SC42050 Literature Assignment
Reusing policies

Gabriel Lopes

The Reinforcement Learning framework (RL) can look for optimal policies when the reward function and the environment are stationary. However, in robotics, stationarity is too strong of an assumption. Can one reuse previously learned control policies to improve learning of new policies when the environment or the rewards changes? Please read (Fernández and Veloso, 2006) and carefully answer the following questions:

1. Do a small literature review and present the basic concepts of Reinforcement Learning, including the update rules for the Q-learning algorithm.

2. Show that the π-reuse exploration strategy in equation (2) does not change the convergence properties of the Q-learning algorithm.

3. As it stands, do you believe the presented PRQ-learning algorithm can be readily implemented in a real robot navigation task? Explain you answer.

4. Try to enumerate the main problems of the current RL framework when faced with real robotic applications (i.e. continuous state spaces, action spaces and time)

References

Fernández, F. and Veloso, M. (2006). Probabilistic policy reuse in a reinforcement learning agent. In *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multi-Agent Systems*, pages 720–727.