QUANTUM REINFORCEMENT LEARNING

Anonymous authors
Paper under double-blind review

ABSTRACT

With the rapid development of quantum technology, it has been confirmed that it can surpass the speed of traditional computing in some fields. Quantum advantage can also be manifested in the field of machine learning. We reviewed many current papers related to quantum reinforcement learning. We discuss in depth how quantum reinforcement learning is implemented and core techniques. quantum reinforcement learning (QRL) method is proposed by combining quantum theory and reinforcement learning (RL).The field of quantum reinforcement learning actually includes two aspects: One is use quantum properties to help reinforcement learning, the other is using reinforcement learning to help quantum circuit design. We have completed agent training for several classic games using quantum reinforcement learning methods, and the superiority and feasibility of the simulation experiments were evaluated. The QRL algorithm can be used in many aspects such as finance, industrial simulation, mechanical control, quantum communication, and quantum circuit optimization. We take a look at the field of quantum reinforcement learning and make bold predictions that many applications in the future will benefit from the development of this technology.

1 INTRODUCTION

Quantum machine learning has emerged as a promising paradigm that could accelerate machine learning calculations. This quantum technology may enhance quantum computation and communication, as well as machine learning, via the fruitful marriage between these previously unrelated fifieldsLamata (2021).We re-design multi-agent reinforcement learning (MARL) based on the unique characteristics of quantum neural networks (QNNs),We theoretically prove the convergence of angle training under the angle-to-pole regularization, and by simulation corroborate the effectiveness of QM2ARL in achieving high reward and fast convergence, as well as of the pole memory in fast adaptation to a time-varying environmentYun et al. (2022).For a broad family of two-qubit unitary gates that are important for quantum simulation of many-electron systems, we improve the control robustness by adding control noise into training environmentsfor reinforcement learning agents trained with trusted-region-policy-optimization.Meanwhile,Our results open a venue for wider applications in quantum simulation, quantum chemistry and quantum supremacy tests using near-term quantum devicesNiu et al. (2019).We introduce a technique combining reinforcement learning, optimisation algorithm and a custom effifffifficient simulation of quantum optics experiments to automate the design of photonic setups maximizing a given function of the measurement statistics.These confifigurations might be helpful to facilitate a fifirst implementation of DIQKD with photonic devices and for future developments targeting improved performancesValcarce et al. (2022). A novel quantum reinforcement learning (QRL) method is proposed by combining quantum theory and reinforcement learning (RL).The present work is also an effective exploration on the application of quantum computation to artifificial intelligence.

A quantum reinforcement learning protocol in the presence of thermal dissipation is introduced and analyzedDong et al. (2008),Olivera-Atencio et al. (2022).Then,this work examines secret key rates of key distribution based on quantum repeaters in a broad parameter space of the communication distance and coherence time of the quantum memoriesRei & van Loock (2022).Quantum annealing algorithms belong to the class of metaheuristic tools, applicable for solving binary optimization problems.we present a way to partially embed both Monte Carlo policy iteration for fifinding an optimal policy on random observations, as well as how to embed n sub-optimal state-value functions for approximating an improved state-value function given a policy for fifinite horizon games with discrete state spaces on a D-Wave 2000Q quantum processing unit (QPU)Neukart et al. (2018). It is known that at low depth, certain locality constraints of QAOA limit its performance. To go beyond these limitations, a non-local variant of QAOA, namely recursive QAOA (RQAOA), was proposed to improve the quality of approximate solutionsPatel et al. (2022). We introduce a model-free and deep learning-based approach to effificiently implement realistic Bayesian quantum metrology tasks accomplishing all the relevant challenges, without relying on any a-priori knowledge on the system.This work represents an important step towards fully artificial intelligence-based quantum metrologyCimini et al. (2022). Recent works show that parameterized quantum circuits (PQCs) can be used to solve challenging reinforcement learning (RL) tasks with provable learning advantages. In this work, we introduce EQAS-PQC, an evolutionary quantum architecture search framework for PQC-based models, which uses a population-based genetic algorithm to evolve PQC architectures by exploring the search space of quantum operationsDing & Spector (2022). Based on the RIMp, an algorithmic robustness-infifidelity measure (ARIM) is developed to quantify the expected robustness and fifidelity of controllers found by a control algorithmKhalid et al. (2022). Based on the training of the network over numerous preparing tasks, we could investigate how to prepare a certain single- or two-qubit target state from arbitrary initial states in semiconductor double quantum dots with only a few discrete control pulses by leveraging the deep reinforcement learningHe et al. (2021). An ideal quantum heat engine operates at high power, high effifffificiency, and high stability. Here we propose such a general framework to identify Pareto-optimal cycles for driven quantum heat engines that trade-offff power, effifffificiency, and flfluctuationsErdman et al. (2022). With quantum computers still under heavy development, already numerous quantum machine learning algorithms have been proposed for both gate-based quantum computers and quantum annealers.We can extend this work based on quantum Boltzmann machines, by allowing for any number of agents. It can improve the learning compared to classical methodsNeumann et al. (2020).Quantum annealing algorithms belong to the class of metaheuristic tools, applicable for solving binary optimization problems. we present a way to partially embed both Monte Carlo policy iteration for fifinding an optimal policy on random observations, as well as how to embed n sub-optimal state-value functions for approximating an improved state-value function given a policy for fifinite horizon games with discrete state spaces on a D-Wave 2000Q quantum processing unit (QPU). And it shows that quantum-enhanced Monte Carlo policy evaluation allows for finding equivalent or better state-value functions for a given policyNeukart et al. (2018).*Deep Learning* Goodfellow et al. (2016) A multi-agent reinforcement learning (MARL) architecture combining both paradigms has been proposedMller et al. (2021). This novel algorithm, which utilizes Quantum Boltzmann Machines (QBMs) for Q-value approximation has out performed regular deep reinforcement learning in terms of time-steps needed to converge. Similar to classic DQNs, we add an experience replay buffer and use different networks for approximating the target and policy values. The experimental results show that learning becomes more stable and enables agents to fifind optimal policies in grid-domains with higher complexity.

2 Review of Quantum Computing and Reinforcement Learning

Quantum algorithms take a new approach to these complex problems by creating multidimensional spaces in which patterns linking individual data points appear. Classical computers cannot create these computational spaces, so they cannot find these patterns.

2.1 Quantum Computing Fundamentals

Qubit and Its Properties
Unit of quantum computing:
one qubit $|\psi >= \alpha|0 > +\beta \mid 1 >$, with $\alpha^2 + \beta^2 = 1$

Superposition : qubit $|\psi\rangle$ is in superposition of $\mid 0 >$ and $|1\rangle$, when measure $|\psi\rangle$, it will Collapse into $\mid 0 >$ w. p. α^2 and $\mid 1 >$ w . p. β^2

Parallelism: the transformation U : $|\psi, 0 > \longrightarrow |\psi, F(\psi) >= \alpha|0, F(0) > +\beta|1, f(1) >$ (quantum black box)evaluates $\mathrm{F(x)}$ for two values simultaneously.

Scaling up to n - qubit system:

$$|\phi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes \ldots |\psi_n\rangle = \sum_{x=00\ldots0}^{11\ldots1} C_x |x\rangle$$

$$U \sum_{x=00\ldots0}^{11\ldots1} C_x |x, 0\rangle = \sum_{x=00\ldots0}^{11\ldots1} C_x U |x, 0\rangle = \sum_{x=00\ldots0}^{11\ldots1} C_x |x, f(x)\rangle$$

Common Quantum Gates Hadamard gate (single-qubit gate) transforms clustering state into to uniform superposed state

$$H \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad H|0\rangle \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$$

Phase gate (single-qubit gate):

$$U_{phase} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\varphi} \end{pmatrix}$$

CNOT gate (two-qubit gate):

$$\mathrm{CNOT} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

3 QUANTUM REINFORCEMENT LEARNING

3.1 PRINCIPLE AND MECHANISM

Quantum reinforcement learning The same interaction-based learning process, whose interaction process can be divided into classical interaction mode and quantum interaction mode. Classical interaction mainly includes the interaction between classical The classical interaction consists of the interaction between the classical task environment and the quantumized intelligent body, and the interaction between the classical intelligent body and the The interaction process between classical intelligence and quantitative intelligence, and the interaction process between classical intelligence and quantitative task environment.

3.2 QUANTUM DQN'S EXPERIMENTS AND ANALYSIS

To test the reliability of the quantum reinforcement learning approach, we tested it on several reinforcement learning base environments and achieved good speedup and model convergence results.

we define the classical pre-processing (*encoding*) layer, which takes the classical inputs $\vec{s} = (s_0, s_1, s_2, s_3)$, multiplies them by a trainable parameters $\vec{w} = (w_0, w_1, w_2, w_3)$, and then applies an $\arctan$ on it, thus leading to the overall mapping:

$$\vec{s} \to \vec{s}' = (s'_0, s'_1, s'_2, s'_3) \quad \text{with} \quad s'_i = \arctan(w_i \cdot s_i)\,.$$

The Q-values $Q(s, a)$ of action a given a state s is given by:

$$Q(s, a) = w_a \cdot \frac{\langle \mathbf{0} | U_\theta(\vec{s})^\dagger O_a U_\theta(\vec{s}) | \mathbf{0} \rangle + 1}{2}$$

where $U_\theta(\vec{s})$ denotes the quantum neural network (which is a function of the input state s and of the variational parameters θ), and w_a denotes a new trainable weight, one for each of the actions.

Number of Shots	Mean Score	Standard deviation	Number of Episodes
1024	115	49.82	100
2048	193	28.01	100
4096	195	11.20	100
8192	199.4	0.283	100

In order to go from the vector of probabilities $\vec{p}$ to the expectation values $\langle Z_0 Z_1 \rangle$ and $\langle Z_2 Z_3 \rangle$, one can do the following. Be $|\psi\rangle = U_\theta(\vec{s})|0000\rangle$ the state created by the Parametrized Quantum Circuit, on which we want to evaluate the expectation vales. Then

$$\langle Z_j Z_k \rangle = \langle \psi | Z_j Z_k \mathbb{1} | \psi \rangle = \sum_{i=0}^{15} \langle \psi | \underbrace{Z_j Z_k |i\rangle}_{=f_{kj}(i)|i\rangle} \langle i | \psi \rangle = \sum_{i=0}^{15} f_{kj}(i) |\langle i|\psi\rangle|^2 = \sum_{i=0}^{15} f_{jk}(i) \cdot p_i$$

where $f_{jk}(i)$ is given by

$$Z_j Z_k |i\rangle = Z_j Z_k \, |i_3 i_2 i_1 i_0\rangle = (-)^{i_j} (-)^{i_k} \, |i_3 i_2 i_1 i_0\rangle = (-)^{i_j + i_k} |i\rangle \quad \text{(NOTE: Qiskit ordering is being used!)}$$

$$\Longrightarrow \langle Z_j Z_k \rangle = \sum_i (-)^{i_j + i_k} p_i$$

One can check manually that (using Qiskit little endian ordering of the qubits):

$$Z_0 Z_1 \rightarrow \vec{f}_{01} = (1, -1, -1, 1, 1, -1, -1, 1, 1, -1, -1, 1, 1, -1, -1, 1)$$
$$Z_2 Z_3 \rightarrow \vec{f}_{23} = (-1, -1, -1, -1, 1, 1, 1, 1, -1, -1, -1, -1, 1, 1, 1, 1)$$

4 Future Prospects

The application of quantum reinforcement learning to quantum experimental control is expected to form a precise Molecular behavioral simulations will provide insights into chemical production, energy, healthcare, and more.
The rapid adaptability of quantum reinforcement learning will have a significant impact on the field of complex financial analysis and rapid decision-making bring hope.
quantum reinforcement learning combined with cloud computing, it can provide a new vision for intelligent quantum cloud computing.

We are more interested in finance.Financial services typically employ probabilities of performance by markets and portfolios Algorithms composed of and hypotheses to make decisions on investment methods. But due to the ability of statistical algorithms to quickly analyze large-scale data in real time is limited. There are still many problems in the integration of risk and fraud detection. Quantum Computing and Machines. The combination of learning methods can effectively eliminate data blind spots and identify unfounded financial assumptions to avoid losses. Quantum Reinforcement Learning Will Optimize Complexity Problem solving offers promising prospects for portfolio risk in the financial system. Risk optimization and fraud detection give fast and effective results. At the same time, based on the A type of quantum reinforcement learning method that can be used to simulate financial trading systems, and Understanding the impact of risk and uncertainty on financial forecasting models, on investment portfolios. Perform parallel simulations to quickly and effectively optimize trading strategies for fast and stable. The realization of the financial transaction forecasting system provides the possibility.

5 Conclusion

Quantum reinforcement learning is the intersection of reinforcement learning and quantum computing.
The research progress has been made in several aspects.
(i) the powerful computing power provided by the parallelism of quantum computing can be exploited to accelerate the reinforcement learning process. (1) the powerful computing power provided

by the parallelism of quantum computing to accelerate the reinforcement learning process;
(2) based on the properties of quantum mechanisms and the continuous enrichment of quantum (2) Based on the properties of quantum mechanism and the rich quantum algorithms, many researches have proposed new reinforcement learning methods.
(3) new quantum mechanics research methods based on the traditional reinforcement learning algorithms. Quantum reinforcement learning Although the research progress of quantum reinforcement learning is in the initial stage, the existing results have already The research progress of quantum reinforcement learning is still in its initial stage, but the existing results have already brought unlimited vision to many researchers.

Author Contributions

If you'd like to, you may include a section for author contributions as is done in many journals. This is optional and at the discretion of the authors.

Acknowledgments

Use unnumbered third level headings for the acknowledgments. All acknowledgments, including those to funding agencies, go at the end of the paper.

References

Valeria Cimini, Mauro Valeri, Emanuele Polino, Simone Piacentini, Francesco Ceccarelli, Giacomo Corrielli, Nicol Spagnolo, Roberto Osellame, and Fabio Sciarrino. Deep reinforcement learning for quantum multiparameter estimation. *arXiv preprint arXiv:2209.00671*, 2022.

Li Ding and Lee Spector. Evolutionary quantum architecture search for parametrized quantum circuits. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, pp. 2190–2195, 2022.

Daoyi Dong, Chunlin Chen, Hanxiong Li, and Tzyh-Jong Tarn. Quantum reinforcement learning. *Neural Computation*, 18:1207–1220, 2008.

Paolo Andrea Erdman, Alberto Rolandi, Paolo Abiuso, Mart Perarnau-Llobet, and Frank No. Pareto-optimal cycles for power, efficiency and fluctuations of quantum heat engines using reinforcement learning. *arXiv preprint arXiv:2207.13104*, 2022.

Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. *Deep learning*, volume 1. MIT Press, 2016.

Run-Hong He, Rui Wang, Shen-Shuang Nie, Jing Wu, Jia-Hui Zhang, and Zhao-Ming Wang. Deep reinforcement learning for universal quantum state preparation via dynamic pulse control. *EPJ Quantum Technology*, 8(1):29, 2021.

Irtaza Khalid, Carrie A Weidner, Edmond A Jonckheere, Sophie G Shermer, and Frank C Langbein. Statistically characterising robustness and fidelity of quantum controls and quantum control algorithms. *arXiv preprint arXiv:2207.07801*, 2022.

Lucas Lamata. Quantum reinforcement learning with quantum photonics. In *Photonics*, volume 8, pp. 33. MDPI, 2021.

Tobias Mller, Christoph Roch, Kyrill Schmid, and Philipp Altmann. Towards multi-agent reinforcement learning using quantum boltzmann machines. *arXiv preprint arXiv:2109.10900*, 2021.

Florian Neukart, David Von Dollen, Christian Seidel, and Gabriele Compostella. Quantum-enhanced reinforcement learning for finite-episode games with discrete state spaces. *Frontiers in physics*, 5:71, 2018.

Niels MP Neumann, Paolo BUL de Heer, Irina Chiscop, and Frank Phillipson. Multi-agent reinforcement learning using simulated quantum annealing. In *International Conference on Computational Science*, pp. 562–575. Springer, 2020.

Murphy Yuezhen Niu, Sergio Boixo, Vadim N Smelyanskiy, and Hartmut Neven. Universal quantum control through deep reinforcement learning. *npj Quantum Information*, 5(1):1–8, 2019.

Mara Laura Olivera-Atencio, Lucas Lamata, Manuel Morillo, and Jes Casado-Pascual. Quantum reinforcement learning in the presence of thermal dissipation. *arXiv preprint arXiv:2208.06450*, 2022.

Yash J Patel, Sofiene Jerbi, Thomas Bäck, and Vedran Dunjko. Reinforcement learning assisted recursive qaoa. *arXiv preprint arXiv:2207.06294*, 2022.

Simon Daniel Rei and Peter van Loock. Deep reinforcement learning for key distribution based on quantum repeaters. *arXiv preprint arXiv:2207.09930*, 2022.

Xavier Valcarce, Pavel Sekastki, Elie Gouzien, Alexey Melnikov, Jean-Daniel Bancal, and Nicolas Sangouard. Automated design of quantum optical experiments for device-independent quantum key distribution. *arXiv preprint arXiv:2209.06468*, 2022.

Won Joon Yun, Jihong Park, and Joongheon Kim. Quantum multi-agent meta reinforcement learning. *arXiv preprint arXiv:2208.11510*, 2022.

A Appendix

You may include other additional sections here.