

SC42050 Literature Assignment

Policy Gradient Reinforcement Learning for Autonomous Robots

Gabriel Lopes

Autonomous navigation in cluttered environments is a complex task due to the difficulty of modeling the environment, the robot's sensing capabilities, and its actuation. Partially Observable Markov Decision Processes (POMDP) provide a rich framework for tackling this problem. Please read (Grudic et al., 2003) and carefully answer the following questions:

1. Do a small literature review and present the notion of POMDP. Show how you can write the Bellman optimality equation in this framework.
2. How does the POMDP changes the structure of the reinforcement learning problem against the standard Markov Decision Process (MDP)? How would one write the reward function in equation (2) for a standard MDP?
3. How well do you think the proposed implementation could cope with non-stationary environments? (e.g. people walking in front of the robot, doors opening and closing, etc.)
4. Try to enumerate the main problems of reinforcement learning when faced with POMDP.

References

Grudic, G. Z., Kumar, V., and Ungar, L. (2003). Using policy gradient reinforcement learning on autonomous robot controllers. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 406–411.