



## Perspective:

# Prospects for multi-agent collaboration and gaming: challenge, technology, and application\*

Yu LIU<sup>†1</sup>, Zhi LI<sup>†‡2</sup>, Zhizhuo JIANG<sup>2</sup>, You HE<sup>1</sup>

<sup>1</sup>Department of Electronic Engineering, Tsinghua University, Beijing 100084, China

<sup>2</sup>Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China

<sup>†</sup>E-mail: liuyu77360132@126.com; zhilizl@sz.tsinghua.edu.cn

Received Feb. 14, 2022; Revision accepted Apr. 23, 2022; Crosschecked May 9, 2022; Published online May 20, 2022

<https://doi.org/10.1631/FITEE.2200055>

Recent years have witnessed significant improvement of multi-agent systems for solving various decision-making problems in complex environments and achievement of similar or even better performance than humans. In this study, we briefly review multi-agent collaboration and gaming technology from three perspectives, i.e., task challenges, technology directions, and application areas. We first highlight the typical research problems and challenges in the recent work on multi-agent systems. Then we discuss some of the promising research directions on multi-agent collaboration and gaming tasks. Finally, we provide some focused prospects on the application areas in this field.

## 1 Introduction

Multi-agent systems, which aim to understand complex environments and adaptively make decisions to compete and coordinate (Vinyals et al., 2019) with humans or other agents, have become increasingly hot topics in both academia and industry. Actually, thanks to tremendous success in

machine learning and reinforcement learning, recent years have witnessed sensational advances in many prominent agent decision-making problems, such as robotic control (Polydoros and Nalpantidis, 2017), Go game playing (Silver et al., 2016, 2017), and autonomous driving (Grigorescu et al., 2020).

For a long time, researchers focused mainly on a single agent to learn environments and take autonomous actions to change the environment (Jennings et al., 1998). These early researchers always treated this problem as a Markov decision process (Puterman, 1994) and applied transition planning methods, such as Monte-Carlo tree search (Coulom, 2007), to optimize the Markov decision process. A principled mathematical framework for adaptive agent learning is reinforcement learning (Arulkumaran et al., 2017). Actually, reinforcement learning has achieved great successes in many tasks (Arulkumaran et al., 2017). However, traditional reinforcement learning methods still have great computational complexity challenges (Arulkumaran et al., 2017). These complex challenges can be overcome by the powerful representation ability of deep neural networks. For example, in the Go game, it is difficult for traditional artificial intelligence (AI) to evaluate all board positions and moves because of the enormous search space (Silver et al., 2016). However, with the representation of deep learning, AI system can achieve superior performance (Silver

<sup>‡</sup> Corresponding author

\* Project supported by the National Key R&D Program of China (No. 2021YFA0715202) and the National Natural Science Foundation of China (Nos. 61790550 and 62022092)

ORCID: Yu LIU, <https://orcid.org/0000-0002-5216-3181>; Zhi LI, <https://orcid.org/0000-0002-8061-7486>

© Zhejiang University Press 2022

et al., 2016, 2017). Consequently, deep reinforcement learning has attracted much attention from many agent application areas, such as business management (Betancourt and Chen, 2021) and industry control (Spielberg et al., 2017).

Recently, with the rapid development of AI technologies, more and more researchers have started to pay much attention to multi-agent systems. Actually, multi-agent systems are of great challenges to involve the participation of more than one single agent (Zhang KQ et al., 2021). To begin with, multi-agent environments are much more sophisticated than single-agent environments, because the actions of other agents may make the environment unstable. This situation always makes it difficult for researchers to model the real-world multi-agent environment and construct a reliable simulation environment. Then, considering multi-agent system characteristics, agent communications and cooperations are two major challenges for multi-agent collaborations. On one hand, communication is a big problem for different agents who share the environment perception and action commands. On the other hand, how to make all the agents collaborate to achieve the final targets is a non-trivial task. Last but not least, multi-agent systems are always applied to competitive environments in many applications. Along this line, the return of agents in competitive environments is usually a zero-sum operation (Zhang KQ et al., 2021). This means that multi-agent systems need to take competitive environments into consideration, which also brings some major challenges.

In this study, we provide a focused look at the multi-agent collaboration and gaming task, from research challenges to technical directions, and to application areas. Although several surveys have reviewed many aspects of multi-agent systems (Busoni et al., 2008; Hernandez-Leal et al., 2017; Oroojlooy and Hajinezhad, 2019; Nguyen et al., 2020; Zhang KQ et al., 2021), recent surveys focus on either a sub-domain (Hernandez-Leal et al., 2017; Oroojlooy and Hajinezhad, 2019) such as cooperative multi-agent systems (Oroojlooy and Hajinezhad, 2019) or specific methods such as deep reinforcement learning (Nguyen et al., 2020). Different from prior work, in this study, we spotlight mainly several new research directions that are comparatively under-explored in existing reviews and hope to suggest some insightful ideas for future multi-agent collabo-

ration and gaming research areas. More specifically, we first introduce some open issues and challenges in this area. Then, we provide a related outlook on technical directions that may bring some insightful thinking to these research challenges. Finally, we bring some prospects of the application areas for multi-agent collaboration and gaming. We hope this paper can provide a quick overview of multi-agent study with a special focus on agent collaboration and gaming.

## 2 Task challenges

The research on multi-agent systems has a long history (Hoen et al., 2005; Busoni et al., 2008), but there are still many open issues. In this section, we provide a selective overview of three major problems in multi-agent collaboration and gaming.

### 2.1 Multi-agent environment construction

Generally, multi-agent environment problems have been ignored by prior reviews. Many multi-agent systems studies rely mainly on totally virtual environments or simulated environments. Previous efforts usually construct virtual environments from games and develop multi-agent evaluation platforms. Arcade learning environment (Atari) (Belle-mare et al., 2013) was first developed to evaluate general agents in Atari 2600 game environments. ViZ-Doom (Kempka et al., 2016) first presents a semi-realistic three-dimensional (3D) evaluation platform in a shooter video game environment. There are also many other multi-agent evaluation platforms, such as MuJoCo (Todorov et al., 2012), Minecraft (Johnson et al., 2016; Tessler et al., 2017), DeepMind Lab (Beattie et al., 2016), OpenAI Gym (Brockman et al., 2016), FAIR TorchCraft (Synnaeve et al., 2016), and Botzone (Zhou et al., 2018). In recent years, researchers have shown special interest in real-time strategy games, such as StarCraft (Shao et al., 2019; Vinyals et al., 2019) and Dota (Berner et al., 2019). Although these platforms have greatly advanced the development of multi-agent technologies and evaluations, there is still a huge gap between these virtual environments and real-world applications.

Consequently, it is an urgent challenge to construct a more open and real environment and then bridge the gap between multi-agent computational

environments and real-world scenes. Actually, both academia and industry have paid much attention to this issue and made several attempts. For example, the RoboCup (<https://www.robocup.org/>) founded in 1996 was a successful attempt at constructing a real collaborative and competitive environment for multi-agent systems. The DJI hosted an annual intercollegiate robot competition, namely RoboMaster (<https://www.robomaster.com/>), and introduced a confrontational 5-on-5 MOBA-style robot combat, which is another great attempt at a real environment for multi-agent systems. These attempts have made important progress on constructing real-world multi-agent collaborative and competitive environments. However, these attempts in real-world multi-agent environments are still limited. On one hand, the existing environments are focused mainly on a specific task. On the other hand, these environments are still in a closed setting with limited time and space. In the real world, tasks for agents are more complex and various. The multi-agent environments are always with open settings, which means that the time and space are without limitations. In other words, it is still an open issue with great challenges to construct a more realistic multi-agent environment.

## 2.2 Multi-agent collaboration

Different from single-agent systems, multi-agent systems naturally consist of multiple agents with various sensors. Therefore, the collaboration of different agents is a major task that cannot be ignored. Indeed, in a complex environment, a single agent could not obtain comprehensive environment information. All the agents need to collaborate with each other to achieve a global perception of the environment and make cooperative decisions. In this subsection, we will discuss two challenges in multi-agent collaboration: multi-agent communication and collaborative perception.

### 2.2.1 Multi-agent communication

Multi-agent communication is a classical and important research topic in multi-agent systems (Georgeff, 1988; Wang RD et al., 2020). With the rapid development of multi-agent reinforcement learning, the communication problem can be divided into two major questions: who needs to be communicated with and what needs to be communi-

cated (Wang RD et al., 2020). Actually, these two major questions are always addressed as scheduling learning tasks and communication protocols (Wang RD et al., 2020). For scheduling learning tasks, recent researchers have shown special interest in multi-agent reinforcement learning with scheduling methods. These methods can be grouped into two classes by the scheduler selections. Several studies proposed gating mechanisms to control agent communications (Jiang and Lu, 2018; Singh et al., 2018; Kim et al., 2019; Mao et al., 2019), whereas other researchers focused on determining the importance of different agent messages and used bi-directional recurrent neural networks (RNNs) (Peng et al., 2017) or the attention mechanism (Das et al., 2019) to learn an adaptive weighted scheduler. For communication protocols, most studies have proposed an end-to-end framework to coordinate message exchanges from different agents (Foerster et al., 2016; Lazaridou et al., 2017; Mordatch and Abbeel, 2018).

In spite of the great progress of prior multi-agent communication studies, there are still many challenges that need to be deeply explored. For example, content redundancy is an urgent problem for scheduling methods. In real multi-agent systems, the bandwidth is always limited (Wang RD et al., 2020), which makes it difficult to transmit a large-scale message between two agents. How to construct a more effective multi-agent communication mechanism is still an open issue in real-world applications, such as the Internet of Vehicles (Li ZY et al., 2021).

### 2.2.2 Collaborative perception

In addition to the communication challenges, the collaborative perception of multi-agent systems is another essential problem. Unfortunately, the exploration of this task is limited (Liu et al., 2020a, 2020b). Different from multi-agent communication, perception focuses mainly on the communication message processing rather than the communication scheduling or protocol. Although single-agent perception has achieved great success in recent years, it is a non-trivial problem to obtain multi-agent perception over single-agent perception. First, a single-agent can hardly perceive comprehensive environments. Then, how to integrate different agents and achieve a precise fused perception is an under-explored issue (Li YM et al., 2021). Second, the transmission bandwidth is a great challenge in both

multi-agent communication (Wang RD et al., 2020) and collaborative perception (Liu et al., 2020a; Li YM et al., 2021). The reason is that communication is an important pre-task for perceptions. Third, considering the various task settings, the sensors in different agents may be different, such as the camera and laser radar (LIDAR) in robotics. Therefore, the fusion of these different sensor data from agents is another important but unexplored problem.

### 2.3 Multi-agent gaming

Recent years have witnessed great improvement in AI of real-time strategy games (Vinyals et al., 2019). Multi-agent gaming, which is one of the major tasks in these areas, has attracted significant research attention. Different from multi-agent collaboration, multi-agent gaming requires an understanding of adversary behaviors in addition to competitive environments and the ability to adaptively make decisions to achieve the task targets or obtain higher scores than competitors. Along this line, multi-agent gaming has two main challenges: competitive environment modeling and competitive decision making.

#### 2.3.1 Competitive environment modeling

For multi-agent gaming, one of the major challenges is understanding competitive environments. Multi-agent competitive environments can be grouped mainly into two classes, i.e., perfect information games and imperfect information games. In perfect information games, the agents can observe all the environments and states of other players in time. On the contrary, imperfect information games present another common competitive environment, in which agents cannot know all the moves already made by the opponent. For example, the Go game (Silver et al., 2016, 2017) is a two-player perfect information game and StarCraft (Shao et al., 2019; Vinyals et al., 2019) is a multi-player imperfect information game. As mentioned above, the biggest difference between competitive and cooperative multi-agent environments is the introduction of competitors. That also brings great challenges. On one hand, the introduction of competitors brings more environmental uncertainty. On the other hand, in multi-agent gaming, environment modeling needs to consider the situation of competitors. Especially for the imperfect information game, the limited ex-

ploration of the environment makes it difficult to predict the behaviors and intent of competitors.

Indeed, most multi-agent gaming methods, especially deep reinforcement learning methods, usually model the competitive environment as a zero-sum game (Barron, 2013; Leonardos et al., 2021). However, in real-world applications, such as power control (Mei et al., 2017), the competitive settings are always general-sum games. Unfortunately, general-sum multi-agent gaming is still largely under-explored (Lin et al., 2019; Mazumdar et al., 2020; Neumeyer et al., 2021).

#### 2.3.2 Competitive decision making

In multi-agent competitive settings, the agent behaviors are much more complex than those in cooperative settings. An agent always needs to predict the behavior of competitors and incorporate an understanding of current environments before creating its own strategy, which makes it difficult to find an equilibrium selection strategy. Prior research has focused mainly on designing convergent algorithms to model the complex behavior of competitive game theoretic settings (Rakhlin and Sridharan, 2013; Balduzzi et al., 2018; Bailey and Piliouras, 2019). However, the competitive decision-making task is still under-explored and presents significant challenges. First, considering complex competitive environments, the different agents have diverse task settings, such as offense, defense, and protection in confrontation games. These task settings should adapt to dynamic competitive environments. Understanding tasks and adaptively selecting agent roles in the environment is one of the key challenges for agent decision making. Second, different from collaborative environments, the introduction of competitors makes understanding and forecasting the purpose of competitors another problem. Third, to achieve the final win, the balance of short- and long-term returns is an important issue that must be considered. In conclusion, these challenges make multi-agent autonomous decisions a non-trivial problem which is still under-explored.

## 3 Technology directions

Based on the research challenges in multi-agent collaboration and gaming, in this section, we discuss some technical areas that may promote

developments of future research in multi-agent systems.

### 3.1 Digital environment construction

As mentioned earlier, most previous works were based on virtual environments, such as games, to develop multi-agent collaboration and gaming platforms. This situation creates a huge gap between virtual environments and real-world applications. To bridge this gap, a more realistic digital environment using real-world scenarios needs to be constructed. Fortunately, there has been tremendous success in computer vision and computer graphics. Computational simulation technologies, such as Digital Twin (Tao et al., 2019) and Metaverse (Dionisio et al., 2013), have been developed to construct 3D environments that can reflect real objects and physical user interactions. These technologies can also add some new perspectives in multi-agent environment construction to further promote application of multi-agent collaboration and gaming in real-world scenarios.

### 3.2 Multi-modal perception and collaboration

In real multi-agent systems, perception of the environment is based mainly on the sensors in the agents. For a comprehensive understanding of the environment, the single agent is always equipped with multiple different sensors to collect multi-modal data, such as the optical signal, electromagnetic signal, and radar signal. Each type of sensor has the specific ability to perceive the environment and the states of other agents. For example, the radar sensor can continuously work in complex monitoring environments, such as foggy or strong light environments. The vision sensors can reflect intuitively and help humans quickly understand the agent's perceptions, which is important for human-machine collaborations. The multi-modal learning and fusion (Baltrušaitis et al., 2019) technologies are important for building an accurate perception of multi-agent systems. Moreover, considering the complexity of agent tasks, the different agents in one multi-agent system could have various perception targets. This means that the collaboration of multiple perceptions is another research direction for multi-agent systems.

### 3.3 Multi-task learning and gaming

In a confrontational and gaming environment, it is important to understand gaming tasks and competitive environments for multi-agent systems. Agent tasks are always changed with dynamic environments, especially with the unknown competitor states in imperfect information gaming. Along this line, transfer learning (Zhuang et al., 2021) and multi-task learning (Zhang Y and Yang, 2018) are considerable methods for improving the generalization performance and adaptively transferring the knowledge from one task to another task in multi-agent gaming. Real responses and data samples are difficult to obtain in many real scenarios, such as military applications. Transfer learning could be an efficient way to transfer a trained model in a simulated environment and apply it in a realistic environment. Considering the large-scale state-action space, hierarchical reinforcement learning (Nachum et al., 2018) is another potential research direction to disentangle the complex multi-agent gaming problem. Inverse reinforcement learning (Arora and Doshi, 2021) is also a potential way to learn from expert trajectories and deal with the reward dilemma designed in the complex multi-agent gaming task.

## 4 Application areas

Since the success of multi-agent systems and reinforcement learning, many researchers have developed multi-agent collaboration and gaming methods in various application areas. In this section, we look at multi-agent collaboration and gaming application areas.

Indeed, a prominent application area is the control of swarm robotic systems (Hüttenrauch et al., 2019) or unmanned aerial vehicles (Tso et al., 1999; Wang YN et al., 2022). These control problems are always based on multi-agent collaboration and gaming methods and require an understanding of the dynamic environment and automatic decisions to achieve complex tasks in competitive or non-competitive settings. Another application area of wide attention is the game AI (Silver et al., 2017; Vinyals et al., 2019). Games have attracted tremendous research attention in recent years and great success has been achieved in related AI methods. Games are good platforms for evaluating multi-agent



collaboration and gaming methods because of the well-defined environments and rewards. However, as mentioned above, how to bridge the gap between virtual game environments and realistic applications is still an open issue. There are also many industrial applications, such as energy management (Lagorse et al., 2010), urban traffic control (Wang Y et al., 2020), and sport AI (Cañizares et al., 2017).

## 5 Conclusions

In this study, we presented the prospects for multi-agent system research with a special focus on agent collaboration and gaming tasks. We briefly introduced some open issues and task challenges from three major perspectives: the multi-agent environment, collaboration, and gaming. Then we provided a related outlook for the technology directions that may create some research challenge insights. Finally, we discussed the outlook for the multi-agent collaboration and gaming application areas.

### Contributors

Yu LIU designed the research. Zhi LI drafted the paper. Yu LIU helped organize the paper. Zhizhuo JIANG and You HE revised and finalized the paper.

### Compliance with ethics guidelines

Yu LIU, Zhi LI, Zhizhuo JIANG, and You HE declare that they have no conflict of interest.

### References

- Arora S, Doshi P, 2021. A survey of inverse reinforcement learning: challenges, methods and progress. *Artif Intell*, 297:103500. <https://doi.org/10.1016/j.artint.2021.103500>
- Arul Kumaran K, Deisenroth MP, Brundage M, et al., 2017. Deep reinforcement learning: a brief survey. *IEEE Signal Process Mag*, 34(6):26-38. <https://doi.org/10.1109/MSP.2017.2743240>
- Bailey JP, Piliouras G, 2019. Multi-agent learning in network zero-sum games is a Hamiltonian system. *Proc 18<sup>th</sup> Int Conf on Autonomous Agents and Multiagent Systems*, p.233-241.
- Balduzzi D, Racanière S, Martens J, et al., 2018. The mechanics of  $n$ -player differentiable games. *Proc 35<sup>th</sup> Int Conf on Machine Learning*, p.354-363.
- Baltrušaitis T, Ahuja C, Morency LP, 2019. Multimodal machine learning: a survey and taxonomy. *IEEE Trans Patt Anal Mach Intell*, 41(2):423-443. <https://doi.org/10.1109/TPAMI.2018.2798607>
- Barron EN, 2013. *Game Theory: an Introduction*. John Wiley & Sons, Hoboken, USA.
- Beattie C, Leibo JZ, Teplyaev D, et al., 2016. DeepMind Lab. <https://arxiv.org/abs/1612.03801v2>
- Bellemare MG, Naddaf Y, Veness J, et al., 2013. The arcade learning environment: an evaluation platform for general agents. *J Artif Intell Res*, 47:253-279. <https://doi.org/10.1613/jair.3912>
- Berner C, Brockman G, Chan B, et al., 2019. Dota 2 with large scale deep reinforcement learning. <https://arxiv.org/abs/1912.06680>
- Betancourt C, Chen WH, 2021. Deep reinforcement learning for portfolio management of markets with a dynamic number of assets. *Expert Syst Appl*, 164:114002. <https://doi.org/10.1016/j.eswa.2020.114002>
- Brockman G, Cheung V, Pettersson L, et al., 2016. OpenAI Gym. <https://arxiv.org/abs/1606.01540>
- Busoniu L, Babuska R, De Schutter B, 2008. A comprehensive survey of multiagent reinforcement learning. *IEEE Trans Syst Man Cybern Part C*, 38(2):156-172. <https://doi.org/10.1109/TSMCC.2007.913919>
- Cañizares PC, Merayo MG, Núñez M, et al., 2017. A multi-agent system architecture for statistics managing and soccer forecasting. *Proc 2<sup>nd</sup> IEEE Int Conf on Computational Intelligence and Applications*, p.572-576. <https://doi.org/10.1109/CIAPP.2017.8167282>
- Coulom R, 2007. Efficient selectivity and backup operators in Monte-Carlo tree search. *Proc 5<sup>th</sup> Int Conf on Computers and Games*, p.72-83. [https://doi.org/10.1007/978-3-540-75538-8\\_7](https://doi.org/10.1007/978-3-540-75538-8_7)
- Das A, Gervet T, Romoff J, et al., 2019. TarMAC: targeted multi-agent communication. *Proc 36<sup>th</sup> Int Conf on Machine Learning*, p.1538-1546.
- Dionisio JDN, Burns WGIII, Gilbert R, 2013. 3D virtual worlds and the metaverse: current status and future possibilities. *ACM Comput Surv*, 45(3):34. <https://doi.org/10.1145/2480741.2480751>
- Foerster JN, Assael YM, de Freitas N, et al., 2016. Learning to communicate with deep multi-agent reinforcement learning. *Proc 30<sup>th</sup> Int Conf on Neural Information Processing Systems*, p.2145-2153.
- Georgeff MP, 1988. Communication and interaction in multi-agent planning. In: Bond AH, Gasser L (Eds.), *Distributed Artificial Intelligence*. Morgan Kaufmann Publishers Inc., San Francisco, USA, p.200-204.
- Grigorescu S, Trasnea B, Cocias T, et al., 2020. A survey of deep learning techniques for autonomous driving. *J Field Robot*, 37(3):362-386. <https://doi.org/10.1002/rob.21918>
- Hernandez-Leal P, Kaisers M, Baarslag T, et al., 2017. A survey of learning in multiagent environments: dealing with non-stationarity. <https://arxiv.org/abs/1707.09183v1>
- Hoehn PJ, Tuyls K, Panait L, et al., 2005. An overview of cooperative and competitive multiagent learning. *Proc 1<sup>st</sup> Int Conf on Learning and Adaption in Multi-Agent Systems*, p.1-46. [https://doi.org/10.1007/11691839\\_1](https://doi.org/10.1007/11691839_1)
- Hüttenrauch M, Šošić A, Neumann G, 2019. Deep reinforcement learning for swarm systems. *J Mach Learn Res*, 20(54):1-31.
- Jennings NR, Sycara K, Wooldridge M, 1998. A roadmap of agent research and development. *Auton Agent Multi-Agent Syst*, 1(1):7-38. <https://doi.org/10.1023/A:1010090405266>

- Jiang JC, Lu ZQ, 2018. Learning attentional communication for multi-agent cooperation. Proc 32<sup>nd</sup> Int Conf on Neural Information Processing Systems, p.7265-7275.
- Johnson M, Hofmann K, Hutton T, et al., 2016. The Malmo platform for artificial intelligence experimentation. Proc 25<sup>th</sup> Int Joint Conf on Artificial Intelligence, p.4246-4247.
- Kempka M, Wydmuch M, Runc G, et al., 2016. ViZDoom: a doom-based AI research platform for visual reinforcement learning. Proc IEEE Conf on Computational Intelligence and Games, p.1-8. <https://doi.org/10.1109/CIG.2016.7860433>
- Kim D, Moon S, Hostallero D, et al., 2019. Learning to schedule communication in multi-agent reinforcement learning. <https://arxiv.org/abs/1902.01554>
- Lagorse J, Paire D, Miraoui A, 2010. A multi-agent system for energy management of distributed power sources. *Renewab Energy*, 35(1):174-182. <https://doi.org/10.1016/j.renene.2009.02.029>
- Lazaridou A, Peysakhovich A, Baroni M, 2017. Multi-agent cooperation and the emergence of (natural) language. <https://arxiv.org/abs/1612.07182>
- Leonardos S, Piliouras G, Spendlove K, 2021. Exploration-exploitation in multi-agent competition: convergence with bounded rationality. <https://arxiv.org/abs/2106.12928>
- Li YM, Ren SL, Wu PX, et al., 2021. Learning distilled collaboration graph for multi-agent perception. <https://arxiv.org/abs/2111.00643v2>
- Li ZY, Yuan Q, Luo GY, et al., 2021. Learning effective multi-vehicle cooperation at unsignalized intersection via bandwidth-constrained communication. Proc IEEE 94<sup>th</sup> Vehicular Technology Conf, p.1-7.
- Lin XM, Adams SC, Beling PA, 2019. Multi-agent inverse reinforcement learning for certain general-sum stochastic games. *J Artif Intell Res*, 66:473-502. <https://doi.org/10.1613/jair.1.11541>
- Liu YC, Tian JJ, Glaser N, et al., 2020a. When2com: multi-agent perception via communication graph grouping. Proc IEEE/CVF Conf on Compute Vision and Pattern Recognition, p.4105-4114.
- Liu YC, Tian JJ, Ma CY, et al., 2020b. Who2com: collaborative perception via learnable handshake communication. Proc IEEE Int Conf on Robotics and Automation, p.6876-6883.
- Mao HY, Gong ZB, Zhang ZC, et al., 2019. Learning multi-agent communication under limited-bandwidth restriction for Internet packet routing. <https://arxiv.org/abs/1903.05561>
- Mazumdar E, Ratliff LJ, Jordan MI, et al., 2020. Policy-gradient algorithms have no guarantees of convergence in linear quadratic games. Proc 19<sup>th</sup> Int Conf on Autonomous Agents and Multiagent Systems, p.860-868.
- Mei SW, Wei W, Liu F, 2017. On engineering game theory with its application in power systems. *Contr Theory Technol*, 15(1):1-12. <https://doi.org/10.1007/s11768-017-6186-y>
- Mordatch I, Abbeel P, 2018. Emergence of grounded compositional language in multi-agent populations. <https://arxiv.org/abs/1703.04908>
- Nachum O, Gu SX, Lee H, et al., 2018. Data-efficient hierarchical reinforcement learning. Proc 32<sup>nd</sup> Int Conf on Neural Information Processing Systems, p.3307-3317.
- Neumeyer C, Oliehoek FA, Gavril DM, 2021. General-sum multi-agent continuous inverse optimal control. *IEEE Robot Autom Lett*, 6(2):3429-3436. <https://doi.org/10.1109/LRA.2021.3060411>
- Nguyen TT, Nguyen ND, Nahavandi S, 2020. Deep reinforcement learning for multiagent systems: a review of challenges, solutions, and applications. *IEEE Trans Cybern*, 50(9):3826-3839. <https://doi.org/10.1109/TCYB.2020.2977374>
- Oroojlooy A, Hajinezhad D, 2019. A review of cooperative multi-agent deep reinforcement learning. <https://arxiv.org/abs/1908.03963>
- Peng P, Wen Y, Yang YD, et al., 2017. Multiagent bidirectionally-coordinated nets: emergence of human-level coordination in learning to play StarCraft combat games. <https://arxiv.org/abs/1703.10069>
- Polydoros AS, Nalpantidis L, 2017. Survey of model-based reinforcement learning: applications on robotics. *J Intell Robot Syst*, 86(2):153-173. <https://doi.org/10.1007/s10846-017-0468-y>
- Puterman ML, 1994. Markov Decision Processes: Discrete Stochastic Dynamic Programming. John Wiley & Sons, Hoboken, USA.
- Rakhlin A, Sridharan K, 2013. Optimization, learning, and games with predictable sequences. Proc 26<sup>th</sup> Int Conf on Neural Information Processing Systems, p.3066-3074.
- Shao K, Zhu YH, Zhao DB, 2019. StarCraft micromanagement with reinforcement learning and curriculum transfer learning. *IEEE Trans Emerg Top Comput Intell*, 3(1):73-84. <https://doi.org/10.1109/TETCI.2018.2823329>
- Silver D, Huang A, Maddison CJ, et al., 2016. Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587):484-489. <https://doi.org/10.1038/nature16961>
- Silver D, Schrittwieser J, Simonyan K, et al., 2017. Mastering the game of Go without human knowledge. *Nature*, 550(7676):354-359. <https://doi.org/10.1038/nature24270>
- Singh A, Jain T, Sukhbaatar S, 2018. Learning when to communicate at scale in multiagent cooperative and competitive tasks. <https://arxiv.org/abs/1812.09755>
- Spielberg SPK, Gopaluni RB, Loewen PD, 2017. Deep reinforcement learning approaches for process control. Proc 6<sup>th</sup> Int Symp on Advanced Control of Industrial Processes, p.201-206. <https://doi.org/10.1109/ADCONIP.2017.7983780>
- Synnaeve G, Nardelli N, Auvolat A, et al., 2016. TorchCraft: a library for machine learning research on real-time strategy games. <https://arxiv.org/abs/1611.00625>
- Tao F, Zhang H, Liu A, et al., 2019. Digital Twin in industry: state-of-the-art. *IEEE Trans Ind Inform*, 15(4):2405-2415. <https://doi.org/10.1109/TII.2018.2873186>
- Tessler C, Givony S, Zahavy T, et al., 2017. A deep hierarchical approach to lifelong learning in minecraft. Proc 31<sup>st</sup> AAAI Conf on Artificial Intelligence, p.1553-1561.
- Todorov E, Erez T, Tassa Y, 2012. MuJoCo: a physics engine for model-based control. Proc IEEE/RSJ Int Conf on Intelligent Robots and Systems, p.5026-5033.

- Tso KS, Tharp GK, Zhang W, et al., 1999. A multi-agent operator interface for unmanned aerial vehicles. Proc Gateway to the New Millennium. Proc 18<sup>th</sup> Digital Avionics Systems Conf, Article 6.A.4.  
<https://doi.org/10.1109/DASC.1999.821969>
- Vinyals O, Babuschkin I, Czarnecki WM, et al., 2019. Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature*, 575(7782):350-354.  
<https://doi.org/10.1038/s41586-019-1724-z>
- Wang RD, He X, Yu RS, et al., 2020. Learning efficient multi-agent communication: an information bottleneck approach. Proc 37<sup>th</sup> Int Conf on Machine Learning, p.9908-9918.
- Wang Y, Cheng ZS, Xiao M, 2020. UAVs' formation keeping control based on multi-agent system consensus. *IEEE Access*, 8:49000-49012.  
<https://doi.org/10.1109/ACCESS.2020.2979996>
- Wang YN, Xu T, Niu X, et al., 2022. STMARL: a spatio-temporal multi-agent reinforcement learning approach for cooperative traffic light control. *IEEE Trans Mob Comput*, 21(6):2228-2242.  
<https://doi.org/10.1109/TMC.2020.3033782>
- Zhang KQ, Yang RZ, Başar T, 2021. Multi-agent reinforcement learning: a selective overview of theories and algorithms. In: Vamvoudakis KG, Wan Y, Lewis FL, et al. (Eds.), *Derya Cansever Handbook of Reinforcement Learning and Control*. Springer, Cham, p.321-384.  
[https://doi.org/10.1007/978-3-030-60990-0\\_12](https://doi.org/10.1007/978-3-030-60990-0_12)
- Zhang Y, Yang Q, 2018. An overview of multi-task learning. *Nat Sci Rev*, 5(1):30-43.  
<https://doi.org/10.1093/nsr/nwx105>
- Zhou HY, Zhang HF, Zhou YS, et al., 2018. Botzone: an online multi-agent competitive platform for AI education. Proc 23<sup>rd</sup> Annual ACM Conf on Innovation and Technology in Computer Science Education, p.33-38.  
<https://doi.org/10.1145/3197091.3197099>
- Zhuang FZ, Qi ZY, Duan KY, et al., 2021. A comprehensive survey on transfer learning. *Proc IEEE*, 109(1):43-76.  
<https://doi.org/10.1109/JPROC.2020.3004555>