Introduction to Quantum Multi-Agent Reinforcement Learning: Concepts and Applications

Yeryeong Cho†, Seok Bin Son†, Soohyun Park§, and Joongheon Kim†
†Department of Electrical and Computer Engineering, Korea University, Seoul, Republic of Korea
§Division of Computer Science, Sookmyung Women's University, Seoul, Republic of Korea
E-mail: joyena0909@korea.ac.kr, lydiasb@korea.ac.kr, soohyun.park@sookmyung.ac.kr, joongheon@korea.ac.kr

***Abstract*—Advancements in quantum computer hardware and software have paved the way for the application of quantum computing across various fields. With the development of computers with thousands of qubits, the power of quantum is increasing. Because of efficiency in high action dimensions, quantum computing has advantages in multi-agent fields where the amount of data is huge. At the same time, reinforcement learning (RL) has gained prominence for its potential in unknown environments. Consequently, the integration of quantum computing and multi-agent reinforcement learning (MARL), known as quantum multi-agent reinforcement learning (QMARL), is proposed. This paper introduces the overall concept and applications of quantum computing, MARL, and QMARL. It also shows the current technology trends depending on the application. In conclusion, the paper presents challenges and future research directions for QMARL and shows that QMARL can be utilized in various applications for excellent performance.**

I. Introduction

Early quantum computers were developed to simulate quantum physics but later expanded into broader areas. With the advances in quantum computers, quantum computing is being used in various fields. Quantum computers utilize quantum bits (qubits) to achieve significant efficiency gains and increase processing speed. The efficient management and reliability of processing large data volumes have garnered significant attention in fields such as artificial intelligence (AI) where data volume is huge. Especially in mobility and networks, many reinforcement learning (RL) algorithms have been proposed for their advantages in uncertain environments. However, as RL has limitations when the action dimension is extremely high, quantum algorithms have been proposed to manage large amounts of data efficiently. Moreover, the application of quantum computing in multi-agent reinforcement learning (MARL), which is named to quantum MARL (QMARL), opens new avenues for dynamic environments, leveraging quantum mechanics to optimize decision-making processes across multiple agents. Finally, this paper proposes an overview and applications of QMARL-based algorithms.

Corresponding Author: Soohyun Park and Joongheon Kim
This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2024-RS-2024-00436887) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation).

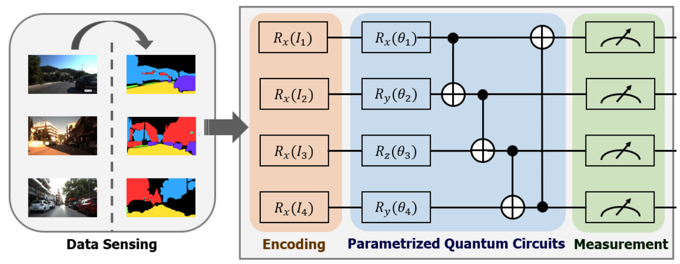

Fig. 1: The structure of QNN.

II. Quantum Neural Network

A. Characteristics of Quantum Computing

Quantum is composed of qubits, which differ from classical bits (cbits). Compared to a cbit, which is simply a 1 or 0, a qubit is expressed as a probability. This allows a qubit to represent both 0 and 1 simultaneously, enabling a small number of qubits to represent a large amount of data. Based on the two-dimensional Hilbert space, the qubit state ψ is represented as follows,

$$\psi = \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} = \alpha_0|0\rangle + \alpha_1|1\rangle, \tag{1}$$

where α_0 and α_1 are the probabilities of 0 and 1, respectively. Because $|0\rangle$ and $|1\rangle$ have a unitary norm, the two elements satisfy $\|\psi\|_2^2 = \langle\psi|\psi\rangle = |\alpha_0|^2 + |\alpha_1|^2 = 1$. This is a superposition of quantum states, which can represent two states at the same moment. Another characteristic of quantum is entanglement. Given the entangled state $|\psi_A\psi_B\rangle = \frac{1}{\sqrt{2}}(|\psi_{A0}\psi_{B0}\rangle + |\psi_{A1}\psi_{B1}\rangle)$, one of following cases occurs with equal probability. If $|\psi_A\rangle$ is observed as 0, then $|\psi_A\psi_B\rangle$ becomes $|01\rangle$. On the other hand, if $|\psi_A\rangle$ is observed as 1, then $|\psi_A\psi_B\rangle$ becomes $|10\rangle$. Because of this feature, when a qubit in the entangled state is measured, the state of another qubit is determined without measurement.

B. Quantum Neural Network

Quantum neural networks (QNNs) are constructed with artificial neural networks and quantum computing, utilizing quantum circuits simultaneously. Quantum circuits are basic operations that manipulate the states of qubits and correspond to logical circuits such as AND, OR, and NOT in classical

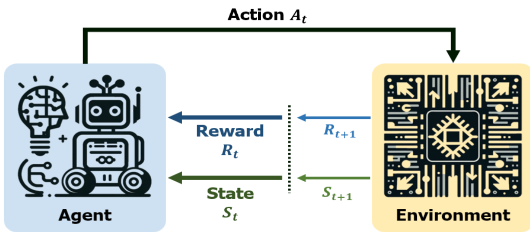

Fig. 2: The progress of Reinforcement learning.

computing, but can take advantage of quantum such as superposition and entanglement. As shown in Fig. 1, the input data is sensed by multiple sensors and encoded [1] to convert classical data into quantum data, which can be used by the QNN. Subsequently, it is transformed through a parameterized quantum circuit (PQC). Lastly, it is decoded into the proper output data. The advantage of using QNN is demonstrated in the reduction of model parameters and low memory consumption [2]. It also demonstrates performance which varies with the structure of the PQC. In addition, QNN can be used in various areas such as network communications, resource allocation, caching networks, and video scheduling [3].

III. Quantum Multi-Agent Reinforcement Learning

A. Basics of Multi-Agent Reinforcement Learning

MARL is an RL approach to multi-agent systems that solves complex problems by breaking them down into smaller tasks. RL algorithm is stated in a Markov decision process (MDP), a mathematical framework for solving problems that require sequential decision-making, presented by $\{S, A, P, R, \gamma\}$. Each element consists of a set of states S, a set of actions A, and a discounted value γ for future reward's value. The state transition function P represents the probability that the next state will be s' when action a is taken from the current state s at time step t, as denoted as (2),

$$P^a_{ss'} = \mathbb{P}[S_{t+1} = s' | S_t = s, A_t = a] \tag{2}$$

and lastly, the reward R can be formulated as (3),

$$R^a_s = \mathbb{E}[R_{t+1} | S_t = s, A_t = a] \tag{3}$$

which means the reward for choosing action a in state s. Therefore, the overall structure of RL is presented in Fig. 2.

In MARL, each agent tries to achieve maximum reward by choosing the optimal behavior in a given environment based on its state. The final goal is for each agent to learn the optimal policy, either individually or collectively, to achieve a common goal or perform a specific task. MARL has advantages in mobility, such as cooperative navigation or manipulation in robotics, optimal route planning in transportation systems, and optimization of energy distribution systems. An optimal trajectory planning algorithm is proposed for multiple unmanned aerial vehicles (UAVs), using a multi-agent deep deterministic policy gradient (MADDPG) to minimize risks caused by large action dimensions [4]. In addition, a MARL-based resource allocation algorithm for UAVs is suggested by using a Q-learning algorithm [5]. Especially, the paper demonstrates energy efficiency in uncertain environments through simulations with multiple UAVs. In robotics, a collision-avoiding algorithm is proposed to avoid collisions between each of the agents in multiple mobile robots [6]. However, due to uncertainty about its applicability in the real world, experiments in heterogeneous environments were required. A multi-robot system architecture composed of cooperative searching and local map merging is proposed [7], with experiments conducted in the real world.

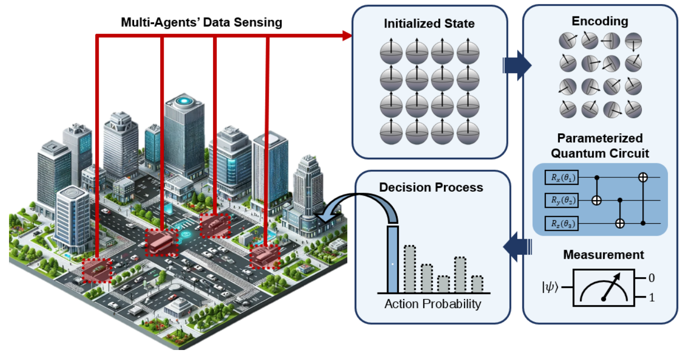

Fig. 3: The structure of QMARL.

B. Quantum Multi-Agent Reinforcement Learning

QMARL is a research field that applies the principles of quantum computing to MARL. It utilizes quantum characteristics like entanglement and superposition to ameliorate the computational limitations and efficiency issues of traditional MARL as shown in Fig. 3. QMARL attempts to use the fundamental principles of quantum mechanics to improve the performance of RL algorithms and more effectively solve optimization problems. When controlling multiple agents, it can efficiently minimize computational complexity by leveraging the computational superiority of quantum mechanics [8]. Furthermore, this characteristic can expedite the training process and facilitate the resolution of more intricate decision-making challenges with large action dimensions [9]. Because of the many advantages of quantum mechanics, applications in mobility with large action dimensions are getting more attention. Effective management of large-scale autonomous mobility systems is demonstrated without the exponential increase in computational complexity associated with traditional MARL algorithms [10]. An algorithm composed of a quantum actor and a quantum centralized critic demonstrates resource efficiency, reducing the number of qubits from 2^n to N [11]. The paper showed that the algorithm outperforms traditional MARL when the number of agents is exponentially large such as autonomous aerial networks composed of multiple UAVs. Another QMARL algorithm demonstrates efficient resource allocation performance that can be applied to a microgrid (MG), the self-sufficient electricity system in smart cities [12]. In addition to various applications, the quantum meta MARL algorithm is proposed by applying meta-learning to quickly adapt and optimize for specific tasks or environments, trying to improve the performance of QMARL by utilizing other methods [13]. A centralized training and decentralized execution (CTDE) QMARL framework is proposed by designing a variational quantum circuit [14]. Moreover, an algorithm that provides stabilized learning and a more optimal policy even in larger dimensions is proposed by adding an experience replay buffer and an additional target network [15]. Consequently, QMARL has great applicability to various fields that require many agents and high action dimensions such as network communication, resource allocation, and route optimization [16].

IV. Challenges and Future Directions

This section introduces the challenges of QMARL. QMARL algorithms have been applied and studied in various fields, however, there are many challenges [17].

- Quantum hardware limitation: Quantum hardware has developed overwhelmingly compared to the past, including a 1000-qubit quantum chip. However, to encode classical data of multi-agent systems to quantum data, the quantum hardware is still insufficient [18].
- Interoperability with existing classical Systems: Various quantum algorithms have been developed and prove their effectiveness. However, current QMARL algorithms cannot fully guarantee their applicability with existing classical systems, due to complex data structures and memory [19].
- Low accessibility: Despite the existence of open-source platforms such as Qiskit, Cirq, and Forest, quantum computing has lower accessibility to quantum algorithms compared to existing algorithms. These quantum platforms are less accessible than existing platforms such as TensorFlow or PyTorch in terms of application and verification as a multi-agent system.
- Limitations of qubit utilization: Current quantum technology cannot maintain the quantum state, making it difficult to preserve quantum information for a long time. Consequently, it limits the number of available qubits which restricts computational capabilities, and the possibility of errors occurring can be high [20].

V. Concluding Remarks

This paper introduces the basic concepts of quantum and presents the applications and limitations of QMARL. With the recent advances in software and hardware of quantum computers, algorithms have been proposed in various fields. There is a clear difference in accuracy and learning speed depending on the QMARL algorithms. In addition, since quantum computing reduces the action dimension in the logarithmic scale, it has the advantage of efficient learning with less memory when fused with MARL. The paper introduces recent research through papers in various fields such as networks, robotics, and transportation. This paper provides an overview of the concepts and applications of QMARL. By suggesting challenges that need to be improved in QMARL, it provides several discussions on future research directions.

References

[1] O. Schumann, M. Hahn, N. Scheiner, F. Weishaupt, J. Tilly, J. Dickmann, and C. Wöhle, "The radarscenes data set," 2022. [Online]. Available: https://www.kaggle.com/datasets/aleksandrdubrovin/the-radarscenes-data-set

[2] Y. Li, M. Tian, G. Liu, C. Peng, and L. Jiao, "Quantum optimization and quantum learning: A survey," *IEEE Access*, vol. 8, pp. 23 568–23 593, Jan. 2020.

[3] Y. Kwak, W. J. Yun, S. Jung, and J. Kim, "Quantum neural networks: Concepts, applications, and challenges," in *Proc. IEEE Int'l Conf. Ubiquitous and Future Networks (ICUFN)*, Jeju Island, Republic of Korea, Aug. 2021, pp. 413–416.

[4] L. Wang, K. Wang, C. Pan, W. Xu, N. Aslam, and L. Hanzo, "Multi-agent deep reinforcement learning-based trajectory planning for multi-uav assisted mobile edge computing," *IEEE Transactions on Cognitive Communications and Networking*, vol. 7, no. 1, pp. 73–84, Mar. 2021.

[5] J. Cui, Y. Liu, and A. Nallanathan, "Multi-agent reinforcement learning-based resource allocation for UAV networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 2, pp. 729–743, Feb. 2020.

[6] P. Long, T. Fan, X. Liao, W. Liu, H. Zhang, and J. Pan, "Towards optimally decentralized multi-robot collision avoidance via deep reinforcement learning," in *Proc. IEEE Int'l Conf. Robotics and Automation (ICRA)*, Brisbane, Australia, May 2018, pp. 6252–6259.

[7] H. Xie, D. Zhang, X. Hu, M. Zhou, and Z. Cao, "Autonomous multirobot navigation and cooperative mapping in partially unknown environments," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–12, Nov. 2023.

[8] E. Andrés, M. P. Cuéllar, and G. Navarro, "Efficient dimensionality reduction strategies for quantum reinforcement learning," *IEEE Access*, vol. 11, pp. 104 534–104 553, Sept. 2023.

[9] H. Baek, S. Park, and J. Kim, "Logarithmic dimension reduction for quantum neural networks," in *Proc. ACM Int'l Conf. Information and Knowledge Management (CIKM)*, New York, USA, Oct. 2023, p. 3738–3742.

[10] S. Park, J. P. Kim, C. Park, S. Jung, and J. Kim, "Quantum multi-agent reinforcement learning for autonomous mobility cooperation," *IEEE Communications Magazine*, vol. 62, no. 6, Jun. 2024.

[11] S. Park and J. Kim, "Quantum reinforcement learning for large-scale multi-agent decision-making in autonomous aerial networks," in *Proc. IEEE VTS Asia Pacific Wireless Communications Symposium (APWCS)*, Taiwan, Aug. 2023, pp. 1–4.

[12] R. Yan, Y. Wang, Y. Xu, and J. Dai, "A multiagent quantum deep reinforcement learning method for distributed frequency control of islanded microgrids," *IEEE Transactions on Control of Network Systems*, vol. 9, no. 4, pp. 1622–1632, Dec. 2022.

[13] W. J. Yun, J. Park, and J. Kim, "Quantum multi-agent meta reinforcement learning," in *Proc. AAAI Conf. Artificial Intelligence*, Washington, DC, USA, Feb. 2023, pp. 11 087–11 095.

[14] W. J. Yun, Y. Kwak, J. P. Kim, H. Cho, S. Jung, J. Park, and J. Kim, "Quantum multi-agent reinforcement learning via variational quantum circuit design," in *Proc. Int'l Conf. Distributed Computing Systems (ICDCS)*, Bologna, Italy, Jul. 2022, pp. 1332–1335.

[15] T. Müller, C. Roch, K. Schmid, and P. Altmann, "Towards multi-agent reinforcement learning using quantum boltzmann machines," in *Proc. Int'l Conf. Agents and Artificial Intelligence (ICCART)*, Virtual, Feb. 2022, pp. 121–130.

[16] B. Narottama and S. Y. Shin, "Federated quantum neural network with quantum teleportation for resource optimization in future wireless communication," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 11, pp. 14 717–14 733, Nov. 2023.

[17] W. Yu and J. Zhao, "Quantum multi-agent reinforcement learning as an emerging AI technology: A survey and future directions," in *Proc. Int'l Conf. Computer and Applications (ICCA)*, Bangkok, Thailand, Nov. 2023, pp. 1–7.

[18] S. R. Hasan, M. Z. Chowdhury, M. Saiam, and Y. M. Jang, "Quantum communication systems: Vision, protocols, applications, and challenges," *IEEE Access*, vol. 11, pp. 15 855–15 877, Feb. 2023.

[19] T. M. Khan and A. Robles-Kelly, "Machine learning: Quantum vs classical," *IEEE Access*, vol. 8, pp. 219 275–219 294, Jul. 2020.

[20] H. Qi, L. Wang, C. Gong, and A. Gani, "A survey on quantum data mining algorithms: challenges, advances and future directions," *Quantum Information Process*, vol. 23, no. 74, pp. 1–42, Feb. 2024.