LITERATURE REVIEW - QRL

A) Methodologies & Approaches

There are 2 main classifications of QRL methodologies, one based on the classical-Quantum trade-off and the other based on the Gate-Annealing Quantum methodologies.

In the former classification, the QRL algorithms can again be broadly classified into two categories:

1. **Classical-quantum hybrid algorithms:** These algorithms use classical RL techniques to train quantum circuits that represent the agent's policy or value function and are a popular approach to QRL due to their compatibility with existing RL techniques. Some of the most used algorithms in this category include:
    a. **Deep Q-Learning (DQN):** DQN is a value-based RL algorithm that uses a deep neural network to approximate the Q-function, which represents the expected future reward for taking a particular action in a given state.
    b. **Soft Actor-Critic (SAC):** SAC is a policy-based RL algorithm that combines the advantages of actor-critic methods with the stability of maximum entropy reinforcement learning.
    c. **Proximal Policy Optimization (PPO):** PPO is a policy gradient RL algorithm that uses a trust region approach to ensure that the policy updates are small and do not significantly change the agent's behavior.

2. **Fully quantum algorithms:** These algorithms use quantum computing throughout the learning process, including the representation of the agent's policy and the computation of the reward function. Fully quantum algorithms are still in their early stages of development, but they have the potential to offer significant advantages over classical-quantum hybrid algorithms. Some of the most promising fully quantum algorithms include:
    a. **Quantum Policy Gradient (QPG):** QPG is a policy gradient RL algorithm that uses quantum circuits to represent the agent's policy and the reward function.
    b. **Quantum Value Iteration (QVI):** QVI is a value-based RL algorithm that uses quantum circuits to represent the Q-function and the environment dynamics.
    c. **Variational Quantum Policy Gradient (VQPG):** VQPG uses a variational quantum circuit to represent the policy and updates the circuit parameters using policy gradient methods.

By comparing the performance of these scenarios of QRL Algorithms, which depends on a variety of factors, including the specific task, the size of the quantum circuit, and the noise level in the quantum system, we find that, classical-quantum hybrid algorithms have been shown to perform well on small-scale QRL tasks, while fully quantum algorithms have the potential to offer significant advantages on larger-scale tasks.

Mentioning the latter classification of QRL Algorithms, not much is explored yet, but on the basis of existing literature, one can also classify the Algorithms on the following criteria (Neumann et al., 2023).

1. **Gate-based quantum RL:** This approach uses quantum gates to implement the RL algorithm. Quantum gates can be used to perform operations such as state preparation, measurement, and unitary evolution.

2. **Annealing-based quantum RL:** This approach uses quantum annealing to find the optimal policy. Quantum annealing is a heuristic optimization technique that can be used to solve combinatorial optimization problems.

This specific sect of classification arose to mostly give insights into speed, efficiency and robustness for a given tasks, to define the above mentioned:

Speed: Quantum algorithms can be exponentially faster than classical algorithms for certain tasks. This can lead to significant speedups in RL training. Different methodologies lead to different tasks speeding up within a certain bound in the QRL algorithm.

Efficiency: Quantum algorithms can be more efficient than classical algorithms in terms of memory usage and energy consumption. Various QRL algorithms lead various efficiency factors based on the task and the compute at hand.

Robustness: Quantum algorithms can be more robust to noise and errors than classical algorithms (If built meticulously) and this also plays a crucial factor in deciding the methodology for a certain task.

To evaluate the performance of different QRL algorithms, researchers have developed standardized benchmarks. These benchmarks typically involve tasks such as controlling quantum systems, optimizing quantum circuits, and solving quantum games.

Several studies have compared the performance of different QRL algorithms on various benchmarks and the choice of algorithm depends on the specific QRL task and the available computational resources. These studies have shown that, DQNs can achieve good performance on small-scale QRL problems while SAC and PPO are more suitable for larger-scale problems and can handle continuous action spaces.

B) Quantum Agents in Gym

Parametrized Quantum Policies

(Skolik et al., 2022) and (Jerbi et al., 2021) introduce quantum agents and parametrized quantum policies for reinforcement learning. They demonstrate the potential of these approaches in solving classical benchmarking tasks and even show an advantage over classical algorithms in certain cases. To subtly elaborate the key-aspects of these papers:

1. **Quantum Agents:** Hybrid systems that combine quantum computations with classical learning algorithms.
2. **Parametrized Quantum Circuits (PQCs):** Quantum circuits with trainable parameters that define a hypothesis family for learning tasks.
3. **RAW-PQC** and **SOFTMAX-PQC:** Two families of PQC-based policies with different output mechanisms.

These specific entities contribute in the following way:

1. **Solving Benchmarking Tasks:** Quantum agents can achieve comparable performance to classical deep neural networks in standard reinforcement learning environments.

2. **Theoretical Learning Advantage:** PQCs can solve certain learning tasks that are intractable to classical models under the Discrete Logarithm Problem assumption.
3. **Efficient Policy Sampling and Training:** Policies can be sampled and trained efficiently using noisy estimates of expectation values.

C) Challenges and Scope

Quantum RL is still a relatively new field, and there are numerous challenges that need to be addressed before it can be widely used. These challenges include:

1. **Hardware limitations:** Current quantum computers are still small and noisy at the hardware level, which limits the size and complexity of RL problems that can be solved.
2. **Algorithm development:** New quantum RL algorithms need to be developed to take advantage of the unique properties of quantum systems.
3. **Software tools:** Software tools need to be developed to make it easier to implement and use quantum RL algorithms. Thanks to TFQ and the upcoming QML (Qiskit Machine Learning) we currently can take baby steps in this field, and we hope more support can be offered in this scope (Broughton et al., 2021).
4. **Theoretical Understanding:** The theoretical foundations of QRL are still under development, and it is not fully understood when and why quantum algorithms outperform classical counterparts.

Future research directions in quantum RL include:

1. **Hybrid quantum-classical algorithms:** Combining quantum and classical RL techniques to exploit the strengths of both approaches.
2. **Development of New QRL Algorithms:** Exploring novel RL algorithms specifically designed for quantum environments, such as algorithms that leverage quantum entanglement or quantum coherence.
3. **Integration with Quantum Error Correction:** Developing techniques to mitigate the effects of noise and errors in quantum systems, ensuring the robustness of QRL algorithms.
4. **Theoretical Analysis of QRL:** Establishing theoretical guarantees for the convergence and optimality of QRL algorithms, providing a deeper understanding of their performance and limitations.
5. **Novel quantum architectures:** Exploring new quantum hardware architectures that are better suited for RL applications.

D) Applications of QRL

QRL has the potential to be applied to a wide range of problems, including:

1. **Quantum Control:** QRL algorithms can be used to optimize quantum control sequences for tasks such as state preparation, quantum gate implementation, and quantum error correction.
2. **Quantum Game Playing:** QRL has been used to develop quantum strategies for games such as Go and chess, demonstrating the potential for quantum computing to enhance decision-making in complex environments.

3. **Quantum Simulation:** QRL can be used to simulate quantum systems more efficiently than classical methods, enabling the study of complex quantum phenomena and the development of new quantum algorithms.
4. **Quantum Chemistry:** QRL can also be applied to quantum chemistry problems such as molecular energy estimation and quantum state preparation, offering the potential for more accurate and efficient simulations.
5. **Quantum Optimization:** QRL has been applied to combinatorial optimization problems, where quantum circuits are used to represent candidate solutions and quantum operators are used to optimize the solutions.
6. **Quantum Finance:** QRL can be explored for financial trading, where quantum circuits can model complex market dynamics and optimize trading strategies.
7. **Quantum Operations:** QRL can also be used in the specific operations research domain for solving complex scheduling and routing problems, along with many more.

In conclusion, this comprehensive review underscores the promising trajectory of quantum reinforcement learning (QRL), a burgeoning field at the convergence of quantum computing and reinforcement learning principles. Central to the advancement of QRL are Variational Quantum Algorithms (VQAs), which play a pivotal role in approximating intricate functions such as value and policy functions. Notably, architectural decisions regarding structure, data encoding techniques, and readout operators are paramount for optimizing VQA performance within the QRL framework.

While both value-based and policy-based QRL algorithms have demonstrated encouraging results across various reinforcement learning tasks, it is imperative to address formidable challenges such as decoherence, scalability, and computational overhead to fully harness the potential of QRL.

Continued research endeavors are diligently focused on surmounting these challenges, paving the way for innovative solutions and exploring the expansive applicability of QRL in addressing real-world problems. As the field progresses, advancements in overcoming these hurdles will undoubtedly propel QRL into a transformative force, revolutionizing decision-making processes across diverse domains.

REFERENCES

1. Broughton, M., Verdon, G., McCourt, T., Martinez, A. J., Yoo, J. H., Isakov, S. V., Massey, P., Halavati, R., Niu, M. Y., Zlokapa, A., Peters, E., Lockwood, O., Skolik, A., Jerbi, S., Dunjko, V., Leib, M., Streif, M., Von Dollen, D., Chen, H., … Mohseni, M. (2021). *TensorFlow Quantum: A Software Framework for Quantum Machine Learning* (arXiv:2003.02989; Version 2). arXiv. http://arxiv.org/abs/2003.02989
2. Jerbi, S., Gyurik, C., Marshall, S. C., Briegel, H. J., & Dunjko, V. (2021). *Parametrized quantum policies for reinforcement learning* (arXiv:2103.05577). arXiv. http://arxiv.org/abs/2103.05577
3. Neumann, N. M. P., de Heer, P. B. U. L., & Phillipson, F. (2023). Quantum reinforcement learning. *Quantum Information Processing*, *22*(2), 125. https://doi.org/10.1007/s11128-023-03867-9

4. Skolik, A., Jerbi, S., & Dunjko, V. (2022). Quantum agents in the Gym: A variational quantum algorithm for deep Q-learning. *Quantum*, *6*, 720. https://doi.org/10.22331/q-2022-05-24-720