Novel Reinforcement Learning Methods in Collaborative Environments

Rao, Ashish (School: Cupertino High School)
Narayanan, Tejas (School: Cupertino High School)
Sarkar, Bidipta (School: Cupertino High School)

Reinforcement learning (RL) is an emerging field with numerous applications in a wide variety of fields, such as robotics, supply chain and network optimization, marketing, and more. Our research focuses on developing improved RL algorithms and applying them to domains involving complex interactions between multiple intelligent systems. We propose a novel algorithm which employs a Bayesian method of parameter space exploration to solve reinforcement learning problems. We parameterize policies as neural networks, and a Gaussian process is used to learn the expected return of a policy given the parameters. The system is trained by updating the parameters in the directions suggested by the Gaussian Process to maximize the expected return and explore important new states. Our new, Bayesian method allows for us to use information about the uncertainty of our estimates to facilitate better exploration of new parameters and states, and allows us to more accurately model the expected return of a particular policy than current state of the art methods which involve lower-bound approximations. Our method is applied to three challenging robotic simulations where we achieve gains of 8%, 20%, and 33% over current methods. We also observe good performance in cooperative and competitive environments that require communication between multiple intelligent systems; we see an improvement of 41% over current methods on one such environment. The improved performance of our methods will enable the development of intelligent systems that tackle significantly more complex problems in robotics and other fields. Our novel method of Bayesian parameter space exploration can be theoretically shown to result in less conservative updates towards better policies.

Awards Won:
Third Award of $1,000