

Learning in multi-agent systems to solve scheduling problems: a systematic literature review

*Aprendizaje en sistemas multi-agentes para resolver problemas
de scheduling: una revisión sistemática de la literatura*

Gabriel Icarte-Ahumada^{1*}  <https://orcid.org/0000-0002-1997-0053>

Johan Montoya¹  <https://orcid.org/0009-0009-5138-348X>

Zhangyuan He²  <https://orcid.org/0000-0003-2070-1819>

Recibido 5 de abril de 2024, aceptado 28 de mayo de 2024

Received: April 05, 2024 Accepted: May 28, 2024

ABSTRACT

Scheduling problems are ubiquitous in various domains, requiring efficient allocation of resources and coordination of tasks to optimize performance and meet desired objectives. Traditional approaches to scheduling often face challenges when dealing with complex and dynamic environments. In recent years, multi-agent systems have emerged as a promising paradigm for addressing scheduling problems. This paper presents a comprehensive survey of learning in multi-agent systems to solve scheduling problems. One hundred twenty-one articles were retrieved from the Scopus and WOS databases, 55 of which were reviewed and analyzed in depth. The results indicate that Reinforcement Learning (RL) is the learning model used in the reviewed articles. Our analysis also identified a tendency to combine two or more RL algorithms to be applied. Furthermore, most of the articles focus on solving dynamic scheduling problems in the manufacturing, wireless and communication network industries.

Keywords: Multi-agent systems, learning, scheduling problems.

RESUMEN

Los problemas de scheduling están presentes en varios dominios y requieren una asignación eficiente de recursos y coordinación de tareas para optimizar el rendimiento y cumplir los objetivos deseados. Los enfoques tradicionales para resolver este tipo de problemas a menudo enfrentan desafíos cuando se trata de entornos complejos y dinámicos. En los últimos años, los sistemas multi-agentes han surgido como un paradigma prometedor para abordar los problemas de scheduling. Este artículo presenta un estudio exhaustivo del aprendizaje en sistemas multi-agente para resolver problemas de scheduling. Se recuperaron un total de 121 artículos de las bases de datos Scopus y WOS, 55 de los cuales fueron revisados y analizados en profundidad. Los resultados indican que el Aprendizaje por Refuerzo (RL) es el modelo de aprendizaje utilizado en los artículos revisados. Nuestro análisis también identificó que existe una tendencia a combinar dos o más algoritmos de RL para su aplicación. Además, la mayoría de los artículos se centran en resolver problemas de scheduling dinámicos de la industria manufacturera y de redes inalámbricas y de comunicación.

Palabras clave: Sistemas Multi-agente, aprendizaje, problemas de scheduling.

¹ Universidad Arturo Prat. Facultad de Ingeniería y Arquitectura. Iquique, Chile.

E-mail: gicarte@unap.cl; johanstation15@gmail.com

² Shenzhen University. School of Urban Planning and Design. Shenzhen, China. E-mail: he.zhangyuan@pku.edu.cn

* Autor de correspondencia: gicarte@unap.cl

INTRODUCTION

Scheduling problems involve efficiently assigning resources such as time, personnel, and equipment to tasks or activities to optimize performance measures like completion time, resource utilization, and overall productivity [1]. Traditional centralized scheduling approaches rely on a single decision-maker to allocate resources, making it challenging to manage real-world environments' dynamic and uncertain nature. An alternative method to overcome the difficulties of centralized methods to solve scheduling problems is a multi-agent system.

Multi-agent systems (MAS) are computational systems composed of multiple autonomous agents that interact with each other and their environment to achieve individual and collective goals [2]. These systems find applications in various domains, including robotics, economics, social networks, and transportation, among others. In the context of scheduling problems, MAS offers a flexible and scalable approach to address the complex allocation of limited resources to perform a set of tasks or activities while considering constraints and objectives [3].

Learning plays a crucial role in enabling agents within MAS to acquire knowledge, adapt their behaviors, and enhance their decision-making abilities over time. By applying learning algorithms, agents can explore the solution space and dynamically adjust their actions to changing conditions and uncertainties. The combination of multi-agent systems, learning, and scheduling problems holds great potential for addressing the complexities of real-world scheduling scenarios. MAS's distributed decision-making capabilities allow for improved coordination and resource allocation while learning mechanisms empower agents to adapt and optimize their scheduling strategies [4]. This integration can lead to enhanced efficiency, robustness, and adaptability in scheduling processes, ultimately resulting in improved task completion times, better resource utilization, and overall performance.

In summary, multi-agent systems offer a promising approach to tackling scheduling problems by enabling distributed decision-making and adaptive strategies. By incorporating learning mechanisms, agents within these systems can improve their individual and

collective performance over time. This combination of MAS, learning, and scheduling problem-solving opens up exciting research opportunities and practical applications in various domains.

Developing an effective MAS to solve scheduling problems requires a clear understanding of the weaknesses and limitations of current MAS applied to solve scheduling problems that must be overcome. Thus, the objective of this paper is to present a comprehensive survey of learning in multi-agent systems to solve scheduling problems and the gaps encountered. To achieve this objective, a systematic literature review has been carried on based on the proposed procedure by [5]. As part of this work, research has been done on how MAS has been applied, the type of programming problems solved, and the learning methods used.

The remainder of this article is organized as follows. Section "Methodology" describes the procedure applied in this work. Section "Results" presents the obtained results. The section "Discussion" examines the obtained results based on a main research question and two sub-research questions. Finally, the section "Conclusions" provides conclusions and future work.

METHODOLOGY

A systematic review is a type of review that uses repeatable methods to search, evaluate, and synthesize scientific research on a specific topic or subject in a scientific manner [5]. It involves identifying and selecting scientific articles for the research and considering different ways to reduce bias. This selection ensures that reliable results with quality assessments are obtained through evidence and synthesis of the findings. Based on this, a systematic review was conducted, considering the following stages: review planning, review execution, and review reporting. The methodology used in this work is based on the proposed procedure by [5]. This section provides details on review planning, and the results and discussion section correspond to the review execution. Finally, this paper corresponds to the review reporting.

The review planning aimed to establish a set of elements that would guide and support the process. The most important aspects considered to achieve

this were the review objective, research questions, source search strategy, selection criteria, and the information to be extracted. These elements were considered fundamental to ensure the review had a clear focus and was carried out efficiently. Each of these aspects is detailed below.

Research question

The review's objective is to establish an empirical understanding of how learning has been applied to improve decision-making in MAS to solve scheduling problems and thus identify areas of relevant research and future work to be done.

There are four models in machine learning: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning (RL). Supervised learning involves training algorithms using labeled data to predict output labels for new instances. Unsupervised learning focuses on discovering hidden patterns or clusters in unlabeled data. Semi-supervised learning combines labeled and unlabeled data for training, which is useful when obtaining labeled data is challenging. Reinforcement learning employs an agent interacting with an environment, learning optimal actions through rewards and penalties. These types –supervised, unsupervised, semi-supervised, and reinforcement learning– offer distinct approaches to solving machine learning problems [6], [7]. In this context, the following research question was formulated:

RQ: How do agents in a Multi-agent System learn to improve their decisions in scheduling problems?

Two specific research sub-questions (SRQs) were added for a more specific study, which can be answered using the items collected from the main research question. In this way, the analysis of the selected articles generated concrete results that contributed to understanding the existing knowledge and maintained a consistent focus throughout the research process. The first sub-research question is related to the specific learning method, and the second concerns the scheduling problem.

There are different techniques and algorithms for each machine learning type. Decision trees, support vector machines (SVM), and artificial neural networks (ANNs) [7] are common algorithms used in supervised learning [7]. Popular unsupervised

learning techniques include clustering algorithms like k-means clustering and hierarchical clustering, and dimensionality reduction methods such as principal component analysis (PCA) [6]. Several popular algorithms are utilized in semi-supervised learning, including Self-Training, Co-Training, Multi-View Learning, Generative Models, and Manifold Regularization. In case of reinforcement learning some popular algorithms are Q-Learning and Deep Q-Networks. In this context, the first sub-research question is formulated:

SRQ1: What are multi-agent systems' most used learning techniques to solve scheduling problems?

Scheduling problems can be classified into various categories based on different characteristics and constraints. According to [1], one common classification scheme distinguishes scheduling problems based on the type of resources involved, such as single-machine, parallel-machine, or multi-machine problems. Another classification criterion considers the objective function, which can be related to minimizing the makespan, total completion time, or lateness. Furthermore, scheduling problems can be categorized based on the complexity level, including deterministic, stochastic, or online problems. Other classifications consider additional factors such as precedence constraints, setup times, and release dates. Each category presents unique challenges and requires specific algorithms and techniques for efficient scheduling solutions. In this context, we formulate the second sub-research question:

SRQ2: Which scheduling problems are most addressed with learning by multi-agent systems?

Recent and relevant studies were examined to answer these questions and provide an overview of trends and advances in using multi-agent systems to solve scheduling problems. This review will provide a better understanding of how agents learn, the algorithms used, and the most commonly addressed scheduling problems.

Data sources

The search strategy focused on identifying the appropriate sources and critical terms to find relevant information on the topic of multi-agent systems in scheduling problems. Therefore, two electronic databases, Scopus and Web of Science, were mainly

used, due to the large number of scientific journals indexed in both platforms. These databases are widely used in scientific research and offer many articles and studies related to the topic.

A specific search string for each database was used to obtain scientific articles related to the context of multi-agent systems in programming problems. These strings differ due to the differences in the structure of both search engines, but they still maintain the semantics of what is to be searched. In the Scopus website was used the string “Title: ((multiagent) OR (multi agent) OR (multi-agent)) AND scheduling AND learning.” On the Web of Science website was used the string “TI = (((multiagent) OR (multi agent) OR (multi-agent)) AND scheduling AND learning).” With these search strings, it was possible to find articles and studies on using multi-agent systems in scheduling problems and their learning.

Inclusion and exclusion criteria

When selecting studies for review, it is important to consider inclusion and exclusion criteria to ensure that only relevant and current information is considered to answer the research questions posed [5]. Inclusion and exclusion criteria are rules used to determine whether or not a study meets the necessary requirements to be included in the research. For this reason, the following criteria were considered in the selection of studies:

- Articles published in English and Spanish language.
- Articles from journals and conferences.

Data retrieval

The following information was extracted from the articles selected for the in-depth analysis.

- Year of publication.
- Learning model.
- Algorithm applied.
- Form of interaction between agents to learn (independent or cooperative or centralized).
- Types of scheduling problems addressed.
- Area of application.

The next stage in the systematic literature review is executing the plan mentioned above, the results of which are shown in the following section.

RESULTS

Through the execution of the described plan, a total of 182 articles were obtained from the SCOPUS and WOS databases. This systematic literature review considered published articles up to July 2023. After eliminating duplicate articles, 121 articles were obtained. Figure 1 shows the number of articles published per year. From 2020, an upward trend in the number of articles published can be seen.

Another important piece of information is to know which conference or journal has the highest number of articles. Table 1 shows the number of published papers at each journal or conference.

From 2022 to 2023, the most significant publications (37 and 22 articles, respectively) were concentrated

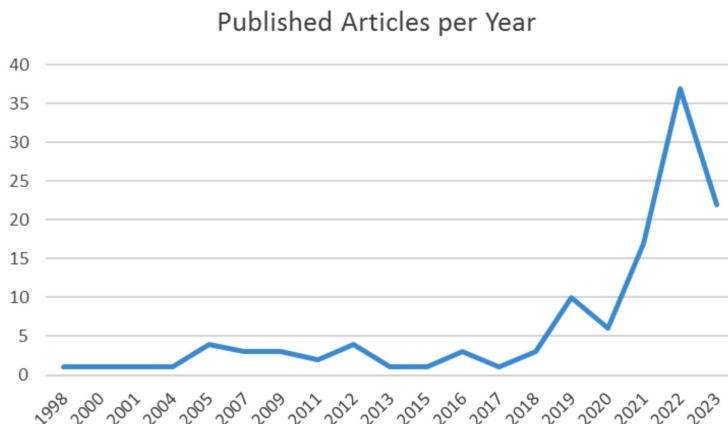


Figure 1. Published articles per year.

Table 1. Number of published articles per journal or conference.

Journal/Conference	Number of Published Articles
Journal of Manufacturing Systems.	5
Robotics and Computer-Integrated Manufacturing.	4
Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).	4
Procedia CIRP	4
IEEE Internet of Things Journal.	3
Electronics (Switzerland).	3
Energies.	3
IEEE Access.	3
IEEE Transactions on Vehicular Technology.	3
Applied Energy.	2
Future Generation Computer Systems.	2
Journal of Intelligent Manufacturing.	2
Communications in Computer and Information Science.	2
IEEE Transactions on Network and Service Management.	2
Wireless Personal Communications.	1
Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science).	1
Jisuanji Jicheng Zhizao Xitong/Computer Integrated Manufacturing Systems, CIMS.	1
Computers and Industrial Engineering.	1
Proceedings - 26th Euromicro International Conference on Parallel, Distributed, and Network-Based Processing, PDP 2018.	1
Computers and Operations Research.	1
International Journal of Hydrogen Energy.	1
Concurrency and Computation: Practice and Experience.	1
5th International Conference on Natural Computation, ICNC 2009.	1
Conference Record - Asilomar Conference on Signals, Systems and Computers.	1
Computer Communications.	1
Dianli Xitong Zidonghua/Automation of Electric Power Systems.	1
Proceedings of 2021 IEEE 2nd International Conference on Information Technology, Big Data and Artificial Intelligence, ICIBA 2021.	1
Doklady Mathematics.	1
International Journal of Computer Networks and Communications.	1
2016 6th International Conference on Computer and Knowledge Engineering, ICCKE 2016.	1
International Journal of Operations and Quantitative Management.	1
2017 IEEE 8th International Conference on Intelligent Computing and Information Systems, ICICIS 2017.	1
Journal of Computer Science & Technology.	1
Energy Conversion and Management.	1
Journal of the Operational Research Society.	1
Entropy.	1
Lecture Notes in Electrical Engineering.	1
Expert Systems with Applications.	1
Proceedings - 2022 Chinese Automation Congress, CAC 2022.	1
CAAI Transactions on Intelligence Technology.	1
Proceedings - 4th International Conference on MultiAgent Systems, ICMAS 2000.	1
2020 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids, SmartGridComm 2020.	1
Chinese Control Conference, CCC.	1
IEEE Communications Letters.	1
19th IEEE International Symposium on Parallel and Distributed Processing with Applications, 11th IEEE International Conference on Big Data and Cloud Computing, 14th IEEE International Conference on Social Computing and Networking and 11th IEEE International Conference on Sustainable Computing and Communications, ISPA/BDCLOUD/SocialCom/SustainCom 2021.	1

Journal/Conference	Number of Published Articles
IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC.	1
International Journal of Electrical Power and Energy Systems.	1
IEEE Geoscience and Remote Sensing Letters.	1
International Journal of Industrial Engineering Computations.	1
2021 IEEE Global Communications Conference, GLOBECOM 2021 - Proceedings.	1
IoTDI 2019 - Proceedings of the 2019 Internet of Things Design and Implementation.	1
IEEE Robotics and Automation Letters.	1
Journal of Computational Information Systems.	1
IEEE Sensors Journal.	1
APSIPA Transactions on Signal and Information Processing.	1
IEEE Transactions on Broadcasting.	1
Journal of Physics: Conference Series.	1
IEEE Transactions on Cybernetics.	1
Knowledge-Based Systems.	1
Mathematical Problems in Engineering.	1
AAAI Fall Symposium - Technical Report.	1
Midwest Symposium on Circuits and Systems.	1
Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST.	1
AAAI Workshop - Technical Report.	1
Proceedings - 2005 IEEE/WIC/ACM International Conference on Intelligent Agent Technology, IAT'05.	1
MobiWac'07 - Proceedings of the 5th ACM International Workshop on Mobility Management and Wireless Access.	1
IEEE Transactions on Neural Networks and Learning Systems.	1
Proceedings - 17th IEEE International Conference on Machine Learning and Applications, ICMLA 2018.	1
IEEE Transactions on Systems, Man, and Cybernetics: Systems.	1
Proceedings - 2019 2nd International Conference on Artificial Intelligence for Industries, AI4I 2019.	1
Proceedings of the 2012 4th World Congress on Nature and Biologically Inspired Computing, NaBIC 2012.	1
Proceedings - 2022 IEEE 19th International Conference on Mobile Ad Hoc and Smart Systems, MASS 2022.	1
Proceedings of the 34th Chinese Control and Decision Conference, CCDC 2022.	1
Proceedings - 30th European Conference on Modelling and Simulation, ECMS 2016.	1
Processes.	1
Proceedings - International Conference on Multi Agent Systems, ICMA 1998.	1
2023 IEEE 2nd International Conference on Electrical Engineering, Big Data and Algorithms, EEBDA 2023.	1
Proceedings of 2023 IEEE 12th Data Driven Control and Learning Systems Conference, DDCLS 2023.	1
Simulation Modelling Practice and Theory.	1
Sustainability (Switzerland).	1
Proceedings of the IEEE Sensor Array and Multichannel Signal Processing Workshop.	1
IET Smart Grid.	1
Reliability Engineering and System Safety.	1
IFAC Proceedings Volumes (IFAC-PapersOnline).	1
Sensors.	1
International Arab Journal of Information Technology.	1
Sustainability.	1
International Conference on Electrical, Computer, Communications and Mechatronics Engineering, ICECCME 2021.	1
Water Supply.	1
International Conference on Human System Interaction, HSI.	1
Zhongguo Dianji Gongcheng Xuebao/Proceedings of the Chinese Society of Electrical Engineering.	1
International Journal of Advanced Manufacturing Technology.	1
International Journal of Aerospace Engineering.	1

in those years, and the articles included in those years were selected for further review. Thus, 59 articles were selected. Finally, 55 articles were available, which were read and analyzed in depth. The list of the 55 articles and the data obtained are shown in Table 2.

DISCUSSION

This section shows a discussion of the 55 selected articles from the perspective of the main research question and the two sub-research questions.

RQ1: How do agents in a multi-agent system learn to improve their decisions in scheduling problems?

Reinforcement Learning (RL) is the learning model used in multi-agent systems to solve scheduling problems. All reviewed articles used this learning model. When RL is used in a multi-agent system, it is called Multi-agent Reinforcement Learning (MRL). RL is preferred over other learning models for application in multi-agent systems due to its ability to handle the complexity and uncertainty inherent in such environments [63]. In multi-agent

Table 2. Selected articles and data retrieval.

Paper	Publication Year	Type of ML	Algorithm	Independent/ Cooperative/ Centralized	Scheduling Classification	Application
[8]	2023	Reinforcement Learning	Deep Q-Network	Cooperative	Job Shop Scheduling	Manufacturing
[9]	2022	Reinforcement Learning	Policy Gradient Method	Centralized	Dynamic Scheduling	Electric Vehicle
[10]	2022	Reinforcement Learning	Bayesian Reinforcement Learning	Independent	N/A	Wireless Communications and Networks
[11]	2022	Reinforcement Learning	Deep Q-Network	Centralized	Dynamic Scheduling	Manufacturing
[12]	2022	Reinforcement Learning	Proximal Policy Optimization	Centralized	Flow Shop Scheduling	Manufacturing
[13]	2022	Reinforcement Learning	Deep Q-Network	Centralized	Dynamic Scheduling	Manufacturing
[14]	2022	Reinforcement Learning	Deep Q-Network	Independent	Flow Shop Scheduling	Manufacturing
[15]	2022	Reinforcement Learning	Proximal Policy Optimization	Cooperative	Job Shop Scheduling	Manufacturing
[16]	2022	Reinforcement Learning	Q-Learning	Cooperative	Workflow Scheduling	Cloud/Edge Computing
[17]	2022	Reinforcement Learning	Q-Learning	Cooperative	Distributed Scheduling	Wireless Communications and Networks
[18]	2022	Reinforcement Learning	Deep Q-Network	Independent	Workflow Scheduling	Energy
[19]	2022	Reinforcement Learning	Actor-Critic	Independent	Dynamic Scheduling	Wireless Communications and Networks
[20]	2022	Reinforcement Learning	Actor-Critic	Centralized	Dynamic Scheduling	Electric Vehicle
[21]	2022	Reinforcement Learning	Deep Q-Network	Centralized	Dynamic Scheduling	Energy
[22]	2022	Reinforcement Learning	Actor-Critic	Cooperative	Job Shop Scheduling	Manufacturing
[23]	2022	Reinforcement Learning	Q-Learning	Cooperative	Dynamic Scheduling	Wireless Communications and Networks
[24]	2022	Reinforcement Learning	Q-Learning	Cooperative	Schedule-based	Wireless Communications and Networks
[25]	2022	Reinforcement Learning	Actor-Critic	Independent	Job Shop Scheduling	Manufacturing
[26]	2022	Reinforcement Learning	Deep Q-Network	Centralized	Job Shop Scheduling	Manufacturing
[27]	2022	Reinforcement Learning	Hybrid	Cooperative	Dynamic Scheduling	Energy
[28]	2022	Reinforcement Learning	Q-Learning	Cooperative	N/A	Cloud/Edge Computing
[29]	2022	Reinforcement Learning	Hybrid	Cooperative	Dynamic Scheduling	Energy
[30]	2022	Reinforcement Learning	Deep Q-Network	Centralized	Dynamic Scheduling	Cloud/Edge Computing
[31]	2022	Reinforcement Learning	Deep Neural Network	Cooperative	N/A	Wireless Communications and Networks
[32]	2022	Reinforcement Learning	Hybrid	Independent	Dynamic Scheduling	Transversal
[33]	2022	Reinforcement Learning	Hybrid	Cooperative	Dynamic Scheduling	Cloud / Edge Computing

Paper	Publication Year	Type of ML	Algorithm	Independent/ Cooperative/ Centralized	Scheduling Classification	Application
[34]	2022	Reinforcement Learning	Hybrid	Cooperative	N/A	Wireless Communications and Networks
[35]	2022	Reinforcement Learning	Proximal Policy Optimization	Centralized	Dynamic Scheduling	Transport
[36]	2022	Reinforcement Learning	Hybrid	Cooperative	N/A	Energy
[37]	2022	Reinforcement Learning	Q-Learning	Centralized	Dynamic Scheduling	Transport
[38]	2022	Reinforcement Learning	Hybrid	Centralized	Dynamic Scheduling	Electric Vehicle
[39]	2022	Reinforcement Learning	Hybrid	Cooperative	N/A	Wireless Communications and Networks
[40]	2022	Reinforcement Learning	Actor-Critic	Centralized	N/A	Wireless Communications and Networks
[41]	2022	Reinforcement Learning	Actor-Critic	Cooperative	N/A	Electric Vehicle
[42]	2022	Reinforcement Learning	Policy Gradient Method	Cooperative	Dynamic Scheduling	Energy
[43]	2022	Reinforcement Learning	Hybrid	Cooperative	N/A	Energy
[44]	2022	Reinforcement Learning	Q-Learning	Independent	N/A	Satellite Network
[45]	2023	Reinforcement Learning	Deep Q-Network	Centralized	Distributed Scheduling	Cloud / Edge Computing
[46]	2023	Reinforcement Learning	Hybrid	Cooperative	Dynamic Scheduling	Manufacturing
[47]	2023	Reinforcement Learning	Deep Deterministic Policy Gradient	Cooperative	Dynamic scheduling	Electric Vehicle
[48]	2023	Reinforcement Learning	Deep Q-Network	Centralized	Job Shop Scheduling	Manufacturing
[49]	2023	Reinforcement Learning	N/A	Decentralized	Dynamic Scheduling	Manufacturing
[50]	2023	Reinforcement Learning	N/A	N/A	N/A	Wireless Communications and Networks
[51]	2023	Reinforcement Learning	Deep Deterministic Policy Gradient	Independent	Dynamic Scheduling	Energy
[52]	2023	Reinforcement Learning	Neural Network	Cooperative	Dynamic Scheduling	Wireless Communications and Networks
[53]	2023	Reinforcement Learning	Actor-Critic	Centralized	Distributed Scheduling	Mining
[54]	2023	Reinforcement Learning	Deep Q-Network	Centralized	Dynamic Job shop scheduling	Manufacturing
[55]	2023	Reinforcement Learning	Proximal Policy Optimization	Centralized	N/A	Electrical Engineering
[56]	2023	Reinforcement Learning	Deep Q-Network	Cooperative	Job Shop Scheduling	Manufacturing
[57]	2023	Reinforcement Learning	Actor-Critic	Independent	Dynamic Scheduling	Electric Vehicle
[58]	2023	Reinforcement Learning	Hybrid	Cooperative	Dynamic Scheduling	Manufacturing
[59]	2023	Reinforcement Learning	Deep Q-Network	Centralized	Dynamic Scheduling	Cloud / Edge Computing
[60]	2023	Reinforcement Learning	Proximal Policy Optimization	Independent	Distributed Scheduling	Wireless Communications and Networks
[61]	2023	Reinforcement Learning	Deep Deterministic Policy Gradient	Cooperative	Dynamic Scheduling	Water Distribution Networks
[62]	2023	Reinforcement Learning	Q-learning	Cooperative	Distributed Scheduling	Hydrogen Refueling Stations

systems, where agents interact and influence each other's outcomes, reinforcement learning provides a natural framework for decision-making under partial observations and dynamic settings. It allows agents to learn optimal strategies through interaction and adapt their policies to changing conditions. Other learning models, like supervised learning, require labeled data, which can be challenging to obtain in dynamic multi-agent scenarios. Additionally,

reinforcement learning's capacity for online learning and generalization to new situations makes it suitable for capturing the intricate dynamics and interactions between agents in multi-agent systems.

During a learning process in a multi-agent system, agents can work cooperatively, independently, or centralized. In a cooperative approach, agents work together, share information, and collaborate

to learn. In an independent approach, each agent learns individually without direct interactions with other agents, which can be useful when goals are divergent or communication is limited. In the centralized approach, one agent performs the learning process and transfers what is learned to the other agents. Forty-five percent of the reviewed articles use a cooperative approach to learning, while 33% use a centralized approach and 19% use an independent approach. In practice, the choice between a cooperative, independent, or centralized approach will depend on the application domain, available resources, and the specific objectives of the multi-agent system. In this context, we believe that future work could include determining a set of criteria to guide the choice between cooperative, independent, or centralized learning.

SRQ1: What are the most used learning techniques in multi-agent systems to solve scheduling problems?

Reinforcement learning (RL) employs several algorithms for autonomous agents to learn optimal decision-making strategies through trial and error. The most well-known RL algorithms are Q-learning, Deep Q-Networks (DQN), Proximal Policy Optimization (PPO), and Actor-Critic algorithms. These algorithms have been widely applied in different domains, such as robotics, gaming, and autonomous systems, to address complex decision-making problems.

The reviewed articles show that Deep Q-Network (DQN), Q-learning, and Actor-Critic are the most used learning techniques in multi-agent systems for solving scheduling problems. Deep Q-Network was used in 23.64% of the articles. Most of these articles show a centralized use of the technique, i.e., one agent learns and transfers what it learns to the other agents. For example, in [11], a MAS is proposed in which each task is considered an agent, and the MAS is trained under a CTDE (Centralized Training Distributed Execution) architecture. A double DQN-based algorithm was applied in [13] to solve a scheduling problem in a manufacturing company. After centralized learning, the agents could make decisions based on local observations.

Regarding Q-learning, 14.55% of reviewed articles used this technique. Most of these articles show a cooperative use of this technique. For example, in [28], a multi-agent collaborative deep reinforcement learning-based distributed scheduling algorithm is

proposed. This algorithm uses graph attention neural networks to solve task-scheduling problems in a Multi-access Edge Computing scenario. In [24], a distributed cooperative multi-agent reinforcement learning approach based on the self-schedule scheme for channel assignment in schedule-based wireless sensor networks is proposed.

Regarding the Actor-Critic technique, 14.55% of reviewed articles also used this technique. Here, most of the articles used both centralized and cooperative learning techniques. For example, [40] proposed a multi-agent deep reinforcement learning algorithm to help vehicles select appropriate radio resources and reduce packet collisions. In [57], a multi-agent Actor-Critic algorithm based on a non-cooperative game is proposed to achieve adequate control of electric vehicle charging stations in large-scale road networks.

It is also observed that several articles (20%) use a combination of learning techniques. Here, most of the articles use the technique for cooperative learning. For instance, in [38], a multi-agent reinforcement learning model with a deep Q-network (DQN) and a deep deterministic policy gradient is proposed to jointly optimize the electric vehicles routing and scheduling decisions in the transportation and power systems. In [43], a data-driven-based multi-agent proximal policy optimization algorithm is proposed for the optimal scheduling of multi-park Integrated Energy Systems. The proposed solution also includes the application of the Actor-Critic technique.

A common problem several authors of the reviewed articles pointed out is scalability. As the number of agents increases, the complexity of reinforcement learning methods and computational demands also grow significantly. Therefore, more research is needed to improve the execution times of learning algorithms in real-world situations.

SRQ2: Which scheduling problems are most addressed with learning by Multi-agent Systems?

The scheduling problems most commonly addressed by multi-agent systems with learning capabilities are Dynamic Scheduling and Job Shop Scheduling. 45.45% of the reviewed articles addressed dynamic scheduling scenarios, while 12.73% addressed job-shop scheduling scenarios. 21.82% of the articles

did not explicitly state what type of scheduling problem they addressed, as they focused more on proposed solution's details than the scheduling problem. The remaining 20% of the articles are distributed in other types of scheduling problems.

A large number of articles use agents as digital twins, i.e., agents represent real-world elements such as machines (robots, vehicles), tasks (travel, operations), and resources (fuel, materials). However, a few articles use agents more functionally, e.g., as an algorithm that searches for a solution in a part of the solution space. For example, in [49], some agents represent real-world elements and other algorithms to which the solution space is distributed, thus allowing faster solution finding. In [58], an agent creates the schedule for all the workers in a company.

Regarding the application area, 27% of the articles were applied in the manufacturing industry. Wireless communication accounted for 22%, followed by 15% of articles related to the energy industry, 11% to electric vehicle charging, and 11% to cloud computing. The remaining articles are distributed in different areas, such as transportation, mining, and satellite networks.

CONCLUSIONS

This paper presents the results of a systematic literature review on how agents in a multi-agent system learn to solve scheduling problems. In all, 55 articles were analyzed in depth. This methodology empirically answered the main research question and the two secondary questions.

The main findings point out that Reinforcement Learning (RL) is the learning model used in the reviewed articles. There is a tendency to combine two or more RL algorithms to apply them in the proposed solutions; most of the papers use these proposed solutions to solve dynamic scheduling issues in the manufacturing and wireless communication network industries.

Research in multi-agent reinforcement learning (MARL) has made significant progress in recent years; however, key research gaps remain to be addressed. Firstly, as the number of agents increases, the complexity and computational demands of

MARL methods also grow significantly. Developing scalable MARL algorithms that can effectively handle large numbers of agents is essential for real-world applications with numerous interacting entities. Secondly, there is a need to explore MARL in more complex and realistic environments, such as real-world simulations or physical systems. Addressing these research gaps will contribute to advancing MARL and its applicability in various domains.

REFERENCES

- [1] P. Brucker, *Scheduling algorithms*, 7th ed. New York, USA: Springer Berlin Heidelberg, 2007.
- [2] M. Wooldridge, "Agent-based software engineering," *IEE Proceedings-Software*, vol. 144, no. 1, pp. 26-37, 1997, doi: 10.1049/ip-sen:19971026
- [3] M.J. Wooldridge, *An introduction to multiagent systems*, 2nd ed. West Sussex, England: John Wiley & Sons, 2009.
- [4] Y. Shoham, and K. Leyton-Brown, *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*, 1st ed. New York, USA: Cambridge University Press, 2009.
- [5] B. Kitchenham, "Procedures for performing systematic reviews," Keele, UK: Keele University, 2004.
- [6] C. Bishop, *Pattern recognition and machine learning*, 1st ed. New York, USA: Springer, 2006.
- [7] S. Russell and P. Norvig, *Artificial intelligence: A modern approach*, 3rd ed. New Jersey, USA: Pearson Upper Saddle River, 2016.
- [8] J. Zhang, Z. He, W. Chan, and C. Chow, "DeepMAG: Deep reinforcement learning with multi-agent graphs for flexible job shop scheduling," *Knowledge-Based Systems*, vol. 259, 2023, doi: 10.1016/j.knosys.2022.110083.
- [9] K. Park and I. Moon, "Multi-agent deep reinforcement learning approach for EV charging scheduling in a smart grid," *Applied Energy*, vol. 328, 2022, doi: 10.1016/j.apenergy.2022.120111.
- [10] M. Naeem, A. Coronato, Z. Ullah, S. Bashir, and G. Paragliola, "Optimal user scheduling in multi antenna system using multi agent reinforcement learning," *Sensors*, vol. 22, no. 21, Nov. 2022, doi: 10.3390/s22218278.
- [11] X. Wang, L. Zhang, Y. Liu, F. Li, Z. Chen, C. Zhao, and T. Bai, "Dynamic scheduling

- of tasks in cloud manufacturing with multi-agent reinforcement learning,” *Journal of Manufacturing Systems*, vol. 65, pp. 130-145, 2022, doi: 10.1016/j.jmsy.2022.08.004.
- [12] X. Wang, L. Zhang, T. Lin, C. Zhao, K. Wang, and Z. Chen, “Solving job scheduling problems in a resource preemption environment with multi-agent reinforcement learning,” *Robotics and Computer-Integrated Manufacturing*, vol. 77, 2022, doi: 10.1016/j.rcim.2022.102324.
- [13] D. Johnson, G. Chen, and Y. Lu, “Multi-agent reinforcement learning for real-time dynamic production scheduling in a robot assembly cell,” *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 7684-7691, 2022, doi: 10.1109/LRA.2022.3184795.
- [14] M. Wang, J. Zhang, P. Zhang, L. Cui, and G. Zhang, “Independent double DQN-based multi-agent reinforcement learning approach for online two-stage hybrid flow shop scheduling with batch machines,” *Journal of Manufacturing Systems*, vol. 65, pp. 694-708, 2022, doi: 10.1016/j.jmsy.2022.11.001.
- [15] Y. Zhang, H. Zhu, D. Tang, T. Zhou, and Y. Gui, “Dynamic job shop scheduling based on deep reinforcement learning for multi-agent manufacturing systems,” *Robotics and Computer-Integrated Manufacturing*, vol. 78, Dec. 2022, doi: 10.1016/j.rcim.2022.102412.
- [16] H. Gholami and M. T. Rezvan, “A cooperative multi-agent offline learning algorithm to scheduling IoT workflows in the cloud computing environment,” *Concurrency and Computation*, vol. 34, no. 22, 2022, doi: 10.1002/cpe.7148.
- [17] Y. Lu, K. Wang, and E. He, “Many-to-Many Data Aggregation Scheduling Based on Multi-Agent Learning for Multi-Channel WSN,” *Electronics*, vol. 11, no. 20, 2022, doi: 10.3390/electronics11203356.
- [18] J. Lu, P. Mannion, and K. Mason, “A multi-objective multi-agent deep reinforcement learning approach to residential appliance scheduling,” *IET Smart Grid*, vol. 5, no. 4, pp. 260-280, 2022, doi: 10.1049/stg2.12068.
- [19] C. Jiang *et al.*, “Attention-shared multi-agent actor-critic-based deep reinforcement learning approach for mobile charging dynamic scheduling in wireless rechargeable sensor networks,” *Entropy*, vol. 24, no. 7, 2022, doi: 10.3390/e24070965.
- [20] M. Alqahtani, M. J. Scott, and M. Hu, “Dynamic energy scheduling and routing of a large fleet of electric vehicles using multi-agent reinforcement learning,” *Computers & Industrial Engineering*, vol. 169, 2022, doi: 10.1016/j.cie.2022.108180.
- [21] Y. Wang, D. Qiu, and G. Strbac, “Multi-agent deep reinforcement learning for resilience-driven routing and scheduling of mobile energy storage systems,” *Applied Energy*, vol. 310, 2022, doi: 10.1016/j.apenergy.2022.118575.
- [22] X. Jing, X. Yao, M. Liu, and J. Zhou, “Multi-agent reinforcement learning based on graph convolutional network for flexible job shop scheduling,” *Journal of Intelligent Manufacturing*, vol. 35, pp. 75-93, 2022, doi: 10.1007/s10845-022-02037-5.
- [23] N.N. Sirhan and M. Martinez-Ramon, “Cognitive radio resource scheduling using multi-agent q-learning for LTE,” *International Journal of Computer Networks and Communications*, vol. 14, no. 2, pp. 77-95, 2022, doi: 10.5121/ijcnc.2022.14205.
- [24] M. Sahraoui, A. Bilami, and A. Taleb-Ahmed, “Schedule-based cooperative multi-agent reinforcement learning for multi-channel communication in wireless sensor networks,” *Wireless Personal Communications*, vol. 122, pp. 3445-3465, 2022, doi: 10.1007/s11277-021-09094-8.
- [25] P. Burggräf, J. Wagner, T. Saßmannshausen, D. Ohrndorf, and K. Subramani, “Multi-agent-based deep reinforcement learning for dynamic flexible job shop scheduling,” *Procedia CIRP*, 2022, vol. 112, pp. 57-62, doi: 10.1016/j.procir.2022.09.024.
- [26] J. Popper, and M. Ruskowski, “Using multi-agent deep reinforcement learning for flexible job shop scheduling problems,” *Procedia CIRP*, vol. 112, pp. 63-67, 2022, doi: 10.1016/j.procir.2022.09.039.
- [27] G. Zhang, W. Hu, D. Cao, Z. Zhang, Q. Huang, Z. Chen, and F. Blaabjerg, “A multi-agent deep reinforcement learning approach enabled distributed energy management schedule for the coordinate control of multi-energy hub with gas, electricity, and freshwater,” *Energy Conversion and Management*, vol. 255, 2022, doi: 10.1016/j.enconman.2022.115340.

- [28] Y. Li, J. Li, and J. Pang, "A graph attention mechanism-based multiagent reinforcement-learning method for task scheduling in edge computing," *Electronics*, vol. 11, no. 9, 2022, doi: 10.3390/electronics11091357.
- [29] G. Gao, Y. Wen, and D. Tao, "Distributed energy trading and scheduling among microgrids via multiagent reinforcement learning," *IEEE Transactions in Neural Networks and Learning Systems*, vol. 34, no. 12, pp. 10638-10652, 2022, doi: 10.1109/TNNLS.2022.3170070.
- [30] F. Jiang, L. Dong, K. Wang, K. Yang, and C. Pan, "Distributed resource scheduling for large-scale MEC systems: A multiagent ensemble deep reinforcement learning with imitation acceleration," *IEEE Internet of Things Journal*, vol. 9, no. 9, pp. 6597-6610, 2022, doi: 10.1109/JIOT.2021.3113872.
- [31] M. Yuan, Q. Cao, M. Pun, and Y. Chen, "Fairness-oriented user scheduling for bursty downlink transmission using multi-agent reinforcement learning," *APSIPA Transactions on Signal and Information Processing*, vol. 11, no. 1, 2022, doi: 10.1561/116.00000028.
- [32] H. Cheng, R. Song, L. Xu, D. Zhang, and S. Xu, " H^∞ consensus design and online scheduling for multiagent systems with switching topologies via deep reinforcement learning," *International Journal of Aerospace Engineering*, vol. 2022, 2022, Art. no. 2650632, doi: 10.1155/2022/2650632.
- [33] X. Zhao and C. Wu, "Large-scale machine learning cluster scheduling via multi-agent graph reinforcement learning," *IEEE Transactions on Network and Service Management*, vol. 19, no. 4, pp. 4962-4974, 2022, doi: 10.1109/TNSM.2021.3139607.
- [34] J. Lee, D. Niyato, Y.L. Guan, and D.I. Kim, "Learning to schedule joint radar-communication with deep multi-agent reinforcement learning," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 1, pp. 406-422, 2022, doi: 10.1109/TVT.2021.3124810.
- [35] I.K. Minashina, R.A. Gorbachev, and E. M. Zakharova, "Scheduling in multiagent systems using reinforcement learning," *Doklady Mathematics*, vol. 106, no. Suppl. 1, pp. S70-S78, 2023, doi: 10.1134/S1064562422060175.
- [36] Z. Zuo, Z. Li, and Y. Wang, "Multi-agent deep reinforcement learning for microgrid energy scheduling," in *41st Chinese Control Conference (CCC)*, Hefei, China, July 25-27, 2022, pp. 6184-6189, doi: 10.23919/CCC55666.2022.9901844.
- [37] Y. Shen, L. Duan, Q. Zhu, Z. Su, and G. Zhang, "Multiagent Q-learning for multicrew dynamic scheduling and routing in road network restoration," in *34th Chinese Control and Decision Conference (CCDC)*, Hefei, China, August 15-17, 2022, pp. 1217-1222, doi: 10.1109/CCDC55256.2022.10033659.
- [38] Y. Wang, D. Qiu, and G. Strbac, "Multi-agent reinforcement learning for electric vehicles joint routing and scheduling strategies," in *2022 IEEE 25th International Conference on Intelligent Transportation Systems*, Macau, China, October, 2022, pp. 3044-3049, doi: 10.1109/ITSC55140.2022.9921744.
- [39] K. Yang, D. Li, C. Shen, J. Yang, S. Yeh, and J. Sydir, "Multi-agent reinforcement learning for wireless user scheduling: performance, scalability, and generalization," in *2022 56th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, USA, October 31- November 02, 2022, pp. 1169-1174, doi: 10.1109/ieeconf56349.2022.10051992.
- [40] B. Gu, W. Chen, M. Alazab, X. Tan, and M. Guizani, "Multiagent reinforcement learning-based semi-persistent scheduling scheme in C-V2X mode 4," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 11, pp. 12044-12056, 2022, doi: 10.1109/TVT.2022.3189019.
- [41] J. Zhang *et al.*, "Multi-AGV scheduling based on hierarchical intrinsically rewarded multi-agent reinforcement learning," in *2022 IEEE 19th International Conference on Mobile Ad Hoc and Smart Systems (MASS)*, Denver, CO, USA, October 19-23, 2022, pp. 155-161, doi: 10.1109/MASS56207.2022.00028.
- [42] Y. Zhang, Q. Yang, D. An, D. Li, and Z. Wu, "Multistep multiagent reinforcement learning for optimal energy schedule strategy of charging stations in smart grid," *IEEE Transactions on Cybernetics*, vol. 53, no. 7, pp. 4292-4305, 2022, doi: 10.1109/TCYB.2022.3165074.
- [43] F. Meng, H. Wang, L. Xu, C. Yuan, H. Wang, and Y. Niu, "Optimal scheduling of integrated energy system based on multi-agent reinforcement learning," in *2022 China*

- Automation Congress (CAC), Xiamen, China, November 25-27, 2022, pp. 1179-1184, doi: 10.1109/cac57257.2022.10055447.
- [44] H. Zhu, Z. Wang, D. Li, and Q. Guo, "Satellite staring beam scheduling strategy based on multi-agent reinforcement learning," in *Wireless and Satellite Systems. WiSATS 2021. Virtual Event. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, China, Q. Guo, W. Meng, M. Jia, and Z. Wang, Eds. August 2, 2022, pp. 23-34, doi: 10.1007/978-3-030-93398-2_3.
- [45] X. Nie, Y. Yan, T. Zhou, X. Chen, and D. Zhang, "A delay-optimal task scheduling strategy for vehicle edge computing based on the multi-agent deep reinforcement learning approach," *Electronics*, vol. 12, no. 7, 2023, doi: 10.3390/electronics12071655.
- [46] A. F. İnal, Ç. Sel, A. Aktepe, A. K. Türker, and S. Ersöz, "A multi-agent reinforcement learning approach to the dynamic job shop scheduling problem," *Sustainability*, vol. 15, no. 10, 2023, doi: 10.3390/su15108262.
- [47] N. Kaewdornhan, C. Srithapon, R. Liemthong, and R. Chatthaworn, "Real-time multi-home energy management with EV charging scheduling using multi-agent deep reinforcement learning optimization," *Energies*, vol. 16, no. 5, 2023, doi: 10.3390/en16052357.
- [48] X. Zhu, J. Xu, J. Ge, Y. Wang, and Z. Xie, "Multi-task multi-agent reinforcement learning for real-time scheduling of a dual-resource flexible job shop with robots," *Processes*, vol. 11, no. 1, 2023, doi: 10.3390/pr11010267.
- [49] J.B.H. C. Didden, Q.V. Dang, and I.J.B.F. Adan, "Decentralized learning multi-agent system for online machine shop scheduling problem," *Journal of Manufacturing Systems*, vol. 67, pp. 338-360, 2023, doi: 10.1016/j.jmsy.2023.02.004.
- [50] Y. Zou, H. Yin, Y. Zheng, and F. Dressler, "Multi-agent reinforcement learning enabled link scheduling for next generation Internet of Things," *Computer Communications*, vol. 205, pp. 35-44, 2023, doi: 10.1016/j.comcom.2023.04.006.
- [51] P. Rokhforoz, M. Montazeri, and O. Fink, "Safe multi-agent deep reinforcement learning for joint bidding and maintenance scheduling of generation units," *Reliability Engineering and System Safety*, vol. 232, 2023, doi: 10.1016/j.res.2022.109081.
- [52] I.S. Comsa *et al.*, "Improved quality of online education using prioritized multi-agent reinforcement learning for video traffic scheduling," *IEEE Transactions on Broadcasting*, vol. 69, no. 2, pp. 436-454, 2023, doi: 10.1109/TBC.2023.3246815.
- [53] Z. Feng, G. Liu, L. Wang, Q. Gu, and L. Chen, "Research on the multiobjective and efficient ore-blending scheduling of open-pit mines based on multiagent deep reinforcement learning," *Sustainability*, vol. 15, no. 6, 2023, doi: 10.3390/su15065279.
- [54] R. Liu, R. Piplani, and C. Toro, "A deep multi-agent reinforcement learning approach to solve dynamic job shop scheduling problem," *Computers & Operations Research*, vol. 159, 2023, doi: 10.1016/j.cor.2023.106294.
- [55] C. Yang, J. Zhang, F. Lin, L. Wang, W. Jiang, and H. Zhang, "Combining forecasting and multi-agent reinforcement learning techniques on power grid scheduling task," in *2023 IEEE 2nd International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA)*, Changchun, China, February 24-26, 2023, pp. 1576-1580. doi: 10.1109/EEBDA56825.2023.10090669.
- [56] Z. Qin, D. Johnson, and Y. Lu, "Dynamic production scheduling towards self-organizing mass personalization: A multi-agent dueling deep reinforcement learning approach," *Journal of Manufacturing Systems*, vol. 68, pp. 242-257, 2023, doi: 10.1016/j.jmsy.2023.03.003.
- [57] L. Fu, T. Wang, M. Song, Y. Zhou, and S. Gao, "Electric vehicle charging scheduling control strategy for the large-scale scenario with non-cooperative game-based multi-agent reinforcement learning," *International Journal of Electrical Power and Energy Systems*, vol. 153, 2023, doi: 10.1016/j.ijepes.2023.109348.
- [58] Y. Liu, J. Fan, L. Zhao, W. Shen, and C. Zhang, "Integration of deep reinforcement learning and multi-agent system for dynamic scheduling of re-entrant hybrid flow shop considering worker fatigue and skill levels," *Robotics and Computer-Integrated*

- Manufacturing*, vol. 84, Dec. 2023, doi: 10.1016/j.rcim.2023.102605.
- [59] Y. Zhang, R. Li, Y. Zhao, R. Li, Y. Wang, and Z. Zhou, "Multi-agent deep reinforcement learning for online request scheduling in edge cooperation networks," *Future Generation Computer Systems*, vol. 141, pp. 258-268, 2023, doi: 10.1016/j.future.2022.11.017.
- [60] L. Niu *et al.*, "Multiagent meta-reinforcement learning for optimized task scheduling in heterogeneous edge computing systems," *IEEE Internet of Things Journal*, vol. 10, no. 12, pp. 10519-10531, 2023, doi: 10.1109/JIOT.2023.3241222.
- [61] S. Hu, J. Gao, and D. Zhong, "Multi-agent reinforcement learning framework for real-time scheduling of pump and valve in water distribution networks," *Water Supply*, vol. 23, no. 7, pp. 2833-2846, 2023, doi: 10.2166/ws.2023.163.
- [62] Y. Jiang, J. Liu, and H. Zheng, "Optimal scheduling of distributed hydrogen refueling stations for fuel supply and reserve demand service with evolutionary transfer multi-agent reinforcement learning," *International Journal of Hydrogen Energy*, vol. 54, pp. 239-255, 2023, doi: 10.1016/j.ijhydene.2023.04.128.
- [63] M. Tan, "Multi-agent reinforcement learning: Independent vs. Cooperative agents," in *ICML'93: Proceedings of the tenth international conference on machine learning*, Amherst, MA, USA, July 27-29, 1993, pp. 330-337, doi : 10.5555/3091529.