CS239 Project Proposal

Anton Apostolatos (antonaf@cs.stanford.edu)

January 20, 2016

1 Outline and Approach

A major roadblock in multicopter drone collision avoidance is finding the optimal decision when there are multiple obstacles. Current research excels at finding a close-to-optimal solution to avoid a given obstacle. The issue lies when there are multiple of these and an agent has to find the optimal solution taking each obstacle into account. Our research will investigate dynamically merging Markov decision processes and finding the optimal policy for a composite MDP. There are two avenues of exploration that we will investigate: finding the optimal policy for each component MDP and then merging these into a composite policy, or creating a composite MDP and then finding a good policy in the context of this task.

2 Resources

We will be utilizing DJI's Matrice 100 quadcopters. These come equipped with a guidance vision system, composed of five sensing modules that continuously feed data to a central processing board on the drone. Each of these modules is composed of a pair of stereo cameras and ultrasonic sensors. For testing we also will be utilizing DJI's Simulator which creates a virtual 3D environment for drone flight on a defined terrain, allowing us to perform data analysis from flight data.

3 Testing and Measuring Success

There are multiple stages of testing. The first involves lab testing the SDK to get a full grasp of the performance of the guidance modular system with the help of the simulation. Then we can move on to a manual test, where we fly the drones by hand and record the responses generated by the algorithms, making sure that these are appropriate. We can then test using hard-coded interactions, whereupon we design and stage certain scenarios (such as holding various objects at various places) and record the drone's chosen action.

Quantifying the success of our agent will be tricky, as we are working with very valuable hardware and equipment that cannot withstand extensive testing. A proposed method is to compare the actions chosen by the programmed agent to those made by a seasoned UAV pilot in different orchestrated scenarios, and see whether or not their decision making concurs and, if not, analyzing why not.

4 Existing Work

I will be working with Eric Mueller and Mykel Kochenderfer at the Stanford Intelligent Systems Laboratory (SISL). Our project will incorporate and build on Eric's PhD dissertation which focuses on multicopter collision avoidance. His research is focused primarily on avoiding single object. We will be extending his research for multi-object collision avoidance.

I will also be building on existing work that focuses on combining MDPs. Of note, I will be working off of Singh and Cohen's paper "How to dynamically merge Markov decision processes"[2], Sprague and Ballard's GM-Sarsa(0) algorithm as explained in "Multiple-Goal Reinforcement Learning with Modular Sarsa(0)" [3], and Schneider et al.'s "Distributed value functions"[1].

References

[1] J. Schneider, W.-K. Wong, A. Moore, and M. Riedmiller. Distributed value functions. In *ICML*, pages 371–378. Citeseer, 1999.

[2] S. Singh, D. Cohn, et al. How to dynamically merge markov decision processes.

[3] N. Sprague and D. Ballard. Multiple-goal reinforcement learning with modular sarsa (0). In *IJCAI*, pages 1445–1447. Citeseer, 2003.