

# Asynchronous Methods for Deep Reinforcement Learning

Lokesh Mohanty

Department of Computational and Data Sciences  
Indian Institute of Science

RL 2023



- Develop a simple and lightweight framework for deep reinforcement learning that uses asynchronous gradient descent for optimizing deep neural network controllers.
- The authors aim to show that parallel actor-learners have a stabilizing effect on training, allowing all four methods to successfully train neural network controllers.
- The best performing method, an asynchronous variant of actor-critic, surpasses the current state-of-the-art on the Atari domain while training for half the time on a single multi-core CPU instead of a GPU.
- Furthermore, they show that asynchronous actor-critic succeeds on a wide variety of continuous motor control problems as well as on a new task of navigating random 3D mazes using visual input.

# Proposed methodology

- Simple and lightweight framework for deep reinforcement learning that uses asynchronous gradient descent for optimizing deep neural network controllers.
- The authors present asynchronous variants of four standard reinforcement learning algorithms, namely Q-learning, SARSA, actor-critic, and deterministic policy gradient, and show that parallel actor-learners have a stabilizing effect on training.
- The proposed framework allows stable training of neural networks through reinforcement learning with both value-based and policy-based methods, off-policy as well as on-policy methods, and in discrete as well as continuous domains.
- The authors also show that the proposed asynchronous algorithms train faster than DQN trained on an Nvidia K40 GPU when trained on the Atari domain using 16 CPU cores, with A3C surpassing the current state-of-the-art in half the training time



- The proposed asynchronous variants of standard reinforcement learning algorithms are able to train neural network controllers on a variety of domains in a stable manner.
- The authors demonstrate that the proposed framework allows stable training of neural networks through reinforcement learning with both value-based and policy-based methods, off-policy as well as on-policy methods, and in discrete as well as continuous domains.
- When trained on the Atari domain using 16 CPU cores, the proposed asynchronous algorithms train faster than DQN trained on an Nvidia K40 GPU, with A3C surpassing the current state-of-the-art in half the training time.
- Furthermore, asynchronous actor-critic succeeds on a wide variety of continuous motor control problems as well as on a new task of navigating random 3D mazes using visual input.



- One possible extension is to investigate the effect of different hyperparameters on the performance of the proposed asynchronous algorithms.
- Another possible extension is to explore the use of more sophisticated exploration strategies, such as intrinsic motivation or curiosity-driven learning, to improve the sample efficiency of the algorithms.
- The authors also suggest that their framework could be extended to incorporate other types of neural network architectures, such as convolutional neural networks or recurrent neural networks, and that it could be applied to other domains beyond those considered in this work.
- Finally, they suggest that their framework could be used for multi-agent reinforcement learning problems by extending it to allow for communication between agents.



# Conclusion

- Presents a simple and lightweight framework for deep reinforcement learning that uses asynchronous gradient descent for optimizing deep neural network controllers.
- The authors demonstrate that parallel actor-learners have a stabilizing effect on training, allowing all four methods to successfully train neural network controllers.
- The best performing method, an asynchronous variant of actor-critic, surpasses the current state-of-the-art on the Atari domain while training for half the time on a single multi-core CPU instead of a GPU.
- Furthermore, they show that asynchronous actor-critic succeeds on a wide variety of continuous motor control problems as well as on a new task of navigating random 3D mazes using visual input.
- The proposed framework allows stable training of neural networks through reinforcement learning with both value-based and policy-based methods, off-policy as well as on-policy methods, and in discrete as well as continuous domains

— *Thank You* —

<sup>1</sup>Kate Donahue and Jon Kleinberg, Optimality and Stability in Federated Learning: A Game-theoretic Approach, NeurIPS 2021

