Thesis Proposal

Rethinking the Design of Coordinated, User-Centric AI Systems for Deployment in Transportation

Rex Chen

May 2024

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Thesis Committee:

Fei Fang	(Co-chair)	Carnegie Mellon University
Norman Sadeh	(Co-chair)	Carnegie Mellon University
Sean Qian		Carnegie Mellon University
Matteo Pozzi		Carnegie Mellon University
Peter Stone		The University of Texas at Austin

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Societal Computing.

Abstract

As traffic demand continues to increase globally, improving the efficiency and safety of the interconnected network of transportation systems around the world has become an increasingly critical priority. Various AI technologies have been designed to this end, and many have achieved good performance in simulation-based evaluations. However, a disconnect between theory and practice exists: few state-of-the-art AI technologies have been deployed to actually help resolve these challenges in the real world. One important cause of this disconnect is that these AI technologies have made unrealistic simplifying assumptions, which have made them unable to address the pain points of human stakeholders. In this thesis, I propose to answer research questions related to how AI technologies can be better designed for deployment by addressing four common challenges: uncertainty in underlying and observed levels of demand; coordination between individuals and systems; interpretability and controllability for complicated decision-making algorithms; and heterogeneity among end-users and deployment contexts. My completed, in-progress, and proposed work tackles these challenges through the lens of two key problem domains, traffic signal control and gig driving. Ultimately, the goal of this thesis is to design AI systems which are capable of being physically deployed and creating tangible impacts in these domains.

Contents

Chapter 1

Introduction

Transportation is an important area of social impact, because it is the driving force that helps move people and goods between different contexts of productivity. In 2015, the United Nations recognised that improving the economic, social, and environmental sustainability of transportation systems is a key component of achieving global "peace and prosperity". Specifically, transportation falls under two of the UN's 17 Sustainable Development Goals — Goal 9, Industry, Innovation, and Infrastructure, and Goal 11, Sustainable Cities and Communities [138].

Despite the importance of transportation as a domain, it also suffers from systemic inefficiencies that are becoming increasingly critical challenges. Schrank et al. [119] measured the average delay of commuters in the United States and found that it more than doubled from 1982 to 2019, with comparable increases in both metropolitan areas with small ($< 500\,000$) and large ($> 3\,000\,000$) populations. Even though the COVID-19 pandemic reduced this delay substantially (especially in larger metropolitan areas), they found that traffic trends in late 2020 point to a rapid resurgence in congestion back to 2019 levels. The nature of traffic is also being reshaped by various emerging trends: on-demand ridesharing and food delivery platforms have enhanced mobility while adding a large volume of short-range trips [56, 79], connected and autonomous vehicles have created new possibilities for coordination while introducing safety concerns [47, 81], and remote and hybrid work have shifted the activity and travel patterns of workers [18, 60].

At the same time, various emerging technologies based on artificial intelligence (AI) are poised to help address these challenges. In my thesis, I explore applications of AI to two prominent problems in transportation: traffic signal control (Section 2.1) and gig driving (Section 2.2). Both have been well-studied by the AI community, with AI algorithms showing strong capabilities in simulations (e.g., [20, 88]). However, a disconnect between theory and practice exists. Although these problems are also crucial for industry stakeholders, state-of-the-art AI technologies have seldom achieved real-world deployment because these stakeholders are reluctant to adopt them over simpler deployed solutions. This disconnect can be partially attributed to how AI technologies often aim to solve problems divorced from the practical contexts in which they are to be applied. Therefore, these technologies are unable to efficiently address the actual pain points of stakeholders.

What, then, are these pain points? In my thesis, I view these pain points as instances of broader challenges that have been inadequately addressed by prior work. My work seeks to help address outstanding research questions pertaining to the following four areas of deployment challenges:

- **Uncertainty** (Section 2.3.1). How can AI technologies in transportation achieve robust levels of performance when the level of demand may be spatiotemporally variable, and may not necessarily even be observable in a noise-free fashion? This challenge arises in the transition from simulation to deployment. Prior work in traffic signal control and gig driving has either abstracted the effects of uncertainty out of the problem, or modelled this uncertainty in ways that are not reflective of real-world conditions. If not accounted for, this could degrade the performance of AI algorithms from what has been seen in simulated environments.
- **Coordination** (Section 2.3.2). How can AI technologies in transportation reason about the complex interactions between individual agents to achieve socially optimal outcomes? This challenge arises inherently in the design process. Prior work in traffic signal control and gig driving has commonly made the assumption of full decentralisation, where self-interested agents (either AIs or humans) are meant to emergently achieve the socially optimal outcome through maximising their individual utility functions. However, it is difficult to guarantee the behaviour that will be exhibited by such agents in practice.
- **Interpretability** (Section 2.3.3). How can AI technologies in transportation present their decisions to human stakeholders in a way that builds trust without compromising system performance? This challenge arises both in an ex ante fashion during design (i.e., interpretability) and in an ex post fashion after deployment (i.e., explainability). Prior work in traffic signal control and gig driving has designed various heuristics to explain and guarantee the behaviour of AI systems, but these solutions still fall short in terms of aligning with the mental models of stakeholders (e.g., traffic engineers or gig drivers).
- **Heterogeneity** (Section 2.3.4). How can AI technologies in transportation be broadly applied in the presence of variation between end-users and deployment contexts, and how might this be hindered by assumptions made during their design? This challenge arises during the generalisation of a single AI system's deployment across many environments. Prior work in traffic signal control and gig driving has abstracted out much of this variation by performing evaluation with small sets of environments and homogeneous agents. However, differences in these dimensions can have important impacts on system outcomes.

In the remainder of this thesis proposal, I will begin with a review of related work on traffic signal control and gig driving in Chapter 2. Next, in Chapter 3 and Chapter 4, I will describe specific projects that I have initiated or planned to address instances of these challenges in traffic signal control and gig driving. Lastly, for the proposed projects in Chapter 4, I will outline a timeline for completing this work over the year to come in Chapter 5.

Chapter 2

Related Work

Many applications of AI technologies to transportation systems have been developed in past work. These technologies have been applied to ground transportation (including basic infrastructure, vehicle control, fleet management, and public transit) [42, 86, 106], aviation [34, 110], and maritime transport [100], solving a gamut of problems ranging from prediction, forecasting, and pattern recognition to control, planning, and optimisation [64]. This thesis focuses on two ground transportation problems, traffic signal control (Section 2.1) and gig driving (Section 2.2). The complexity, scale, and importance of these problems makes them a suitable proving ground for the design of AI technologies in transportation. I also believe that insights about deploying AI for social impact can also readily generalise from these challenging problems to other domains.

2.1 Traffic Signal Control

Traffic signals are a fundamental component of transportation infrastructure that helps to improve the efficiency and safety of vehicular traffic. Thus, *traffic signal control* (TSC) is one of the two transportation problems that I consider. Even though congestion is fundamentally caused by insufficient capacity for high traffic volumes on roadways [109], retiming traffic signals can significantly alleviate this issue. In 2004, [26] summarised reports showing that 75% of traffic signals in the United States could be made more efficient, and that a reduction of 15-20% in delay and an increase of 8-23% in capacity can be achieved through retiming. Improvements in congestion, in turn, can lead to economic and environmental benefits. [119] estimated that, in 2020, American commuters collectively incurred a travel delay of 4.3 billion hours, an opportunity cost of $101 billion, and excess emissions of 18 million tons of greenhouse gases due to congestion.

Formally, the problem of TSC aims to find a signal plan for each intersection in a road network that minimises the delay of vehicles as they travel through the network. At each intersection, every *approach* (roadway) is split into *lanes*. Vehicles traversing the intersection can follow different *movements*, each of which is defined by one incoming and one outgoing lane [51, 149, 165]. A *phase* is a combination of green lights permitting travel through a subset of movements. A signal plan for an intersection is defined by a sequence of phases and the time allocated to each phase. For efficiency, pairs of non-conflicting phases are often signalled simultaneously (e.g., westbound/eastbound left turns, northbound left turn/through) [72, 102, 148].

Historically, three main types of algorithmic approaches have been taken to TSC. In *fixed-time control*, a single plan or a small number of time-of-day plans are optimized based on historical traffic volumes. These plans most often loop through fixed phase sequences in cycles, with the *cycle length*, *splits* (proportions of time allocated to each phase), and *offsets* (timing offsets between adjacent intersections to achieve coordination) being the key parameters. In *actuated control*, a fixed-time plan is adjusted based on detector inputs (such as vehicle presence data from loop detectors) through a fixed set of logical rules. Finally, in *adaptive control*, signal plans are adjusted based on more complex optimization algorithms, such that the mapping from detector inputs to plans can vary over time [37, 51]. I focus on adaptive signal control in this thesis.

2.1.1 Deployed Algorithms

Industry-developed adaptive TSC methods have generally adhered closely to the control-flow of fixed-time control systems. This can be attributed to the fact that the installation and operation costs for wholesale replacement are seen as a barrier to the adoption of adaptive TSC [121]. Consequently, deployed adaptive TSC algorithms focus on making incremental changes to the cycles, splits, and offsets of predefined time-of-day plans. This design philosophy was pioneered by the SCOOT and SCATS algorithms in the 1970s [128], but it persisted through the Federal Highway Administration (FHWA)'s efforts in developing the ACS Lite system in the 2000s [121] and is still followed by industry-leading systems such as Econolite's Edaptive [36]. These systems generally rely on local search and optimisation techniques to determine cycles, splits, and offsets [30, 31], and focus more on leveraging domain-specific heuristics than striving for optimality. For instance, deployed offset optimisation methods are based on the link-pivoting combination method of [30], which greedily chooses offsets for additional intersections in an ordering based on the road network topology. Thus, while these algorithms cause minimal disruption to existing workflows, it is not clear that they are leveraging the full potential of adaptive signal plan optimisation.

2.1.2 AI Algorithms

While adaptive TSC algorithms based on scheduling (e.g., the SURTRAC algorithm of [126]) have achieved limited deployment, in this thesis I focus on a class of algorithms that has attracted significant attention from AI researchers but (to my knowledge) has hitherto remained undeployed. *Reinforcement learning* (RL) algorithms have achieved superhuman performance in contexts involving high-dimensional state and action spaces and real-time decision-making, including card and video games [15, 124, 139]. Consequently, they have also been applied to TSC, where they have also significantly improved performance metrics over fixed-time and actuated methods [20, 166].

As shown in Figure 2.1, RL methods learn state-dependent signal control policies through trial-and-error interactions with an environment (often a *traffic simulator*). A RL agent learns to control an intersection by receiving as input the intersection's state (e.g., the number of queueing vehicles in each lane), taking a signalling action, and receiving a reward (e.g., the reduction in the number of queueing vehicles). State-of-the-art RL algorithms for TSC include those of [20, 89, 93]. These are deep Q-network (DQN)-based algorithms that execute acyclic plans [102], which have no fixed phase sequences but instead freely choose the phase for the next time step. While RL algorithms can freely optimise to achieve good performance, they thus significantly depart from the status quo. In this thesis, I aim to design and deploy RL algorithms capable of bridging this gap.

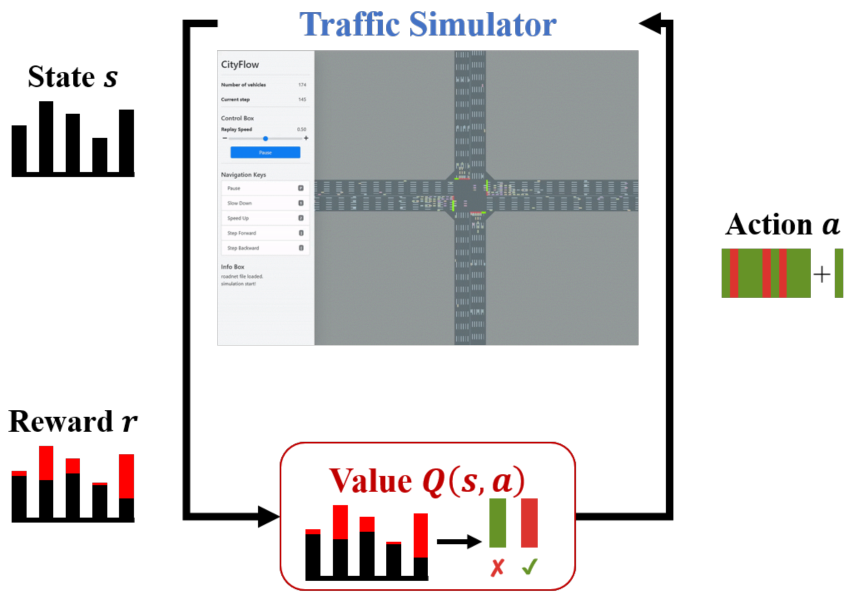


Figure 2.1: Training loop for a reinforcement learning-based traffic signal control algorithm.

2.2 Gig Driving

Moving away from transportation infrastructure to the traffic that uses it, the other of the two transportation problems that I consider will be broadly referred in this thesis to as *gig driving*. Gig drivers are independent contractors who use their personal vehicles to provide on-demand transportation of either passengers (*ridesourcing*) and restaurant or grocery orders (*food delivery*) from app platforms such as Uber, Lyft, DoorDash, and Instacart. I consider ridesourcing and food delivery as a unified problem, not only due to similarities in their problem formulations [59] but also because drivers view them interchangeably [76] (e.g., Uber and its food delivery business, Uber Eats, are mutually cannibalistic platforms [28]).

Unlike TSC, gig driving involves interactions between three distinct groups of stakeholders: the customers who request rides or deliveries, the drivers who service these requests, and the platforms which mediate their interactions. Customers initiate by sending their requests and payments directly to the platforms. The platforms then attempt to match the set of requests with the set of available drivers. Each driver is then shown a list of requests that the platform recommends, and is free to accept or decline each one. If they accept, they fulfill the request; if they decline, the platform falls back to an alternative assignment. Lastly, platforms distribute payments to drivers. Thus, from the perspective of the platforms, this is an iterative weighted matching problem [59].

Although the realisation of this model is a relatively recent phenomenon, gig driving has experienced explosive growth [55] due to the benefits that it provides. For customers, gig driving provides a convenient alternative to both private vehicles and public transport [27, 61, 116, 120]. For drivers, gig driving offers them the flexibility and autonomy to choose when and where to work [21, 76]. Yet, gig driving has also introduced new challenges. Negative interactions between customers and drivers degrade the experiences of both parties [57, 113, 129]. For drivers, the promise of autonomy is also hindered by the opacity of platforms' assignment, pricing, and evaluation mechanisms [53, 76, 125, 133, 153], leading to volatility and systemic inequities in driver outcomes [33, 95]. Lastly, gig driving has created societal externalities. At least in the short term, gig driving will continue to contribute to increased congestion, as these platforms have not just shifted demand away from other transport modes but also created new demand altogether [56, 116].

2.2.1 Deployed Algorithms

In practice, gig platforms extend the basic problem formulation to account for fluctuations in supply and demand. One of the most prevalent strategies adopted by these platforms is *dynamic pricing*, referred to as "surge pricing" by Uber, "Prime Time" by Lyft, "Peak Pay" by DoorDash, and "Blitz Pricing" by Postmates [76, 77]. By setting higher prices during peak hours, platforms balance demand by disincentivising customers from making requests, while also incentivising drivers to reposition to high-demand regions and accept requests in exchange for bonuses [85, 104]. Precise implementation details for these mechanisms are generally scant. [104] describes Lyft's "escrow" mechanism, in which convex optimisation problems are solved to prospectively allocate customers' expected Prime Time payments equitably among drivers. However, it is likely that deployed algorithms incorporate various heuristics based on business requirements.

Just as external researchers have little visibility into the inner workings of dynamic pricing algorithms, drivers themselves must contend with both information asymmetries and gamification mechanisms that they perceive to be unfair [161]. Within their limited control, drivers have responded to the algorithmic management exerted by platforms in various ways: switching between platforms [76], engaging in discussion forums [90, 161], and even colluding to induce artificial surges [136]. Nevertheless, the incentives created by these opaque pricing mechanisms still create tangible effects on drivers, including inequitable increases in the level of competition and requisite effort [95]. How are drivers to navigate this complex ecosystem? In this thesis, I take a driver-centric approach in responding to these challenges with novel AI-driven solutions.

2.2.2 AI Algorithms

One line of work in AI for gig driving has focused on gig platforms by designing mechanisms for driver dispatching and payment. These mechanisms aim to achieve various desirable properties such as welfare maximisation and incentive compatibility (i.e., guaranteeing that drivers will accept dispatched trips) [14, 46, 88, 111]. In particular, incentive compatibility is enforced by ensuring that prices offset drivers' opportunity costs [46, 88]. However, these mechanisms generally make unrealistic simplifying assumptions. For instance, drivers in reality rarely have perfect information about even dispatched trips (e.g., drop-off locations are often obscured) [161], much less the opportunity costs of declining them. Even works grounded in particular platforms (such as [104]) fail to consider the complexities that arise from interactions between platforms: how can the opportunity costs of drivers change when they entail the option to switch to a different platform?

Simultaneously, a complementary line of work has focused on understanding and responding to the needs of drivers. [161] conducted focus groups with ridesourcing drivers to envision possible improvements to platforms; one of their core findings was a need for native in-platform data-driven insights that would eliminate the need for third-party tools (e.g., mileage tracking apps). Based on this, [160] designed "data probes" to help drivers understand how their work patterns interact with platforms' management practices. [70], meanwhile, created a measurement suite that quantitatively analyses the dynamic pricing strategies of platforms. Yet, given this abundance of information, how should drivers make decisions? In this thesis, I combine these algorithmic and user-centred perspectives to afford drivers more control in their decision-making.

2.3 Common Challenges

For a 2022 paper reviewing the progress that has been made to date on deploying RL-based algorithms for TSC [24], I outlined four engineering challenges that would arise from deploying RL-based TSC as part of end-to-end physical systems. In the following sections, I will describe how these challenges — uncertainty (Section 2.3.1), coordination (Section 2.3.2), interpretability (Section 2.3.3), and heterogeneity (Section 2.3.4) — exist generally in transportation systems, and specifically within the contexts of TSC and gig driving. My thesis aims to address ways in which the literature has fallen short in responding to these challenges.

2.3.1 Uncertainty

One of the central challenges for AI technologies in transportation is that the level of demand and supply is not constant over time and space. Transportation systems are not closed [144], as the number of end-users can vary according to recurrent (e.g., rush hours) and non-recurrent (e.g., public events) factors, the latter of which makes traffic prediction difficult. For TSC, [4] showed that the performance of fixed RL policies degrades in previously unseen contexts. Either a diverse set of training contexts and a rich feature representation must be available during training [4], or the RL algorithm must be capable of domain adaptation [99, 103]. For gig driving, both the flexibility of driver schedules [21] and the on-demand nature of customer requests [5] introduce significant volatility into platform dynamics, even if drivers do acquire experientially acquire intuitions and patterns for their work [76, 160]. In both problems, AI techniques have been evaluated using static or synthetic datasets that likely fail to capture the full impacts of demand and supply variability.

Even observing current traffic rather than predicting future traffic can be difficult when AI technologies lack a complete picture of a transportation system. For TSC, although most RL algorithms assume perfect state observability [102], deployed policies must adapt to the varying strengths and weaknesses of physical detectors [134]. For instance, loop detectors are vulnerable to wear and overcount large vehicles [49], while camera detectors undercount in inclement conditions [112, 131]. While this can be addressed through domain randomisation techniques [45, 99, 132], the protocols used by prior work have not been calibrated to real-world conditions. For gig driving, the information asymmetry imposed by gig platforms limits visibility into real-time system conditions for both drivers and third-party tools alike [70, 161]. Prior tools have abstracted the effects of this uncertainty away from drivers, but how would exposing it affect their behaviour?

2.3.2 Coordination

To achieve socially optimal outcomes in transportation, AI technologies must be able to reason about interactions between many different agents. In TSC, the signalling actions taken by individual intersections are not independent, but will affect the traffic state of upstream and downstream intersections. In gig driving, the choices taken by individual drivers to accept or reject rides will likewise shift the overall distribution of demand and supply. While various AI algorithms make the assumption of decentralisation between agents in transportation, the resulting behaviour lacks guarantees. Specifically, theoretical guarantees regarding the ability of decentralised agents to achieve coordination do not exist, except when a target state visitation distribution is given [141]. For TSC, although RL policies can emergently lead to *green progression*, or unobstructed movement through traffic corridors via sequences of green lights [145, 146]. However, this behaviour is still less synchronised than signals deployed in practice [1]. For gig driving, while mechanisms such as those of [46, 88] are designed to be incentive-compatible for individual, utility-maximising drivers, they fail to provide concrete incentives for drivers to coordinate to satisfy demand [104].

Introducing hierarchical organisation can help groups of agents to coordinate on achieving higher-level objectives [3, 156]. Both TSC and gig driving are domains with inherent hierarchical structure. For TSC, various RL algorithms have used the graph structure of road networks for coordination — either emergently learning hierarchical representations using graph neural networks [101, 143, 147, 159], or explicitly introducing higher-level controller agents that coordinate intersection agents by setting subgoals or reward signals [2, 32, 89, 158, 164]. However, this work has not focused on the practicalities of deploying such systems, including the distribution of computation between hardware at different intersections. For gig driving, different platforms can themselves be viewed as agents directing the actions of driver agents. In turn, competitive dynamics between platforms [28, 67] influence their dispatching mechanisms. While differences between platforms have been considered by driver-centric work [160, 161], they have not been considered from the perspective of coordinating drivers at the platform level through mechanism design.

2.3.3 Interpretability

AI technologies for transportation will be applied to established systems instead of evolving along with the deployment context. To ensure that stakeholders can trust AI-driven decisions, these decisions must be communicated in a manner that is compatible with their existing mental models. For TSC, the deep neural network-based acyclic signal plans common in the RL literature [102] are a far cry from the heuristics-based cycle-offset-split signal plans used by traffic engineers. A minority of existing RL methods output variable phase durations within fixed cycle lengths [68, 91, 155, 158, 122], but they are either heuristic or intractable. For gig driving, the needs and mental models of drivers are a significant factor in trust. For instance, [19, 76, 160] reported that drivers rejected platforms' nudges to "surge chase" upon finding misalignment in their expectations about dynamic pricing with reality. However, there has been no work on how these considerations factor into gig drivers' interactions with third-party tools.

Both TSC and gig driving are also domains with elevated stakes, where wrong decisions made by AI technologies can have significant impacts on stakeholders. Controls and constraints must thus be placed as safeguards against adverse outcomes. For TSC, failing to set appropriate constraints on signal plans can lead to safety issues. The potential unpredictability of acyclic RL-based signal plans elevates the regulatory risk for municipalities, as evidenced by a precedent for tort cases [135]. Existing RL methods for TSC rely either on reward shaping [38, 50, 80, 82, 157] or postprocessing [35, 98] to enforce constraints, but neither strikes an ideal balance between safety guarantees and performance optimality. For gig driving, failing to accurately estimate potential demand can lead to opportunity costs in terms of drivers' monetary earnings and time. While [54, 88] have examined this issue from the mechanism design perspective, the question of how to communicate assurances about potential earnings to drivers remains an open problem. In both contexts, there is a trade-off between providing guarantees through constraints and providing performance through flexibility; the key problem is how to strike this balance so as to satisfy stakeholders.

2.3.4 Heterogeneity

Variation between individual end-users can impact how they respond to decisions made by AI technologies in transportation. For TSC, classes of vehicles such as trucks [52], public transit [150], and gig driving vehicles [56] all have disparate impacts on traffic patterns. Even within vehicle classes, factors such as driving speed, reaction time, and braking sharpness affect how individual drivers respond to signals such as yellow lights [69]. However, the majority of traffic simulations used for RL training abstract away inter-vehicle variation, and many of them ignore pedestrian traffic [102, 162]. How do these varying assumptions impact RL training? For gig driving, drivers also have a variety of motivations and habits. Some rely on it as a primary source of income, while others view it as supplemental income [88, 154]; some drive full-time, while others drive part-time [13, 76, 88]. Although some mechanisms incentivise drivers to report their preferences while maintaining theoretical guarantees [111], it is likely in practice that these guarantees will need to be relaxed due to the complexities of drivers' preferences and constraints.

Similarly, the performance of AI technologies can vary between deployment contexts, particularly if they make assumptions about properties of their environments that apply in some deployments but not in others. For TSC, the most common strategy for RL is to train and evaluate on a single road network, and often using only a single benchmark algorithm [92, 102]. Although benchmarks collecting widely-used simulations have been created [7, 92], they have received limited attention outside of their originating research groups. If a RL policy is to be deployed with minimal finetuning in new contexts, it is not clear how well existing algorithms would generalise. For gig driving, it has been consistently recognised that the market entry of platforms has disparate impacts in different cities depending on their socioeconomic makeup [33, 48] and existing transportation infrastructure [48, 120]. Beyond the spatiotemporal locality considered in prior work [88], there is likely also a role for these dimensions in influencing how drivers respond to a single mechanism across cities, as has been found by [104] in their empirical evaluations.

Chapter 3

Completed and In-Progress Work

3.1 Uncertainty

Framing Uncertainty in AI Decision Aids for Gig Drivers

Gig drivers must contend with information asymmetry imposed by platforms (Section 2.2.1). Data-driven insights can provide drivers with additional visibility into how their activity impacts their earnings and other outcomes; such insights are provided by third-party applications such as Gridwise and Stride [58], as well as the tools designed by [70, 160]. However, even with these predictive tools, the onus of planning a driver's daily activity still falls on themselves. There is a lack of prescriptive tools, deployed or otherwise, that can directly support the planning process. Here, AI-driven *decision aids* can help by both recommending decisions to drivers and estimating how good the outcomes of following those decisions will be.

Uncertainty (see Section 2.3.1), however, is a barrier to the effectiveness of such decision aids. In the gig driving context, two sources of uncertainty prevent users from determining whether decision aids will be able to help them make better decisions, i.e., those that enable drivers to maximise their earnings. (1) Decision aids must be able to predict how earnings are affected by complex interactions with other users and systems [115], including variability in demand and supply. (2) Third-party decision aids will also be subject to the same information asymmetry as the drivers, and thus must estimate earnings based on historical data or other proxies. As part of under-review work [25], I conducted a user study to investigate how the framing of uncertainty impacts drivers' interactions with such a decision aid — a schedule recommendation tool.

I designed this schedule recommendation tool to consist of two modules. (1) An **estimation module** prospectively predicts the earnings that drivers can expect during each hour of each day of the week, which could be computed by a machine learning model or directly averaged from historical data. To focus on the effects of the tool's design, I took the latter approach for this work. (2) A **scheduling module** takes the estimated earnings and drivers' availability constraints as inputs to produce an optimal set of working times that leads to the highest possible earnings. I also designed a front-end interface that would allow drivers both to enter their availability constraints and to view the schedule generated by the tool.

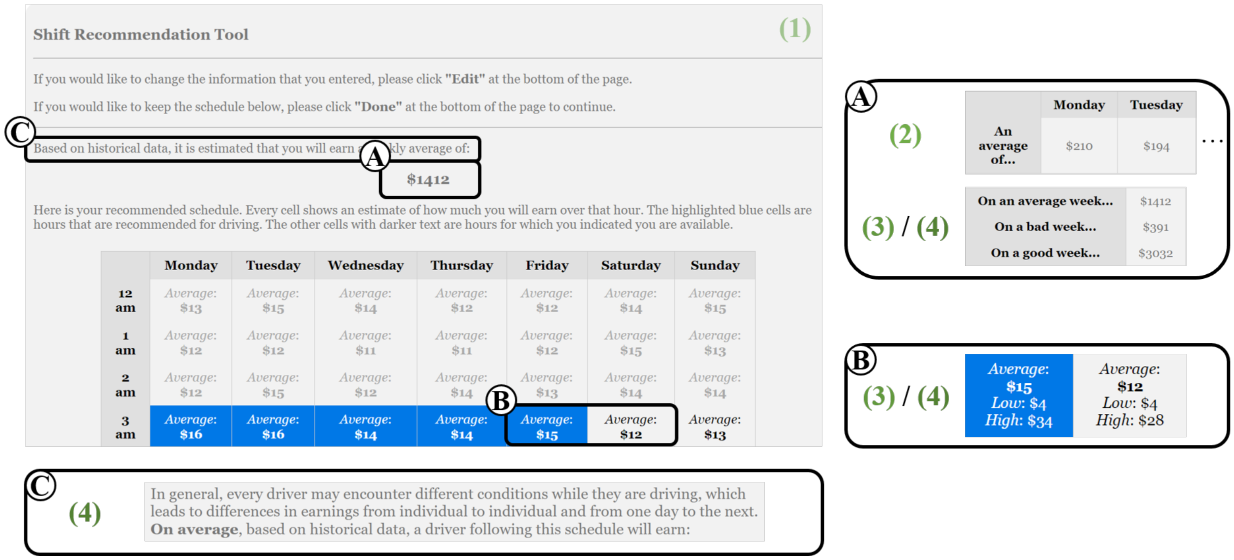


Figure 3.1: Comparison of the four conditions for the presentation of uncertainty in estimates by the schedule recommendation tool.

Starting from the basic design of the front-end interface, I created four conditions varying how the tool presented its earnings estimates (Figure 3.1). In two of these four conditions, I exposed uncertainty by showing mean historical earnings as point estimates along with pessimistic (10th percentile of historical data) and optimistic (90th percentile of historical data) estimates.

(1) **Reference condition**. Users were only shown their mean weekly estimated earnings and their mean estimated earnings for each hour during the week.

(2) **Daily estimates**. To assess the effect of increasing granularity but not error rate visibility, users were shown their mean estimated earning for each day instead of their mean estimated weekly earning. Mean estimated earnings were still presented for each hour during the week.

(3) **Ranged estimates**. To assess the effect of increasing error rate visibility, participants were shown mean, pessimistic, and optimistic estimated earnings for the entire week, as well as for each hour during the week.

(4) **Ranged and hedged estimates**. To assess the effect of hedging language on error rate visibility, the textual description of the estimates was changed from Condition 3. Instead of "Based on historical data, it is estimated that you will earn", Condition 4 states "On average, based on historical data, a driver following this schedule will earn".

For the user study, my goal was to assess how users' perceptions of these four conditions impact their *trust* in, *reliance* on, and *compliance* with the schedule recommendation tool, filling a gap identified in Section 2.3.3. I consider trust to be an internal attitude involving "the willingness of a person to be vulnerable to the actions of an AI decision aid"; reliance to be the external expression of that attitude in allowing the decision aid to recommend decisions; and compliance to be the external expression of following the decision aid's recommended decisions [127]. Past work in human-AI interaction has assessed users' trust in AI under a variety of contexts [137], and this work has firmly established that it is impacted by the design of AI systems [16, 29, 65, 71, 74, 117]. However, prior studies of trust have largely relied on hypothetical or low-stakes scenarios. Even studies of deployed systems [12, 142] have been limited to observational studies that did not compare multiple designs. Gig driving is a representative context that captures several key challenges that AI decision aids must contend with in real-world deployments: uncertainty, longitudinal interactions (since users must repeatedly rely on the tool to plan their activity each day), and elevated stakes (since choosing to follow or not follow the tool's recommendations will have material impacts on the financial well-being of users).

The user study involved a longitudinal trial in which participants interacted with the schedule recommendation tool for 7 days (Figure 3.2). To ensure the ecological validity of the trial, I recruited 51 real gig drivers from the user base of the driver assistant app Gridwise across four cities (Los Angeles, New York, Chicago, and Houston), and used historical data from these cities to estimate earnings. On **Day 0**, participants completed **Survey 1**, a series of formative questions, before they interacted with the schedule recommendation tool for the first time to enter their constraints and receive a schedule for the coming week. Then, after participants saw this schedule for the first time, they completed a pre-interaction measurement of trust in **Survey 2**. On each of **Day 1** to **Day 7**, participants received their recommended schedule for that particular day, and were free to incorporate the schedule into their regular driving activity for that day as they saw fit. At the end of each day, they completed another measurement of trust in **Survey 3**. Finally, on Day 7, participants completed a final, post-interaction measurement of trust in **Survey 4**.

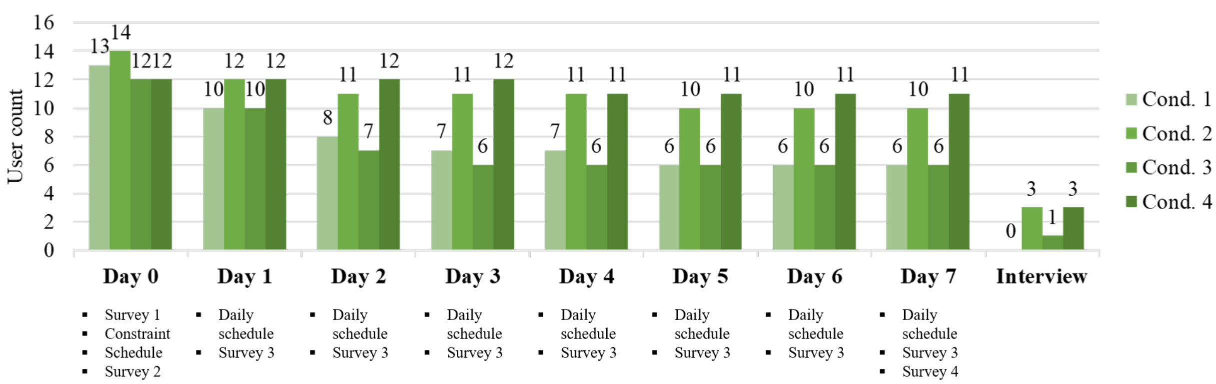


Figure 3.2: Retention statistics for the longitudinal trial, decomposed by the condition presented to participants for the schedule recommendation tool.

Despite the relatively small sample size, I observed significant differences in trust and reliance in the schedule recommendation tool between the framings of uncertainty in the four conditions. Conditions 2 and 4 had the lowest levels of daily (Survey 3) trust, yet significantly higher daily (Survey 3) reliance than Conditions 1 and 3. This was in spite of the fact that Conditions 1 and 3 had higher pre-interaction (Survey 2) trust. Therefore, adding either granularity or hedging to the estimates may have helped participants to maintain their reliance on the tool for decision-making, while calibrating their perceptions of its trustworthiness to be realistic but not necessarily optimal [43, 118]. If the nominal purpose of an AI decision aid is to support the user's own goals, then such a notion of "appropriate reliance" may be a more suitable design goal than strict compliance [118]. Meanwhile, Condition 3 had significantly higher pre-interaction (Survey 2) trust than Condition 1, suggesting that adding ranges improved trust. While this result contradicts prior work [108], I attributed it to the more task-aligned presentation of uncertainty in this tool. However, the effect of Condition 3 diminished over longitudinal interactions. Furthermore, Condition 4 had significantly lower pre-interaction (Survey 2) and daily (Survey 3) trust than Condition 3. This suggests that the addition of hedging language may have increased error visibility to the extent that it offset the participants' perceived utility of the ranges. Beyond the effects of the conditions, I also found more generally that the extent to which participants perceived the tool's estimated earnings as being accurate impacted their daily trust and reliance.

Lastly, I also supplemented my quantitative analyses with qualitative findings from interviews that I conducted with 7 participants. While the participants largely felt that the time they spent interacting the tool was worthwhile, they also reported a diversity of motivations, routines, and experiences with the tool. Even within the same condition, participants varied in how they incorporated the tool's recommendations into their planning, and in how they reacted to the outcomes of their reliance. This points to a need to consider the heterogeneity of gig drivers (see Section 2.3.4) when designing AI decision aids for them — the ways in which these decision aids present and explain their recommendations must be aligned with their diverse needs and perceptions.

3.2 Coordination

Designing Cyclic Traffic Signal Controllers with Multi-Agent RL

As outlined in Section 2.1.2, RL algorithms for TSC are generally acyclic, which is a drastic departure from the cycle-split-offset plans commonly deployed in real-world road networks. This makes it more difficult for stakeholders to trust the behaviour of these algorithms, and also would prevent an end-to-end signal control solution such as Econolite's Edaptive from directly programming a signal controller with its outputs. Furthermore, this also means that coordination between intersections (see Section 2.3.2) is learned emergently rather than enforced as a constraint. It is easy for a cycle-split-offset plan to achieve green progression by first synchronising cycle lengths along a corridor, and then adjusting offsets based on intersection distances and real-time traffic demand [1]. Strategies for enforcing fixed-length cyclic plans in the RL literature involve either postprocessing logit vectors to convert them into valid splits [155, 158] or encoding valid splits as discrete actions [68, 91, 122], but these algorithms have not achieved state-of-the-art performance and — in the case of the latter — are unlikely to scale well.

In ongoing work, I am modifying *MPLight* [20], one of the state-of-the-art RL algorithms for TSC. This algorithm is fully decentralised, and trains a single Q-network for all agents using an intersection topology-invariant embedding of the intersection agents' state observations [165]. Owing to this parameter sharing, their algorithm has n times fewer parameters and n times more samples to train from for an n-agent environment. Consequently, it is able to scale up performantly to a large road network of 2 510 intersections [20]. However, the algorithm cannot be readily deployed in the real world because it does not respect the following signal plan constraints:

- **Minimum green times.** Each phase must be signalled for at least a minimum amount of time, so as to accomplish three purposes: (1) to ensure that drivers have time to react and safely traverse the intersection; (2) to provide pedestrians; and (3) to clear built-up queues [72]. While MPLight's action interval [20] effectively sets a uniform minimum green of 10 s for each phase, the FHWA's recommended times differ based on roadway and phase types.
- **Maximum green times.** Each phase must also be signalled for at most a maximum amount of time, so as to ensure that demand for other phases is not excessively delayed [72]. Although it is conceivable that MPLight could emergently learn to "force off" excessively long phases, this behaviour cannot be guaranteed during execution.
- **Cycle lengths.** To enable green progression, an intersection must share cycle lengths with other intersections in the same arterial or grid network [72]. Since MPLight policies signal myopically without considering past signalling actions, it has no notion of a cyclic policy.
- **Phase sequences.** To make traffic signals predictable for drivers, the sequence of phases that is signalled must be kept as consistent as possible (see also Section 4.2). Although Edaptive and other deployed signal control methods do enable phase sequence changes based on the time of day [72], these are much less frequent than those of MPLight's acyclic policy.
- **Predetermined splits.** To ensure that stakeholders are able to inspect and predict the upcoming behaviour of the signalling policy, the splits for the entire phase sequence must be predetermined before they are signalled. This is a capability that is supported by Edaptive and deployed signal control methods (see Figure 3.3a), but not by MPLight.

The implementation of the first four of these constraints is relatively intuitive, due in part to MPLight's topology-invariant nature. At execution time, MPLight policies take the action with the highest state-action (Q) value among those that are valid for the intersection. Tracking these constraints and restricting the set of valid actions accordingly allows them to be enforced. However, the last of these constraints — predetermined splits — is more difficult to implement, because MPLight policies are incapable of outputting splits. My proposed solution, which I am currently implementing, involves *imitation learning*: training a surrogate RL policy to output splits that mimic those of a constrained MPLight policy as closely as possible. This split-based policy could be directly deployed in place of MPLight, which would result in greater stakeholder trust.

Currently, my evaluation of the constrained MPLight algorithm uses a 36-intersection road network based on the city of Strongsville, Ohio. This road network includes two arterial corridors: Pearl Road (US 42) and Royalton Road (Ohio SR 82); their intersection is shown in Figure 3.3b. Strongsville is an ideal real-world locality for benchmarking RL algorithms for TSC due to (1) the extensive detector equipment that provides high-fidelity traffic data; (2) the challengingly high volumes of traffic that lead to congestion and related crashes, due to connections between these roads and interstate I-71 [40]; and (3) the current use of Edaptive to perform coordination. Figure 4.2 shows that, in a simulation based on 5 pm rush-hour traffic, the naïve MPLight algorithm improves the average queue length per intersection over a fixed-time plan by 28%. Adding action-based constraints worsens the queue length, but still improves over the fixed-time plan by 18%. Results for the imitation learning-based RL algorithm are currently still forthcoming.

(a) Phase sequence and splits output by a historical run of Econolite's Edaptive system.

(b) SUMO [6] traffic simulation of Pearl Road (US 42) & Royalton Road (SR 82), overlaid on a satellite image.

Figure 3.3: Traffic conditions in Strongsville, Ohio.

3.3 Interpretability

Coordinating Decision Tree Surrogates for Multi-Agent RL

Deep neural networks are difficult for humans to interpret due to the large number of parameters that are not well-aligned with human intuition [163]. Thus, reliance on neural network policies is another pitfall of state-of-the-art RL algorithms for TSC, as noted in Section 2.3.3. However, this is another area where imitation learning can be helpful. Instead of training a surrogate neural network policy, the original neural network could be distilled into an alternative model with a more human-interpretable structure while preserving as much of the original behaviour as possible. One class of models that holds promise for this goal is *decision trees* (DTs), which have been found to be significantly more interpretable than either neural networks or decision rules [123].

[11] introduced *VIPER*, an algorithm which extends the no-regret imitation learning framework DAGGER [114] to learning DT policies. DAGGER iteratively uses the original neural network policy to provide expert action labels for a student policy. In the first iteration, the expert collects a set of trajectories that it provides to the student along with labels; in subsequent iterations, the student instead collects new trajectories for the expert to label. This ensures that the student policy can still perform well when its interactions with the environment lead it into previously unseen portions of the state space [114]. VIPER adds a resampling procedure where it upweights critical states, i.e., those where taking the incorrect action could lead to a significant loss in the Q-value. [66] applied VIPER to the TSC setting as part of a mixed autonomous supervision framework, in which DTs extracted from RL policies can fall back to hand-engineered DTs in unsafe states.

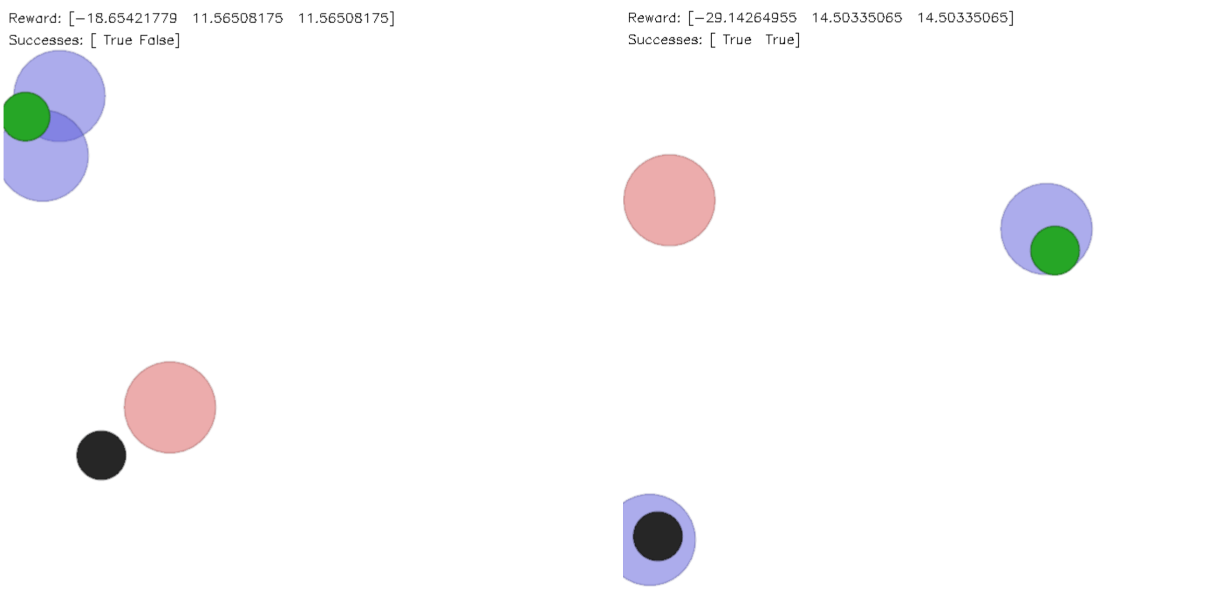
However, VIPER — and by extension the approach of [66] — is limited to single-agent settings, and thus its notion of critical states fails to capture the non-stationarity that arises from the types of complex multi-agent interactions [84] that characterise the TSC domain. In 2022, [97] extended the VIPER algorithm into multi-agent settings with two algorithms: *IVIPER* and *MAVIPER*.

IVIPER is analogous to VIPER in the single-agent setting; each student abstracts away the other DT students into being part of the environment. Specifically, it separately trains a DT for each agent using trajectories collected by the DT student for that agent against the expert policies for all other agents. Meanwhile, MAVIPER aims to achieve coordination by having each student reason about the learning of the other DT students. It collects trajectories by rolling out all student DTs at once. During each iteration, it jointly grows the student DTs one level at a time, and it focuses the training process by discarding samples where many student DTs mispredict the expert actions (i.e., the students fail to coordinate). In ongoing work, I am collaborating with the lead authors, Stephanie Milani and Zhicheng Zhang, to extend these algorithms into the TSC setting.

Our goal is to generate DT surrogates for RL-based TSC policies that are both performant (in terms of reward) and scalable (in terms of runtime), so that stakeholders can efficiently use them to reason about the behaviour of RL policies. Neither IVIPER nor MAVIPER achieves this goal; the lack of coordination in IVIPER degrades its performance, while various aspects of [97]'s implementation of MAVIPER hinders it from being truly scalable. We have envisioned and implemented a number of algorithmic ideas that could lead to the best of both worlds:

- **Sampling**. To upweight critical states, MAVIPER computes a Q-value weight for each experience in the collected trajectories. MAVIPER weight for each agent has two key differences from the single-agent VIPER weight: (1) it is given by the maximum possible loss in Q-value if the agent's entire team deviates from the expert action profile; and (2) it is taken as an expectation over action profiles of agents in opposing teams. Since each agent in the TSC setting has approximately 10 actions, [97]'s approach of enumerating all team and opponent actions is likely to be intractable as the number of agents scales up. Instead, we sample a subset of team and opponent actions to compute the weights. Experiments have shown that this change, surprisingly, improves not only runtime but also reward in some cases.
- **Feature selection**. Another critical bottleneck in MAVIPER is feature selection during the DT fitting process, which requires the algorithm to iterate through all features and feature values in the collected trajectories. Providing this subroutine with a subset of features to focus on would greatly improve its runtime. Experiments have also shown that it is possible for MAVIPER students to reach 100% of expert performance with feature masks handcrafted based on domain knowledge. We have also implemented two simple but more general solutions: (1) using SHAP values [87] as a measure of feature importance in the single-agent setting; and (2) greedily constructing a set of features based on how new features improve the reward obtained by IVIPER students. However, these solutions focus on feature importances in the single-agent setting, and are thus liable to miscoordinate. We are thus developing an alternative solution that combines feature masking with joint DT training.
- **Sequential training**. By analogy to meta-algorithms for multi-agent RL such as [75]'s policy-space response oracles, we believe that one promising approach to coordinating the training of feature masks between agents is to allow the agents to best-respond to each other. Specifically, student 1 would train its mask by interacting with the experts, student 2 would train its mask by interacting with the masked student 1 and the experts for other agents, and so on. We are experimenting with this best response-based approach not only to train feature masks, but also to train the DT students themselves.

For our experimental evaluation, we are currently relying on the multi-agent particle world environments introduced by [84], which are a popular benchmark for reinforcement learning tasks. We specifically focus on the physical deception ("adversary"), cooperative navigation ("spread"), and predator-prey ("tag") environments. Although they may be divorced from the motivating context of TSC, the optimal coordinated behaviour is well-defined for agents in these environments. For instance, consider the physical deception environment in Figure 3.4. The goal of the defender agents (in purple) is to deceive the attacker agent (in red) from reaching a target landmark (in green). This can be achieved by the defender agents coordinating to divide attention between the target (green) and non-target (black) landmarks, which MAVIPER (Figure 3.4b) but not IVIPER (Figure 3.4a) is able to achieve. We aim to ensure that our proposed methods retain these coordination capabilities. However, we will extend our evaluation to TSC environments so as to demonstrate their utility for producing interpretable surrogates of RL-based signalling policies.



(a) DT students trained with IVIPER fail to coordinate and converge on the same landmark.

(b) DT students trained with MAVIPER coordinate by dividing attention between landmarks.

Figure 3.4: Behaviour of DT surrogates in the physical deception environment of [84].

3.4 Heterogeneity

Assessing Traffic Simulators' Distributional Equivalence

One frequently-overlooked aspect in the training loop of RL algorithms for TSC is the traffic simulator that is used as a training environment. While these algorithms would be ideally trained through interactions with physical detectors and controllers, collecting data in this manner may result in sigificant efficiency and safety costs [44]. Traffic simulators serve as a safe sandbox that can generate large quantities of realistic training data [10]. However, if we seek to transfer the performance of RL algorithms as seen in simulations into deployment, their modelling assumptions must be sufficiently realistic that they capture scenarios likely to be encountered by RL policies in physical deployments (see Section 2.3.4). This means that real-world validation of traffic simulations is an important task. However, a chicken-and-egg problem exists: traffic simulators are meant to replace large-scale data collection, but large-scale data collection is needed to validate traffic simulators. Nevertheless, comparisons between simulators take a partial step towards this goal by verifying that they lead to equivalent outcomes.

In work published at the 2023 Winter Simulation Conference [22], I conducted such a comparison between two simulators commonly used to train RL algorithms for TSC: *SUMO* [6] and *CityFlow* [162]. Both of these are microscopic simulators, which model the behaviour of individual vehicles [9]. Yet, the two simulators exist at opposite ends of a spectrum. SUMO is a granular simulator that supports detailed simulation definitions, but is less efficient due to being single-threaded, while CityFlow is a more abstract simulator that provides a minimal interface for RL training, but is more efficient due to being multi-threaded. Are these simulators interchangeable for training RL policies? While [162] compared the system-level measure of mean travel time between the simulators, my work assessed whether the vehicle-level measures generated by the simulators (e.g., queue lengths and other direct inputs to RL) are distributionally equivalent.

My experimental design is shown in Figure 3.5. For each replication, I fixed a set of various parameters shared between the two simulators. Next, I ran a simulation in both SUMO and CityFlow and computed the mean of vehicle-level measures over the simulation run: (1) the RMSE of per-vehicle travel time; (2) the RMSE of per-vehicle delay; (3) the RMSE and KL divergence of per-lane vehicle counts; (4) the RMSE and KL divergence of per-lane queued vehicle counts; (5) the RMSE of per-vehicle speed; and (6) the RMSE of per-vehicle acceleration. After running 9 replications for each experimental cell, I performed a t-test to assess whether the measures significantly differed between the simulators. Between cells, I varied two categories of parameters.

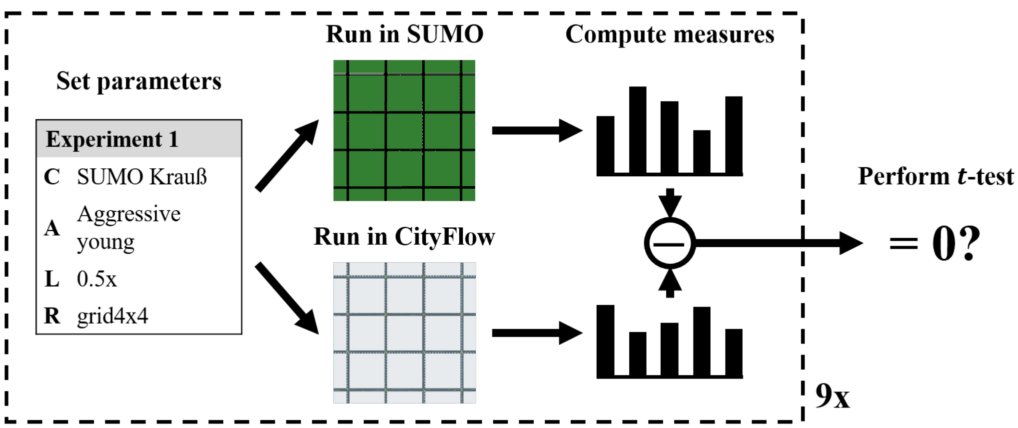


Figure 3.5: Experimental procedure for assessing the effect of simulation parameters on the effect of distributional equivalence between outcome measures in SUMO [6] and CityFlow [162].

First, I co-varied parameters relating to driver behavioural models in both simulators, so as to assess the effect of introducing heterogeneity not just among the environment but also among individual road users (see Section 2.3.4). SUMO and CityFlow model individual vehicles as making decisions for three types of tasks: (1) car-following, (2) intersection navigation, and (3) lane-changing. In my experiments, I varied the following dimensions of driver behaviour.

- **Car-following models** ("C" in Figure 3.5) model how vehicles determine the speeds at which to travel while unobstructed, following behind a lead vehicle, or stopping at an obstacle. Following the taxonomy of [78], I considered two types of car-following models: *collision avoidance models*, which maintain safe stopping distances at all times; and *action point models*, which vary between different modes of behaviour. Both SUMO and CityFlow use a collision avoidance model from [73]. I also extended CityFlow to add several collision avoidance [140, 151] and action point [96] models that had previously only been implemented in SUMO. I varied between six car-following model in my experiments.
- **Car-following parameters** ("A" in Figure 3.5) such as maximum acceleration and deceleration are shared between different car-following models, but have varying effects on each model. To limit the parameter space, I co-varied five parameters between six settings taken from a taxonomy of "aggressiveness types" introduced by [17]. These settings represent how car-following behaviour varies between young, middle-aged, and old drivers, as well as between aggressive and courteous drivers.
- **Lane-changing parameters** ("L" in Figure 3.5) influence simulated vehicles' decisions about whether a sufficiently large gap exists between the lead and lag vehicles on the destination lane to safely complete the manoeuvre. Both simulators use rule-based models that instantaneously move vehicles to the destination lane if such a gap exists, but differ in implementation details. I varied a "gap tolerance factor", a multiplicative factor that scales the size of the computed safe gap, between five levels based on empirical data from [130].

Second, I co-varied the overall scale of the simulation by using different road networks taken from the benchmark dataset of [7]. In one set of experiments, I evaluated the effect of increasing traffic intensity using two synthetic 4×4 grid networks (shown in Figure 3.5). In another set of experiments, I evaluated the effect of increasing the size of the road network by using simulations of one and seven intersections, taken from a road network based on Ingolstadt, Germany [83].

The fundamental conclusion from my experiments was that the low-level measures generated by SUMO and CityFlow are not distributionally equivalent, even if their system-level measures may be. One-sample t-tests indicated that all of the RMSE and KL divergence measures were significantly different from 0. Introducing additional heterogeneity into the simulation also impacted the lack of distributional equivalence between the simulators. For car-following models, the models did not impact different measures in the same ways; for instance, the model of [140] led to larger differences in speed and acceleration but smaller differences in vehicle counts. For simulation scale, increasing traffic intensity and simulation scale both significantly increased differences between the two simulators. While these findings pertain only to the inputs to RL — it is unclear whether RL algorithms would be able to recover similar policies given observations from these two simulators — they still suggest that the design of simulations is an important aspect of the RL training pipeline that cannot be taken for granted.

Chapter 4

Proposed Work

4.1 Coordination

Learning Equilibria for Game-Theoretic Models of Gig Driving

I noted in Section 2.3.2 that it is difficult for self-interested drivers to coordinate on fulfilling demand such that social welfare is maximised. Part of this arises from the information asymmetry inherent in gig platforms (Section 2.2.1): drivers can only observe the gig requests and dynamic prices provided to them by platforms, and have no visibility into or agency over how other drivers will respond to this information. [104] describes how Lyft's platform circumvents this issue by providing fixed, personalised bonuses to drivers. But how can coordination be achieved from the drivers' perspective? Specifically, how could an AI decision aid help drivers to maximise their potential earnings based on predicted fares and bonuses (Section 3.2), when other drivers — both users and non-users of the decision aid — will also be relying on similar information?

As a first step towards addressing this problem, I propose to start from a theoretical model. If gig drivers are modelled as agents in an n-player game (assumed to be homogeneous for now), where the actions entail repositioning to different zones based on dynamic pricing, and the demand in each zone is modelled as a consumable resource, then this problem can be viewed as a *congestion game*. Similar formulations have been previously adopted by [62, 63], although their work was not grounded in real-world data. *Nash equilibria* in such games correspond to the solution that could be achieved by self-interested drivers. They are a subset of *correlated equilibria*, which is the set of all solutions achievable by agents conditioning their strategies on the observation of a random signal. A shared, public signal yields Nash equilibria [8]; if each agent observes a private signal, however, it is possible to achieve correlated equilibria with higher social welfare. In this context, the AI decision aid's recommendations can be modelled as the signal.

Some past work has recognised the difficulty of solving for equilibria in congestion games; finding Nash equilibria is PLS-complete [41], and — although finding *any* correlated equilibrium can be done in polynomial time — finding social welfare-maximising correlated equilibria is NP-hard [105]. An alternative approach is to learn these equilibria emergently through RL. In 2021, I collaborated with James Cunningham, Justin Kiefel, and Chun Kai Ling to design several RL algorithms for learning correlated equilibria in sequential extensions of normal-form games [23]. We created a multi-agent RL environment based on the normal-form Chicken game, a classic motivating example for the utility of correlated equilibria. In our experimental evaluation, we found that the MADDPG algorithm of [84] allowed agents to achieve social welfares corresponding to non-trivial correlated equilibria (MADDPG and MMADDPG in Figure 4.1). However, we discontinued this line of work due to negative results for some of the other algorithms that we designed.

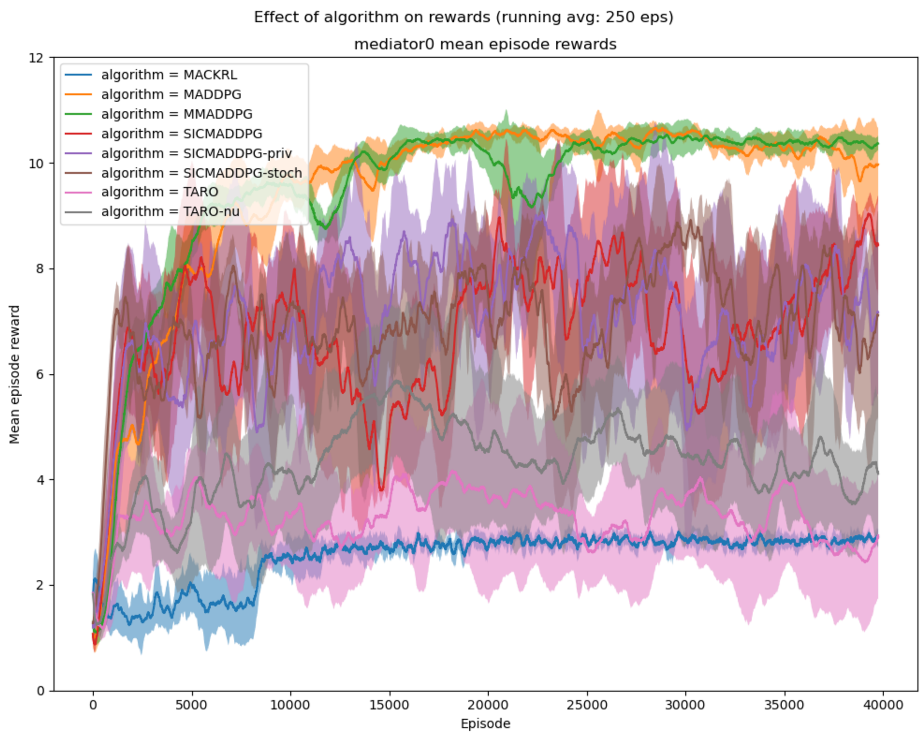


Figure 4.1: Mediator rewards (i.e., agent social welfare) for various RL algorithms in the multi-agent Chicken environment [23]. The maximum possible social welfare in the game is 10.5, which is approached by the MADDPG and MMADDPG algorithms (≈ 10).

Now, I propose to continue this line of work anew in the congestion game setting, by using RL algorithms to learn correlated equilibria in environments based on classic congestion games as well as the gig driving domain. Our results in the normal-form setting suggest that RL can also achieve non-trivial correlated equilibria for congestion games. However, compared to the normal-form setting, the optimal policy is less obvious in this setting, and thus during evaluation the RL algorithms would instead be benchmarked against potentially suboptimal correlated equilibria found by the algorithm of [105]. Not only would this proposed work thus lead to theoretically innovative contributions, but it could also be combined with the work presented in Section 3.1 to design practical AI decision aids that give gig drivers personalised yet coordinated recommendations.

4.2 Interpretability

Optimising Traffic Signal Phase Sequences with Multi-Agent RL

The experimental evaluation in Section 3.2 demonstrated that unconstrained RL algorithms for TSC are able to significantly improve performance over the existing fixed-time signal plan for the Strongsville, Ohio road network, but also that versions of these algorithms modified to respect signal plan constraints can still achieve nontrivial improvements. However, the constraints implemented in Section 3.2 are more restrictive than those imposed by Econolite's Edaptive system. In a real-world deployment of an RL algorithm for TSC, stakeholders may desire to relax some of these constraints in exchange for performance improvements (see Section 2.3.3).

Figure 4.2 shows the effect of removing individual constraints (while retaining the other constraints) on the performance of the constrained MPLight algorithm for the Strongsville road network. First, removing the minimum green constraint considerably degrades MPLight's performance. I conjecture that the algorithm is unable to learn a performant policy due to (1) rapid switching between phases due to a temporally brittle policy and (2) extension of the final phase to meet the cycle length constraint. Next, removing either the maximum green constraint or the cycle length constraints does not have a significant impact on performance, considering the level of variability between random seeds. Lastly, removing the phase sequence constraint results in significant gains in performance. In this proposed work, I will focus on the last of these constraints.

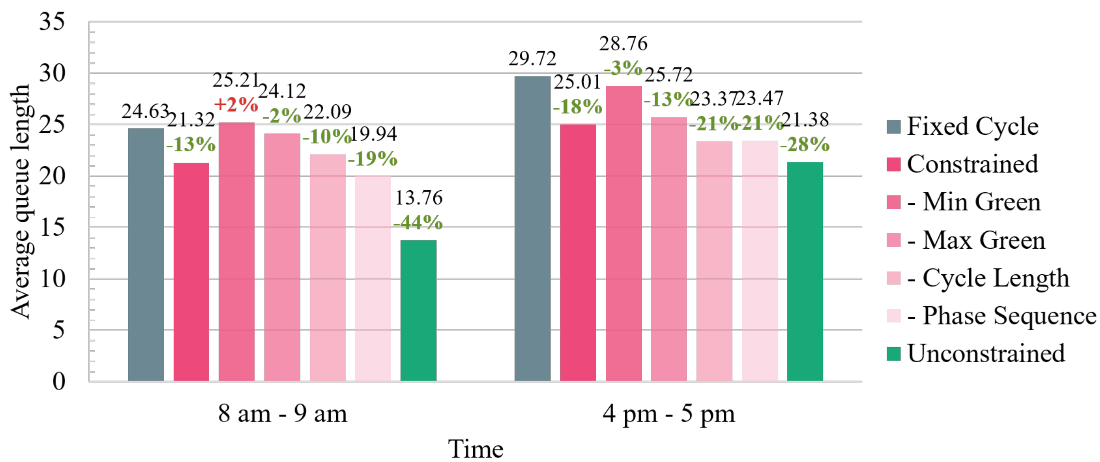


Figure 4.2: Effect of removing individual constraints on the test performance of the constrained MPLight algorithm on the Strongsville road network, based on traffic from May 17, 2023 at 5 pm.

Removing the phase sequence constraint entails allowing the RL algorithm to freely signal phases in any order. This would hinder drivers from predicting the next phase, which may cause safety issues if drivers begin accelerating before what they believe is the start of the green light for their movement (as observed by [152]). Therefore, in practice, the phase sequence should still be kept static — but it is not necessarily the case that the currently-implemented phase sequence in a given road network is optimal. Indeed, Edaptive is capable of heuristically optimising phase sequences, which usually involves conversion between *leading turn* (i.e., signalled for an extended time before the straight-through movement) and *lagging turn* (i.e., signalled for an extended time after the straight-through movement) phases in a base time-of-day plan. Other potential phase sequence modifications include skipping and combining phases.

Preliminary results show that, with fixed but randomly initialised phase sequences, the queue length achieved by the constrained MPLight algorithm can vary by up to 23% across random seeds. I anticipate that, even when phase sequences are kept fixed, there is still room for improvement from optimising them. How, then, could phase sequence optimisation be performed in a non-heuristic manner? A brute-force solution would entail training one RL policy for each possible permutation of phases. [94] proposed a simulated annealing algorithm, but combining their method with RL would also require training to convergence. I propose to solve phase sequence optimisation using a *branch-and-bound* procedure, in which information from a small set of RL trajectories can be used as a branching heuristic to select phases. This is complementary to recent work which models variable selection in branch-and-bound as an RL problem [39, 107], but their methods cannot be directly applied because the reward here is also generated by an RL policy.

4.3 Heterogeneity

Evaluating RL for Traffic Signal Control in the Real World

Ultimately, the goal of the work presented in Sections 3.2, 3.4 and 4.2 is to design a RL algorithm for TSC which can replace the heuristic optimisation routines in existing solutions such as Edaptive. But can such an algorithm be actually deployed? Over the past two years, I have been collaborating with Econolite, Path Master (a distributor of traffic control technologies in Ohio, Western Pennsylvania, Kentucky, and West Virginia); and the Traffic Management Centre (TMC) in Strongsville, Ohio, which currently uses Edaptive to control the Pearl Road and Royalton Road corridors. The work outlined in these sections has all been motivated by their pain points.

My "stretch goal" for this thesis is to integrate my theoretical and practical contributions in RL for TSC into a field test in Strongsville over the Pearl Road and Royalton Road corridors (Figure 3.3b). It would have two stages: (1) an offline evaluation where the RL policy simulates signalling actions using live detector data; and (2) an online evaluation where the RL policy's actions directly control traffic signals. I expect that some of the performance improvements delivered by RL in simulations of Strongsville can also translate to reality. However, a much more significant barrier will be ensuring that various stakeholders — particularly Strongsville's TMC personnel and contracted traffic engineers — will be able to trust RL-generated signal plans (see Section 2.3.3).

Currently, the stakeholders' most prominent concern is that the traffic patterns of our simulation generated from detector data are inconsistent with real-world traffic in Strongsville. While this does not diminish the capabilities of the constrained MPLight algorithm as demonstrated in Section 3.2, it does raise questions regarding whether these results can directly generalise to the real world (see Section 2.3.4). I am collaborating with an undergraduate student, Karen Wu, to calibrate the simulation; along with my algorithm design work, I anticipate that this will move the stakeholders to become increasingly amenable to a field test.

Chapter 5

Timeline

To conclude this thesis proposal, I now propose a timeline for completing the in-progress and proposed work introduced in Chapters 3 and 4.

- **Apr 2024**: **Thesis Proposal**
- **May–Aug 2024**

  Evaluating RL for Traffic Signal Control in the Real World (Section 4.3, Part I)

  Over the coming summer, I will conduct some preliminary, exploratory work to prepare for a field trial of RL algorithms for TSC in Strongsville, Ohio, by collaborating with two undergraduate students, Karen Wu and Nick Fettig. I will keep working with Karen to calibrate the Strongsville simulation. Despite the high-resolution detector data that is available from Strongsville, systematic undercounting issues mean that we will likely need to combine computer vision and optimisation techniques to generate consistent traffic flows. Meanwhile, I will work with Nick to expand my review paper on RL for TSC [24], combining an updated literature review with further stakeholder conversations to understand what developments must be made before stakeholders can trust RL-based signalling policies. Overall, the goal will be to ensure that the remainder of my in-progress and proposed work aligns with the needs of Econolite, Path Master, and the Strongsville TMC.
- **Sep–Dec 2024**

  Coordinating Decision Tree Surrogates for Multi-Agent RL (Section 3.3)

  In the fall, I will work with Stephanie Milani and Zhicheng Zhang to progress our work with MAVIPER to a publishable state. Although we have experimented extensively with baselines, some additional effort will be required to implement the best response-based sequential imitation learning algorithm that we envision. To demonstrate that it can strike a reasonable trade-off between performance and scalability, we will also thoroughly evaluate it on both [84]'s particle world and TSC environments. Our contributions will thus be primarily methodological in nature. To publish this work, we are aiming for a submission to a conference similar to ECML.

- **Sep–Dec 2024**

  Designing Cyclic Traffic Signal Controllers with Multi-Agent RL (Section 3.2)

  Optimising Traffic Signal Phase Sequences with Multi-Agent RL (Section 4.2)

  At the same time, I will continue to extend the MPLight algorithm to incorporate the remaining constraints that are needed to guarantee its behaviour to stakeholders: phase sequence optimisation (using a branch-and-bound algorithm) and split predetermination (using imitation learning). Between these two types of constraints, I anticipate that the former will lead to more significant methodological contributions, while the latter will lead to more significant practical contributions. My goal is to integrate them into a single publication describing the innovations needed to make MPLight deployable. To publish this work, I am aiming for a submission to a conference similar to COMPASS.

- **Jan–Apr 2025**

  Learning Equilibria for Game-Theoretic Models of Gig Driving (Section 4.1)

  Following the conclusion of my in-progress work, I plan to turn my attention fully towards this methodological project. Although I have some algorithmic ideas as well as an existing codebase that I can start from, the mixed results of my prior work suggest that this project will likely involve more exploration than the remainder of my in-progress and proposed work. To publish this work, I am aiming for a submission to a conference similar to AAAI.

The remainder of the timeline is contingent upon our progress towards building stakeholder trust on the TSC project. If Econolite, Path Master, and the Strongsville TMC all trust the simulated evaluations of my RL algorithms enough such that a field test is feasible, I will extend the timeline of my PhD by 6 months so that I can bring this work to completion:

- **May–Nov 2025**: **Stretch Goal**

  Evaluating RL for Traffic Signal Control in the Real World (Section 4.3, Part II)

  Working together with these stakeholders, I will translate the algorithmic developments introduced in Sections 3.2, 3.4 and 4.2 into a practical signal control algorithm that can be deployed end-to-end through live connections to Strongsville's infrastructure. This will likely require significant iteration with the stakeholders to ensure that they remain confident in the efficiency and safety of the algorithm at every step, hence the longer time estimate. However, if successful, this work has the potential to make significant and concrete social impacts.

- **Dec 2025**: **Thesis Defence**

Otherwise, I will follow the originally-planned timeline, and defer a field test to future work:

- **May 2025**: **Thesis Defence**

Bibliography

[1] Montasir Abbas, Darcy Bullock, and Larry Head. 2001. Real-Time Offset Transitioning Algorithm for Coordinating Traffic Signals. *Transportation Research Record* 1748 (2001), 26–39.

[2] Monireh Abdoos and Ana L.C. Bazzan. 2021. Hierarchical traffic signal optimization using reinforcement learning and traffic prediction with long-short term memory. *Expert Systems with Applications* 171 (2021), 114580.

[3] Sanjeevan Ahilan and Peter Dayan. 2019. Feudal Multi-Agent Hierarchies for Cooperative Reinforcement Learning. arXiv:1901.08492

[4] Lucas N. Alegre, Ana L.C. Bazzan, and Bruno C. da Silva. 2021. Quantifying the impact of non-stationarity in reinforcement learning-based traffic signal control. *PeerJ Computer Science* 7 (2021), e575.

[5] Tal Altshuler, Yaniv Altshuler, Rachel Katoshevski, and Yoram Shiftan. 2019. Modeling and Prediction of Ride-Sharing Utilization Dynamics. *Journal of Advanced Transportation* 2019 (2019), 6125798.

[6] Pablo Alvarez Lopez, Michael Behrisch, Laura Bieker-Walz, Jakob Erdmann, Yun-Pang Flötteröd, Robert Hilbrich, Leonhard Lücken, Johannes Rummel, Peter Wagner, and Evamarie Wießner. 2018. Microscopic Traffic Simulation using SUMO. In *Proceedings of the 21st International Conference on Intelligent Transportation Systems (ITSC '18)*. Institute of Electrical and Electronics Engineers, Piscataway, USA, 2575–2582.

[7] James Ault and Guni Sharon. 2021. Reinforcement learning benchmarks for traffic signal control. In *Proceedings of the 35th Conference on Neural Information Processing Systems, Datasets and Benchmarks Track (NeurIPS '21)*. NeurIPS, Virtual, 1–11.

[8] Robert J. Aumann. 1987. Correlated Equilibrium as an Expression of Bayesian Rationality. *Econometrica* 55, 1 (1987), 1–18.

[9] Jaume Barceló. 2010. Models, Traffic Models, Simulation, and Traffic Simulation. In *Fundamentals of Traffic Simulation*. Springer, New York, USA, 1–62.

[10] J. Barceló, E. Codina, J. Casas, J.L. Ferrer, and D. García. 2004. Microscopic Traffic Simulation: A Tool for the Design, Analysis and Evaluation of Intelligent Transport Systems. *Journal of Intelligent and Robotic Systems* 41 (2004), 173–203.

[11] Osbert Bastani, Yewen Pu, and Armando Solar-Lezama. 2018. Verifiable Reinforcement

Learning via Policy Extraction. In *Proceedings of the 32nd Conference on Neural Information Processing Systems (NeurIPS '18)*. NeurIPS, Montréal, Canada, 2494–2504.

[12] Emma Beede, Elizabeth Baylor, Fred Hersch, Anna Iurchenko, Lauren Wilcox, Paisan Ruamviboonsuk, and Laura M. Vardoulakis. 2020. A Human-Centered Evaluation of a Deep Learning System Deployed in Clinics for the Detection of Diabetic Retinopathy. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. ACM, Honolulu, USA, 1–12.

[13] Thor Berger, Carl Benedikt Frey, Guy Levin, and Santosh Rao Danda. 2020. Uber happy? Work and well-being in the 'Gig Economy'. *Economic Policy* 34, 99 (2020), 429–477.

[14] Kostas Bimpikis, Ozan Candogan, and Daniela Saban. 2019. Spatial Pricing in Ride-Sharing Networks. *Operations Research* 67, 3 (2019), 744–769.

[15] Noam Brown and Tuomas Sandholm. 2017. Superhuman AI for heads-up no-limit poker: Libratus beats top professionals. *Science* 359, 6374 (2017), 418–424.

[16] Eleanor R. Burgess, Ivana Jankovic, Melissa Austin, Nancy Cai, Adela Kapuścińska, Suzanne T. Currie, J. Marc Overhage, Erika S. Poole, and Jofsh Kaye. 2023. Healthcare AI Treatment Decision Support: Design Principles to Enhance Clinician Adoption and Trust. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. ACM, Hamburg, Germany, 1–19.

[17] José António Capela Dias, Penousal Machado, and Francisco Câmara Pereira. 2013. Simulating the Impact of Drivers' Personality on City Transit. In *Proceedings of the 13th World Conference on Transport Research (WCTR '13)*. World Conference on Transport Research Society, Leeds, UK, 1–13.

[18] Nicholas S. Caros, Xiaotong Guo, Yunhan Zheng, and Jinhua Zhao. 2023. The impacts of remote work on travel: insights from nearly three years of monthly surveys. arXiv:2303.06186

[19] Ngai Keung Chan and Lee Humphreys. 2018. Mediatization of Social Space and the Case of Uber Drivers. *Media and Communication* 6, 2 (2018), 29–38.

[20] Chacha Chen, Hua Wei, Nan Xu, Guanjie Zheng, Ming Yang, Yuanhao Xiong, Kai Xu, and Zhenhui Li. 2020. Toward A Thousand Lights: Decentralized Deep Reinforcement Learning for Large-Scale Traffic Signal Control. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI '20)*. AAAI, New York, USA, 3414–3421.

[21] M. Keith Chen, Peter E. Rossi, Judith A. Chevalier, and Emily Oehlsen. 2019. The Value of Flexible Work: Evidence from Uber Drivers. *Journal of Political Economy* 127, 6 (2019), 2735–2794.

[22] Rex Chen, Kathleen M. Carley, Fei Fang, and Norman Sadeh. 2023. Purpose in the Machine: Do Traffic Simulators Produce Distributionally Equivalent Outcomes for Reinforcement Learning Applications?. In *Proceedings of the 2023 Winter Simulation Conference (WSC '23)*. ACM, San Antonio, USA, 1842–1853.

[23] Rex Chen, James Cunningham, and Justin Kiefel. 2021. *Learning to Signal Correlated Equilibria with Multi-Agent Reinforcement Learning*. Technical Report. Carnegie Mellon

University. 16 pages.

[24] Rex Chen, Fei Fang, and Norman Sadeh. 2022. The Real Deal: A Review of Challenges and Opportunities in Moving Reinforcement Learning-Based Traffic Signal Control Systems Towards Reality. In *Proceedings of the 12th International Workshop on Agents in Traffic and Transportation (ATT '22)*. CEUR, Vienna, Austria, 1–21.

[25] Rex Chen, Ruiyi Wang, Fei Fang, and Norman Sadeh. 2024. Missing Pieces: How Framing Uncertainty Impacts Longitudinal Trust in AI Decision Aids — A Gig Driver Case Study. arXiv:2404.06432

[26] S.M. Chin, O. Franzese, D.L. Greene, H.L. Hwang, and R.C. Gibson. 2004. *Temporary losses of highway capacity and impacts on performance: Phase 2*. Technical Report ORNL/TM-2004/209. Oak Ridge National Laboratory. 107 pages.

[27] Peter Christensen and Adam Osman. 2023. *The Demand for Mobility: Evidence from an Experiment with Uber Riders*. Working Paper 31330. NBER. 48 pages.

[28] Hyuck David Chung, Yue Maggie Zhou, and Christine Choi. 2022. When Uber Eats its Own Business, and That of its Competitors Too. *Academy of Management Proceedings* 2022, 1 (2022), 15263.

[29] Simon Danner, Matthias Pfromm, and Klaus Bengler. 2020. Does Information on Automated Driving Functions and the Way of Presenting It before Activation Influence Users' Behavior and Perception of the System? *Information* 11, 1 (2020), 54.

[30] Christopher M. Day and Darcy M. Bullock. 2011. Computational Efficiency of Alternative Algorithms for Arterial Offset Optimization. *Transportation Research Record* 2259 (2011), 37–47.

[31] Christopher M. Day and Darcy M. Bullock. 2011. *Optimization of Offsets and Cycle Length Using High Resolution Signal Event Data*. Working Paper SPR-3409. Joint Transportation Research Program. 36 pages.

[32] Taylor de O. Antes, Ana L.C. Bazzan, and Anderson Rocha Tavares. 2022. Information upwards, recommendation downwards: reinforcement learning with hierarchy for traffic signal control. *Procedia Computer Science* 201 (2022), 24–31.

[33] Arjan de Ruijter, Oded Cats, and Hans van Lint. 2024. Ridesourcing platforms thrive on socio-economic inequality. *Scientific Reports* 14 (2024), 7371.

[34] Augustin Degas, Mir Riyanul Islam, Christophe Hurter, Shaibal Barua, Hamidur Rahman, Minesh Poudel, Daniele Ruscio, Mobyen Uddin Ahmed, Shahina Begum, Md Aquif Rahman, Stefano Bonelli, Giulia Cartocci, Gianluca Di Flumeri, Gianluca Borghini, Fabio Babiloni, and Pietro Aricó. 2022. A Survey on Artificial Intelligence (AI) and eXplainable AI in Air Traffic Management: Current Trends and Development with Future Research Trajectory. *Applied Sciences* 12, 3 (2022), 1295.

[35] Wenlu Du, Junyi Ye, Jingyi Gu, Jing Li, Hua Wei, and Guiling Wang. 2023. SafeLight: A Reinforcement Learning Method toward Collision-Free Traffic Signal Control. In *Proceedings of the 37th AAAI Conference on Artificial Intelligence, Special Track on Safe and*

Robust AI (AAAI '23). AAAI, Washington, DC, USA, 14801–14810.

[36] Econolite. 2023. *Centracs Edaptive datasheet*. Document CNTRC-EDPTV 11.2023. Econolite Product Resource Library. 2 pages.

[37] Myungeun Eom and Byung-In Kim. 2020. The traffic signal control problem for intersections: a review. *European Transport Research Review* 12 (2020), 50.

[38] Mohamed Essa and Tarek Sayed. 2020. Self-learning adaptive traffic signal control for real-time safety optimization. *Accident Analysis & Prevention* 146 (2020), 105713.

[39] Marc Etheve, Zacharie Alès, Côme Bissuel, Olivier Juan, and Safia Kedad-Sidhoum. 2020. Reinforcement Learning for Variable Selection in a Branch and Bound Algorithm. In *Proceedings of the 2020 International Conference on Integration of Constraint Programming, Artificial Intelligence, and Operations Research (CPAIOR '20)*. Springer, Virtual, 176–185.

[40] Euthenics and TranSystems. 2023. *Preliminary Feasibility Study: CUY/MED Traffic Study*. Technical Report 116069. City of Strongsville. 759 pages.

[41] Alex Fabrikant, Christos Papadimitriou, and Kunal Talwar. 2004. The complexity of pure Nash equilibria. In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing (STOC '04)*. ACM, Chicago, USA, 604–612.

[42] Nahid Parvez Farazi, Bo Zou, Tanvir Ahamed, and Limon Barua. 2021. Deep reinforcement learning in transportation research: A review. *Transportation Research Interdisciplinary Perspectives* 11 (2021), 100425.

[43] Andrea Ferrario and Michele Loi. 2022. How Explainability Contributes to Trust in AI. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22)*. ACM, Seoul, South Korea, 1457–1466.

[44] Deepeka Garg, Maria Chli, and George Vogiatzis. 2019. Traffic3D: A Rich 3D-Traffic Environment to Train Intelligent Agents. In *Proceedings of the 19th International Conference on Computational Science (ICCS '19)*. Springer, New York, USA, 749–755.

[45] Deepeka Garg, Maria Chli, and George Vogiatzis. 2022. Fully-Autonomous, Vision-based Traffic Signal Control: from Simulation to Reality. In *Proceedings of the 21th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS '22)*. IFAAMAS, Auckland, New Zealand, 454–462.

[46] Nikhil Garg and Hamid Nazerzadeh. 2022. Driver Surge Pricing. *Management Science* 68, 5 (2022), 3219–3235.

[47] Timothy Geary and David Danks. 2019. Balancing the Benefits of Autonomous Vehicles. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society (AIES '19)*. ACM, Honolulu, USA, 1–6.

[48] Steven R. Gehrke. 2020. Uber service area expansion in three major American cities. *Journal of Transport Geography* 86 (2020), 102752.

[49] David Gibson, Milton K. (Pete) Mills, and Doug Rekenthaler Jr. 1998. Staying in The Loop: The Search for Improved Reliability of Traffic Sensing Systems Through Smart Test Instruments. *Public Roads* 62, 2 (1998), 47–51.

[50] Yaobang Gong, Mohamed Abdel-Aty, Jinghui Yuan, and Qing Cai. 2020. Multi-Objective reinforcement learning approach for improving safety at intersections with adaptive traffic signal control. *Accident Analysis & Prevention* 144 (2020), 105655.

[51] Robert L. Gordon and Warren Tighe. 2005. *Traffic Control Systems Handbook*. Federal Highway Administration.

[52] Lance R. Grenzeback, William R. Reilly, Paul O. Roberts, and Joseph R. Stowers. 1990. Urban Freeway Gridlock Study: Decreasing the Effects of Large Trucks on Peak-Period Urban Freeway Congestion. *Transportation Research Record* 1256 (1990), 16–26.

[53] Kathleen Griesbach, Adam Reich, Luke Elliott-Negri, and Ruth Milkman. 2019. Algorithmic Control in Platform Food Delivery Work. *Socius* 5 (2019), 1–15.

[54] Harish Guda and Upender Subramanian. 2019. Your Uber Is Arriving: Managing On-Demand Workers Through Surge Pricing, Forecast Communication, and Worker Incentives. *Management Science* 65, 5 (2019), 1995–2014.

[55] Jonathan V. Hall and Alan B. Krueger. 2018. An Analysis of the Labor Market for Uber's Driver-Partners in the United States. *ILR Review* 71, 3 (2018), 705–732.

[56] Chengzheng Hang, Zhenfei Liu, Yujing Wang, Caiyi Hu, Yuelong Su, and Zhenning Dong. 2019. Sharing diseconomy: impact of the subsidy war of ride-sharing companies on urban congestion. *International Journal of Logistics Research and Applications* 22, 5 (2019), 491–500.

[57] Benjamin V. Hanrahan, Ning F. Ma, and Chien Wen Yuan. 2017. The Roots of Bias on Uber. In *Proceedings of the 15th European Conference on Computer-Supported Cooperative Work (ECSCW '17)*. EUSSET, Sheffield, UK, 1–17.

[58] Rie Helene (Lindy) Hernandez, Qiurong Song, Yubo Kou, and Xinning Gui. 2024. "At the end of the day, I am accountable": Gig Workers' Self-Tracking for Multi-Dimensional Accountability Management. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24)*. ACM, Honolulu, USA, 1–20.

[59] Hannah Horner, Jennifer Pazour, and John E. Mitchell. 2021. Optimizing driver menus under stochastic selection behavior for ridesharing and crowdsourced delivery. *Transportation Research Part E: Logistics and Transportation Review* 153 (2021), 102419.

[60] Zhiran Huang, Becky P.Y. Loo, and Kay W. Axhausen. 2023. Travel behaviour changes under Work-from-home (WFH) arrangements during COVID-19. *Travel Behaviour and Society* 30 (2023), 202–211.

[61] Hyeonjun Hwang, Clifford Winston, and Jia Yan. 2020. *Measuring the Benefits of Ridesharing Services to Urban Travelers: The Case of The San Francisco Bay Area*. Working Paper 70. Hutchins Center. 18 pages.

[62] Adrianto Ravi Ibrahim, Ahmet Cetinkaya, and Masako Kishida. 2022. Modeling Heterogeneous Transportation Services by Two-Stage Congestion Games. In *Proceedings of the 2022 European Control Conference (ECC '22)*. IEEE, London, UK, 2117–2123.

[63] Tatsuya Iwase and Takahiro Shiga. 2016. Pure Nash Equilibrium and Coordination of Play-

ers in Ride Sharing Games. arXiv:1604.00710

[64] Lakshmi Shankar Iyer. 2021. AI enabled applications towards intelligent transportation. *Transportation Engineering* 5 (2021), 100083.

[65] Maia Jacobs, Jeffrey He, Melanie F. Pradier, Barbara Lam, Andrew C. Ahn, Thomas H. McCoy, Roy H. Perlis, Finale Doshi-Velez, and Krzysztof Z. Gajos. 2021. Designing AI for Trust and Collaboration in Time-Constrained Medical Decisions: A Sociotechnical Lens. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. ACM, Yokohama, Japan, 1–14.

[66] Vindula Jayawardana, Anna Landler, and Cathy Wu. 2021. Mixed Autonomous Supervision in Traffic Signal Control. In *Proceedings of the 2021 IEEE 24th International Conference on Intelligent Transportation Systems (ITSC '21)*. IEEE, Indianapolis, USA, 1767–1773.

[67] Shan Jiang, Le Chen, Alan Mislove, and Christo Wilson. 2018. On Ridesharing Competition and Accessibility: Evidence from Uber, Lyft, and Taxi. In *Proceedings of the 2018 World Wide Web Conference (WWW '18)*. ACM, Lyon, France, 863–872.

[68] Junchen Jin and Xiaoliang Ma. 2019. A Multi-Objective Agent-Based Control Approach With Application in Intelligent Traffic Signal System. *IEEE Transactions on Intelligent Transportation Systems* 20, 10 (2019), 3900–3912.

[69] Min-Wook Kang, Moynur Rahman, and Joyoung Lee. 2020. Determination and Utilization of Dilemma Zone Length and Location for Safety Assessment of Rural High-Speed Signalized Intersections. *Transportation Research Record* 2674, 4 (2020), 272–280.

[70] Hassan Ali Khan, Muhammad Shahzad, Hassan Iqbal, and Guoliang Jin. 2022. RMS: Removing Barriers to Analyze the Availability and Surge Pricing of Ridesharing Services. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. ACM, New Orleans, USA, 1–18.

[71] Rafal Kocielnik, Saleema Amershi, and Paul N. Bennett. 2019. Will You Accept an Imperfect AI? Exploring Designs for Adjusting End-user Expectations of AI Systems. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, Glasgow, UK, 1–14.

[72] Peter Koonce, Lee Rodegerdts, Kevin Lee, Shaun Quayle, Scott Beaird, Cade Braud, Jim Bonneson, Phil Tarnoff, and Tom Urbanik. 2008. *Traffic Signal Timing Manual*. Federal Highway Administration.

[73] Stefan Krauß. 1998. *Microscopic Modeling of Traffic Flow: Investigation of Collision Free Vehicle Dynamics*. Ph. D. Dissertation. German Aerospace Center, Cologne.

[74] Johannes Kunkel, Tim Donkers, Lisa Michael, Catalin-Mihai Barbu, and Jürgen Ziegler. 2019. Let Me Explain: Impact of Personal and Impersonal Explanations on Trust in Recommender Systems. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, Scotland, UK, 1–12.

[75] Marc Lanctot, Vinicius Zambaldi, Audrunas Gruslys, Angeliki Lazaridou, Karl Tuyls, Julien Perolat, David Silver, and Thore Graepel. 2017. A Unified Game-Theoretic Approach to

Multiagent Reinforcement Learning. In *Proceedings of the 31st Conference on Neural Information Processing Systems (NeurIPS '17)*. NeurIPS, Long Beach, USA, 4193–4206.

[76] Min Kyung Lee, Danyel Kusbit, Evan Metsky, and Laura Dabbish. 2015. Working with Machines: The Impact of Algorithmic and Data-Driven Management on Human Workers. In *Proceedings of the 2015 CHI Conference on Human Factors in Computing Systems (CHI '15)*. ACM, Seoul, South Korea, 1–13.

[77] Yanzhe (Murray) Lei, Stefanus Jasin, Jingyi Wang, Houtao Deng, and Jagannath Putrevu. 2020. Dynamic Workforce Acquisition for Crowdsourced Last-Mile Delivery Platforms. arXiv:3532844

[78] Yongfu Li and Dihua Sun. 2012. Microscopic car-following model for the traffic flow: the state of the art. *Journal of Control Theory and Applications* 10 (2012), 133–143.

[79] Ziru Li, Chen Liang, Yili Hong, and Zhongju Zhang. 2022. How Do On-demand Ridesharing Services Affect Traffic Congestion? The Moderating Role of Urban Compactness. *Production and Operations Management* 31, 1 (2022), 239–258.

[80] Lyuchao Liao, Jierui Liu, Xinke Wu, Fumin Zou, Jengshyang Pan, Qi Sun, Shengbo Eben Li, and Maolin Zhang. 2020. Time Difference Penalized Traffic Signal Timing by LSTM Q-Network to Balance Safety and Capacity at Intersections. *IEEE Access* 8 (2020), 80086–80096.

[81] Todd Litman. 2023. *Autonomous Vehicle Implementation Predictions: Implications for Transport Planning*. Technical Report. Victoria Transport Policy Institute. 49 pages.

[82] Ying Liu, Lei Liu, and Wei-Peng Chen. 2017. Intelligent Traffic Light Control Using Distributed Multi-agent Q Learning. In *Proceedings of the 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC '17)*. IEEE, Yokohama, Japan, 1–8.

[83] Silas C. Lobo, Stefan Neumeier, Evelio M. G. Fernandez, and Christian Facchi. 2020. InTAS — The Ingolstadt Traffic Scenario for SUMO. In *Proceedings of the 2020 SUMO User Conference (SUMO '20)*. German Aerospace Center, Cologne, 1–20.

[84] Ryan Lowe, Yi Wu, Aviv Tamar, Jean Harb, Pieter Abbeel, and Igor Mordatch. 2017. Multi-Agent Actor-Critic for Mixed Cooperative-Competitive Environments. In *Proceedings of the 31st Conference on Neural Information Processing Systems (NeurIPS '17)*. NeurIPS, Long Beach, USA, 6379–6390.

[85] Alice Lu, Peter Frazier, and Oren Kislev. 2018. Surge Pricing Moves Uber's Driver Partners. arXiv:3180246

[86] Kai Lu, Jiangtao Liu, Xuesong Zhou, and Baoming Han. 2020. A Review of Big Data Applications in Urban Transit Systems. *IEEE Transactions on Intelligent Transportation Systems* 22, 5 (2020), 2535–2552.

[87] Scott M. Lundberg and Su-In Lee. 2017. A Unified Approach to Interpreting Model Predictions. In *Proceedings of the 31st Conference on Neural Information Processing Systems (NeurIPS '17)*. NeurIPS, Long Beach, USA, 4765–4774.

[88] Hongyao Ma, Fei Fang, and David C. Parkes. 2022. Spatio-Temporal Pricing for Rideshar-

ing Platforms. *Operations Research* 70, 2 (2022), 1025–1041.

[89] Jinming Ma and Feng Wu. 2020. Feudal Multi-Agent Deep Reinforcement Learning for Traffic Signal Control. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS '20)*. IFAAMAS, Auckland, New Zealand, 816–824.

[90] Ning F. Ma, Chien Wen Yuan, Moojan Ghafurian, and Benjamin V. Hanrahan. 2018. Using Stakeholder Theory to Examine Drivers' Stake in Uber. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, Montréal, Canada, 1–12.

[91] Zian Ma, Chengcheng Xu, Yuheng Kan, Maonan Wang, and Wei Wu. 2021. Adaptive Coordinated Traffic Control for Arterial Intersections based on Reinforcement Learning. In *Proceedings of the 2021 IEEE 24th International Conference on Intelligent Transportation Systems (ITSC '21)*. IEEE, Indianapolis, USA, 2562–2567.

[92] Hao Mei, Xiaoliang Lei, Longchao Da, Bin Shi, and Hua Wei. 2023. Libsignal: an open library for traffic signal control. *Machine Learning* 114 (2023), 1–37.

[93] Hao Mei, Junxian Li, Bin Shi, and Hua Wei. 2023. Reinforcement learning approaches for traffic signal control under missing data. In *Proceedings of the 32nd International Joint Conference on Artificial Intelligence (IJCAI '23)*. ACM, Macao, 2261–2269.

[94] Silvio Memoli, Giulio E. Cantarella, Stefano de Luca, and Roberta Di Pace. 2017. Network signal setting design with stage sequence optimisation. *Transportation Research Part B: Methodological* 100 (2017), 20–42.

[95] Wei Miao, Yiting Deng, Wei Wang, Yongdong Liu, and Christopher S. Tang. 2023. The effects of surge pricing on driver behavior in the ride-sharing market: Evidence from a quasi-experiment. *Journal of Operations Management* 69, 5 (2023), 794–822.

[96] Vicente Milanés and Steven E. Shladover. 2014. Modeling cooperative and autonomous adaptive cruise control dynamic responses using experimental data. *Transportation Research Part C: Emerging Technologies* 48 (2014), 285–300.

[97] Stephanie Milani, Zhicheng Zhang, Nicholay Topin, Zheyuan Ryan Shi, Charles Kamhoua, Evangelos E. Papalexakis, and Fei Fang. 2022. MAVIPER: Learning Decision Tree Policies for Interpretable Multi-agent Reinforcement Learning. In *Proceedings of the 2022 European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD '22)*. ACM, Grenoble, France, 251–266.

[98] Arthur Müller, Vishal Rangras, Georg Schnittker, Michael Waldmann, Maxim Friesen, Tobias Ferfers, Lukas Schreckenberg, Florian Hufen, Jürgen Jasperneite, and Marco Wiering. 2021. Towards Real-World Deployment of Reinforcement Learning for Traffic Signal Control. In *Proceedings of the 20th IEEE International Conference on Machine Learning and Applications (ICMLA '21)*. IEEE, Pasadena, USA, 507–514.

[99] Arthur Müller and Matthia Sabatelli. 2023. Bridging the Reality Gap of Reinforcement Learning based Traffic Signal Control using Domain Randomization and Meta Learning. In

Proceedings of the 2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC '23). IEEE, Bilbao, Spain, 5271–5278.

[100] Ziaul Haque Munim, Mariia Dushenko, Veronica Jaramillo Jimenez, Mohammad Hassan Shakil, and Marius Imset. 2020. Big data and artificial intelligence in the maritime industry: a bibliometric review and future research directions. *Maritime Policy & Management* 47, 5 (2020), 577–597.

[101] Tomoki Nishi, Keisuke Otaki, Keiichiro Hayakawa, and Takayoshi Yoshimura. 2018. Traffic Signal Control Based on Reinforcement Learning with Graph Convolutional Neural Nets. In *Proceedings of the 2018 IEEE 21st International Conference on Intelligent Transportation Systems (ITSC '18)*. IEEE, Maui, USA, 877–883.

[102] Mohammad Noaeen, Atharva Naik, Liana Goodman, Jared Crebo, Taimoor Abrar, Zahra Shakeri Hossein Abad, Ana L.C. Bazzan, and Behrouz Far. 2022. Reinforcement learning in urban network traffic signal control: A systematic literature review. *Expert Systems with Applications* 199 (2022), 116830.

[103] Mojtaba Norouzi, Monireh Abdoos, and Ana L. C. Bazzan. 2021. Experience classification for transfer learning in traffic signal control. *The Journal of Supercomputing* 77 (2021), 780–795.

[104] Hao Yi Ong, Daniel Freund, and Davide Crapis. 2021. Driver Positioning and Incentive Budgeting with an Escrow Mechanism for Ride-Sharing Platforms. *INFORMS Journal on Applied Analytics* 51, 5 (2021), 373–390.

[105] Christos H. Papadimitriou and Tim Roughgarden. 2008. Computing Correlated Equilibria in Multi-Player Games. *J. ACM* 55, 3 (2008), 1–29.

[106] Darsh Parekh, Nishi Poddar, Aakash Rajpurkar, Manisha Chahal, Neeraj Kumar, Gyanendra Prasad Joshi, and Woong Cho. 2022. A Review on Autonomous Vehicles: Progress, Methods and Challenges. *Electronics* 11, 14 (2022), 2162.

[107] Christopher W. F. Parsonson, Alexandre Laterre, and Thomas D. Barrett. 2023. Reinforcement Learning for Branch-and-Bound Optimisation Using Retrospective Trajectories. In *Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI '23)*. AAAI, Washington, DC, USA, 4061–4069.

[108] Snehal Prabhudesai, Leyao Yang, Sumit Asthana, Xun Huan, Q. Vera Liao, and Nikola Banovic. 2023. Understanding Uncertainty: How Lay Decision-makers Perceive and Interpret Uncertainty in Human-AI Decision Making. In *Proceedings of the ACM Conference on Intelligent User Interfaces 2023 (IUI '23)*. ACM, Sydney, Australia, 379–396.

[109] Md. Mokhlesur Rahman, Pooya Najaf, Milton Gregory Fields, and Jean-Claude Thill. 2022. Traffic congestion and its urban scale factors: Empirical evidence from American urban areas. *International Journal of Sustainable Transportation* 16, 5 (2022), 406–421.

[110] Pouria Razzaghi, Amin Tabrizian, Wei Guo, Shulu Chen, Abenezer Taye, Ellis Thompson, Alexis Bregeon, Ali Baheri, and Peng Wei. 2022. A Survey on Reinforcement Learning in Aviation Applications. arXiv:2211.02147

[111] Duncan Rheingans-Yoo, Scott Duke Kominers, Hongyao Ma, and David C. Parkes. 2019. Ridesharing with driver location preferences. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI '19)*. ACM, Macao, 557–564.

[112] Avery Rhodes, Darcy M. Bullock, James R. Sturdevant, and Zachary Thomas Clark. 2005. *Evaluation of Stop Bar Video Detection Accuracy at Signalized Intersections*. Technical Report FHWA/IN/JTRP-2005/28. Joint Transportation Research Program, Indiana Department of Transportation and Purdue University. 418 pages.

[113] Brishen Rogers. 2015. The Social Costs of Uber. *University of Chicago Law Review* 82 (2015), 85–102.

[114] Stéphane Ross, Geoffrey J. Gordon, and J. Andrew Bagnell. 2011. A Reduction of Imitation Learning and Structured Prediction to No-Regret Online Learning. In *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics (AISTATS '11)*. PMLR, Fort Lauderdale, USA, 627–635.

[115] Stephen Russell, Ira S. Moskowitz, and Adrienne Raglin. 2017. Human Information Interaction, Artificial Intelligence, and Errors. In *Autonomy and Artificial Intelligence: A Threat or Savior?* Springer, Berlin, 71–101.

[116] Bruce Schaller. 2021. Can sharing a ride make for less traffic? Evidence from Uber and Lyft and implications for cities. *Transport Policy* 102 (2021), 1–10.

[117] Nicolas Scharowski, Michaela Benk, Swen J. Kühne, Léane Wettstein, and Florian Brühlmann. 2023. Certification Labels for Trustworthy AI: Insights From an Empirical Mixed-Method Study. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23)*. ACM, Chicago, USA, 248–262.

[118] Max Schemmer, Niklas Kühl, Carina Benz, Andrea Bartos, and Gerhard Satzger. 2023. Appropriate Reliance on AI Advice: Conceptualization and the Effect of Explanations. In *Proceedings of the 28th International Conference on Intelligent User Interfaces (IUI '23)*. ACM, Sydney, Australia, 410–422.

[119] David Schrank, Luke Albert, Bill Eisele, and Tim Lomax. 2021. *2021 Urban Mobility Report*. Technical Report. Texas A&M Transportation Institute. 70 pages.

[120] Susan Shaheen and Nelson Chan. 2016. Mobility and the Sharing Economy: Potential to Facilitate the First- and Last-Mile Public Transit Connections. *Built Environment* 42, 4 (2016), 573–588.

[121] Steven G. Shelby, Darcy M. Bullock, Doug Gettman, Raj S. Ghaman, Ziad A. Sabra, and Nils Soyke. 2008. An Overview and Performance Evaluation of ACS Lite — A Low Cost Adaptive Signal Control System. In *Proceedings of the 87th TRB Annual Meeting (TRB '08)*. TRB, Washington, DC, USA, 1–17.

[122] Yang Shi, Zhenbo Wang, Tim J. LaClair, Chieh (Ross) Wang, Yunli Shao, and Jinghui Yuan. 2023. A Novel Deep Reinforcement Learning Approach to Traffic Signal Control with Connected Vehicles. *Applied Sciences* 13, 4 (2023), 2750.

[123] Andrew Silva, Taylor Killian, Ivan Rodriguez Jimenez, Sung-Hyun Son, and Matthew Gom-

bolay. 2020. Optimization Methods for Interpretable Differentiable Decision Trees in Reinforcement Learning. In *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics (AISTATS '20)*. PMLR, Palermo, Italy, 1855–1865.

[124] David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis. 2016. Mastering the game of Go with deep neural networks and tree search. *Nature* 529 (2016), 484–489.

[125] Anubha Singh, Patricia Garcia, and Silvia Lindtner. 2023. Old Logics, New Technologies: Producing a Managed Workforce on On-Demand Service Platforms. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. ACM, Hamburg, Germany, 1–15.

[126] Stephen Smith, Gregory Barlow, Xiao-Feng Xie, and Zack Rubinstein. 2013. Smart urban signal networks: Initial application of the SURTRAC adaptive traffic signal control system. In *Proceedings of the 23rd International Conference on Automated Planning and Scheduling (ICAPS '13)*. PKP, Rome, Italy, 434–442.

[127] Elizabeth Solberg, Magnhild Kaarstad, Maren H. Rø Eitrheim, Rossella Bisio, Kine Reegard, and Marten Bloch. 2022. A Conceptual Model of Trust, Perceived Risk, and Reliance on AI Decision Aids. *Group & Organization Management* 47, 2 (2022), 187–222.

[128] Aleksandar Stevanovic, Cameron Kergaye, and Peter T. Martin. 2009. SCOOT and SCATS: A Closer Look into Their Operations. In *Proceedings of the 88th TRB Annual Meeting (TRB '09)*. TRB, Washington, DC, USA, 1–17.

[129] Erose Sthapit and Peter Björk. 2019. Sources of value co-destruction: Uber customer perspectives. *Tourism Review* 74, 4 (2019), 780–794.

[130] Daniel (Jian) Sun and Alexandra Kondyli. 2010. Modeling Vehicle Interactions during Lane-Changing Behavior on Arterial Streets. *Computer-Aided Civil and Infrastructure Engineering* 25 (2010), 557–571.

[131] Srinivasa Sunkari, Apoorba Bibeka, Nadeem Chaudhary, and Kevin Balke. 2019. *Impact of Traffic Signal Controller Settings on the Use of Advanced Detection Devices*. Technical Report FHWA/TX-18/0-6934-R1. Texas A&M Transportation Institute. 67 pages.

[132] Kai Liang Tan, Anuj Sharma, and Soumik Sarkar. 2020. Robust Deep Reinforcement Learning for Traffic Signal Control. *Journal of Big Data Analytics in Transportation* 2, 3 (2020), 263–274.

[133] Zhi Ming Tan, Nikita Aggarwal, Josh Cowls, Jessica Morley, Mariarosaria Taddeo, and Luciano Floridi. 2021. The ethical debate about the gig economy: A review and critical analysis. *Technology and Society* 65 (2021), 101594.

[134] Pankaj P. Tasgaonkar, Rahul Dev Garg, and Pradeep Kumar Garg. 2020. Vehicle Detection and Traffic Estimation with Sensors Technologies for Intelligent Transportation Systems.

Sensing and Imaging 21 (2020), 29.

[135] Larry W. Thomas. 2014. *Effect of MUTCD on Tort Liability of Government Transportation Agencies*. Legal Research Digest 63. National Cooperative Highway Research Program. 99 pages.

[136] Manish Tripathy, Jiaru Bai, and H. Sebastian (Seb) Heese. 2022. Driver collusion in ride-hailing platforms. *Decision Sciences* 54, 4 (2022), 434–446.

[137] Takane Ueno, Yuto Sawa, Yeongdae Kim, Jacqueline Urakami, Hiroki Oura, and Katie Seaborn. 2022. Trust in Human-AI Interaction: Scoping Out Models, Measures, and Methods. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems (CHI EA '22)*. ACM, New Orleans, USA, 1–7.

[138] United Nations. 2015. *Sustainable Transport*. United Nations Department of Economic and Social Affairs. `https://sdgs.un.org/topics/sustainable-transport`

[139] Oriol Vinyals, Igor Babuschkin, Wojciech M. Czarnecki amd Michaël Mathieu, Andrew Dudzik, Junyoung Chung, David H. Choi, Richard Powell, Timo Ewalds, Petko Georgiev, Junhyuk Oh, Dan Horgan, Manuel Kroiss, Ivo Danihelka, Aja Huang, Laurent Sifre, Trevor Cai, John P. Agapiou, Max Jaderberg, Alexander S. Vezhnevets, Rémi Leblond, Tobias Pohlen, Valentin Dalibard, David Budden, Yury Sulsky, James Molloy, Tom L. Paine, Caglar Gulcehre, Ziyu Wang, Tobias Pfaff, Yuhuai Wu, Roman Ring, Dani Yogatama, Dario Wünsch, Katrina McKinney, Oliver Smith, Tom Schaul, Timothy Lillicrap, Koray Kavukcuoglu, Demis Hassabis, Chris Apps, and David Silver. 2019. Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature* 575 (2019), 350–354.

[140] Peter Wagner. 2008. Action point models of human driving behaviour. In *Proceedings of the 2008 Traffic Simulation Workshop*. Institute of Highway Engineering and Transport Planning, Graz University of Technology, Graz, Austria, 1.

[141] Caroline Wang, Ishan Durugkar, Elad Liebman, and Peter Stone. 2023. DM^2: Decentralized Multi-Agent Reinforcement Learning via Distribution Matching. In *Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI '23)*. AAAI, Washington, DC, USA, 11699–11707.

[142] Dakuo Wang, Liuping Wang, Zhan Zhang, Ding Wang, Haiyi Zhu, Yvonne Gao, Xiangmin Fan, and Feng Tian. 2021. "Brilliant AI Doctor" in Rural Clinics: Challenges in AI-Powered Clinical Decision Support System Deployment. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. ACM, Yokohama, Japan, 1–18.

[143] Min Wang, Libing Wu, Jianxin Li, and Liu He. 2021. Traffic Signal Control With Reinforcement Learning Based on Region-Aware Cooperative Strategy. *IEEE Transactions on Intelligent Transportation Systems* 23, 7 (2021), 6774–6785.

[144] Xiaoyu Wang, Baher Abdulhai, and Scott Sanner. 2023. A Critical Review of Traffic Signal Control and A Novel Unified View of Reinforcement Learning and Model Predictive Control Approaches for Adaptive Traffic Signal Control. In *Handbook on Artificial Intelligence and Transport*. Edward Elgar, Northampton, USA, Chapter 17, 482–532.

[145] Hua Wei, Chacha Chen, Kan Wu, Guanjie Zheng, Zhengyao Yu, Vikash Gayah, and Zhenhui Li. 2019. Deep Reinforcement Learning for Traffic Signal Control along Arterials. In *Proceedings of the 1st Workshop on Deep Reinforcement Learning for Knowledge Discovery (DRL4KDD '19)*. ACM, Anchorage, USA, 1–7.

[146] Hua Wei, Chacha Chen, Guanjie Zheng, Kan Wu, Vikash Gayah, Kai Xu, and Zhenhui Li. 2019. PressLight: Learning Max Pressure Control to Coordinate Traffic Signals in Arterial Network. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '19)*. ACM, Anchorage, USA, 1290–1298.

[147] Hua Wei, Nan Xu, Huichu Zhang, Guanjie Zheng, Xinshi Zang, Chacha Chen, Weinan Zhang, Yanmin Zhu, Kai Xu, and Zhenhui Li. 2019. CoLight: Learning Network-level Cooperation for Traffic Signal Control. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management (CIKM '19)*. ACM, Beijing, China, 1913–1922.

[148] Hua Wei, Guanjie Zheng, Vikash Gayah, and Zhenhui Li. 2019. A Survey on Traffic Signal Control Methods. arXiv:1904.08117

[149] Hua Wei, Guanjie Zheng, Vikash Gayah, and Zhenhui Li. 2021. Recent Advances in Reinforcement Learning for Traffic Signal Control: A Survey of Models and Evaluation. *ACM SIGKDD Explorations Newsletter* 22, 2 (2021), 12–18.

[150] Keji Wei, Vikrant Vaze, and Alexandre Jacquillat. 2021. Transit Planning Optimization Under Ride-Hailing Competition and Traffic Congestion. *Transportation Science* 56, 3 (2021), 725–749.

[151] Rainer Wiedemann. 1974. Simulation des straßenverkehrsflußes. *Publications of the Institute for Transportation, University of Karlsruhe* 8 (1974), 1–42.

[152] Wisinee Wisetjindawat, Motohiro Fujita, Masayoshi Tomomatsu, and Koji Noda. 2017. Impact of Different Patterns of Red Signal Countdown Timer on Drivers' Startup Behavior. *Journal of the Eastern Asia Society for Transportation Studies* 12 (2017), 1619–1636.

[153] Alex J Wood, Mark Graham, Vili Lehdonvirta, and Isis Hjorth. 2019. Good Gig, Bad Gig: Autonomy and Algorithmic Control in the Global Gig Economy. *Employment and Society* 33, 1 (2019), 56–75.

[154] Qingjun Wu, Hao Zhang, Zhen Li, and Kai Liu. 2019. Labor control in the gig economy: Evidence from Uber in China. *Journal of Industrial Relations* 61, 4 (2019), 574–596.

[155] Huan Yang, Han Zhao, Yu Wang, Guoqiang Liu, and Danwei Wang. 2022. Deep Reinforcement Learning Based Strategy For Optimizing Phase Splits in Traffic Signal Control. In *Proceedings of the 2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC '22)*. IEEE, Macau, 2329–2334.

[156] Jiachen Yang, Igor Borovikov, and Hongyuan Zha. 2020. Hierarchical Cooperative Multi-Agent Reinforcement Learning with Skill Discovery. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS '20)*. IFAAMAS, Auckland, New Zealand, 1566–1574.

[157] Bingquan Yu, Jinqiu Guo, Qinpei Zhao, Jiangfeng Li, and Weixiong Rao. 2020. Smarter and Safer Traffic Signal Controlling via Deep Reinforcement Learning. In *Proceedings of the 29th ACM International Conference on Information and Knowledge Management (CIKM '20)*. ACM, Virtual, 3345–3348.

[158] Jing Zeng, Jie Xin, Ya Cong, Jiancong Zhu, Yihao Zhang, Weihao Jiang, and Shiliang Pu. 2022. HALight: Hierarchical Deep Reinforcement Learning for Cooperative Arterial Traffic Signal Control with Cycle Strategy. In *Proceedings of the 2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC '22)*. IEEE, Macau, 479–485.

[159] Zheng Zeng. 2021. GraphLight: Graph-based Reinforcement Learning for Traffic Signal Control. In *Proceedings of the 6th International Conference on Computer and Communication Systems (ICCCS '21)*. IEEE, Las Vegas, USA, 645–650.

[160] Angie Zhang, Alexander Boltz, Jonathan Lynn, Chun-Wei Wang, and Min Kyung Lee. 2023. Stakeholder-Centered AI Design: Co-Designing Worker Tools with Gig Workers through Data Probes. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. ACM, Hamburg, Germany, 1–18.

[161] Angie Zhang, Alexander Boltz, Chun-Wei Wang, and Min Kyung Lee. 2022. Algorithmic Management Reimagined For Workers and By Workers: Centering Worker Well-Being in Gig Work. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. ACM, New Orleans, USA, 1–20.

[162] Huichu Zhang, Siyuan Feng, Chang Liu, Yaoyao Ding, Yichen Zhu, Zihan Zhou, Weinan Zhang, Yong Yu, Haiming Jin, and Zhenhui Li. 2019. CityFlow: A Multi-Agent Reinforcement Learning Environment for Large Scale City Traffic Scenario. In *Proceedings of the 2019 World Wide Web Conference (WWW '19)*. ACM, San Francisco, USA, 3620–3624.

[163] Yu Zhang, Peter Tiňo, Aleš Leonardis, and Ke Tang. 2021. A Survey on Neural Network Interpretability. *IEEE Transactions on Emerging Topics in Computational Intelligence* 5, 5 (2021), 726–742.

[164] Pengqian Zhao, Yuyu Yuan, and Ting Guo. 2022. Extensible Hierarchical Multi-Agent Reinforcement-Learning Algorithm in Traffic Signal Control. *Applied Sciences* 12, 24 (2022), 12783.

[165] Guanjie Zheng, Yuanhao Xiong, Xinshi Zang, Jie Feng, Hua Wei, Huichu Zhang, Yong Li, Kai Xu, and Zhenhui Li. 2021. Learning Phase Competition for Traffic Signal Control. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management (CIKM '19)*. ACM, Beijing, China, 1963–1972.

[166] Guanjie Zheng, Xinshi Zang, Nan Xu, Hua Wei, Zhengyao Yu, Vikash Gayah, Kai Xu, and Zhenhui Li. 2019. Diagnosing Reinforcement Learning for Traffic Signal Control. arXiv:1905.04716