

The Application of Multi-Agent Reinforcement Learning in Path Planning for Automated Parking

Jiawei Guo

*Institute of Problem Solving, Hainan International College, Communication University of China,
Tianjin, China
guojiawei114@outlook.com*

Abstract. With the advancement of artificial intelligence technologies, autonomous vehicles are progressively emerging as a transformative transportation solution. However, decision-making complexities and dynamic interaction challenges in complex environments continue to impede the realization of fully autonomous driving. As an emerging paradigm, Multi-Agent Reinforcement Learning (MARL) offers an innovative perspective to address these challenges. This paper provides a systematic review of recent research advancements in MARL applications for automated parking: beginning with an exposition of the technical foundations of automated parking and MARL theoretical principles; proceeding to a detailed analysis of core algorithmic implementation mechanisms—including Q-learning and Deep Q-Networks (DQN)—and their specific applications in parking path planning; subsequently synthesizing mainstream metric systems for performance evaluation; and concluding with an examination of critical challenges such as learning efficiency and safety limitations, alongside prospective research directions. This review aims to clarify developmental trajectories within the field while highlighting MARL's significant potential for constructing intelligent, high-efficiency automated parking solutions.

Keywords: Automated Parking, Path Planning, Multi-Agent Reinforcement Learning (MARL), Deep Reinforcement Learning.

1. Introduction

Automated parking technology stands as a significant subfield within autonomous driving and has emerged as a key research area in recent years [1]. The technology relies on sensors and tracking systems to identify viable parking spaces and subsequently plan an optimal vehicle trajectory. The necessity for advancing this technology is underscored by the projection that future transportation systems will incorporate fully autonomous vehicles. Presently, semi-autonomous features are already prevalent, and a higher degree of vehicle autonomy is anticipated in the near future. The foundational ambition of autonomous driving is to establish a coordinated, interconnected traffic ecosystem that enhances safety and reduces travel time. Within this context, automated parking addresses a critical and high-demand practical need [2]. The annual increase in vehicle ownership has intensified parking congestion, especially in high-traffic locations like hospitals and tourist

destinations. Consequently, substantial investment in both time and technological development for automated parking remains essential [3].

The application of Reinforcement Learning (RL) to Multi-Agent Systems (MAS) has given rise to Multi-Agent Reinforcement Learning (MARL). At its core, MARL investigates how multiple agents can effectively accomplish tasks through interaction and cooperation, thereby improving their collective capabilities. Within the MAS framework, individual agents operate with distinct states, actions, and reward functions, working collaboratively toward a shared objective [4]. This introduces complex aspects absent in single-agent reinforcement learning, as multi-agent reinforcement learning must account for dynamic interactions and mutual influences among agents. Within multi-agent reinforcement learning environments, agents may employ diverse strategies, such as cooperation, competition, or coordination—which can be either learned through experience or derived from predefined rules to achieve globally optimal solutions.

2. MARL for automated parking path planning

The automated parking technology process is primarily divided into two aspects: the perception aspect and the path planning aspect. The perception aspect mainly constructs environmental representations through hardware equipment such as sensors; enhancement relies on hardware enhancement. The planning aspect is based on how to compute optimal vehicle trajectories. This process fundamentally belongs to an algorithm optimization problem. Within the path planning aspect, Multi-Agent Reinforcement Learning (MARL) is one prominent area receiving considerable attention from the research community in recent years. It primarily aims to achieve globally optimal paths. Over the past few decades, the research community has developed various MARL methodological frameworks. MARL has precisely evolved step by step based on these foundational aspects.

Figure 1 shows the chronological progression of the relevant technological developments.

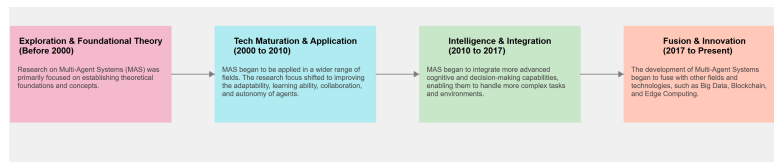


Figure 1. Development timeline of multi-agent reinforcement learning methodologies [4]

2.1. Technical pipeline and foundations

Multi-agent reinforcement learning algorithms have been applied multiple times in automated parking path planning. As one of the core operational components of automated parking, the methodological system of path planning has evolved from geometry-based rule-driven approaches to learning-based paradigms. Compared with traditional schemes that directly apply algorithms such as A* (these algorithms struggle to simulate human parking behavior in complex environments), reinforcement learning (RL) enhances the generation capability of potential paths through autonomous learning mechanisms [5]. Reinforcement learning effectively resolves these limitations by enabling agents to be able to learn optimal paths through direct interaction with the environment. The subsequent emergence of Q-learning represents a foundational advancement in RL methodologies. As a core algorithmic framework, Q-learning employs a Q-value table (Q-table) to store expected rewards for state-action pairs. Through iterative updates to this table, agents progressively converge toward optimal policies [6]. Nevertheless, Q-learning exhibits significant

limitations when confronted with the high-dimensional state spaces characteristic of real-world operational environments.

To address this, deep learning has been integrated with RL, leading to the development of Deep Q-Networks (DQN). DQN replaces the Q-table with a deep neural network to approximate the Q-value function, enabling it to handle complex, high-dimensional inputs such as images from a camera [7]. DQN also introduces an experience replay mechanism, which stores past transitions in a memory buffer and samples from it randomly during training. This breaks the correlation between consecutive samples and stabilizes the learning process. These model-free approaches are popular; however, Model-Based Reinforcement Learning (MBRL), where the agent first learns a model of the environment's dynamics, also exists and can be combined with model-free methods to improve sample efficiency [8].

2.2. Specific applications and methods

Researchers have developed specialized MARL frameworks to tackle the unique challenges of automated parking, such as collision avoidance and efficient coordination. These approaches are demonstrated in several key studies:

To manage a fleet of vehicles in an automated valet system, some studies employ a Multi-Agent Deep Deterministic Policy Gradient (MADDPG) algorithm. The reward function in these models is typically designed to penalize collisions and long parking times, successfully improving both parking efficiency and the overall success rate compared to non-cooperative methods. A similar MADDPG model, using centralized training and decentralized execution, has also been applied to solve complex multi-agent path planning for ships and aircraft carriers [9]. To address congestion at entry and exit points, Nishio et al. developed a decentralized cooperation mechanism where each vehicle acts as a DQN agent [10]. A key innovation was a shared "experience" mechanism that allowed agents to learn from the trajectories of others, significantly reducing average waiting and parking times.

For the congestion problem at entry and exit points, Nishio et al. applied a decentralized cooperation mechanism where each vehicle operates as a DQN [10]. It allows agents to learn from the trajectories of other vehicles, significantly reducing the average waiting and parking times.

Unlike the method that directly applies multi-agent reinforcement learning. Also there are improved versions of traditional algorithms, the hybrid application of traditional algorithms with modern new technologies. Its effect also shows advantages. The research community also pays attention to the combination of traditional algorithms and modern control. The HOPE planner proposed by JiangLi's team combines classical geometric algorithms with deep reinforcement learning strategies, mainly applying the mechanism of action mask to overcome the problem that geometric methods cannot adapt in complex environments. It achieved extremely high data in parking success rate [11]. Similarly, Zhu et al. through the integration of Hybrid A* algorithm and Model Predictive Control (MPC), the MPC control algorithm provides future state information and real-time output of chassis control parameters, ensuring the vehicle completes the automated parking task along the dynamically calculated path, significantly improving path planning accuracy compared to traditional methods [2].

Reinforcement Learning (RL) also has applications in long-distance searching for optimal parking spaces. For example, Khalid et al. proposed the RL-LAVP algorithm—this reinforcement learning algorithm is specially used to guide vehicles to optimal parking locations [12]. Their algorithm increases the search distance while reducing energy consumption. It has an effective solution for the problem of tight parking spaces in urban areas.

3. Mainstream evaluation metrics

The evaluation of automated parking systems is a multi-faceted process that requires a comprehensive set of metrics. These standards can be broadly categorized into three main areas: functional performance, safety, and user experience.

3.1. Functional performance

This dimension evaluates the system's core capability to efficiently complete parking tasks. Key metrics include: **Parking Space Recognition Rate**: Quantifies the system's accuracy in detecting and classifying various parking spaces [13]. **Parking Success Rate**: Defined as the ratio of successful parking attempts, including statistics on trajectory adjustment counts. **Pose Accuracy**: Assesses the precision of the vehicle's final pose, with emphasis on: **Lateral deviation** (distance from the ideal centerline) **Angular yaw difference**. **Efficiency**: Measured through the following parameters: **Task completion time** (from initiation to parking completion) **Planned path smoothness** [5]. This structured translation maintains academic rigor, ensures technical terminology consistency, and adheres to standard research formatting conventions.

3.2. Safety

Safety is the most critical aspect of any autonomous function. Evaluation metrics focus on the system's ability to operate without causing collisions. This includes the **Emergency Braking** and **Obstacle Avoidance** capability, which assesses responsiveness to sudden obstacles. In MARL simulations, this is often quantified by the collision rate among agents. **Minimum Clearance Control** measures the ability to maintain a safe distance from all surrounding obstacles, a parameter heavily dependent on sensor precision [14]. Another important factor is the **Self-Recovery Capability**, which is the system's ability to resume the parking process automatically after a temporary interruption, such as resolving a deadlock situation between two or more vehicles.

3.3. User experience

This dimension measures how intuitive, comfortable, and trustworthy the system is from the driver's perspective. **Interaction Convenience** refers to the ease of activating and monitoring the function through interfaces like touchscreens or voice commands. **Ride Comfort** is determined by the smoothness of the vehicle's motion, including controlled acceleration, deceleration, and steering to avoid abrupt movements [15]. The **Level of Automation** is also crucial, ranging from L2 systems requiring driver supervision to L4 systems that offer fully autonomous, "mind-off" operation, which directly impacts the user's trust and cognitive load.

4. Existing limitations and future outlook

Despite the significant progress in applying MARL to automated parking, several formidable challenges remain that must be addressed to enable widespread, reliable deployment. These challenges span learning efficiency, environmental complexity, and the socio-technical aspects of trust and safety. One of the most significant hurdles is learning efficiency. MARL algorithms often require a vast amount of training data and computational resources to converge to an effective policy. This is because the learning process of other agents causes each agent's decision-making to continuously evolve, forming a dynamic learning environment. The sample inefficiency of many

reinforcement learning (RL) algorithms means that training in real-world complex scenarios is often impractical and potentially hazardous.

Closely related is the challenge of learning in complex environments. In multi-agent settings, from the perspective of any individual agent, the environment is complex because the policies of other agents are constantly changing during the learning process. This violates the stationarity assumption inherent in traditional reinforcement learning algorithms, leading to unstable learning dynamics [16]. Algorithms capable of effectively modeling and adapting to the behaviors of other agents—such as those based on opponent modeling or the centralized training with decentralized execution (CTDE) paradigm—are a key focus of current research [17]. Another critical issue is the lack of explainability. Most advanced Multi-Agent Reinforcement Learning (MARL) models, particularly those based on deep neural networks, operate as "black boxes," making it extremely difficult to interpret the rationale behind an agent's specific decision at a given moment. This lack of transparency poses a significant barrier to certification procedures and public trust, especially in safety-critical applications such as autonomous driving. Understanding the behavioral logic of a MARL-based parking system is crucial if it fails. Developing Explainable AI (XAI) methods specifically tailored to MARL systems represents an important future research direction [18].

Furthermore, the safety of learning-based agent systems is also a concern. Formal verification methods and techniques for certifying the safety of learning-based systems are required to guarantee that agents will not take unsafe actions. Future research will likely involve integrating learning-based approaches with traditional rule-based safety controllers to provide a robust safety net.

5. Conclusions

This article details the application of multi-agent reinforcement learning in automated parking. It is divided into two types: one introduces applications in path planning, the other introduces applications in long-distance parking space search. It then explains fundamental principles and developmental progression, starting from the most primitive geometric algorithms, gradually progressing to the A* algorithm, then to ordinary reinforcement learning (Q-learning), and further to current multi-agent reinforcement learning. It describes deficiencies of original algorithms. Nevertheless, some researchers combine traditional algorithms with modern technologies, achieving favorable outcomes using hybrid approaches. Subsequently, it introduces evaluation criteria for parking systems and existing challenges, such as insufficient learning efficiency and urgent demand for safety robustness. These limitations indicate core directions for future research. Regarding path planning, the research community still needs to focus on developing algorithms with higher sample efficiency, constructing interpretable and verifiable models, and designing robust safety protocols capable of handling unknown edge cases. In conclusion, the exploration of MARL in automated parking still remains filled with many unknowns, requiring diligent efforts from researchers to drive change.

References

- [1] Lan, W., Zhang, W., & Luo, J. (2017). Design of RFID indoor positioning system based on phase difference ranging. *Transducer Microsyst. Technol.*, 36(10), 85-91.
- [2] Zhu, C., Luo, Q., Xie, K., & Cao, W. (2024, November). Autonomous Parking Path Planning and Control for Unmanned Vehicles Based on Hybrid A* Algorithm and MPC Algorithm. In *2024 5th International Conference on Artificial Intelligence and Computer Engineering (ICAICE)* (pp. 1143-1147). IEEE.
- [3] Saxena, S., Isukapati, I. K., Smith, S. F., & Dolan, J. M. (2019, October). Multiagent sensor fusion for connected & autonomous vehicles to enhance navigation safety. In *2019 IEEE Intelligent Transportation Systems Conference*

- (ITSC) (pp. 2490-2495). IEEE.
- [4] Wang, R., & Dong, Q. (2024). A review of multi-agent reinforcement learning based on learning mechanism. *Chinese Journal of Engineering Science*, 46(7), 1251-1268.
 - [5] Liu, W., Li, Z., Li, L., & Wang, F. Y. (2017). Parking like a human: A direct trajectory planning solution. *IEEE Transactions on Intelligent Transportation Systems*, 18(12), 3388-3397.
 - [6] Watkins, C. J. C. H. (1989). Learning from delayed rewards.
 - [7] Yang, S., Shan, Z., Ding, Y., & Li, G. (2021). A review of deep reinforcement learning research. *Computer Engineering*, 47(12), 19-29.
 - [8] Aditya, M. O., Sujatha, C. N., Adithya, J., & Kiran, B. S. R. P. S. (2023, April). Automated valet parking using double deep Q learning. In *2023 International Conference on Advances in Electronics, Communication, Computing and Intelligent Information Systems (ICAECIS)* (pp. 259-264). IEEE.
 - [9] Shang, Z., Mao, Z., Zhang, H., & Xu, M. (2022, June). Collaborative path planning of multiple carrier-based aircraft based on multi-agent reinforcement learning. In *2022 23rd IEEE International Conference on Mobile Data Management (MDM)* (pp. 512-517). IEEE.
 - [10] Nishio, T., Zhao, M., Anzai, T., Kojima, K., Okada, K., & Inaba, M. (2021, May). Fixed-root aerial manipulator: Design, modeling, and control of multilink aerial arm to adhere foot module to ceilings using rotor thrust. In *2021 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 283-289). IEEE.
 - [11] Jiang, M., Li, Y., Zhang, S., Chen, S., Wang, C., & Yang, M. (2025). Hope: A reinforcement learning-based hybrid policy path planner for diverse parking scenarios. *IEEE Transactions on Intelligent Transportation Systems*.
 - [12] Khalid, M., Aslam, N., & Wang, L. (2020, October). A reinforcement learning based path guidance scheme for long-range autonomous valet parking in smart cities. In *2020 IEEE eighth international conference on communications and networking (ComNet)* (pp. 1-7). IEEE.
 - [13] Lee, H., Kim, S., Park, S., Jeong, Y., Lee, H., & Yi, K. (2017, June). AVM/LiDAR sensor based lane marking detection method for automated driving on complex urban roads. In *2017 IEEE Intelligent Vehicles Symposium (IV)* (pp. 1434-1439). IEEE.
 - [14] Liu, H., Luo, S., & Lu, J. (2019). Method for adaptive robust four-wheel localization and application in automatic parking systems. *IEEE Sensors Journal*, 19(22), 10644-10653.
 - [15] Naranjo, J. E., Gonzalez, C., Garcia, R., & De Pedro, T. (2008). Lane-change fuzzy control in autonomous vehicles for the overtaking maneuver. *IEEE Transactions on Intelligent Transportation Systems*, 9(3), 438-450.
 - [16] Lanctot, M., Zambaldi, V., Gruslys, A., Lazaridou, A., Tuyls, K., Pérolat, J., ... & Graepel, T. (2017). A unified game-theoretic approach to multiagent reinforcement learning. *Advances in neural information processing systems*, 30.
 - [17] Foerster, J., Assael, I. A., De Freitas, N., & Whiteson, S. (2016). Learning to communicate with deep multi-agent reinforcement learning. *Advances in neural information processing systems*, 29.
 - [18] Gilpin, L. H., Bau, D., Yuan, B. Z., Bajwa, A., Specter, M., & Kagal, L. (2018, October). Explaining explanations: An overview of interpretability of machine learning. In *2018 IEEE 5th International Conference on data science and advanced analytics (DSAA)* (pp. 80-89). IEEE.