Optimizing Reinforcement Learning Through Dynamic Environment Manipulation

Patel, Mihir (School: Thomas Jefferson High School for Science and Technology)

Reinforcement learning (RL) is an AI/machine learning technique capable of learning complex Markov decision process state-action behaviors from thousands of simulation to learn advanced decision making skills and can surpass all human/computer benchmarks in problems such as chess and Go. However, RL agents are notoriously slow to train. For example, it took Google over 40 days to train AlphaGo, the first algorithm to best humans in Go. This research aims to accelerate training time for reinforcement learning through dynamic environment manipulation. Currently, RL has primarily been trained against the hardest opponent, such as the best opposing chess bot in existence. However, if RL trained against something closer to its level, such as a beginner in chess, and the difficulty scaled to match its ability, faster training time might be observed. This would mimic patterns in humans whereby we learn faster by having a “sparring partner” that is slightly above our level. This research provides a mathematical conjecture explaining this phenomenon, produces a construct to test it in, and proves several properties of this scenario. Experimental results on standard benchmark problems show 1.596 times higher rewards and 8.61% higher area under curves for reward over time. This work yields accelerated training times and reduced computational demands on the entire field as a whole, allowing greater access and success with RL in a wide variety of fields that require complex control and logic skills from self-driving cars to stock trading to humanoid robots.

Awards Won:
Second Award of $2,000
Association for the Advancement of Artificial Intelligence: Honorable Mention