Report on DDPG Algorithm

It is known that DQN is able to solve problems with discrete action space (i.e., low dimensions), by combining a Q-Network and a Deep Neural Network, to approximate a continuous state observation and non-continuous actions; yet when faced with continuous action space, DQN fails to work well. If we come up with a solution to discretize continuous actions there would be an exponential growth of dimensions. Take human arms (with 3 joints) as an example; we could use {-k, 0, k} as three individual sets to approximate human-arm actions with 7 degrees. That would lead to a distribution 2187(3^7) probabilities of action space, which not only slows the approximation of action by a great deal, but also takes a heavy toll on converging to the best training result, by yielding to unstable performance and divergence during training.

As a result, DPG algorithm was introduced to solve the problem instead, with a main goal to optimize pi, the policy.