

Dear Editor,

I have chosen to submit my manuscript “Fitted Policy Iteration for a POMDP Peg-In-Hole search task” to the Journal of Robotics and Autonomous Systems.

Currently applying RL solvers to robotic applications in which the state space is partially observable in addition to being continuous (actions as well) is still problematic. Applying traditional stochastic exploration methods during the policy search optimization works best when the policy is initialized close the optimum as most policy search methods guarantee local optimality. Policy search methods work well when the number of parameters are few when compared to the state space.

Our manuscript addresses the exploration problem and the local optimality of traditional policy search methods. We present a novel RL methodology for learning POMDP policies for continuous belief-state and action spaces. By using Programming by Demonstration (PbD) we initially gather a set of demonstrations which contains both exploration and exploitative behavior. From this we learn a belief-space value function through fitted policy evaluation. We subsequently use the value function to learn a Gaussian Mixture Model policy. We apply our approach to learn a peg-in-hole search task in which the state space uncertainty is non-Gaussian and multi-modal. We further demonstrate that the learned policy can generalize without needing any autonomous rollouts traditionally used in RL.

Using PbD in combination with Fitted Policy Iteration we are able to learn a POMDP policy for a peg-in-hole search task in which the uncertainty is high. We show that it is possible to learn policies with a large amount of parameters in continuous belief-state and action space. We think this work is thus applicable to robotic community at large.

This work is currently not under review or submission in any other journal.

Guillaume de Chambrier
on behalf of the co-authors