THE REINFORCEMENT LEARNING FOR AUTONOMOUS ACCELERATORS COLLABORATION

A. Santamaria Garcia *, L. Scomparin, C. Xu
Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
S. Hirlaender, S. Pochaba, Paris Lodron Universität Salzburg, Salzburg, Austria
A. Eichler, J. Kaiser, Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany
M. Schenk, CERN, Geneva, Switzerland

Abstract

Reinforcement learning (RL) is a unique learning paradigm that is particularly well-suited to tackle complex control tasks, can deal with delayed consequences, and can learn from experience without an explicit model of the dynamics of the problem. These properties make RL methods extremely promising for applications in particle accelerators, where the dynamically evolving conditions of both the particle beam and the accelerator systems must be constantly considered. While the time to work on RL is now particularly favorable thanks to the availability of high-level programming libraries and resources, its implementation in particle accelerators is not trivial and requires further consideration. In this context, the Reinforcement Learning for Autonomous Accelerators (RL4AA) international collaboration was established to consolidate existing knowledge, share experiences and ideas, and collaborate on accelerator-specific solutions that leverage recent advances in RL. Here we report on two collaboration workshops, RL4AA'23 and RL4AA'24, which took place in February 2023 at the Karlsruhe Institute of Technology and in February 2024 at the Paris-Lodron Universität Salzburg.

THE MISSION OF THE RL4AA COLLABORATION

Machine learning (ML) has significantly increased in popularity over recent years within the particle accelerator community. However, RL remains relatively unknown, as indicated by the low number of related publications, shown in Fig. 4. This is partly due to the complex design of these algorithms and the substantial time required to understand, engineer, and deploy them.

Given the increasingly stringent beam parameters and performance metrics in frontier particle accelerators, the precise control and real-time optimization of beam parameters in a dynamically changing environment will be crucial for the efficient operation of future facilities [1]. In this context, the primary aim of the RL4AA collaboration is to consolidate RL efforts in the particle accelerator community by establishing a unified platform that:

1. Connects RL enthusiasts within the particle accelerator community to foster collaborative projects across institutions and facilitates interaction with other RL experts for the exchange of ideas.
2. Educates on both fundamental and advanced RL concepts and demonstrates practical applications in accelerators, offering valuable resources such as programming tutorials, lectures, and educational events.
3. Facilitates discussions on the challenges of developing and deploying RL algorithms in particle accelerators and other large-scale infrastructures.

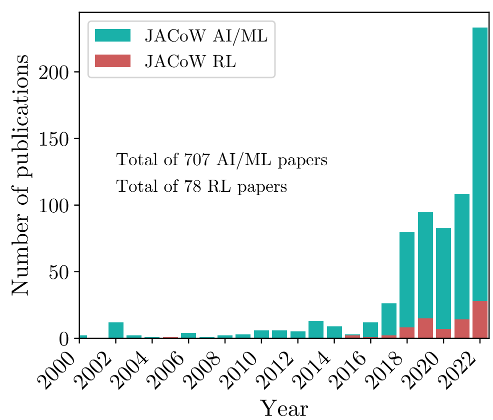

Figure 1: Number of publications featuring the terms "artificial intelligence" or "machine learning" versus "reinforcement learning" in their abstracts, sourced from the JACoW database.

These objectives are realized through annual workshops, details of which are discussed in the following sections. The communication and dissemination platforms provided by the RL4AA collaboration include a website [2] that aggregates news, relevant links, and RL-related publications; a GitHub organization [3] that compiles the RL programming tutorials offered by the collaboration; a Discord server [4] for general announcements, meetings, and broader community engagement; and a YouTube channel with recorded talks [5].

CHALLENGES IN RL

The challenges of applying RL algorithms to particle accelerators are consistent with those encountered during their deployment in any real-world system. The most relevant challenges include:

- **Partial observability**: This refers to situations where the state is not directly observable and must be inferred

* andrea.santamaria@kit.edu

from limited and/or noisy observations, a common scenario in real-world environments. For example, during accelerator operation, complete phase space information is unavailable, relying on partial observations or derived quantities such as beam position readings and synchrotron radiation measurements. Some measurements are destructive and can disturb the current state, making them sparse and asynchronous compared to non-destructive ones. With partial observability, the agent must effectively infer the necessary information from incomplete data to make optimal decisions, which can make learning stable policies more challenging and computationally intensive.

- **Sample complexity**: This term refers to the number of interactions with the environment required to achieve a certain level of performance during the decision-making process. Improving sample efficiency is crucial when applying RL methods to particle accelerators, as the cost of beam time for gathering real-world interactions can be prohibitive. This depends on the accelerator's repetition rate and the diagnostics available. Reducing sample complexity can decrease the training costs of RL algorithms, enhance their scalability and accessibility, facilitate real-time learning, and even increase their safety.

- **Safety**: In RL, safety ensures that measures are in place to prevent an agent from taking dangerous actions during both training and deployment. For accelerator physics applications, safety involves protecting against injury or loss of life, damage to machine equipment, and loss of beam time. The level of importance placed on safety heavily depends on the accelerator facility's purpose, the beam energy and particle type, and the complexity and extent of the machine interlocking system. Although soft safety using negative rewards has been demonstrated in other fields, such as fusion reactor control [6], hard safety remains an unsolved challenge in the RL community at large, including within the RL for accelerator community. Ensuring safety during training is particularly challenging, making training on certain particle accelerators unfeasible, even with available beam time and sample-efficient RL algorithms.

- **Robustness**: This refers to the ability of an RL algorithm to perform effectively across a variety of environmental variations that were not specifically accounted for during the training phase. This is particularly important when transferring an agent trained on simulations to the real world (sim2real), in the presence of parameter drifts, or more generally, in non-stationary problems. It is closely related to the concept of generalization.

- **Generalization**: This refers to the ability of a trained agent to perform effectively in an environment other than its training environment. Developing generalizable or transferable RL agents is especially valuable, as particle accelerators share common design principles and control tasks.

The RL4AA collaboration's founders actively work on these topics, for example, by developing high-speed, differentiable optics simulations for faster training [7], exploring domain randomization for lattice-agnostic algorithms [8], implementing novel meta-RL solutions [9], designing systems for online training and control on hardware [10], and comparing RL to other ML solutions [11]. Other challenges such as algorithmic stability, theoretical guarantees, and hyperparameter tuning are also relevant and actively considered by the community.

FIRST WORKSHOP: RL4AA'23

The RL4AA collaboration was officially launched with the first RL4AA workshop, which was held at the Karlsruhe Institute of Technology (KIT) on the 20^{th}-21^{st} February 2023 [12]. With 31 registered participants, the event targeted mostly scientists in Germany and Switzerland and focused on connecting the attendees through introductory speed talks about their work in RL and through targeted discussion sessions. These discussion sessions were split in four groups to address the third goal of the RL4AA collaboration in a structured manner, namely:

- **Community**: group to discuss the origins of RL, have a broader perspective on its evolution, and understand the current trends.

- **Modeling, methods, and limitations**: group to discuss the origins and mathematical meaning behind RL problems and methods so that we can use these tools efficiently to solve relevant problems (if solvable at all).

- **Challenges I & II**: groups to address the particular challenges that arise in applying RL to real particle accelerators.

Figure 2: The RL4AA'23 workshop participants on the bridge of the Karlsruhe Research Accelerator (KARA) storage ring at the KIT during the accelerator facility tour.

RL4AA'23 tutorial

The programming tutorial of the RL4AA'23 workshop presented a real tuning task in the Accelerator Research Experiment at SINBAD (ARES) particle accelerator at DESY [13]. The task is performed in a particular section of the accelerator, composed of three quadrupole focusing magnets and two corrector magnets, as shown in Fig. 3. The goal of the tuning task is to adjust the field strength of the quadrupole magnets and the steering angles of the corrector magnets to achieve a target beam chosen by a human operator, defined by its transverse size and position on a diagnostic screen. The tutorial focuses on the importance of reward definition, and walks the user through a series of rewards and their effects on the agent's actions. It uses the high-speed, differentiable optics simulations code Cheetah [7] for faster simulation-based training, and is based on current research, where the algorithm was deployed in the real accelerator [11].

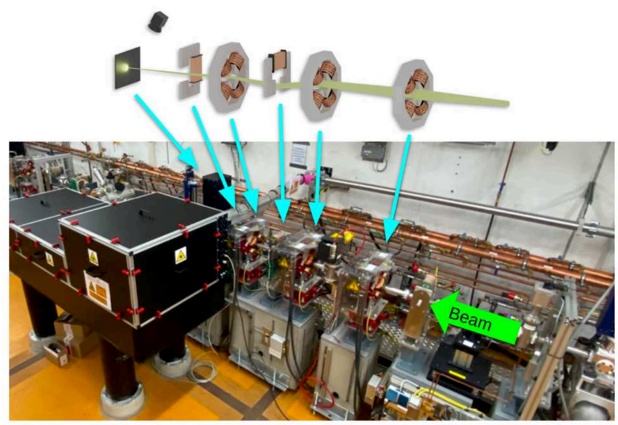

Figure 3: Relevant section of the ARES accelerator used in the RL4AA'23 workshop tutorial.

SECOND WORKSHOP: RL4AA'24

The second RL4AA workshop was held at the Paris Lodron Universität Salzburg on the 5th-7th February 2024 [14]. With 56 registered participants, the event almost doubled in size with respect to the previous year, showing the consolidation of this event. Two keynote speakers outside of the field of accelerators were invited to share their valuable expertise on deploying RL algorithms in real-world systems [15, 16]. The workshop also hosted facility overview talks, contributed talks, student talks, and a poster session. The tutorial was preceeded by an RL crash course and a dedicated lecture, and the event ended with an open discussion about the community with the participants. All talks were recorded and published in the RL4AA YouTube channel [5].

RL4AA'24 tutorial

The programming tutorial of the RL4AA'24 workshop targeted advanced RL users and dealt with a tuning task at the Advanced Proton Driven Plasma Wakefield Acceleration Experiment (AWAKE) beamline at CERN [17]. The goal of

Figure 4: RL4AA'24 workshop participants at the Paris Lodron Universität Salzburg.

the tuning task is minimizing the distance between an initial beam trajectory and a target beam trajectory, as described by the readings of 10 beam position monitors (BPMs), by using the available corrector magnets. The tutorial first shows how to solve this task with a proximal policy optimization (PPO) agent, and then with the model agnostic meta-learning (MAML) algorithm. Meta-learning, also known as "learning to learn", aims to improve the learning process itself by training a model on a variety of tasks so that it can learn new tasks more efficiently. In the tutorial 8 different tasks were used to update a meta policy, where the different tasks corresponded to different lattice configurations (i.e. quadrupole strengths). The tutorial shows that by using meta-learning methods, the RL algorithm can converge in only a few steps. A final part of the tutorial showed an example of model-based reinforcement learning (MBRL), which constructs an internal model of the environment to simulate interactions. This method enhances sample efficiency by reducing the need for direct system interaction. This tutorial is based on current research, where the algorithm was deployed in the real accelerator [18]. More information can be found in [9].

OUTLOOK

The RL4AA collaboration was established in 2023 and has organized two workshops to date. The next RL4AA'25 workshop will be held at DESY in early spring 2025. With these efforts we hope to bring RL to the accelerator community and to develop novel solutions for current and future accelerators.

ACKNOWLEDGEMENTS

A. Santamaria Garcia acknowledges the support of the Institute of Beam Physics and Technology (IBPT) of KIT in the organization of the RL4AA'23 workshop. A. Eichler and J. Kaiser acknowledge support from DESY, Germany, a member of the Helmholtz Association HGF.

REFERENCES

[1] A. Santamaria Garcia *et al.*, "How Can Machine Learning Help Future Light Sources?" *JACoW*, vol. FLS2023, TH3D3, 2024. doi:10.18429/JACoW-FLS2023-TH3D3

[2] *Reinforcement Learning for Autonomous Accelerators community (RL4AA)*. https://rl4aa.github.io/

[3] *Reinforcement Learning for Autonomous Accelerators Github Organization*. https://github.com/RL4AA

[4] *Reinforcement Learning for Autonomous Accelerators Discord server*. https://discord.gg/QtBMqsjWH2

[5] *Reinforcement Learning for Autonomous Accelerators YouTube channel*. www.youtube.com/@RL4AACollaboration

[6] J. Degrave *et al.*, "Magnetic control of tokamak plasmas through deep reinforcement learning," *Nature*, vol. 602, pp. 414–419, 7897 2022. doi:10.1038/s41586-021-04301-9

[7] J. Kaiser, C. Xu, A. Eichler, and A. Santamaria Garcia, "Cheetah: Bridging the Gap Between Machine Learning and Particle Accelerator Physics with High-Speed, Differentiable Simulations," 2024. doi:10.48550/arXiv.2401.05815

[8] C. Xu *et al.*, "Beam trajectory control with lattice-agnostic reinforcement learning," English, in *Proc. IPAC'23*, Venice, Italy, 2023, pp. 4487–4490. doi:10.18429/JACoW-IPAC2023-THPL029

[9] S. Hirlaender *et al.*, "Towards few-shot reinforcement learning in particle accelerator control," in *these proceedings*, Nashville, United States.

[10] L. Scomparin *et al.*, "Preliminary results on the reinforcement learning-based control of the microbunching instability," in *these proceedings*, Nashville, United States.

[11] J. Kaiser *et al.*, "Learning to do or learning while doing: Reinforcement learning and bayesian optimisation for online continuous tuning," 2023. doi:10.48550/arXiv.2306.03739

[12] *1st collaboration workshop on Reinforcement Learning for Autonomous Accelerators (RL4AA'23)*, 2023. https://indico.scc.kit.edu/event/3280/

[13] C. Xu, A. Santamaria Garcia, and J. Kaiser, *Tutorial on Applying Reinforcement Learning to the Particle Accelerator ARES*, version v1.0.1, 2024. doi:10.5281/zenodo.10869763

[14] *2nd collaboration workshop on Reinforcement Learning for Autonomous Accelerators (RL4AA'24)*, 2024. https://indico.scc.kit.edu/event/3746/

[15] A. Raffin, *Designing and Running Real-World RL Experiments*, 2023. https://indico.scc.kit.edu/event/3746/contributions/15475/

[16] F. Berkenkamp, *Towards real-world RL*, 2023. https://indico.scc.kit.edu/event/3746/contributions/16017/

[17] S. Hirlaender, J. Kaiser, C. Xu, and A. Santamaria Garcia, *Tutorial on Meta-Reinforcement Learning and GP-MPC at the RL4AA'24 Workshop*, version v1.0.2, 2024. doi:10.5281/zenodo.10887397

[18] S. Hirlaender, L. Lamminger, G. Zevi Della Porta, and V. Kain, "Ultra fast reinforcement learning demonstrated at CERN AWAKE," *JACoW*, vol. IPAC2023, THPL038, 2023. doi:10.18429/JACoW-IPAC2023-THPL038