PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA

MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

HIGHER SCHOOL OF COMPUTER SCIENCE AND DIGITAL TECHNOLOGIES - BEJAIA

Dissertation Submitted to the Department Of Computer Science

in Partial Fulfillment of the Requirements for

Master's Degree in Computer Science

Specialty: Artificial Intelligence and Data Science

Traffic Management Optimization in the Presence of Automated Vehicles

Submitted By:

Cheima MEZDOUR

Supervised By:

Pr. Nadir Farhi

Dr. Meriem Bouali

Members of Jury:

Dr. MEDJOUDJ Rafik	President	ESTIN
Mrs. HADJOUT Siham	Examiner	ESTIN
Dr. ALKAMA Lynda	Examiner	ESTIN
Dr. LEKEHALI Soumia	Examiner	ESTIN

Academic Year: 2023/2024

DEDICATION

I would like to express my deepest gratitude to my parents, who have been my greatest source of support throughout my academic journey. Their unwavering belief in me and constant encouragement have been invaluable. Without them, I would not be who I am today.

To my family, especially my siblings, thank you for your understanding, patience, sacrifices, and love. Your support has given me the strength and motivation to persevere.

To my best friend, Amani,
who accompanied me every step of the way during this journey and throughout my academic career, thank you for always being there for me.

Thank you to everyone who supported me with companionship, encouragement, and love. You have made this journey enjoyable and have been a source of great strength.

ACKNOWLEDGEMENTS

I want to thank the Almighty God, ALLAH, who has been my unwavering companion and guide throughout my academic career. His guidance has been a constant source of strength for me.

A heartfelt thank you to my family, whose support during these challenging six months has been nothing short of indispensable. Your encouragement, patience, and love have been vital to my perseverance and success. I couldn't have done it without you.

I am incredibly grateful to my Supervisor, Pr. Nadir Farhi, for his invaluable guidance, expertise, and contributions. His advice and support have played a crucial role in the success of this project, and I truly appreciate all his efforts.

Lastly, I want to acknowledge the dedication and hard work of all the school teachers. Your instruction and support have laid a strong foundation for my academic journey. Your efforts have been crucial in shaping my educational path, and I am deeply thankful for all you have done.

ABSTRACT

Traffic congestion on freeways is a significant issue today due to the high volume of vehicles. This congestion impacts traffic flow and can lead to severe delays. Effective traffic management methods and solutions are essential to mitigate this problem and enhance safety for all road users.

One specific issue in freeway traffic congestion is the problem of ramp metering. This occurs when vehicles merge from an entrance ramp onto the main freeway. Any malfunction or inefficiency in the traffic light system at these ramps can cause significant congestion and even lead to accidents.

By implementing a sophisticated decision-making system that follows a real-world network, we can develop better policies to address this issue. This approach can optimize traffic flow, reduce congestion, and improve overall road safety.

We proposed a deep q-learning algorithm to address this problem, which guarantees better learning policies and adapt to changing traffic conditions in real-time, learning optimal actions to take at different states of traffic flow in complex environments.

Deep Q-Networks(DQNs) can be used across multiple ramps, so this can improve the scalability of the model. So during this study, we will present the performance of this algorithm compared to traditional systems in improving the efficiency of traffic control.

Keywords : Traffic congestion, Freeways, Traffic flow, Ramp metering, Traffic management, Deep Q-learning, Decision-making system, Real-world network, Traffic control, Optimization, Scalability, Road safety, Adaptive policies, Traffic conditions

RESUME

La congestion du trafic sur les autoroutes est un problème important aujourd'hui en raison du volume élevé de véhicules. Cette congestion affecte le flux de la circulation et peut entraîner des retards considérables. Des méthodes et solutions efficaces de gestion du trafic sont essentielles pour atténuer ce problème et améliorer la sécurité pour tous les usagers de la route.

Un problème spécifique est celui de la régulation des entrées, qui se produit lorsque les véhicules fusionnent depuis une rampe d'accès vers l'autoroute principale. Toute inefficacité du système de feux à ces rampes peut provoquer une congestion importante et même conduire à des accidents. En mettant en œuvre un système de prise de décision sophistiqué qui suit un réseau réel, nous pouvons développer de meilleures politiques pour résoudre ce problème.

Cette approche peut optimiser le flux de la circulation, réduire la congestion et améliorer la sécurité routière. Nous avons proposé un algorithme de deep Q-learning, qui s'adapte aux conditions de circulation changeantes en temps réel, apprenant les actions optimales à prendre dans différents états de flux de trafic dans des environnements complexes.

Les DQNs, utilisés sur plusieurs rampes, améliorent la scalabilité du modèle. Nous présenterons les performances de cet algorithme par rapport aux systèmes traditionnels dans l'amélioration de l'efficacité du contrôle du trafic.

Mots-clés : Congestion du trafic, Autoroutes, Flux de trafic, Régulation des accès, Gestion du trafic, Réseau de neurones profond, Système de prise de décision, Réseau réel, Optimisation, Scalabilité, Sécurité routière, Politiques adaptatives, Conditions de trafic.

ملخص

تُعَدّ ازدحامات المرور على الطرق السريعة مشكلة كبيرة اليوم بسبب العدد الكبير من المركبات. هذا الازدحام يؤثر على تدفق حركة المرور ويمكن أن يؤدي إلى تأخيرات كبيرة. إن أساليب وحلول إدارة المرور الفعّالة ضرورية للتخفيف من هذه المشكلة وتحسين السلامة لجميع مستخدمي الطريق.

إحدى المشاكل المحددة في ازدحام المرور على الطرق السريعة هي مشكلة التحكم في إشارات الدخول. يحدث ذلك عندما تندمج المركبات من منحدر الدخول مع الطريق السريع الرئيسي. أي خلل أو عدم كفاءة في نظام الإشارات الضوئية عند هذه المنحدرات يمكن أن يتسبب في ازدحام كبير وقد يؤدي حتى إلى وقوع حوادث.

من خلال تنفيذ نظام اتخاذ قرار متطور يتبع شبكة واقعية، يمكننا تطوير سياسات أفضل لمعالجة هذه المشكلة. يمكن لهذا النهج تحسين تدفق حركة المرور، وتقليل الازدحام، وتحسين السلامة على الطرق بشكل عام.

لقد اقترحنا خوارزمية التعلم العميق لمعالجة هذه المشكلة، والتي تضمن سياسات تعلم أفضل والتكيف مع ظروف المرور المتغيرة في الوقت الفعلي، من خلال تعلم الإجراءات المثلى التي يجب اتخاذها في حالات تدفق حركة المرور المختلفة في البيئات المعقدة.

يمكن استخدام الشبكات العصبية العميقة عبر العديد من المنحدرات، مما يمكن أن يحسن من قابلية توسع النموذج. خلال هذه الدراسة، سنعرض أداء هذه الخوارزمية مقارنة بالأنظمة التقليدية في تحسين كفاءة التحكم في حركة المرور.

الكلمات المفتاحية: ازدحام المرور، الطرق السريعة، تدفق حركة المرور، التحكم في إشارات الدخول، إدارة المرور، التعلم العميق، نظام اتخاذ القرار، شبكة واقعية، التحكم في المرور، التحسين، قابلية التوسع، السلامة على الطرق، السياسات التكيفية، ظروف المرور.

Acronyms

A2C	Advantage Actor-Critic
AD	Autonomous Driving
AI	Artificial Intelligence
ALINEA	Asservissement LINéaire d'Entrée Autoroutière
AMOC	Adaptive Multi-Objective Control
ANCONA	Adaptive Nonlinear Control Algorithm
AV	Automated Vehicle
BCQ	Batch-Constrained Q-Learning
CACC	Cooperative Adaptive Cruise Control
C-ITS	Cooperative Intelligent Transportation System
CAV	Connected and Automated Vehicle
CV	Connected Vehicle
DRL	Deep Reinforcement Learning
DQN	Deep Q-Network
GA	Genetic Algorithm
HERO	Heuristic Ramp Metering Coordination
IoT	Internet of Things
ITS	Intelligent Transportation System
MARL	Multi-Agent Reinforcement Learning
MADRL	Multi-Agent Deep Reinforcement Learning
MetaNet	Macroscopic Traffic Network Simulation Environment
ML	Machine Learning
MCTS	Monte Carlo Tree Search
MORL	Multi-Objective Reinforcement Learning
MPC	Model Predictive Control
PPO	Proximal Policy Optimization

PSO	Particle Swarm Optimization
Q-Learning	Quality-Learning
RL	Reinforcement Learning
RL-Agent	Reinforcement Learning Agent
RL-TL	Reinforcement Learning for Traffic Light Control
SARSA	State-Action-Reward-State-Action
SOTA	State Of The Art
SUMO	Simulation of Urban Mobility
SZM	Stochastic Zone Model
TD	Temporal Difference
TL	Traffic Light
TORCS	The Open Racing Car Simulator
TraCI	Traffic Control Interface
TTS	Traffic Theory Simulation
V2V	Vehicle-to-Vehicle Communication
V2X	Vehicle-to-Everything Communication
VISUM	Verkehrsinformationssystem für die urbane Mobilität
VISSIM	Verkehr In Städten - SIMulationsmodell

Contents

List of Figures

List of Tables

General Introduction

In the transportation realm, traffic congestion is a major urban challenge facing cities all around the world. It disrupts the normal flow of roads due to the excessive number of vehicles on a portion of a roadway. This problem has significant economic impacts, with billions of dollars lost annually in wasted time and fuel consumption. Additionally, it causes a loss of productivity, which refers to the reduction in the efficiency and effectiveness of individuals and businesses caused by time spent in traffic.

On the environmental side, traffic congestion has more serious consequences, such as increased emissions and air pollution. Cities known for traffic congestion suffer from these issues, which contribute to climate change and global warming.

Interestingly, some studies have shown that heavy traffic can increase the level of the hormone adrenaline in the blood, leading to more stress. This stress can affect individuals, potentially decreasing their productivity, as well as reducing the time available for personal activities.

One critical area that often exacerbates this problem is freeway congestion, particularly at points where vehicles from ramps merge onto the freeway. This merging process creates a significant bottleneck, leading to slowdowns and increased traffic density.

When vehicles enter a freeway from an on-ramp, they must adjust their speed to merge seamlessly with the traffic flow. This process often disrupts the steady movement of vehicles already on the freeway, causing drivers to brake and adjust their speeds. As a result, a ripple effect occurs, leading to stop-and-go traffic that extends well beyond the merge point.

This merging congestion is particularly problematic during peak hours when the volume of vehicles is at its highest. The increased number of vehicles trying to enter the freeway can overwhelm the capacity of the merge lanes, leading to extended queues on the ramps and further exacerbating congestion on the freeway

The impact of this type of congestion on traffic is significant. Drivers experience longer travel times as they navigate through congested freeway areas. This often leads to frequent braking and acceleration, which increases the likelihood of collisions and accidents near the ramps where vehicles merge onto the freeway. Additionally, if measures are taken to free up the freeway and allow vehicles to move more smoothly, it can cause backups on the ramps. Vehicles will queue up as they wait to merge onto the freeway, potentially leading to congestion and collisions in other traffic areas.

Our primary objective is to develop an intelligent solution to effectively regulate traffic, ensuring a smooth transition of vehicles between the freeway and the ramps. This solution aims to minimize travel time and reduce congestion both on the freeway and the ramps. By addressing these issues, we aim to enhance the overall flow of traffic, decrease the density of vehicles in these critical areas, and maximize the speed at which vehicles can traverse the freeway.

To achieve this, we will use advanced traffic management techniques and technologies. These may include adaptive traffic signal control, ramp metering, and real-time traffic monitoring systems. The goal is to dynamically adjust traffic flow based on current conditions, thereby preventing bottlenecks and ensuring that vehicles can merge onto the freeway without causing significant slowdowns or disruptions.

By optimizing the coordination between freeway traffic and ramp entries, we expect to see a significant improvement in traffic efficiency. This will not only reduce travel times for drivers but also enhance safety by minimizing the risk of accidents caused by sudden braking and frequent stops. Ultimately, our solution will contribute to a more efficient and safer transportation network, benefiting all road users.

This comprehensive approach integrates various technologies to create a responsive and adaptive traffic management system. It underscores the importance of real-time data in making informed decisions to regulate traffic flow. Through the implementation of these strategies, we seek to create a sustainable and effective solution to the persistent problem of traffic congestion in urban environments.

The remainder of this dissertation is structured as follows:

- **Chapter 1: Background**

  This chapter explains the main technical concepts needed for this project. It covers the foundational principles of AI and introduces the specialized tools and methodologies applied throughout the work.

- **Chapter 2: Deep Reinforcement Learning**

  This chapter covers the fundamentals of reinforcement learning, focusing on the transition from RL to DRL while explaining the key DRL algorithms.

- **Chapter 3: State of the Art**

  This chapter reviews previous solutions to the ramp metering problem, covering a range of methods from traditional approaches to the latest proposed methods, along with their variations for managing freeway congestion.

Background

1.1 Introduction

This chapter provides an essential foundation for the technical concepts central to this thesis. We will explore ramp metering, the SUMO simulation tool, and key aspects of traffic management, followed by introductions to Artificial Intelligence (AI), Machine Learning (ML), and Reinforcement Learning (RL). Together, these areas support the methods used to optimize traffic systems, particularly with automated vehicles.

Ramp metering is a traffic management strategy used on highways to regulate the flow of vehicles entering from on-ramps. By controlling the entry rate of vehicles, it aims to minimize congestion, improve traffic flow, and reduce bottlenecks. This concept is significant in this thesis, as our goal is to enhance traffic efficiency through intelligent decision-making algorithms for ramp metering.

SUMO (Simulation of Urban Mobility) is crucial for simulating realistic traffic scenarios, enabling the testing and validation of traffic management strategies. Core traffic management principles further contextualize how these advanced technologies are applied to address real-world challenges.

Finally, AI, ML, and RL provide the intelligence behind data-driven decision-making systems. AI encompasses cognitive tasks, ML focuses on learning from data, and RL enables agents to make optimal decisions through environmental interactions.

By exploring these concepts, this chapter sets the stage for applying these technologies in the subsequent sections of the thesis, ensuring a well-rounded understanding of the foundational elements necessary for this research.

1.2 Ramp Metering in Traffic Management

Ramp metering is a crucial traffic management strategy aimed at improving the flow and efficiency of traffic on highways. By regulating the rate at which vehicles enter the highway from on-ramps, ramp metering helps reduce congestion, minimize stop-and-go conditions, and improve overall traffic throughput. It is especially effective in areas prone to high traffic volumes, where on-ramp congestion often disrupts the mainline flow, leading to delays and increased travel times.

1.2.1 Purpose and Mechanism

The primary purpose of ramp metering is to control the timing of vehicles entering the highway, using traffic signals positioned at on-ramps to release vehicles at regulated intervals. These intervals are based on real-time data on traffic density, speed, and flow on the highway. By spacing out vehicles, ramp metering can prevent sudden spikes in traffic that often cause bottlenecks and turbulence in traffic flow, thus reducing delays and improving overall traffic stability. Figure 1.1 illustrates the mechanism by which vehicles enter the freeway from on-ramps.

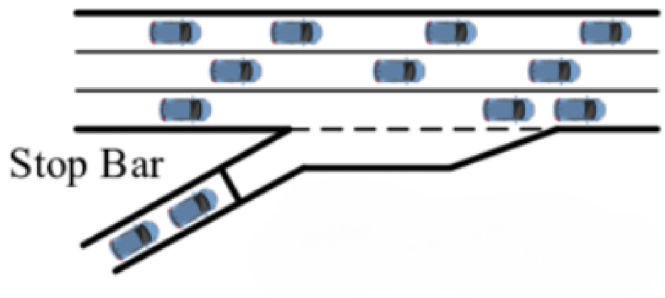

Figure 1.1: Ramp Metering Mechanism

There are several ramp metering strategies, the most widely used are:

- **Fixed-Time Metering**: This uses pre-determined intervals for vehicle release, irrespective of real-time traffic conditions.

- **Demand-Capacity or Occupancy-Based Metering**: This strategy adjusts intervals based on the density of vehicles on the highway, releasing vehicles only when there is sufficient space.

- **ALINEA**: A feedback-based approach that dynamically adjusts the release rate based on real-time measurements of traffic occupancy downstream from the on-ramp.

Each of these approaches serves the same goal but differs in complexity, adaptability, and computational requirements.

1.2.2 Benefits of Ramp Metering

Ramp metering provides several benefits:

- **Reduced Congestion**: By smoothing out traffic flow and preventing bottlenecks, ramp metering reduces the likelihood of congestion near merging points.

- **Increased Safety**: Ramp metering minimizes sudden braking and lane-changing, reducing the risk of accidents and collisions.

- **Improved Travel Times**: Vehicles experience fewer stops and reduced delays, improving overall travel times and reducing fuel consumption.

Studies have shown that ramp metering can reduce travel time by 20-30% in congested corridors, making it an essential strategy for urban and suburban areas with dense traffic networks.

1.2.3 Challenges in Implementation

Despite its advantages, ramp metering faces some challenges:

- **Infrastructure Requirements**: Installing ramp metering systems requires sensors, traffic lights, and communication networks, which can be costly.

- **Public Acceptance**: Ramp metering may initially cause delays at the ramp itself, requiring public education to highlight its long-term benefits.
- **Coordination Across Ramps**: In high-density urban networks, multiple on-ramps may need coordinated metering to prevent overflow from one ramp to another.

1.3 SUMO (Simulation of Urban Mobility)

1.3.1 Introduction

SUMO is an open-source traffic simulation package developed by the German Aerospace Center (DLR) starting in 2001. It aims to support the traffic simulation community by providing a comprehensive, extensible suite of tools for traffic modeling and simulation. The package includes utilities for network import, demand generation, and high-performance simulation of traffic scenarios ranging from single intersections to entire cities [1].

1.3.2 Components of SUMO

1. **Net Import and Generation:** SUMO allows for the import of road networks from various formats such as VISUM, VISSIM, MATsim, OpenStreetMap, and shapefiles. It also supports network generation using tools like netconvert [1].
2. **Demand Modeling:** Tools like od2trips convert origin/destination matrices into vehicle trips, facilitating large-scale traffic simulations [1].
3. **Simulation:** SUMO employs a microscopic simulation approach where each vehicle is modeled with specific parameters such as departure time, route, and vehicle type. The simulation environment is time-discrete and space-

continuous, offering both graphical visualization and command-line batch processing [1].

4. **TraCI (Traffic Control Interface):** This interface allows real-time interaction with the simulation, enabling dynamic adjustments and integration with other simulators [1].

1.3.3 Features of SUMO

1. **Vehicular Communication:** SUMO is extensively used in research on vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication. It supports integration with communication simulators like ns2 and ns3, enabling the simulation of Vehicle-to-Everything V2X applications [1].

2. **Route Choice and Navigation:** The package includes tools like duarouter for route assignment and supports dynamic navigation systems, which helps investigate the impact of real-time traffic information on route choices and traffic flow [1].

3. **Traffic Light Algorithms:** SUMO facilitates the evaluation of traffic light programs and adaptive traffic control algorithms, making it suitable for both intersection-level and network-wide traffic management studies [1].

4. **Surveillance Systems:** With its large-scale simulation capabilities, SUMO supports the evaluation of traffic surveillance systems, including applications like airborne traffic monitoring and GSM-based travel speed observation [1].

1.4 MetaNet Traffic Flow Model

The MetaNet model is a macroscopic traffic flow model developed by Messmer and Papageorgiou in the early 1990s [2]. It is designed to simulate the dynamics of traffic on large-scale motorway networks. The key features of MetaNet include its ability to model traffic flow, density, and speed over time and space using a set of differential equations [3]. These equations are based on the principles of traffic flow theory and are used to predict the evolution of traffic states in response to various control measures such as ramp metering and variable speed limits [4].

1.4.1 Key Components of MetaNet

1. **Density Update Equation:**

   This equation models how vehicle density changes on each road segment over time, considering the inflow and outflow of vehicles [2].

$$\rho_i(t+1) = \rho_i(t) + \frac{T}{L_i}\left(q_{i-1}(t) - q_i(t)\right) \tag{1.1}$$

   Where:

   - $\rho_i(t)$: Density of vehicles on segment i at time t.
   - T: Time step size.
   - L_i: Length of segment i.
   - $q_{i-1}(t)$: Flow into segment i from the previous segment.
   - $q_i(t)$: Flow out of segment i.

2. **Flow Update Equation:**

   This equation relates traffic flow to vehicle density and speed [3].

$$q_i(t) = \rho_i(t) \cdot v_i(t) \tag{1.2}$$

   Where:

   - $q_i(t)$: Traffic flow on segment i.
   - $\rho_i(t)$: Density of vehicles on segment i.
   - $v_i(t)$: Speed of vehicles on segment i.

3. **Speed Update Equation:**

   This equation models how vehicle speed changes based on density and desired speed adjustments [4].

$$v_i(t+1) = v_i(t) + \frac{T}{\alpha}\left(v_{i,\text{desired}}(\rho_i(t)) - v_i(t)\right) \tag{1.3}$$

   Where:

   - $v_i(t)$: Speed on segment i at time t.
   - T: Time step size.
   - α: Parameter that affects how quickly speeds adjust.
   - $v_{i,\text{desired}}(\rho_i(t))$: Desired speed based on the density $\rho_i(t)$.

1.4.2 Advantages of MetaNet

- **Scalability**: MetaNet can handle large networks efficiently, making it suitable for regional or city-wide traffic management [3].
- **Speed**: It offers faster computations compared to microscopic models, which is beneficial for real-time traffic management applications [2].
- **Control Strategy Development**: It is highly effective in testing and optimizing high-level traffic control strategies such as ramp metering and variable speed limits [4].

1.4.3 Integration with SUMO

Combining MetaNet with SUMO leverages the strengths of both models:

- **MetaNet**: provides a macroscopic view, useful for strategic planning and network-wide traffic management [2].
- **SUMO**: offers detailed, microscopic simulations to fine-tune and validate strategies at a granular level [3].

1.5 Artificial Intelligence

Artificial Intelligence (AI) is a scientific field focused on creating machines and computers that can reason, learn, and act in ways typically requiring human intelligence or handle data on a scale beyond human capabilities. AI encompasses a wide range of topics, including computer science, data analysis, statistics, hardware and software engineering, linguistics, neuroscience, and philosophy.

AI varies depending on the techniques employed, but the fundamental factor is data. AI systems learn and improve by being exposed to vast amounts of data, identifying patterns and relationships that might be missed by humans. This exposure allows AI to continuously refine its algorithms, leading to increasingly accurate and efficient performance.

The learning process in AI involves algorithms that guide how AI analyzes and makes decisions. In machine learning, a key area of AI, algorithms learn from labeled or unlabeled data to make predictions or classify information. Deep learning, a more advanced type of AI, uses multi-layered artificial neural networks to process data similarly to the human brain, enabling AI systems to improve at specific tasks like image recognition and language translation [5].

AI is increasingly integrated across sectors, including medicine, finance, agriculture, and notably, transportation. Its impact in transportation is significant, driving progress through sophisticated systems. AI applications in this field include computer vision, object detection, and tracking, which are pivotal in developing self-driving vehicles, autonomous air taxis, and smart highways. Implementing AI in transportation enhances safety and efficiency while transforming how we commute and transport goods [6].

AI is utilized in various applications, each with distinct functionalities:

- **Natural Language Processing**: Enables machines to understand and generate human language, facilitating translation and sentiment analysis [5].

- **Computer Vision**: Focuses on interpreting visual content to identify objects and patterns, extensively used in facial recognition, object detection, and autonomous vehicles [5].

- **Robotics**: Concentrates on the design, construction, and functionality of robots, enhancing automation and efficiency across industries [5].

- **Machine Learning**: Enables computers to learn from data and improve performance over time, applied in predictive analytics, fraud detection, and recommendation systems [5].

Machine Learning (ML), as a subset of AI, deserves particular attention due to its extensive applications and importance in advancing AI technologies. The section that follows will delve deeper into the concepts and methodologies of ML, highlighting its role and impact in modern technology.

1.6 Machine Learning

Machine learning is the study of computer algorithms that improve automatically through experience and data, serving as a subset of artificial intelligence. It involves constructing systems that learn from and make predictions based on data without explicit instructions.

This definition is rooted in the work of researchers like Tom M. Mitchell, who defined it as follows:

> ”A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E.” [7].

1.6.1 Key Components of Machine Learning

Essential elements of machine learning systems include algorithms, models, training data, experience, tasks, and performance measures. Understanding these components is crucial for grasping how machine learning operates [8].

1. **Algorithms**: Procedures or formulas for solving problems through specified actions.
2. **Models**: Representations built by algorithms from data, used for predictions or decisions.
3. **Training Data**: Input-output pairs used to train the model.
4. **Experience (E)**: Data or feedback received during training.
5. **Tasks (T)**: Specific functions the model performs, such as classification or regression.
6. **Performance Measure (P)**: Metrics to evaluate the model's accuracy and effectiveness, such as accuracy or F1 score.

1.6.2 Types of Machine Learning

Machine learning can be categorized into types based on the learning system's nature. As shown in the Figure 1.2, these types include supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning, each type with distinct characteristics and applications [9].

1. **Supervised Learning**: Trained on labeled datasets; tasks include classification and regression.
2. **Unsupervised Learning**: Trained on unlabeled data to find underlying structures; tasks include clustering and association.
3. **Semi-Supervised Learning**: Trained on datasets with both labeled and unlabeled data, useful when fully labeled datasets are costly to obtain.
4. **Reinforcement Learning**: Learns by interacting with an environment, receiving rewards or penalties, focusing on exploration and exploitation.

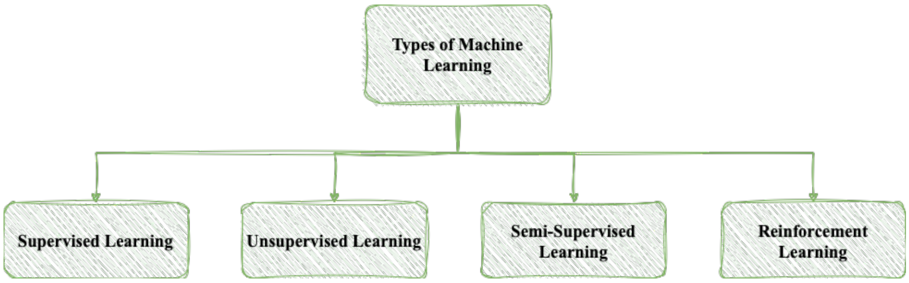

Figure 1.2: The different types of machine learning.

1.7 Conclusion

This chapter provides a crucial foundation for understanding the technical concepts central to this thesis. It has explored the principles of ramp metering, the SUMO simulation tool, and essential traffic management strategies, while also introducing key concepts in Artificial Intelligence (AI), Machine Learning (ML), and Reinforcement Learning (RL). These areas collectively establish the theoretical framework for the subsequent application of AI in optimizing traffic systems. By integrating these technologies, this research sets the stage for advancing traffic management strategies, particularly in the context of automated vehicles, contributing to the development of smarter and more efficient transportation systems.

Deep Reinforcement Learning

2.1 Introduction

This chapter explores the concepts of Deep Reinforcement Learning, a powerful combination of Reinforcement Learning and Deep Learning, which enables autonomous agents to navigate complex environments with high-dimensional input spaces. The chapter begins by providing an overview of Reinforcement Learning, exploring its foundational elements such as policy, reward signals, value functions, and environment models. As traditional RL struggles with large state spaces, the chapter then transitions to the necessity of Deep Learning in RL, detailing how DRL leverages deep neural networks to address these challenges. Finally, it covers key DRL algorithms such as Deep Q-Networks, Policy Gradient methods, and Actor-Critic approaches, illustrating how these methods enhance the agent's decision-making capabilities in complex environments. Through this exploration, the chapter sets the stage for understanding the integration of DRL in traffic optimization systems, particularly in the context of automated vehicles and intelligent transportation networks.

2.2 Deep Reinforcement Learning

Deep Reinforcement Learning combines Reinforcement Learning with Deep Learning, enabling agents to learn effective policies in environments with high-dimensional state spaces. Traditional RL struggles with complex data representations, but DRL leverages deep neural networks to approximate value functions and policies, making it suitable for tasks with high-dimensional inputs like images and text. This integration has led to breakthroughs in areas such as robotics, game playing, and autonomous driving.

2.2.1 Reinforcement Learning

Reinforcement Learning is a machine learning field where an autonomous agent learns to make optimal decisions through interactions with an environment. Beginning without prior knowledge, the agent refines its strategy by taking actions, receiving rewards or penalties, and aiming to maximize cumulative rewards. These interactions, progressing in discrete steps from initial to terminal states, form episodes through which the agent gradually improves its decision-making ability.[10] [11] [12].

2.2.2 Basic Concepts of RL

Reinforcement learning systems consist of several key elements beyond just the agent and the environment as shown in Figure 2.1 [13]:

1. **Policy** : It defines the agent's behavior at any given time and maps perceived states of the environments to actions to be taken in those states. Policy can be simple function, lookup table, or involve extensive computation like a search process. Policies may be stochastic, specifying probabilities for each action [13].

2. **Reward Signal** : The goal of the reinforcement learning problem is defined by the environment, which provides a reward at each time step. The agent's objective is to maximize the cumulative reward over time. This reward signal guides the alteration of the policy based on the actions taken and their outcomes. The rewards may be stochastic functions of the environment's state and the actions taken [13].

3. **Value Function** : The value function in reinforcement learning specifies the long-term desirability of states. It represents the total expected reward an

agent can accumulate from a given state. Compared to immediate rewards, values are more refined and farsighted judgments. Estimating values is more complex than determining rewards, as it involves making predictions over the agent's lifetime [13].

4. **Model of the Environment** : The Model of the Environment in reinforcement learning mimics the behavior of the environment. It allows inferences about future states and rewards based on the current state and action. This model is used for planning future actions by considering possible future situations before they occur. Methods that use models are called model-based, while those relying purely on trial-and-error are called model-free [13].

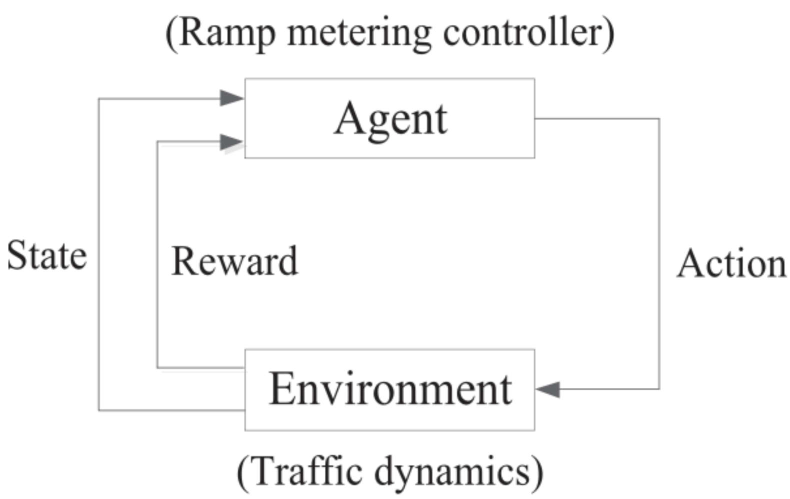

Figure 2.1: The basic framework of reinforcement learning.

These elements form the foundation of reinforcement learning systems, facilitating learning and decision-making through interaction with the environment.

2.2.3 From RL to DRL: The Need for Deep Learning

In complex environments, especially those with high-dimensional input spaces like images, traditional RL techniques struggle due to the curse of dimensionality [11].

Deep Reinforcement Learning (DRL) addresses these challenges by using deep neural networks to approximate the Q-value or policy functions, allowing the agent to effectively process complex data and make decisions [14].

DRL leverages several techniques to improve stability and performance:

- **Experience Replay**: Stores past experiences in a memory buffer and samples them randomly during training, reducing correlation between experiences and enhancing learning stability [14].

- **Target Networks**: In Deep Q-Networks (DQN), a separate target network helps stabilize training by keeping target values consistent over multiple updates [14].

2.2.4 Key DRL Algorithms

Q-Learning

Q-Learning is a popular model-free RL algorithm that aims to learn the value of state-action pairs, enabling the agent to determine an optimal policy for maximizing rewards. The update rule of Q-Learning is defined as follows [15]:

$$Q(s,a) \leftarrow Q(s,a) + \alpha \left[r + \gamma \max_{a'} Q(s',a') - Q(s,a) \right] \tag{2.1}$$

Where :

- $Q(s,a)$: The value of taking action a in state s.

- α: The learning rate, which determines the extent to which new information overrides the old.

- r: The immediate reward received after taking action a from state s.

- γ: The discount factor, representing the importance of future rewards.

- s': The next state after taking action a.

Deep Q-Networks (DQN)

DQN is an extension of Q-Learning using deep neural networks to approximate the Q-function in large state spaces [14]. The update rule in DQN incorporates experience replay and target networks as follows:

$$\theta \leftarrow \theta + \alpha \left[r + \gamma \max_{a'} Q(s', a'; \theta^-) - Q(s, a; \theta) \right] \nabla_\theta Q(s, a; \theta) \tag{2.2}$$

where :

- θ: Parameters of the Q-network.
- θ^-: Parameters of the target network, updated periodically.
- $\max_{a'} Q(s', a'; \theta^-)$: The maximum predicted Q-value for the next state s' across all possible actions a', calculated using the target network with parameters θ^-.
- $Q(s, a; \theta)$: The predicted Q-value for taking action a in state s, based on the current learned parameters θ of the network.
- $\nabla_\theta Q(s, a; \theta)$: The gradient of the Q-value function with respect to the network parameters θ.

Policy Gradient Methods

These methods directly learn the policy by optimizing it with respect to expected cumulative rewards, rather than estimating a value function [16]. The **objective function** for policy gradients is:

$$J(\theta) = \mathbb{E}_{\pi_\theta} \left[\sum_{t=0}^{T} \gamma^t r_t \right] \tag{2.3}$$

while the **Policy Gradient Theorem** is as follows:

$$\nabla_\theta J(\theta) = \mathbb{E}_{\pi_\theta} \left[\nabla_\theta \log \pi_\theta(a|s) Q^{\pi_\theta}(s, a) \right] \tag{2.4}$$

Where :

- θ: Parameters of the policy.
- $\pi_\theta(a|s)$: Probability of taking action a in state s under policy π_θ.
- $Q^{\pi_\theta}(s, a)$: Q-value under policy π_θ.

Actor-Critic Methods

Actor-Critic algorithms combine the benefits of value-based and policy-based approaches by having two main components:

- **Actor**: Learns the policy, deciding which action to take.
- **Critic**: Evaluates actions by estimating the value function, providing feedback to the actor [17].

2.3 Conclusion

In conclusion, this chapter has provided a comprehensive overview of Deep Reinforcement Learning (DRL), detailing its evolution from traditional Reinforcement Learning and its integration with deep neural networks to handle complex, high-dimensional environments. Through the exploration of its foundational principles, we have established how agents can learn optimal decision-making strategies by interacting with their surroundings. The chapter also highlights the advancements DRL offers over traditional RL methods, enabling agents to tackle tasks previously considered computationally infeasible. As this chapter demonstrates, DRL has the potential to significantly enhance the efficiency and effectiveness of traffic management systems, particularly in the context of autonomous vehicles and intelligent transportation networks, paving the way for smarter, more adaptive urban infrastructure.

State Of Art

3.1 Introduction

Ramp metering is a crucial traffic management strategy that controls the flow of vehicles entering highways to improve traffic conditions, prevent congestion, and enhance road efficiency. As urbanization and vehicle numbers grow, managing traffic flow becomes increasingly important. By regulating how vehicles merge onto highways, ramp metering helps reduce bottlenecks, smooth traffic, and maintain safe driving conditions, making it essential for modern road networks.

The evolution of ramp metering reflects the increasing complexity of modern traffic systems and their demands. Early methods, like demand capacity control and the ALINEA algorithm, used simple rule-based systems to manage traffic at specific ramps. While effective in some cases, these traditional approaches struggled with the unpredictable nature of real-world traffic, which can change rapidly due to accidents, weather, or varying traffic volumes.

To address the challenges of traditional ramp metering, traffic management has shifted toward more advanced, data-driven approaches. Researchers have explored intelligent control methods, like fuzzy logic and neural networks, to manage the complexities of traffic systems. However, reinforcement learning (RL) has emerged as the most adaptive and efficient solution, offering greater flexibility in responding to changing traffic conditions.

Reinforcement learning (RL) enables more dynamic and responsive ramp metering by allowing systems to learn and adapt to changing traffic conditions without predefined models. This makes RL ideal for handling the unpredictable nature of traffic. Multi-Agent Reinforcement Learning (MARL) further enhances RL by coordinating multiple ramps to optimize traffic flow across an entire network, rather than just individual points.

This literature review explores the evolution of ramp metering strategies, from early rule-based methods to recent advancements in reinforcement learning and multi-agent systems. It aims to provide a comprehensive understanding of current methodologies, evaluate their effectiveness, and investigate future innovations in traffic management. The review highlights the progress made while addressing the ongoing challenges and opportunities in developing smarter and more efficient road networks.

3.2 Traditional Approaches

3.2.1 Feedback-Based Control (ALINEA and PI-ALINEA)

In the paper [18], the authors discuss the use of traditional feedback-based ramp metering strategies, particularly focusing on the well-established methods of ALINEA and PI-ALINEA. These strategies have been pivotal in traffic management for decades, offering a structured way to control the flow of vehicles onto freeways to prevent congestion and ensure smoother traffic movement.

ALINEA, which stands for Asservissement Linéaire d'entrée Autoroutière, was developed by Papageorgiou et al. In 1991 and has since become a cornerstone in the field of ramp metering. The strategy operates by dynamically adjusting the rate at which vehicles are allowed to merge onto the freeway based on the occupancy levels observed at downstream bottlenecks. The core idea is to keep the freeway occupancy at a critical, yet optimal level, which in turn helps to maximize throughput and minimize the risk of congestion build-up. This method is particularly valued for its simplicity and effectiveness, and it has been widely implemented in real-world scenarios where it has proven successful in reducing traffic delays and improving overall flow [19].

Building on the success of ALINEA, the PI-ALINEA strategy was introduced to tackle the more complex challenge of managing traffic in the presence of distant downstream bottlenecks. PI-ALINEA incorporates a proportional-integral (PI) controller into the original ALINEA framework, allowing it to adjust the ramp metering rate more dynamically by considering not only the current occupancy but also the trends and changes in traffic flow over time. This enhancement enables the system to respond more effectively to varying traffic conditions, providing a more stable and robust control mechanism.

Despite their widespread use and proven effectiveness, both ALINEA and PI-ALINEA have their limitations. These methods are highly dependent on the precise tuning of parameters, which can be a challenge in real-world conditions where traffic patterns can be unpredictable and rapidly changing. Moreover, while these strategies excel at managing traffic at specific bottlenecks, they are inherently localized in their approach. This means that while they can significantly improve traffic conditions at specific points on the freeway, they may not perform as well in more complex scenarios where a network-wide approach to traffic management is required.

The paper by Han et al. in [18] highlights the enduring relevance of these feedback-based methods in ramp metering, while also acknowledging the need for more adaptable and comprehensive approaches to meet the demands of modern traffic systems.

3.2.2 The demand-capacity and occupancy-capacity approaches

Wattleworth (1967) [20] was a pioneering figure in the development of early ramp metering strategies, introducing foundational concepts such as the demand-capacity and occupancy-capacity approaches. His work laid the groundwork for how modern traffic management systems attempt to balance the inflow of vehicles with the available road capacity, a crucial aspect of maintaining efficient traffic flow. The demand-capacity strategy proposed by Wattleworth involves regulating the rate at which vehicles are allowed to enter the freeway based on the difference between the upstream and downstream capacities. This approach aims to prevent congestion by ensuring that the volume of vehicles entering the freeway does not exceed the capacity of the road ahead, thereby avoiding bottlenecks and maintaining smoother traffic flow.

In parallel, the occupancy-capacity approach adjusts the metering rate based on real-time occupancy levels of the mainline downstream of the ramp. This method provides a more responsive control mechanism by directly monitoring the density of traffic on the freeway and modulating the inflow from ramps to prevent the road from becoming oversaturated. These strategies were among the first to systematically address the issue of congestion on freeways, marking a significant advancement in traffic control technology at the time.

However, despite their innovative nature, Wattleworth's methods had limitations. The strategies were relatively simplistic and did not fully account for the dynamic and often unpredictable nature of traffic flow. They were primarily reactive rather than proactive, meaning they responded to current conditions without the capability to anticipate future changes in traffic patterns. This limitation made them less effective in managing complex and rapidly changing traffic environments, where real-time adaptability and anticipation of congestion are crucial. Nonetheless, Wattleworth's work provided a crucial stepping stone for the development of more sophisticated traffic management systems and continues to be recognized for its foundational role in the field.

3.2.3 The Advanced Motorway Optimal Control (AMOC)

Kotsialos and Papageorgiou (2004) [21] conducted an in-depth study on the Advanced Motorway Optimal Control (AMOC) strategy, focusing specifically on its implementation on the A10 ring road in Amsterdam. AMOC is a sophisticated, proactive, network-wide ramp metering strategy that leverages a discrete-time optimal control approach to continuously optimize traffic flow across a freeway network.

By updating the control strategy periodically based on real-time traffic data, AMOC aims to maintain an optimal balance between ramp metering and the mainline traffic conditions. This iterative process ensures that the system adapts dynamically to changing traffic patterns, enhancing the overall efficiency of the road network.

The results from Kotsialos and Papageorgiou's research were highly promising. The application of AMOC on the A10 ring road led to significant improvements in traffic conditions, notably reducing overall travel times and delays. The strategy proved particularly effective during peak periods, where congestion is most severe, by optimizing the use of available road capacity and preventing bottlenecks. This optimization resulted in smoother traffic flow and a more reliable travel experience for road users.

However, the study also highlighted several limitations of the AMOC strategy. The approach relies heavily on high-quality, real-time traffic data to inform its control decisions. This requirement can pose a challenge, as obtaining accurate and timely data can be difficult, especially in large-scale implementations. Additionally, AMOC's computational demands are considerable, necessitating significant resources to solve the optimization problem iteratively. The complexity of implementing AMOC on a broader scale can be daunting, and its effectiveness may be compromised by unpredictable incidents or inaccuracies in traffic data predictions. Despite these challenges, the potential benefits of AMOC in managing freeway congestion and improving traffic flow make it a noteworthy advancement in the field of traffic management.

3.2.4 The ANCONA strategy

Kerner (2006) [22] made a notable contribution to the field of traffic management with the development of the ANCONA strategy, which is rooted in the three-phase traffic theory he pioneered. This theory breaks down traffic flow into three distinct phases: free flow, synchronized flow, and wide moving jams. Kerner's ANCONA strategy leverages these phases to manage and confine congestion spatially by fostering synchronized flow conditions on highways. The central idea behind ANCONA is to create a traffic environment where the transition from free flow to wide moving jams is minimized by maintaining traffic in the more stable synchronized flow phase. By doing so, ANCONA effectively manages the length of queues at on-ramps and prevents the upstream propagation of traffic jams, which are often the precursors to severe congestion.

The application of the ANCONA strategy has shown promising results, particularly in its ability to reduce the frequency and severity of wide moving jams. By keeping traffic within the synchronized flow phase, where vehicles move at a consistent and controlled pace, the strategy enhances the overall stability and manageability of traffic conditions on highways. This stability helps to prevent the sudden breakdown of traffic flow that leads to widespread congestion, thus improving the efficiency of the roadway and the driving experience for motorists.

However, the ANCONA strategy is not without its challenges. The underlying three-phase traffic theory is complex, requiring a deep understanding of traffic dynamics and precise calibration of the system to be effective. The success of ANCONA depends heavily on the ability to accurately predict and manage the transitions between the different traffic phases, a task that can be difficult in the unpredictable conditions of real-world traffic. Additionally, implementing ANCONA requires detailed, high-quality data on traffic patterns, which can be difficult to obtain and analyze in a timely manner. Despite these limitations, Kerner's work on ANCONA represents a significant advancement in traffic management, offering a novel approach to mitigating congestion on busy highways by strategically managing traffic flow phases.

3.2.5 The Stratified Zone Metering (SZM) algorithm

Srivastava (2011) [23] made a significant contribution to traffic management with the introduction of the Stratified Zone Metering (SZM) algorithm, a sophisticated approach designed to manage highway congestion more effectively. This algorithm was developed for and utilized by the Minnesota Department of Transportation (MnDOT) to address the challenges of maintaining smooth traffic flow on heavily congested freeways. The SZM algorithm organizes freeway segments into distinct zones, each of which is monitored for traffic density. Based on the real-time data collected from these zones, the algorithm applies ramp metering rates that are tailored to the specific conditions of each segment, allowing for a more nuanced and targeted control of traffic.

One of the key strengths of the SZM approach is its consideration of critical factors such as bottleneck locations and ramp queue lengths, which are crucial in managing traffic flow effectively. By doing so, the algorithm can prevent the formation of congestion before it becomes problematic, thereby smoothing out traffic conditions across the freeway network. The results of Srivastava's work demonstrated that SZM was particularly effective in areas prone to frequent bottlenecks and where traffic volumes were consistently high, leading to a noticeable improvement in overall traffic flow and a reduction in congestion.

However, the effectiveness of the SZM approach is contingent on the availability of detailed and accurate traffic data, as well as the precise calibration of the metering system. The complexity of the algorithm can present challenges in its implementation, particularly on highways with highly variable traffic patterns where conditions can change rapidly and unpredictably. Additionally, while SZM performs well under normal traffic conditions, it may struggle to maintain efficiency during unexpected traffic situations or incidents, where the lack of flexibility could limit its effectiveness. Despite these challenges, Srivastava's SZM algorithm represents an important advancement in the field of traffic management, offering a structured and data-driven method to alleviate congestion on busy highways.

3.3 Reinforcement Learning (RL) based approaches

3.3.1 Q-Learning

In [24], the authors explore a novel application of Q-learning, a type of model-free reinforcement learning algorithm, to address the challenges of ramp metering on busy motorways. Q-learning stands out as a powerful tool because it allows a system, or agent, to learn the best course of action by interacting directly with its environment. Imagine an agent that observes the current traffic conditions, such as the number of vehicles waiting to enter the freeway and the congestion levels on the main lanes. The agent then makes a decision—perhaps to let more cars merge onto the freeway or to hold them back for a few more seconds. After this decision is made, the agent receives feedback in the form of rewards, which could be a reduction in overall congestion or a smoother flow of traffic. Over time, through many cycles of observation, action, and feedback, the agent becomes increasingly skilled at making decisions that optimize traffic flow and minimize delays.

In [18], Yu Han and colleagues builds on these ideas mentioned in [24] by applying Q-learning not just to local ramp metering but also to coordinated ramp metering scenarios, where multiple ramps work together to manage traffic across a broader network. Their research highlights the adaptability of Q-learning, showing how it can dynamically respond to changing traffic conditions. The authors' simulations reveal that Q-learning can achieve better results than traditional ramp metering methods like ALINEA, which rely on pre-defined rules and lack the ability to learn and adapt. By constantly updating its policy based on real-time data, Q-learning can more effectively reduce congestion and improve the overall efficiency of the freeway network.

What makes Q-learning particularly interesting is its ability to learn from experience without needing a complete model of the environment. This model-free nature allows it to be applied in complex and unpredictable scenarios, where traffic patterns can change rapidly due to accidents, weather conditions, or other unforeseen events. The agent's continuous learning process means it can adapt to these changes and still strive to optimize traffic flow.

However, the implementation of Q-learning is not without its challenges. Both papers point out the issue of the "curse of dimensionality," a common problem in reinforcement learning where the number of possible states and actions becomes so large that it is difficult for the algorithm to efficiently explore all the options. This can slow down learning and make it harder for the agent to find the best policies, especially in large and complex traffic networks. Moreover, Q-learning requires a significant amount of data to learn effectively. In real-world scenarios, gathering such data can be expensive and time-consuming, and the variability in traffic conditions means that the agent needs a robust and diverse set of experiences to generalize well.

Despite these challenges, the use of Q-learning in ramp metering represents a significant advancement in the field of traffic management. By enabling systems to learn and adapt in real-time, Q-learning offers a promising alternative to static control methods, providing a more flexible and responsive approach to managing the complex and ever-changing nature of freeway traffic. The work of these researchers underscores the potential of reinforcement learning to transform how we approach traffic control, moving towards systems that can autonomously improve their performance over time, making our roads safer and less congested.

3.3.2 Deep Reinforcement Learning

In their paper [18], Yu Han and colleagues explore the innovative application of Deep Reinforcement Learning (DRL) to tackle the complex challenge of coordinated ramp metering control, which is essential for modern traffic management. DRL stands out as a significant advancement because it integrates the strengths of deep learning with the adaptive capabilities of reinforcement learning. This powerful combination allows the system to navigate and optimize large, intricate state-action spaces by using neural networks to approximate complex relationships within the data. Techniques such as Deep Q-Networks (DQN) and Batch-constrained deep Q-learning (BCQ) enable DRL models to process vast amounts of high-dimensional data, including real-time traffic patterns, vehicle movements, and congestion levels, allowing the system to make informed and sophisticated decisions about how to control ramp metering rates.

The research by Han and his team demonstrates DRL's potential to revolutionize both local and coordinated ramp metering strategies by enabling these systems to dynamically adapt to fluctuating traffic conditions. Unlike traditional methods like ALINEA, which are effective in simpler scenarios but rely on pre-set rules, DRL systems continuously learn and evolve by interacting with their environment. This ongoing learning process enables DRL to excel in more complex and varied traffic scenarios, making it capable of significantly reducing congestion and improving overall traffic flow efficiency.

The simulations conducted by Han et al. underscore the impressive capabilities of DRL in optimizing traffic management. By leveraging extensive and diverse datasets, these models learn to recognize intricate traffic patterns and adjust their strategies accordingly. For example, in a heavily congested area, the system might limit the number of vehicles entering the freeway to prevent gridlock, while in lighter traffic conditions, it might allow more vehicles to merge, maintaining a steady flow. This adaptability makes DRL particularly promising for urban environments where traffic conditions can change rapidly and unpredictably.

However, DRL implementation in real-world traffic systems presents its own challenges. One of the primary obstacles is the significant computational power required to train these models. Training a DRL model involves processing large datasets and running numerous simulations, which can be both time-consuming and resource-intensive. Additionally, there is a risk of overfitting, where the model becomes so finely tuned to the training data that it struggles to perform well in new, unseen situations. This is a common issue in machine learning, where balancing learning from data and generalizing to new scenarios is crucial.

Another challenge with DRL lies in the complexity of the models. Unlike traditional traffic management systems, which are often rule-based and relatively easy to interpret, DRL models function more like ”black boxes.” This lack of transparency can make it difficult for traffic engineers to understand how the system is making decisions, which poses a significant obstacle when validating the model and ensuring it behaves as expected in all scenarios. The complexity of these models also makes them more challenging to deploy in real-world environments, where rigorous testing and validation are necessary to ensure reliability and effectiveness.

Despite these challenges, the potential benefits of DRL for traffic management are undeniable. As cities continue to grow and traffic networks become more complex, the need for intelligent, adaptive systems that can efficiently manage traffic and reduce congestion will only increase. DRL offers a promising solution, with the ability to learn, adapt, and improve over time, making roads safer, less congested, and more efficient. The work of Han and his colleagues is paving the way for a new era in traffic management, where advanced technologies like DRL play a central role in keeping cities moving smoothly and efficiently.

3.4 Multi-Agent Reinforcement Learning (MARL) Approaches

In [25], Ahmed Fares and Walid Gomaa introduce the concept of Multi-Agent Reinforcement Learning (MARL) as a novel approach to managing ramp metering on freeways. MARL involves multiple reinforcement learning agents, each responsible for controlling a specific ramp in the network. These agents operate either cooperatively or competitively, depending on the specific design of the system, to achieve the overarching goal of optimizing traffic flow across the entire freeway network. Each agent learns to make decisions based on local observations, such as traffic density and flow at its assigned ramp, while also considering the impact of its actions on the overall network. This decentralized approach allows agents to respond quickly to local traffic conditions, but it also requires sophisticated coordination mechanisms to ensure that these local decisions contribute positively to global traffic management goals.

Building on this foundation, the paper [26] explores more advanced methods for achieving coordination among multiple agents. The authors emphasize the importance of communication protocols and shared reward structures in MARL, which are essential for aligning the objectives of individual agents with the overall system performance. By sharing information and rewards, agents can learn not just to optimize their local environment but also to contribute to the efficiency of the entire traffic network.

The results from these studies demonstrate that MARL can effectively manage complex traffic systems, offering a significant advantage over traditional single-agent approaches. The ability to decentralize control means that each agent can make decisions tailored to its specific ramp, leading to more responsive and adaptive traffic management. However, the complexity of coordinating multiple agents introduces significant challenges. The communication and synchronization between agents can be computationally demanding, and designing reward structures that balance both local and global traffic objectives is far from straightforward. As the number of agents in the system increases, scalability becomes a critical issue, potentially limiting the practicality of MARL in very large or highly dynamic traffic networks.

Despite these challenges, the potential of MARL to transform ramp metering and broader traffic management strategies is clear. By enabling a more flexible and distributed approach to traffic control, MARL offers a promising pathway toward more intelligent and adaptive traffic systems that can better cope with the complexities of modern urban environments. The work of these researchers highlights both the opportunities and the hurdles in implementing MARL, paving the way for future advancements in this exciting area of study.

3.5 Hybrid Approaches in Ramp Metering:

Hybrid approaches in ramp metering are strategies that combine traditional traffic control methods with reinforcement learning (RL) techniques to optimize traffic flow more effectively than either approach could achieve on its own. These methods aim to leverage the strengths of both conventional control systems and modern machine learning algorithms, addressing the limitations of each.

3.5.1 Physics-Informed RL

In the paper [18], Yu Han and colleagues introduce an innovative approach that blends physics-based traffic flow models with reinforcement learning (RL) to enhance the effectiveness of ramp metering systems. This strategy represents a significant advancement in traffic management by ensuring that the RL agent benefits from a more structured and informed learning process. Specifically, the RL agent is trained using both historical traffic data and synthetic data generated from a macroscopic traffic model like MetaNet as shown in the Figure 3.1. The integration of a physics-based model serves as a guiding framework for the RL agent, helping it avoid suboptimal policies by anchoring its learning in real-world traffic dynamics.

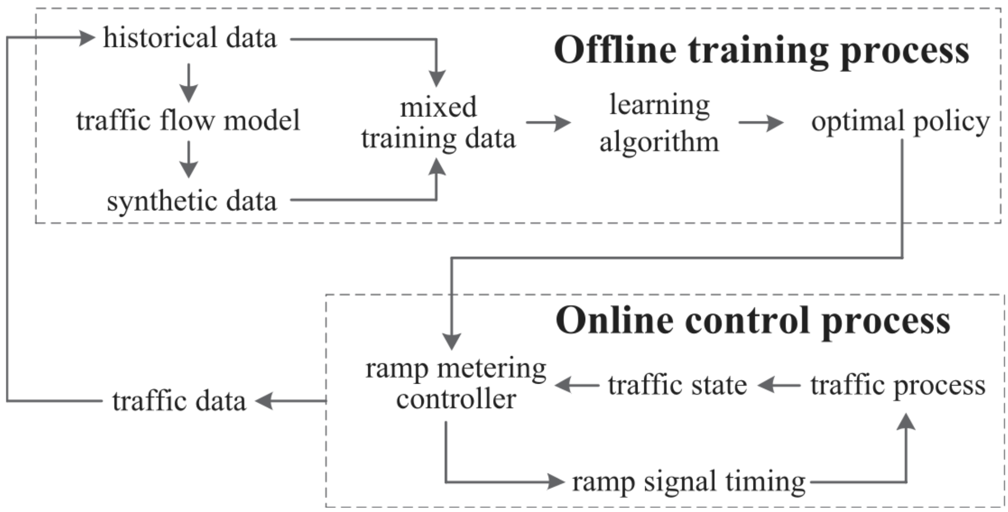

Figure 3.1: The framework of the iterative RL-based ramp metering control strategy.

The methodology involves an iterative training process where the RL model is continuously updated with new batches of data as shown in the Figure 3.2. This dynamic updating allows the system to refine its policy over time, responding to both real-world and model-generated traffic scenarios. By incorporating knowledge from physics-based models, the RL agent can make more informed decisions, leading to a more reliable and effective ramp metering strategy.

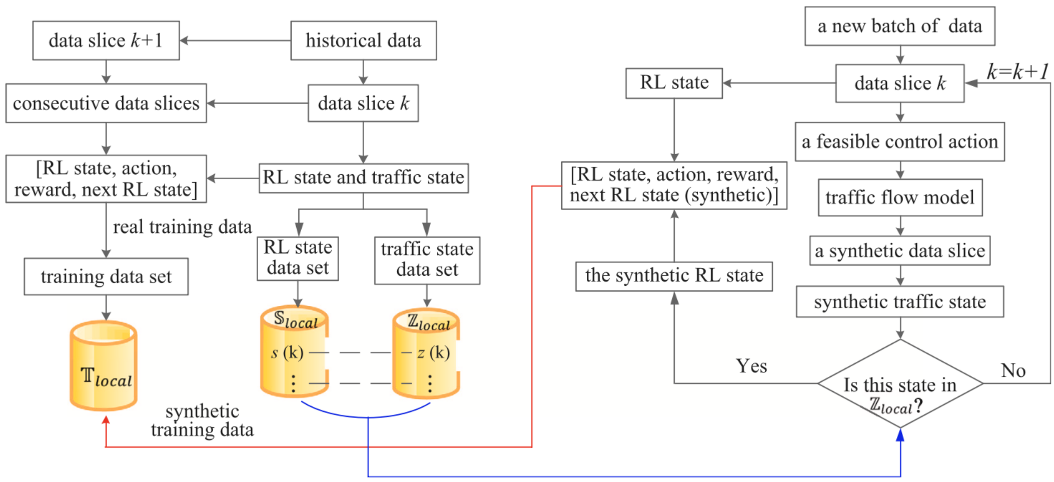

Figure 3.2: The procedure of synthetic data generation in the offline learning of the local ramp metering approach.

The results of Han et al.'s research demonstrate the clear advantages of this physics-informed RL approach. In simulation experiments, it outperforms purely simulation-based RL methods, particularly in metrics like total time spent (TTS) on the freeway and overall traffic throughput. This superior performance is largely attributed to the fact that the RL agent is not solely reliant on simulated data, which can sometimes lead to overfitting or unrealistic policy development. Instead, by grounding its learning in both real and modeled data, the agent can generalize better to real-world traffic conditions, leading to more robust and effective traffic management.

However, the approach can have some challenges. The success of the physics-informed RL strategy is heavily dependent on the accuracy of the underlying physics-based traffic models. If these models do not accurately reflect real-world traffic dynamics, there is a risk that the RL agent may learn policies that are suboptimal or even counterproductive.

Additionally, integrating and synchronizing different types of data—such as real-world observations and model-generated simulations—can be both complex and computationally demanding. This process requires careful calibration to ensure that the various data sources align correctly and work together seamlessly. The need for substantial computational resources adds another layer of challenge, potentially creating a barrier to practical implementation, especially in resource-constrained environments. Overcoming these hurdles is crucial for making this approach viable on a larger scale.

Despite these challenges, the work of Han and his colleagues represents a promising direction in the field of traffic management, showing how hybrid approaches that combine the strengths of traditional modeling and modern machine learning can lead to significant improvements in system performance.

3.5.2 Decentralized and Coordinated Control

Another solution was proposed by Yu Han and colleagues in their paper [18], where they introduced an innovative decentralized and coordinated approach to ramp metering, which leverages the power of multiple reinforcement learning (RL) agents to manage complex freeway networks more efficiently. This approach marks a departure from traditional centralized control systems by dividing the freeway network into smaller, more manageable subsystems, each controlled by its own RL agent. These agents operate independently, making real-time decisions based on local traffic conditions. However, what makes this approach truly effective is the coordination among these agents. By working together, the agents aim to optimize the overall network performance, ensuring that traffic flows smoothly across the entire system.

Coordination in this decentralized system can be achieved through a hierarchical control structure. In this setup, higher-level agents are responsible for managing broader traffic objectives, such as preventing bottlenecks and ensuring overall network efficiency. Meanwhile, lower-level agents focus on more localized tasks, such as adjusting ramp metering rates based on immediate traffic conditions. This hierarchical approach not only reduces the complexity and computational load that typically burden centralized control systems but also allows for greater flexibility and adaptability in responding to dynamic traffic patterns. The procedure of synthetic data generation process of this coordinated ramp metering approach is shown in the Figure 3.3.

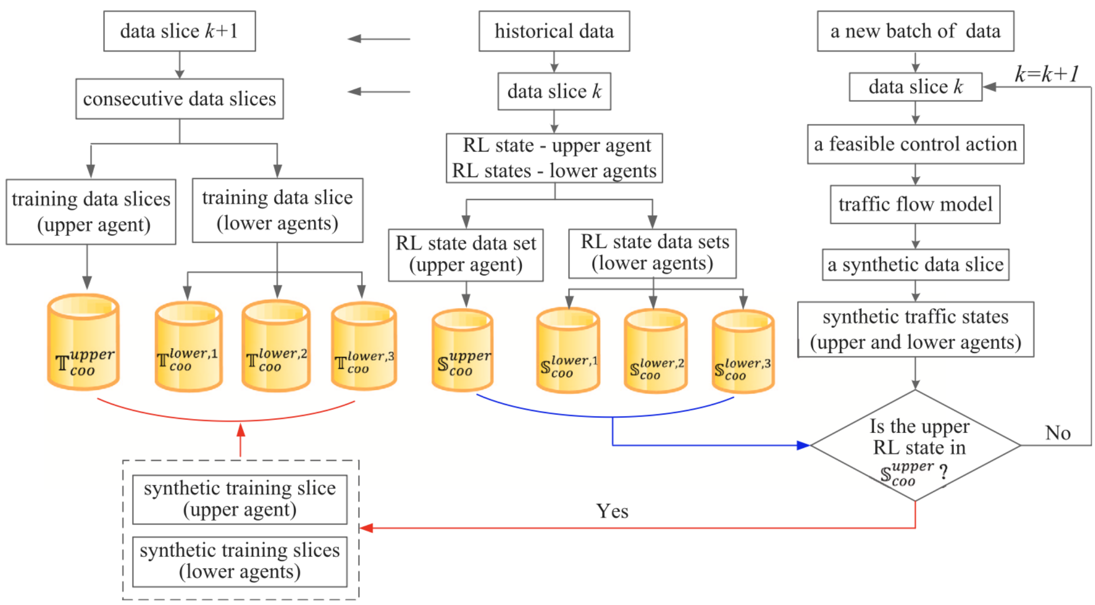

Figure 3.3: The procedure of synthetic data generation process of the coordinated ramp metering approach.

The results of Han et al.'s research highlight the significant advantages of this decentralized and coordinated control system. By breaking down the traffic management problem into smaller, more focused parts, the system can handle large and complex networks with greater scalability and robustness. Each agent's ability to make independent decisions ensures that local traffic conditions are addressed promptly, while the overall coordination ensures that these local actions contribute positively to the broader traffic management goals.

However, this approach is not without its challenges. One of the primary difficulties lies in ensuring effective communication and coordination between the agents, especially as the network becomes more complex and the number of agents increases. Balancing local optimizations with global performance goals can also be tricky, as what benefits one part of the network might negatively impact another. This potential for conflict requires sophisticated algorithms and communication protocols to ensure that all agents are aligned with the overall objectives of the system. Despite these challenges, the decentralized and coordinated approach offers a promising pathway toward more intelligent and adaptive traffic management systems, capable of meeting the demands of increasingly complex urban environments.

3.6 Comparative summary between the ramp metering approaches

Previously, we have presented the main approaches in the field of ramp metering optimization. In the Table 3.1, we will carry out a comparative study of the approaches proposed above according to the following factors:

- **Strategy**: Approach used in paper.
- **Control Method**: How system is controlled.
- **Adaptability to Traffic Conditions**: Responds to traffic changes.
- **Complexity of Implementation**: Difficulty of applying solution.

While in the Table 3.2, we presented the same approaches with the following factors :

- **Paper** : Presents the paper
- **Approach** : Present the approach adopted in the paper
- **Methodology** : Present the methodology adopted in the paper

- **Results and limitations** : Present the Key findings and limitations of each approach
- **Simulator** : The simulator used in each approach.

Table 3.1: Comparison of Ramp Metering Strategies

Strategy	**Control Method**	**Adaptability to Traffic Conditions**	**Complexity of Implementation**
Demand-Capacity Control [20]	Rule-based, balances inflow with capacity	Low, reactive and not adaptive to dynamic changes	Low, simple implementation but limited effectiveness in complex scenarios
ALINEA [19]	Feedback-based, adjusts metering rate based on downstream occupancy	Moderate, effective for specific bottlenecks but limited in network-wide scenarios	Moderate, requires precise tuning and is localized
PI-ALINEA	Proportional-Integral controller added to ALINEA	Moderate, more responsive to traffic trends over time	Moderate, enhanced over ALINEA but still localized
AMOC [21]	Network-wide optimal control using real-time data	High, continuously updates strategies based on real-time conditions	High, requires substantial computational resources and real-time data
ANCONA [22]	Phase-based control focused on maintaining synchronized flow	Moderate, effective in preventing wide moving jams	High, complex implementation with a need for precise calibration
Stratified Zone Metering (SZM) [23]	Zone-based, adjusts metering rates for specific freeway segments	High, tailored control for specific congestion-prone areas	High, requires detailed traffic data and precise calibration
Multi-Agent Reinforcement Learning (MARL) [25]	Decentralized, RL agents control individual ramps	High, agents adapt to local conditions while optimizing network-wide flow	High, complex coordination and communication between agents
Q-Learning [18]	Model-free RL, learns optimal policies through interaction with traffic environment	High, adapts to changing conditions without predefined models	High, faces challenges related to the curse of dimensionality
Deep Reinforcement Learning (DRL) [18]	Combines deep learning with RL to manage complex state-action spaces	Very High, capable of handling intricate traffic scenarios	Very High, computationally demanding and risks of overfitting
Physics-Informed RL [18]	Hybrid, combines traditional traffic models with RL	High, grounded in real-world dynamics with enhanced adaptability	High, success depends on the accuracy of traffic models and data integration
Decentralized Coordinated Control [18]	Multi-agent RL with hierarchical control structure	Very High, flexible and scalable for complex networks	High, challenges in ensuring effective communication among agents

Table 3.2: Comparative Summary of Ramp Metering Approaches

Papers	**Approach**	**Methodology**	**Results and Limitations**	**Simulator**
Han et al. (2022) [18]	**ALINEA and PI-ALINEA**	Feedback-based control adjusting ramp metering rates based on freeway occupancy.	Effective but requires precise tuning and limited to local scenarios.	VISSIM
Wattleworth (1967) [20]	**Demand-Capacity/ Occupancy-Capacity**	Regulates entry based on capacity differences or occupancy levels.	Prevents bottlenecks but less adaptive to rapid changes.	N/A
Kotsialos and Papageorgiou (2004) [21]	**AMOC**	Proactive network-wide optimization of traffic flow.	Improves travel times but requires high-quality data and is computationally intensive.	VISSIM
Kerner (2006) [22]	**ANCONA**	Manages traffic by maintaining synchronized flow.	Reduces traffic jams but requires precise data and complex management.	N/A
Srivastava (2011) [23]	**SZM Algorithm**	Applies metering based on traffic density in defined zones.	Effective in high-traffic areas but relies on detailed data.	N/A
Motorway Ramp-Metering Control (2011) [24], **Han et al. (2022)** [18]	**Q-Learning**	Model-free RL learning optimal actions from traffic state observations.	Adapts well but faces dimensionality challenges and needs extensive data.	TORCS
Han et al. (2022) [18]	**DRL**	Integrates deep learning with RL for managing large state-action spaces.	Superior in simulations but computationally demanding and complex.	TORCS
Fares and Gomaa (2014) [25], **Han et al. (2020)** [26]	**MARL**	Multiple RL agents coordinating actions across a network.	Effective but requires complex coordination and is less scalable.	VISSIM
Han et al. (2022) [18]	**Physics-Informed RL**	Combines RL with physics-based traffic models.	Outperforms simulation-only methods but relies on accurate models and data integration.	VISSIM
Han et al. (2022) [18]	**Decentralized/ Coordinated Control**	Independent RL agents managing smaller subsystems with coordination.	Scalable and robust but requires effective communication.	VISSIM

3.7 Conclusion

In this chapter we have explored the evolution of ramp metering strategies, starting from early feedback-based control systems like ALINEA and PI-ALINEA to the latest advancements in reinforcement learning and hybrid approaches. The development of ramp metering systems has been driven by the need to effectively manage increasing traffic demand and reduce congestion, a challenge that has become more complex with the growth of modern urban environments.

Traditional methods like demand-capacity and occupancy-capacity approaches laid a strong foundation for traffic management but struggled with unpredictable traffic and network-wide control. This led to the development of more advanced solutions, including reinforcement learning (RL) and multi-agent RL, which adapt dynamically to changing traffic conditions and optimize flow through continuous learning and decentralized decision-making. Innovations such as deep RL and physics-informed RL leverage extensive traffic data to create more flexible ramp metering strategies, while hybrid approaches blend RL with traditional models, striking a balance between rule-based systems and adaptive learning. These advancements have significantly improved both localized and network-wide traffic management, enhancing road safety and efficiency.

In conclusion, despite recent breakthroughs, challenges remain in the real-world implementation of advanced ramp metering systems due to data requirements, computational complexity, and scalability limitations. Issues like the ”curse of dimensionality” in RL-based approaches and coordination challenges in MARL systems call for further research. However, the integration of machine learning with traditional traffic control offers a promising path forward. Continued exploration of hybrid, decentralized approaches, improved data collection, and greater computational efficiency will be key to developing adaptive, scalable, and robust traffic management solutions for modern urban networks.

General Conclusion

This thesis has provided a comprehensive exploration of the application of machine learning, particularly reinforcement learning, in the optimization of ramp metering strategies. The research presented highlights the growing relevance of data-driven approaches in traffic management and demonstrates how reinforcement learning can be applied to address the complex and dynamic nature of modern transportation networks.

The background review illustrated the fundamental principles of machine learning and reinforcement learning, introducing key algorithms and their role in traffic management. SUMO and MetaNet, two important simulation tools, were discussed in detail, emphasizing their importance in modeling traffic flows and validating new ramp metering strategies.

The discussion of deep reinforcement learning and its integration into ramp metering strategies showcased the adaptability and effectiveness of these systems in optimizing real-time traffic. Traditional ramp metering methods, such as ALINEA, PI-ALINEA, and demand-capacity approaches, served as a strong foundation for understanding the limitations of static, rule-based systems, while the exploration of reinforcement learning-based approaches, particularly Q-learning and MARL, demonstrated the potential for more flexible and responsive traffic control mechanisms.

Additionally, the analysis of hybrid approaches, such as physics-informed RL and decentralized control systems, provided insight into how the strengths of traditional and modern methods can be combined to create more robust, scalable, and effective ramp metering solutions. These hybrid methods offer a promising pathway for future research and development, as they balance the predictability of traditional models with the adaptability of machine learning techniques.

In conclusion, the integration of advanced machine learning techniques, particularly reinforcement learning and its multi-agent and hybrid extensions, presents a promising direction for future traffic management solutions. As urban traffic networks continue to grow in complexity, the ability to apply adaptive, scalable, and efficient strategies will be crucial in meeting the demands of modern transportation. These advancements offer the potential to revolutionize traffic control systems, leading to safer, more efficient highways and a better quality of life in urban environments.

Bibliography

[1] Michael Behrisch, Laura Bieker-Walz, Jakob Erdmann, and Daniel Krajzewicz. Sumo – simulation of urban mobility: An overview. In *Proceedings of SIMUL 2011, The Third International Conference on Advances in System Simulation*. DLR, 2011. URL https://www.researchgate.net/publication/225022282_SUMO_-_Simulation_of_Urban_MObility_An_Overview.

[2] A. Messmer and M. Papageorgiou. Metanet: A macroscopic simulation program for motorway networks. *Traffic Engineering & Control*, 31:466–470, 1990.

[3] M. Papageorgiou, Hani S. Mahmassani, and Andreas Kotsialos. A review of dynamic traffic assignment models. *European Journal of Operational Research*, 218(2):300–313, 2012.

[4] Andreas Kotsialos, M. Papageorgiou, and C. Diakaki. Traffic flow modeling of large-scale motorway networks using the macroscopic modeling tool metanet. *IEEE Transactions on Intelligent Transportation Systems*, 3(4):282–292, 2002.

[5] Google Cloud. What is artificial intelligence?, 2024. URL https://cloud.google.com/learn/what-is-artificial-intelligence#artificial-intelligence-defined. Accessed on: 2024-07-18.

[6] Irina Kolesnikova. How ai in transportation can improve our everyday lives, 2023. URL https://mindtitan.com/resources/blog/ai-in-transportation/. Accessed on : 2024-07-19.

[7] Tom M. Mitchell. *Machine Learning*. McGraw-Hill, 1997. URL https://www.cs.cmu.edu/~tom/mlbook.html.

[8] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006. URL https://www.springer.com/gp/book/9780387310732.

[9] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. URL https://www.deeplearningbook.org/.

[10] Phil Tabor. Modern reinforcement learning: Deep q learning in pytorch, 2020. URL https://www.udemy.com/course/deep-q-learning-from-paper-to-code. Udemy Course.

[11] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, second edition, 2018. URL http://incompleteideas.net/book/the-book-2nd.html.

[12] Romain Ducrocq. Deep reinforcement learning for traffic signal control in partial detection environments. Master's thesis, Université Gustave Eiffel, Marne-la-Vallée, France, 2021. Master Thesis.

[13] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, 1998. URL http://incompleteideas.net/book/the-book-2nd.html.

[14] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg, and Demis Hassabis. Human-level control through deep reinforcement learning. *Nature*, 518:529–533, 2015. URL https://www.nature.com/articles/nature14236.

[15] Christopher J. C. H. Watkins and Peter Dayan. Q-learning. *Machine Learning*, 8:279–292, 1992. URL https://link.springer.com/article/10.1007/BF00992698.

[16] David Silver, Guy Lever, Nicolas Heess, Thomas Degris, Daan Wierstra, and Martin Riedmiller. Deterministic policy gradient algorithms. *International Conference on Machine Learning (ICML)*, 2014.

[17] John Schulman, Sergey Levine, Philipp Moritz, Michael I Jordan, and Pieter Abbeel. Trust region policy optimization. In *International Conference on Machine Learning (ICML)*, pages 1889–1897, 2015.

[18] Yu Han, Meng Wang, Linghui Li, Claudio Roncoli, Jinda Gao, and Pan Liu. A physics-informed reinforcement learning-based strategy for local and coordinated ramp metering. *Transportation Research Part C: Emerging Technologies*, 137:103584, 2022.

[19] M. Papageorgiou, H. Hadj-Salem, and J.-M. Blosseville. Alinea: A local feedback control law for on-ramp metering. *Transp. Res. Rec.*, 1320(1):58–67, 1991.

[20] John A. Wattleworth. Peak period analysis and control of a freeway system. In *Highway Research Record*, number 157, pages 1–21. National Research Council (USA), Highway Research Board, National Academies Press, 1967.

[21] Andreas Kotsialos and Markos Papageorgiou. Efficiency and equity properties of freeway network-wide ramp metering with amoc. *Transportation Research Part C: Emerging Technologies*, 12:401–420, 2004. URL `10.1016/j.trc.2004.07.007`.

[22] Boris S. Kerner. Probabilistic nature of breakdown phenomenon and on-ramp metering in three-phase traffic theory. In *IFAC Proceedings Volumes*, volume 39, pages 273–278. Elsevier, 2006. doi: 10.3182/20060829-3-NL-2908.00048.

[23] Amit Srivastava. Development of next generation ramp metering algorithm based on freeway density. Master's thesis, University of Minnesota, 2011.

[24] Mohsen Davarynejad, Andreas Hegyi, Jos Vrancken, and Jan van den Berg. Motorway ramp-metering control with queuing consideration using q-learning. In *14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pages 1652–1658. IEEE, 2011.

[25] Ahmed Fares and Walid Gomaa. Multi-agent reinforcement learning control for ramp metering. In *2014 International Conference on Engineering and Technology (ICET)*, pages 1–6. IEEE, 2014.

[26] Jiyuan Tan, Qianqian Qiu, and Weiwei Guo. Coordinated ramp metering control based on multi-agent reinforcement learning. In *2020 35th Youth Academic Annual Conference of Chinese Association of Automation (YAC)*. IEEE, 2020.