Research Foundations Underlying RL Core's Technology

RL Core's team is composed of internationally-recognized leaders in reinforcement learning. The Reinforcement Learning and AI (RLAI) lab at the University of Alberta is one of the top reinforcement learning centers in the world, with more than 10 professors focused on reinforcement learning research. Founders Martha White and Adam White run large research groups investigating fundamental algorithm development in reinforcement learning, several of which have application to real-world systems.

RL Core's research team has particular expertise in **continual reinforcement learning**: learning in deployment, on the facility, without access to a simulator or digital twin. Most research in reinforcement learning relies on simulators, which results in solutions that are not well suited for real-world deployment. Naively applying these approaches would be ineffective because simulators allow practitioners to test options and interactions equivalent to hundreds of years of data. We do not have this luxury when learning in deployment, on a system, in industrial applications.

When moving RL from the lab to industrial application, the critical issues to address are:
- Learning efficiently from limited interaction (learning in weeks rather than years)
- Guaranteeing safety and stability
- Handling (nuisance) hyperparameters in the algorithms

Combining reinforcement learning with powerful neural networks–called **deep reinforcement learning**— helps to address these challenges and is a surprisingly recent development. The rise of deep reinforcement learning can be traced to around 2015. There has been significant progress in the decade since, though many open questions remain.

Efficient learning

Reusing data effectively

Martha White introduced a seminal algorithm for learning decision-making policies using historical data, which has since become the foundation for many deep reinforcement learning algorithms. Martha's work developed an actor-critic algorithm to leverage previously generated data in the off-policy learning setting, enabling agents to learn with minimal interaction.

Martha's lab has a deep focus on sound off-policy actor-critic algorithms, including important follow-up work highlighting and fixing soundness issues with this basic algorithm:

Seminal paper: https://dl.acm.org/doi/10.5555/3042573.3042600
Follow-up theoretical work: https://www.jmlr.org/papers/volume24/21-1350/21-1350.pdf

A new foundation for predictive modeling and long-horizon predictions

Adam's research develops on the idea that algorithms from reinforcement learning (RL) can be used to holistically represent the knowledge of an AI system. More practically, the seminal paper established that RL methods provide a new way of framing and learning long-horizon predictions about complex systems, like those commonly found in industrial control.

Adam's PhD thesis significantly expanded and developed these ideas. The work has formed the basis for dozens of distinct lines of work in reinforcement learning.

- Seminal Horde paper from Adam: https://dl.acm.org/doi/10.5555/2031678.2031726
- Adam's PhD thesis: https://sites.ualberta.ca/~amw8/phd.pdf
- Adam's Follow up work showing how the approach can be scaled to thousands of predictions updated at faster than 1Hz: https://dl.acm.org/doi/abs/10.1177/1059712313511648

Adam and Martha later applied this predictive modelling approach to real-world drinking water treatment data. This paper was based on work with Drayton Valley showing RL methods can outperform classical modelling approaches using noisy, partially observable, high-dimensional data from a real water treatment plant:
https://link.springer.com/article/10.1007/s10994-023-06413-x

Guaranteeing stability

Martha, Adam and Andy have a series of contributions focused on stability in algorithms and rigorous evaluation of their limitations, which is critical to effectively and safely utilizing reinforcement learning in real-world industrial settings. They introduced a sequence of work around more stable reinforcement learning algorithms, which culminated in a definitive journal paper:
https://www.jmlr.org/papers/volume23/21-037/21-037.pdf
and an extension to even more robust losses:
https://dl.acm.org/doi/10.1109/TPAMI.2022.3213503

Martha, Adam, and Andy wrote a comprehensive document on how to effectively run experiments by bringing together years of experience in empirical work in reinforcement learning:
https://jmlr.org/papers/volume25/23-0183/23-0183.pdf

Handling hyperparameters

Calibration models for safe off-site agent specialization

The majority of academic work in reinforcement learning (RL) relies on extensively tuning algorithms in deployments. That is, an RL agent has many tunable knobs that must be set to specialize the solution to the particulars of the problem. This, typically, either requires a digital twin or access to a physical-twin (pilot). Martha and Adam's lab developed an approach that extracts a twin (called a calibration model) based on historical operator data that can be used to safely tune the agent before deployment on the real system:
https://openreview.net/pdf?id=AiOUi3440V