

# QTAccel: A Generic FPGA based Design for Q-Table based Reinforcement Learning Accelerators

Rachit Rajat\*, Yuan Meng\*, Sanmukh Kuppannagari,  
Ajitesh Srivastava, Viktor Prasanna, *University of  
Southern California*  
Rajgopal Kannan, *US Army Research Lab*  
Contact: prasanna@usc.edu

Q-Table based Reinforcement Learning (QRL) is a class of widely used algorithms in AI that work by successively improving the estimates of Q values -- quality of state-action pairs, stored in a table. They significantly outperform Neural Network based techniques when the state space is tractable. Fast learning for AI applications in several domains (e.g. robotics), with tractable 'mid-sized' Q-tables, still necessitates performing substantial rapid updates. State-of-the-art FPGA implementations of QRL do not scale with the increasing Q-Table state space, thus are not efficient for such applications. In this work, we develop a novel FPGA implementation of QRL, scalable to large state spaces and facilitating a large class of AI applications. Our pipelined architecture provides higher throughput while using significantly fewer on-chip resources and thereby supports a variety of action selection policies that covers Q-Learning and variations of bandit algorithms. Possible dependencies caused by consecutive Q value updates are handled, allowing the design to process one Q-sample every clock cycle. Additionally, we provide the first known FPGA implementation of the SARSA (State-Action-Reward-State-Action) algorithm. We evaluate our architecture for Q-Learning and SARSA algorithms and show that our designs achieve a high throughput of up to 180 million Q samples per second.

\*: Equal contribution

**Acknowledgement:** This work was supported by the U.S. National Science Foundation under grants CNS-1643351 and OAC-1911229.

**Keywords:** Reinforcement Learning Accelerator; Artificial Intelligence; FPGA Acceleration; Q Learning

**DOI:** <https://doi.org/10.1145/3373087.3375359>