

Project Proposal CS234

Julien Boussard, Andrew Deveau, Brian Lui

Project Mentor: James Harrison

Risk-Sensitive Decision Making Under Model Uncertainty with Continuous State and Action Spaces

When performing reinforcement learning for robotics, safety is often a key concern and it is accordingly important to be risk-sensitive when learning a policy. One forthcoming paper in this area, which is co-authored by our mentor, James, gives an algorithm that leverages Monte Carlo Tree Search for robust decision making when the dynamics model is unknown and an adversary can choose the agent's prior over the possible models. This paper restricts to the case where the state and action spaces are finite. For our project, we will be experimenting with extending to the case where one or both are continuous. The future work section of the paper sketches an idea of how to do this, and we will be implementing, experimenting with, and perhaps tweaking that idea.

In addition to the paper that we are seeking to extend, we have examined papers on risk-sensitive learning more broadly [SLMP17] [MP17] as well as a recent paper about Monte Carlo Tree Search in continuous POMDPs [SK17]. We expect that the literature on Bayes-Adaptive Markov Decision Process may also be relevant.

Although the long-term goal is to leverage this work on real robots, we will likely evaluate algorithms via their performance, first in some simple simulated environments (possibly from OpenAI Gym) and subsequently in environments relevant to the Autonomous Systems Lab. As in the finite case, we will be chiefly concerned with the (risk-adjusted) rewards obtained by an algorithm and with how quickly it converges in practice. Both of these lend themselves well to graphics and plots of learning curves.

References

- [MP17] Anirudha Majumdar and Marco Pavone. How Should a Robot Assess Risk? Towards an Axiomatic Theory of Risk in Robotics. *CoRR*, abs/1710.11040, 2017.
- [SK17] Zachary Sunberg and Mykel J. Kochenderfer. POMCPOW: an online algorithm for POMDPs with continuous state, action, and observation spaces. *CoRR*, abs/1709.06196, 2017.

- [SLMP17] Sumeet Singh, Jonathan Lacotte, Anirudha Majumdar, and Marco Pavone. Risk-sensitive Inverse Reinforcement Learning via Semi- and Non-Parametric Methods. *CoRR*, abs/1711.10055, 2017.