# Multi-Agent Reinforcement Learning and its practical applications in games: a review of the literature

Sean Marco Bonifacio  
University of Northampton

## Abstract

Multi-agent reinforcement learning (MARL) has achieved tremendous success in various artificial intelligence domains such as the gaming industry. This literature review explores the theoretical foundations, principles, algorithms, applications, challenges, and recent advancements in MARL.

Algorithms such as Q-Learning, Policy-based algorithms, SARSA, and more algorithms have laid the groundwork for MARL by providing approaches to agent decision-making and learning. Q-decomposition methods have also further enhanced MARL by decomposing the global value function into local components, allowing for efficient coordination, and learning among agents (Russel & Zimdars, 2003).

The architectures employed in MARL range from centralised to decentralised, with each influencing the coordination and decision-making processes of agents in the environment. Centralised architectures, in which all agents share a single state representation and make collaborative decisions, provide benefits in terms of global coordination and information exchange. Decentralised architectures, on the other hand, encourage agent autonomy by allowing each agent to have its own state representations and decision-making procedures. These designs, when combined with communication and coordination mechanisms, enables MARL systems to model complex interactions and achieve emergent behaviours.

Beyond the gaming industry, MARL is extensively used in a variety of domains, including multi-robot systems, autonomous vehicles, and social dilemmas. Obstacles such as credit assignment, scalability, and ethical considerations persist, emphasising the importance of continued research and innovation in the field.

It is hoped that this paper will provide valuable insights for game developers and researchers regarding MARL and its challenges and advancements.

## Review of Literature

### Overview of MARL

Multi-agent reinforcement learning (MARL) is a subfield of reinforcement learning that focuses on the behaviour of multiple learning agents interacting in a shared environment. MARL algorithms are designed to address various scenarios, including fully cooperative, fully competitive, and mixed cooperative-competitive environments. These algorithms enable agents to learn optimal strategies through interactions with the environment and other agents.

MARL presents unique challenges and applications across different domains. Challenges include nonstationary, varying learning speeds, scalability issues in deep reinforcement learning, and handling partial observability through centralised learning of decentralized policies. To address these challenges, researchers have developed diverse algorithms and methodologies to enhance the performance of MARL systems (Canese, et al., Multi-Agent Reinforcement Learning: A Review of Challenges and Applications, 2021).

The framework of MARL involves models such as Markov Decision Processes (MDPs), Markov Games, Partially-Observable Markov Decision Processes (POMDPs), and Decentralized Partially-Observable Markov Decision Processes (Dec-POMDPs). Single-agent reinforcement learning algorithms like Q-Learning and REINFORCE serve as foundational components in developing MARL algorithms. These algorithms consist of critics for estimating value functions and actors for updating policy parameters based on feedback from the environment. (Canese, et al., Multi-Agent Reinforcement Learning: A Review of Challenges and Applications., 2021).

MARL has been applied in various domains such as robotics, distributed control, telecommunications, and economics due to its ability to address complex tasks that are challenging to solve with preprogrammed behaviours. The field of MARL continues to evolve with advancements in algorithm design, benchmark environments for multi-agent systems, and practical applications across industries (L. Bus¸oniu & Schutter, 2010).

### Frameworks of MARL expanded

#### Markov Decision Processes (MDPs)

Markov Decision Processes (MDPs) serve as a foundational concept in the realm of Multi-Agent Reinforcement Leaning (MARL), providing a mathematical framework for modelling decision-making processes where outcomes are influenced by both random factors and agent actions. In the context of MARL, MDPs extend to game models, enabling the representation of interactions among multiple agents within a shared environment (Albrecht, Christianos, & Schäfer, 2024).

1. Write about the first provided source (Albrecht, Christianos, & Schafer, 2024). Highlights the extension of the MDP model to incorporate game models where multiple agents interact in shared environments, emphasizing its relevance in MARL.
2. <https://www.researchgate.net/publication/235004620_Reinforcement_Learning_and_Markov_Decision_Processes> - Introduces the intuitions and concepts behind MDPs and their relationship with reinforcement learning algorithms, showcasing their significance in computing optimal behaviours.

# References

Albrecht, S. V., Christianos, F., & Schäfer, L. (2024). *Multi-Agent Reinforcement Learning.* Cambridge, Massachusetts: MIT Press.

Canese, L., Cardarilli, G., Di Nunzio, L., Fazzolari, R., Giardino, D., Re, M., & Spanò, S. (2021). Multi-Agent Reinforcement Learning: A Review of Challenges and Applications. *Applied Sciences* .

Canese, L., Cardarilli, G., Di Nunzio, L., Fazzolari, R., Giardino, D., Re, M., & Spanò, S. (2021). Multi-Agent Reinforcement Learning: A Review of Challenges and Applications. *Applied Sciences*.

L. Bus¸oniu, R. B., & Schutter, B. D. (2010). Multi-agent reinforcement learning: An Overview. In D. Srinivasan, & L. Jain, *Innovations in Multi-Agent Systems and Applications* (pp. 183-221). Berlin, Germany: Springer.

Russel, S., & Zimdars, A. L. (2003). Q-Decomposition for Reinforcement Learning Agents. *Proceedings of the Twentieth Conference on Machine Learning* .