165-1684(94)90029-9.

[141] Shohei Shimizu, Takanori Inazumi, Yasuhiro Sogawa, Aapo Hyvärinen, Yoshinobu Kawahara, Takashi Washio, Patrik O Hoyer and Kenneth Bollen. 'DirectLiNGAM: A Direct Method for Learning a Linear Non-Gaussian Structural Equation Model'. In: *Journal of Machine Learning Research* 12.33 (2011), pp. 1225–1248. URL: http://jmlr.org/papers/v12/shimizu11a.html.

[142] Hui Zou. 'The Adaptive Lasso and Its Oracle Properties'. In: *Journal of the American Statistical Association* 101.476 (2006), pp. 1418–1429. DOI: 10.1198/016214506000000735.

[143] Kunihiro Baba, Ritei Shibata and Masaaki Sibuya. 'Partial Correlation and Conditional Correlation as Measures of Conditional Independence'. In: *Australian & New Zealand Journal of Statistics* 46.4 (2004), pp. 657–664. DOI: 10.1111/j.1467-842X.2004.00360.x.

[144] Peter Spirtes, Christopher Meek and Thomas Richardson. 'An Algorithm for Causal Inference in the Presence of Latent Variables and Selection Bias'. In: *Computation, Causation, and Discovery*. AAAI Press, May 1999. Chap. 6. ISBN: 9780262315821. DOI: 10.7551/mitpress/2006.003.0009.

[145] Biwei Huang, Kun Zhang, Jiji Zhang, Joseph Ramsey, Ruben Sanchez-Romero, Clark Glymour and Bernhard Schölkopf. 'Causal Discovery from Heterogeneous/Nonstationary Data'. In: *Journal of Machine Learning Research* 21.89 (2020). URL: http://jmlr.org/papers/v21/19-232.html.

[146] Gideon Schwarz. 'Estimating the Dimension of a Model'. In: *The Annals of Statistics* 6.2 (1978), pp. 461–464. ISSN: 00905364. URL: http://www.jstor.org/stable/2958889.

[147] David Heckerman, Dan Geiger and David M Chickering. 'Learning Bayesian networks: The combination of knowledge and statistical data'. In: *Machine learning* 20.3 (1995), pp. 197–243.

[148] Jose M Peña, Johan Björkegren and Jesper Tegnér. 'Learning dynamic Bayesian network models via cross-validation'. In: *Pattern Recognition Letters* 26.14 (2005), pp. 2295–2308. ISSN: 0167-8655. DOI: 10.1016/j.patrec.2005.04.005.

[149] John Houston, Guido Zuidhof, Luca Bergamini, Yawei Ye, Long Chen, Ashesh Jain, Sammy Omari, Vladimir Iglovikov and Peter Ondruska. 'One Thousand and One Hours: Self-driving Motion Prediction Dataset'. In: *Proceedings of the 2020 Conference on Robot Learning*. Ed. by Jens Kober, Fabio Ramos and Claire Tomlin. Vol. 155. Proceedings of Machine Learning Research. PMLR, Nov. 2021, pp. 409–418. URL: https://proceedings.mlr.press/v155/houston21a.html.

[150] Robert Krajewski, Julian Bock, Laurent Kloeker and Lutz Eckstein. 'The highD Dataset: A Drone Dataset of Naturalistic Vehicle Trajectories on German Highways for Validation of Highly Automated Driving Systems'. In: *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. 2018, pp. 2118–2125. DOI: 10.1109/ITSC.2018.8569552.

[151] Aapo Hyvärinen and Stephen M Smith. 'Pairwise Likelihood Ratios for Estimation of Non-Gaussian Structural Equation Models'. In: *Journal of Machine Learning Research* 14 (2013), pp. 111–152.

[152] Eric V Strobl. 'Causal discovery under non-stationary feedback'. PhD thesis. Department of Biomedical Informatics, University of Pittsburgh, 2017. ISBN: 978-0-355-41081-5. URL: https://www.proquest.com/dissertations-theses/causal-discovery-under-non-stationary-feedback/docview/2013196949/se-2?accountid=13042.

[153] Carles Balsells Rodas, Ruibo Tu, Yingzhen Li and Hedvig Kjellstrom. 'Causal Discovery from Conditionally Stationary Time Series'. In: *UAI 2022 Workshop on Causal Representation Learning*. 2022. URL: https://openreview.net/forum?id=LIAN4ILH_68.

[154] Peter Spirtes and Kun Zhang. 'Causal discovery and inference: concepts and recent methodological advances'. In: *Applied Informatics* 3.3 (Feb. 2016). ISSN: 2196-0089. DOI: 10.1186/s40535-016-0018-x.

[155] Martin L Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. 1st. USA: John Wiley & Sons, Inc., 1994. ISBN: 0471619779.

[156] Richard E Fikes and Nils J Nilsson. 'STRIPS: A new approach to the application of theorem proving to problem solving'. In: *Artificial Intelligence* 2.3 (1971), pp. 189–208. ISSN: 0004-3702. DOI: 10.1016/0004-3702(71)90010-5.

[157] Drew McDermott, Malik Ghallab, Adele Howe, Craig Knoblock, Ashwin Ram, Manuela Veloso, Daniel Weld and David Wilkins. *PDDL — The Planning Domain Definition Language*. Tech. rep. Technical Report CVC TR98003/DCS TR1165. New Haven, CT, US: Yale Center for Computational Vision and Control, Oct. 1998.

[158] Edwin Pednault. 'Formulating multiagent, dynamic-world problems in the classical planning framework'. In: *Reasoning about actions and plans* (1987), pp. 47–82.

[159] Arquímides Méndez-Molina, Eduardo F Morales and Luis E Sucar. 'CARL: A Synergistic Framework for Causal Reinforcement Learning'. In: *IEEE Access* 11 (2023). DOI: 10.1109/ACCESS.2023.3331728.

[160] Edward K Blum and Sergey V Lototsky. *Mathematics of physics and engineering*. World Scientific Publishing Company, 2006.

[161] Michael L Scott. *Programming Language Pragmatics*. 3rd. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2009. ISBN: 0123745144.

[162] Steven Macenski, Tully Foote, Brian Gerkey, Chris Lalancette and William Woodall. 'Robot Operating System 2: Design, architecture, and uses in the wild'. In: *Science Robotics* 7.66 (2022). DOI: 10.1126/scirobotics.abm6074.

[163] Philip Wadler. 'The essence of functional programming'. In: *Proceedings of the 19th ACM SIGPLAN-SIGACT symposium on Principles of programming languages*. 1992.

[164] Mark P Jones and Luc Duponcheel. *Composing monads*. Technical Report YALEU/DCS/RR-1004. Department of Computer Science. Yale University., 1993.

[165] Charles K Assaad, Emilie Devijver and Eric Gaussier. 'Discovery of extended summary graphs in time series'. In: *Proceedings of the Thirty-Eighth Conference on Uncertainty in Artificial Intelligence*. Ed. by James Cussens and Kun Zhang. Vol. 180. Proceedings of Machine Learning Research. PMLR, Aug. 2022, pp. 96–106. URL: https://proceedings.mlr.press/v180/assaad22a.html.

[166] Søren Wengel Mogensen and Daniel Malinsky. 'Causal Learning for Partially Observed Stochastic Dynamical Systems'. In: *Proceedings of the Thirty-Fourth Conference on Uncertainty in Artificial Intelligence*. UAI'18. Monterey, CA, USA: AUAI Press, 2018.

[167] Joris M Mooij, Sara Magliacane and Tom Claassen. 'Joint Causal Inference from Multiple Contexts'. In: *Journal of Machine Learning Research* 21.99 (2020). URL: http://jmlr.org/papers/v21/17-123.html.

[168] Luca Castri, Sariah Mghames, Marc Hanheide and Nicola Bellotto. 'CAnDOIT: Causal Discovery with Observational and Interventional Data from Time Series'. In: *Advanced Intelligent Systems* 6.12 (2024). DOI: https://doi.org/10.1002/aisy.202400181.

[169] Massimo Guiggiani. *The Science of Vehicle Dynamics: Handling, Braking, and Ride of Road and Race Cars*. Springer, 2018.

[170] Béla G Lipták. *Process Control: Instrument Engineers' Handbook*. Butterworth-Heinemann, 2013.

[171] Donald Richard John Laming. *Information theory of choice-reaction times*. Oxford, UK: Academic Press, 1968. ISBN: 9780124354500.

[172] Tim Schreiter, Tiago Rodrigues de Almeida, Yufei Zhu, Eduardo Gutierrez Maestro, Lucas Morillo-Mendez, Andrey Rudenko, Luigi Palmieri, Tomasz P Kucner, Martin Magnusson and Achim J Lilienthal. 'THÖR-MAGNI: A large-scale indoor motion capture recording of human movement and robot interaction'. In: *The International Journal of Robotics Research* 44.4 (2025), pp. 568–591. DOI: 10.1177/02783649241274794.

[173] Road Safety Authority (Government of Ireland). *The two-second rule*. Mar. 2012. URL: https://web.archive.org/web/20120309213451/http://www.rotr.ie/rules-for-driving/speed-limits/speed-limits_2-second-rule.html.

[174] Gaël Guennebaud, Benoît Jacob et al. *Eigen v3*. 2010. URL: http://eigen.tuxfamily.org.

[175] Tobias Moers, Lennart Vater, Robert Krajewski, Julian Bock, Adrian Zlocki and Lutz Eckstein. 'The exiD Dataset: A Real-World Trajectory Dataset of Highly Interactive Highway Scenarios in Germany'. In: *2022 IEEE Intelligent Vehicles Symposium (IV)*. 2022, pp. 958–964. DOI: 10.1109/IV51971.2022.9827305.

[176] Julian Bock, Robert Krajewski, Tobias Moers, Steffen Runde, Lennart Vater and Lutz Eckstein. 'The inD Dataset: A Drone Dataset of Naturalistic Road User Trajectories at German Intersections'. In: *2020 IEEE Intelligent Vehicles Symposium (IV)*. 2020, pp. 1929–1934. DOI: 10.1109/IV47402.2020.9304839.

[177] Nick Hawes, Christopher Burbridge, Ferdian Jovan, Lars Kunze, Bruno Lacerda, Lenka Mudrova, Jay Young, Jeremy Wyatt, Denise Hebesberger, Tobias Kortner, Rares Ambrus, Nils Bore, John Folkesson, Patric Jensfelt, Lucas Beyer, Alexander Hermans, Bastian Leibe, Aitor Aldoma, Thomas Faulhammer, Michael Zillich, Markus Vincze, Eris Chinellato, Muhannad Al-Omari, Paul Duckworth, Yiannis Gatsoulis, David C Hogg, Anthony G Cohn, Christian Dondrup, Jaime Pulido Fentanes, Tomas Krajnik, Joao M Santos, Tom Duckett and Marc Hanheide. 'The STRANDS Project: Long-Term Autonomy in Everyday Environments'. In: *IEEE Robotics & Automation Magazine* 24.3 (2017), pp. 146–156. DOI: 10.1109/MRA.2016.2636359.

[178] Ricardo Cannizzaro, Michael Groom, Jonathan Routley, Robert Osazuwa Ness and Lars Kunze. 'COBRA-PPM: A Causal Bayesian Reasoning Architecture Using Probabilistic Programming for Robot Manipulation Under Uncertainty'. In: *2025 European Conference on Mobile Robots (ECMR)*. 2025.

[179] Kun Zhang, Jonas Peters, Dominik Janzing and Bernhard Schölkopf. 'Kernel-Based Conditional Independence Test and Application in Causal Discovery'. In: *Proceedings of the Twenty-Seventh Conference on Uncertainty in Artificial Intelligence*. UAI'11. Barcelona, Spain: AUAI Press, 2011, pp. 804–813. ISBN: 9780974903972.

[180] Yoav Benjamini and Yosef Hochberg. 'Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing'. In: *Journal of the Royal Statistical Society. Series B (Methodological)* 57.1 (1995), pp. 289–300. ISSN: 00359246. URL: http://www.jstor.org/stable/2346101.

[181] William S Cleveland and Susan J Devlin. 'Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting'. In: *Journal of the American Statistical Association* 83.403 (1988), pp. 596–610. ISSN: 01621459. URL: http://www.jstor.org/stable/2289282.

[182] Hirotogu Akaike. 'Information Theory and an Extension of the Maximum Likelihood Principle'. In: *Selected Papers of Hirotugu Akaike*. Ed. by Emanuel Parzen, Kunio Tanabe and Genshiro Kitagawa. New York, NY: Springer New York, 1998, pp. 199–213. ISBN: 978-1-4612-1694-0. DOI: 10.1007/978-1-4612-1694-0_15.

[183] Diederik P Kingma and Jimmy Ba. 'Adam: A Method for Stochastic Optimization'. In: *International Conference on Learning Representations*. 2015. URL: https://hdl.handle.net/11245/1.505367.

[184] Joseph Ramsey, Peter Spirtes and Jiji Zhang. 'Adjacency-Faithfulness and Conservative Causal Inference'. In: *Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence*. UAI'06. Cambridge, MA, USA: AUAI Press, 2006, pp. 401–408. ISBN: 0974903922.