QKSA
Quantum Knowledge Seeking Agent

Sarkar, Aritra; Al-Ars, Zaid; Bertels, Koen

DOI
10.1007/978-3-031-19907-3_37

Publication date
2023

Document Version
Final published version

Published in
Artificial General Intelligence - 15th International Conference, AGI 2022, Proceedings

QKSA: Quantum Knowledge Seeking Agent

Aritra Sarkar[1,2(✉)], Zaid Al-Ars[1], and Koen Bertels[2]

[1] Department of Quantum and Computer Engineering, Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Delft, The Netherlands
{a.sarkar-3,z.al-ars}@tudelft.nl
[2] QBee.eu, Leuven, Belgium
{aritra.sarkar,koen.bertels}@qbee.eu

Abstract. In this research, we extend the universal reinforcement learning agent models of artificial general intelligence to quantum environments. The utility function of a classical exploratory stochastic Knowledge Seeking Agent, KL-KSA, is generalized to distance measures from quantum information theory on density matrices. Quantum process tomography (QPT) algorithms form a tractable subset of programs for modeling environmental dynamics. The optimal QPT policy is selected based on a mutable cost function based on algorithmic complexity as well as computational resource complexity. The entire agent design is encapsulated in a self-replicating quine which mutates the cost function based on the predictive value of the optimal policy choosing scheme. Thus, multiple agents with pareto-optimal QPT policies evolve using genetic programming, mimicking the development of physical theories each with different resource trade-offs. This formal framework, termed Quantum Knowledge Seeking Agent (QKSA), is a resource-bounded participatory observer modification to the recently proposed algorithmic information-based reconstruction of quantum mechanics. A proof-of-concept is implemented and available as open-sourced software.

Keywords: Algorithmic information theory · Quantum computing · Reinforcement learning · Mutating quine

1 Introduction

The overwhelming success of deep learning over the last decade is encouraging the revival of research on artificial general intelligence (AGI) from various directions [7,9]. The most mathematically rigorous among these is universal artificial intelligence (UAI) [11]. The agent-environment paradigm of model-based reinforcement learning (RL), is best suited to mimic the interactive learning behavior of biological intelligence. UAI-based RL agents are concisely referred to as universal reinforcement learning (URL) agents. This research examines policies of modeling an unknown environment as the general task assigned to a URL agent.

B. Goertzel et al. (Eds.): AGI 2022, LNAI 13539, pp. 384–393, 2023.
https://doi.org/10.1007/978-3-031-19907-3_37

URL agents have been instrumental in proving asymptotic optimal behavior [20,21] in partially observable environments by merging theoretical concepts in universal automata, algorithmic information theory (AIT) [24] and decision theory. However, the dependence on AIT makes these agents generally uncomputable. While resource-bounded variants have been proposed, these models still remain intractable for real-world applications. Moreover, these resource bounds introduce arbitrary hyper-parameters. To address this issue, we utilize the proposal [37] of embedding RL agents within an evolutionary framework (Evo-RL) to guide the meta-learning for a specific application scenario. In this work, we propose the idea of a resource-bounded evolutionary URL (Evo-URL) for the first time. The framework implements the resource cost function as a genetic program encapsulated by a mutating quine. This work is prompted by the suggestion of UAI systems to eventually play the role of autonomous scientists by recursive self-improvement [35].

The properties of the environment is as crucial for AGI as that of the learning strategy. In its most general form, physical systems should include classical, quantum, and relativistic scenarios. This work addresses the first two cases by defining the environment as an unknown quantum process. The proposed agent uses quantum process tomography (QPT) strategies as a tractable predefined subset of programs for actively learning the environmental model. Limiting the evaluation to this subset of programs alleviates the exponential scaling of the space of programs, which limits UAI's applicability beyond toy models.

The proposed AGI framework, called *Quantum Knowledge Seeking Agent (QKSA)*, models classical and quantum dynamics by merging ideas from AIT, quantum information, constructor theory, and genetic programming. Following the artificial life (or, animat) path to intelligence, a population of classical agents undergoes open-ended evolution (OEE) to explore pareto-optimal ways of modeling the perceptions from a quantum environment. Similar to how AIXI-tl [12] is a resource-bounded active version of Solomonoff universal induction [39], QKSA is a resource-bounded participatory observer [13,44] framework to the recently proposed [26] algorithmic information-based reconstruction of quantum mechanics. QKSA can be applied for simulating and studying aspects of quantum information theory like control automation, multiple observers, course-graining, distance measures, resource complexity trade-offs, etc.

The rest of the article is organized as follows. Section 2 presents the four features of the QKSA model that distinguish it from other similar concepts and models. In Sect. 3 we present the formalization of QKSA's policy. Section 4 concludes the article with suggestive applications.

2 Framework Features

In this section, we present the four distinguishing features of the QKSA framework, (i) representations of general quantum environments, (ii) process tomography algorithms for modeling, (iii) computational resource-bounded algorithmic cost, and (iv) mutating meta-learning hyper-parameter embedded in a quine.

2.1 Representations of General Quantum Environments

The class of environments an agent can model define the bound of its applicability. Solomonoff's theory of universal inductive inference [39] forms the theoretical basis of UAI, and automated scientific modeling in general. In it, the environment is assumed to be computable by a universal Turing machine [42]. The hypothesis size (i.e., the Kolmogorov/algorithmic complexity [17]) is used to proportionally weigh (i.e., the Solomonoff/algorithmic probability) the environmental models for future predictions. The invariance theorem allows any universal automata/language to be used for estimating the hypothesis size, up to a constant overhead.

The active generalization of Solomonoff's induction using Bellman's optimality equation form the basis of URL agents. The agent and the environment interact in turns. At every time step, the agent supplies the environment with an action. The environment then performs some computation and returns a percept to the agent, and the procedure repeats. The environment is modeled as a partially observable Markov decision process. The canonical URL model is the AIXI model [11].

Knowledge Seeking Agents [28] replaces the extrinsic reward function in AIXI with a utility function defined as information gain of the model. Thus, this collapses the exploration-exploitation trade-off to simply exploration, allowing agents to explore the environment in a principled approach. The goal of these agents is to entirely explore their world optimally, form a model, and get a reward for reducing the entropy (uncertainty) in its model from the two components: uncertainty in the agent's beliefs and environmental noise. A particularly interesting case is the KL-KSA [29], which is robust to stochastic noise as the utility function is given as the Kullback-Leibler divergence.

While KL-KSA generalizes over arbitrary countable classes and priors, it cannot intrinsically interpret quantum information. This is because quantum information [27] is a generalization of classical probability theory to the complex domain. It allows richer representations and manipulations of information based on superposition, unitary evolution, interference, entanglement, and projective measurement. QKSA generalizes the probability distribution of KL-KSA to density matrices and the KL divergence to various distance measures on quantum processes.

A brief necessary background of these representations are presented here. Statistical ensembles of N pure quantum states $|\psi\rangle$ are described as a density matrix $\rho = \sum_{k=1}^{N} p_k |\psi\rangle\langle\psi|$, where the probabilities satisfy $0 < p_k \leq 1$ and $\sum_{k=1}^{N} p_k = 1$. A projective measurement of an observable M_m is given by the expectation value $Pr(m) = Tr(M_m \rho)$. Statistics of observable probabilities from quantum measurements can only estimate the density matrix instead of the state, thus fitting the QKSA use case. The unitary U evolution of closed quantum systems is denoted for pure states as, $|\psi'\rangle = U|\psi\rangle$ and for mixed states as $\rho' = U \rho U^{\dagger}$. More generally, a quantum process Φ that transforms a density matrix need not always be unitary. Given classical processes are often irreversible and include measurements, a quantum generalization includes unitary transforms (symmetry transformations of isolated systems), probabilistic logic, measurements and transient interactions with an environment. Thus, quantum processes formalize the time evolution of open quantum systems as linear quantum dynamical maps from the set of density matrices to itself. For a quantum system with an input state ρ_{in} of dimension $n \times n$ and an output state $\rho_{out} = \Phi(\rho_{in})$ of dimension $m \times m$, Φ is a linear superoperator mapping between the space of Hermitian matrices $\Phi : \mathcal{M}_{n\times n} \rightarrow \mathcal{M}_{m\times m}$. There are other equivalent representations of quantum processes like Choi matrix ρ_{Choi}, Kraus operators, Stinespring, Pauli basis Chi matrix χ, Lindbladian, etc. For instance, the Choi matrix ρ_{Choi} is the density matrix obtained after putting half of the maximally entangled state $|\Omega\rangle$ through the channel Φ, while doing nothing on the other half, i.e. if $\Lambda = \sum_{i,j} \frac{1}{2^n}|i\rangle\langle j| \otimes \Phi(|i\rangle\langle j|)$, the $\rho_{Choi} = \Lambda(|\Omega\rangle\langle\Omega|)$ The evolution of a density matrix with respect to the Choi-matrix is given by, $\rho_{out} = \Phi(\rho_{in}) = \mathrm{Tr}_1((\rho_{in}^T \otimes \mathrm{I})\rho_{choi}))$, where Tr_1 is the partial trace over subsystem 1. As a result of the Choi-Jamiolkowski isomorphism, the Choi matrix ρ_{Choi} characterizes the process Φ completely. This isomorphism forms the basis of the channel-state duality in quantum information.

Like classical probability distribution, there are many measures of quantum distances, each with its own application advantage. The QKSA framework allows the user to select a distance metric as part of the experimental setup. The current implementation provides the following distance metrics, Hamming distance, KL divergence, trace distance, Hilbert-Schmidt norm, and Bures distance (fidelity). Users can also define a custom distance measure. A future extension would provide diamond distance, Hellinger distance, quantum Kolmogorov complexity, quantum relative entropy, RÃl'nyi divergence, Bhattacharyya distance, and quantum complexity action [10].

2.2 Process Tomography Algorithms for Modeling

In canonical UAI formalism, the programs are drawn randomly from a prefix code for a universal automata. However, the space of programs grows exponentially, limiting its applicability beyond simple grid-world exploration and games. We restrict this space to a constant number of predefined algorithms provided to the framework. This pragmatic design feature allows us to implement interpretable and tractable URLs.

Characterization of quantum dynamical systems is a fundamental problem in quantum information science. The procedures that achieve this goal are called quantum process tomography. Some examples of these well-developed techniques are: standard QPT [6], entanglement-assisted QPT, direct characterization of quantum dynamics, compressed-sensing QPT, permutation-invariant tomography, self-guided QPT and shadow QPT [23]. Each QPT technique has a different experimental setup and computational resource requirements. These QPT algorithms form the space of programs that QKSA evaluates as candidates for modeling the environment. Intuitively, a QPT algorithm will better predict a quantum environment than a random program. Thus, it allows us to apply the tools of AIT in a practical setting where available expert knowledge can be embedded within the agent. Given computational resource limitations, QKSA is designed to automatically discover the optimal strategy in the available pool of QPT algorithms.

Recent publications study learning in a quantum environment, for e.g., process learning with restricted Boltzmann machines [41], RL based optimizer for variational algorithms [31,43], automated design of experimental quantum optics [19], and, projective simulation (PS) [4,8]. Despite the similarities with QKSA (especially of PS), these approaches are not based on URL. Also unlike [5,33], we do not assume the quantum computational capability of the agent for estimating the AIT metric, in line with the conventional qualia of human intelligence. QKSA is an RL framework to study quantum information and computation via the lens of AIT.

2.3 Computational Resource-Bounded Algorithmic Cost

The algorithmic probability of a candidate model/program is used as a weight for choosing an action and thereby the reward in UAI. However, this also makes such models impractical due to the uncomputability of algorithmic information metrics like algorithmic probability and algorithmic complexity. Being asymptotically computable, URL is thus not a pragmatic algorithmic solution to general RL, and must be simplified in any implementation. In principle, there are an infinite number of programs that can be candidate models of the environment. Also, while evaluating, the programs can enter infinite loops. To circumvent these two issues, modifications are proposed on the agents, like AIXI-tl [12], MC-AIXI$_{(\text{FAC-CTW})}$ and UCAI [14]. These bound the program length and runtime per step to explore a subspace of promising hypotheses that models the interactive behavior registered till the current time step. There arise three issues with this approach:

1. The bounds introduce heuristic hyper-parameters that depend on the available computational resources. Thus, selecting an appropriate value to apply the model for a given use case becomes difficult.
2. The bounds sharply cut off models beyond the specification while keeping the weight for the models within the specification unaffected. So a model that performs well but lies beyond the defined bound may be unreachable.
3. It is possible to trade off these resource bounds with other computational resources, like additional memory.

Using the QKSA platform, it is possible to investigate these issues. In the framework we propose five computational resources, together we call the *LEAST* metric, as an acronym for (program) length, (compute) energy, approximation, (work memory) space and (run) time. Similar algorithmic observables have been suggested in [1]. We provide estimation techniques of the LEAST metric in our implementation, based on state-of-the-art algorithmic information research and general practices in computer engineering. The estimation technique, however can easily be redefined by the user. The estimated metric is used in a two-fold way. Firstly, it is used to qualify the hypothesis for consideration based on upper bounds for each of the five computational resources individually. This is dictated by the available computational resource of the substrate the implementation is executed on, and is similar to the resource-bounded UAI models [12]. These bounds can be included in the list of evolving hyper-parameters to allow QKSA to mutate and adjust autonomously to the available computational resource. After that, the metrics for valid hypotheses are fed to a cost function (a genetic program) that outputs a single positive real value which is used as the weight for the hypothesis (instead of only the length, as in algorithmic probability). We call this the *least action* as a parameterization to optimize the Lagrangian dynamics of computation.

2.4 Mutating Meta-learning Hyper-parameter Embedded in a Quine

There is no unifying cost function that can serve as a metric to trade-off bounds on resources (like space, time, approximation), and possibly cannot exist [13, 30]. In fact, this depends closely on the policy of the agent. For example, a physicist might use simpler Newtonian mechanics instead of complex relativistic mechanics for modeling where the approximations are acceptable. Thus, instead of a single metric, a pareto-optimal frontier on the LEAST metrics maps to models and algorithms that can be used to predict the environment dynamics.

Various research has explored this frontier, considering a few of the LEAST metrics. Some examples are, Levin complexity [22], Bennett's logical depth [2] and pebbling game [3], Schmidhuber's speed prior [36], Wolpert's statistical thermodynamics of Turing machines [16], Zenil's block decomposition method [38], and look-up tables. These resource-bounded metrics are not immune to the no-free-lunch theorems [45] and adversarial cases of environments.

QKSA holistically (yet, subjectively) explores these trade-offs by dynamically adapting the cost function to the environment. The five estimates of the LEAST metrics are given as input to a cost function. We employ evolutionary computation, a population-based trial and error problem solving technique for meta-heuristic or stochastic optimization. More specifically, we use genetic programming (GP) [18]. The cost function itself is a gene represented as a program tree with the leaf nodes as the metrics or constants, and the internal nodes are from a set of essential arithmetic functions. Once QKSA learns an environment optimally or completely fails to do so (i.e. when the learning rate stabilizes), the QKSA self-replicates. The child QKSA has the same source code as the parent, except for a mutation on the cost function that modifies the weights and structure embedded via the cost function gene. Thus the open-ended evolution of the pareto-optimal manifold converges on QPT algorithms which fits well in the available computational resource. The parent QKSA perishes if the prediction of the model fails persistently (i.e., when the rate stabilizes as the strategy fails to learn) or continues to correctly predict environmental interaction and can be inspected to obtain the cost function. Thus, a single QKSA may not have an objective optimal resource trade-off for a static environment, but the population is expected to converge to an optimal policy even for a dynamic environment (provided the dynamics are slower than the learning rate).

The entire agent framework described so far is embedded within a quine. Quines are self-replicating programs that are the software embodiment of constructors, an idea foundational to artificial life [40] and physical theories alike [25]. The Kleene recursion theorem [15] allows any program to be modified such that it (a) replicates its source code, (b) executes an orthogonal payload that serves the same purpose as the original non-quine version. This embellishment on the evolutionary cost function qualifies QKSA as a recursive self-improving system.

3 QKSA Formalism

In this section, we start from the formalism of AIXI and elucidate the changes described in the previous section. For brevity, we omit the mathematical reasoning behind the AIXI, which can be found in [11]. The canonical expectimax equation in UAI is used by the agent to rationalize the choice of a particular action at the current time step. For AIXI [11], it takes the form:

$$a_t = \arg \lim_{m\to\infty} \max_{a_t\in\mathcal{A}} \sum_{e_t\in\mathcal{E}} \ldots \max_{a_m\in\mathcal{A}} \sum_{e_m\in\mathcal{E}} \sum_{k=t}^{m} \gamma_k r_k \sum_{p:U(p;a_{<k})=e_{<k}} 2^{-l(p)}$$

where, a_t is the action at time step t from the action space $\mathcal{A}$, e_k is a perception from the percept space $\mathcal{E}$ defined over the time step span from t to m, γ is a reward discount function, U is an universal automata, p is a program that forms the model/hypothesis for the environment, r_k is the reward signal from the environment, and $l(p)$ is the length of the program p. In the case of KL-KSA [29], the reward for AIXI is generalized to the utility given by, $u_k = u(e_k|ae_{<k}a_k) = Ent(w_\nu|ae_{<k+1}) - Ent(w_\nu|ae_{<k}a_k)$, where $Ent()$ refers to the entropy function and w_ν refers to the agent's credence in the percept distribution ν representing the environment.

The first change is to restrict the search space of programs p to quantum process tomography algorithms, denoted as p_{qpt}. It is important that the QPT algorithm reconstructs and outputs a process representation ρ_k instead of the prediction of the subsequent perception. $\lambda^{e'_k} \in \{0,1\}$ is the probability of the quantum state collapsing to the prediction e'_k made at time step $t-1$. This modification is imperative due to the stochastic nature of individual quantum measurements and the calculation of the utility.

The second change is to replace the length estimate of the $2^{-l(p)}$ factor from the algorithmic probability with the estimate of the evolving cost function c_{est}. The cost function is denoted by c_{least}, i.e. $c_{est} = c_{least}(p_{qpt})$. Thus, the learning part of the equation is:

$$a_t^{QKSA} = \arg \lim_{m\to\infty} \max_{a_t\in\mathcal{A}} \sum_{e'_t\in\mathcal{E}} \lambda^{e'_t} \ldots \max_{a_m\in\mathcal{A}} \sum_{e'_m\in\mathcal{E}} \lambda^{e'_m} \sum_{k=t}^{m} \gamma_k u'_k \sum_{\substack{p_{qpt}:U(p_{qpt};h_k)=\rho_k \\ p_{qpt}:U(p_{qpt};\rho_k;a_k;e'_k)=\lambda^{e'_k}}} 2^{-c_{least}(p_{qpt})}$$

The third change is to define the utility function as a quantum distance measure on the space of quantum processes ρ (defined as the density matrix in the Choi process matrix representation). A higher predicted utility indicates that the current estimate of the quantum process will be updated more significantly based on the perception, thus, a potential knowledge gain for choosing that action.

$$\boxed{u'_t = \Delta(\rho'_{t+1}, \rho_t) = \Delta(U(p_{qpt}; h_t; a_t; e'_t), U(p_{qpt}; h_t))}$$

A detailed description of the QKSA framework and policy is provided in [32]. A full proof-of-concept of the discussed QKSA framework is implemented on Python and Qiskit. It is available as open-source software at the following link: https://github.com/Advanced-Research-Centre/QKSA.

4 Conclusion

In this article, we extended the formalism of UAI to quantum environments by generalizing the KL-KSA to a quantum knowledge seeking agent (QKSA). The environment within the reinforcement learning setup is defined by an unknown quantum circuit that the agent attempts to model using quantum process tomography. A quantum environment prevents the exact prediction of perceptions (as used by AIXI), and a single probability distribution of perception based on the set of actions (as used by KL-KSA). The subjective model is conditioned on the chosen action and is thus represented by the more general density matrix formalism. Any quantum process can be represented as a Choi density matrix, which forms a model of the environmental dynamics. To circumvent the uncomputability of UAI models, we propose to evaluate the algorithmic cost within a set of user-provided programs. This consideration makes the framework more tractable and interpretable. Also, the resource restrictions used in computable UAI models are arbitrary. In our model, these resource bounds are interdependent hyper-parameters whose value and trade-off relations are optimized using genetic programming. Thus, this allows open-ended evolution of the agents for dynamic environments. Each agent can self-replicate as a quine and thus is a recursive self-improving intelligence model.

As part of ongoing research [34], we are applying the QKSA framework as described in this article to study course-graining in multi-observer scenarios and quantum uncomplexity resources. It also has near term applicability in optimizing NISQ era hybrid variational quantum algorithms.

References

1. Baez, J., Stay, M.: Algorithmic thermodynamics. Math. Struct. Comput. Sci. **22**(5), 771–787 (2012)
2. Bennett, C.H.: Logical depth and physical complexity. Citeseer (1988)
3. Bennett, C.H.: Time/space trade-offs for reversible computation. SIAM J. Comput. **18**(4), 766–776 (1989)

4. Briegel, H.J., De las Cuevas, G.: Projective simulation for artificial intelligence. Sci. Rep. **2**(1), 1–16 (2012)
5. Catt, E., Hutter, M.: A gentle introduction to quantum computing algorithms with applications to universal prediction. arXiv preprint arXiv:2005.03137 (2020)
6. Chuang, I.L., Nielsen, M.A.: Prescription for experimental determination of the dynamics of a quantum black box. J. Mod. Opt. **44**(11–12), 2455–2467 (1997)
7. Domingos, P.: The master algorithm: How the quest for the ultimate learning machine will remake our world. In: Basic Books (2015)
8. Dunjko, V., Taylor, J.M., Briegel, H.J.: Advances in quantum reinforcement learning. In: 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 282–287 (2017)
9. Goertzel, B.: Artificial general intelligence: concept, state of the art, and future prospects. J. Artif. Gen. Intell. **5**(1), 1 (2014)
10. Halpern, N.Y., Kothakonda, N.B., Haferkamp, J., Munson, A., Eisert, J., Faist, P.: Resource theory of quantum uncomplexity. arXiv preprint arXiv:2110.11371 (2021)
11. Hutter, M.: Universal Artificial Intellegence. TTCSAES, Springer, Heidelberg (2005). https://doi.org/10.1007/b138233
12. Hutter, M.: Universal algorithmic intelligence: a mathematical top → down approach. In: Artificial General Intelligence, pp. 227–290. Springer (2007). https://doi.org/10.1007/978-3-540-68677-4_8
13. Hutter, M.: A complete theory of everything (will be subjective). Algorithms **3**(4), 329–350 (2010)
14. Katayama, S.: Computable variants of aixi which are more powerful than aixitl. J. Artif. Gen. Intell. **10**(1), 1–23 (2019)
15. Kleene, S.C.: Introduction to metamathematics. North-Holland Publishing Co. (1952)
16. Kolchinsky, A., Wolpert, D.H.: Thermodynamic costs of turing machines. Phys. Rev. Res. **2**(3), 033312 (2020)
17. Kolmogorov, A.N.: Three approaches to the quantitative definition of information. Int. J. Comput. Math. **2**(1–4), 157–168 (1968)
18. Koza, J.R., Koza, J.R.: Genetic programming: on the programming of computers by means of natural selection, vol. 1. MIT press (1992)
19. Krenn, M., Malik, M., Fickler, R., Lapkiewicz, R., Zeilinger, A.: Automated search for new quantum experiments. Phys. Rev. Lett. **116**(9), 090405 (2016)
20. Leike, J., Hutter, M.: Bad universal priors and notions of optimality. In: Conference on Learning Theory, pp. 1244–1259. PMLR (2015)
21. Leike, J., Lattimore, T., Orseau, L., Hutter, M.: Thompson sampling is asymptotically optimal in general environments. arXiv preprint arXiv:1602.07905 (2016)
22. Levin, L.A.: Universal sequential search problems. Probl. Peredachi Informatsii **9**(3), 115–116 (1973)
23. Levy, R., Luo, D., Clark, B.K.: Classical shadows for quantum process tomography on near-term quantum computers. arXiv preprint arXiv:2110.02965 (2021)
24. Li, M., Vitányi, P.: An Introduction to Kolmogorov Complexity and Its Applications. TCS, Springer, New York (2008). https://doi.org/10.1007/978-0-387-49820-1
25. Marletto, C.: Constructor theory of life. J. R. Soc. Interface **12**(104), 20141226 (2015)
26. Mueller, M.P.: Law without law: from observer states to physics via algorithmic information theory. Quantum **4**, 301 (2020)
27. Nielsen, M.A., Chuang, I.: Quantum computation and quantum information (2002)

28. Orseau, L.: Universal knowledge-seeking agents. Theor. Comput. Sci. **519**, 127–139 (2014)
29. Orseau, L., Lattimore, T., Hutter, M.: Universal knowledge-seeking agents for stochastic environments. In: Jain, S., Munos, R., Stephan, F., Zeugmann, T. (eds.) ALT 2013. LNCS (LNAI), vol. 8139, pp. 158–172. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40935-6_12
30. Poland, K., Beer, K., Osborne, T.J.: No free lunch for quantum machine learning. arXiv preprint arXiv:2003.14103 (2020)
31. Rivera-Dean, J., Huembeli, P., Acín, A., Bowles, J.: Avoiding local minima in variational quantum algorithms with neural networks. arXiv preprint arXiv:2104.02955 (2021)
32. Sarkar, A.: Applications of Quantum Computation and Algorithmic Information: for Causal Modeling in Genomics and Reinforcement Learning. Ph. D. thesis, Delft University of Technology (2022)
33. Sarkar, A., Al-Ars, Z., Bertels, K.: Estimating algorithmic information using quantum computing for genomics applications. Appl. Sci. **11**(6), 2696 (2021)
34. Sarkar, A., Al-Ars, Z., Gandhi, H., Bertels, K.: Qksa: quantum knowledge seeking agent-resource-optimized reinforcement learning using quantum process tomography. arXiv preprint arXiv:2112.03643 (2021)
35. Schmidhuber, J.: On learning how to learn learning strategies. Fakultät für Informatik, Technische Universität München, Technical report (1995)
36. Schmidhuber, J.: The speed prior: a new simplicity measure yielding near-optimal computable predictions. In: Kivinen, J., Sloan, R.H. (eds.) COLT 2002. LNCS (LNAI), vol. 2375, pp. 216–228. Springer, Heidelberg (2002). https://doi.org/10.1007/3-540-45435-7_15
37. Silver, D., Singh, S., Precup, D., Sutton, R.S.: Reward is enough. Artif. Intell. **299**, 103535 (2021)
38. Soler-Toscano, F., Zenil, H., Delahaye, J.P., Gauvrit, N.: Calculating kolmogorov complexity from the output frequency distributions of small turing machines. PloS One **9**(5) (2014)
39. Solomonoff, R.J.: A formal theory of inductive inference. Part i. Inf. Control **7**(1), 1–22 (1964)
40. Strannegård, C., Svangård, N., Lindström, D., Bach, J., Steunebrink, B.: The animat path to artificial general intelligence. In: Proceedings of IJCAI-17 Workshop on Architectures for Generality & Autonomy (2017)
41. Torlai, G., Mazzola, G., Carrasquilla, J., Troyer, M., Melko, R., Carleo, G.: Neural-network quantum state tomography. Nat. Phys. **14**(5), 447–450 (2018). https://doi.org/10.1038/s41567-018-0048-5
42. Turing, A.M., et al.: On computable numbers, with an application to the entscheidungsproblem. J. Math. **58**(345–363), 5 (1936)
43. Wauters, M.M., Panizon, E., Mbeng, G.B., Santoro, G.E.: Reinforcement-learning-assisted quantum optimization. Phys. Rev. Res. **2**(3), 033446 (2020)
44. Wheeler, J.A.: At home in the universe. American Institute of Physics (1996)
45. Wolpert, D.H.: The lack of a priori distinctions between learning algorithms. Neural Comput. **8**(7), 1341–1390 (1996)