Adaptive Multi-Agent Path Planning for Distributed UAV Systems

CS 229 Autumn 2017 Project Milestone
Category: Physical Sciences

Lloyd Maza — `lmaza`
Cameron McMillan — `cmac12`

Motivation

In recent years, unmanned aerial vehicles (UAVs) have become increasingly prevalent in the modern world. These systems are used for a wide variety of applications, from recreation and entertainment to warfare and espionage. Generally speaking, UAVs are favored for their low cost and low risk, not to mention their ability to execute tasks in environments which might be inaccessible or dangerous to humans.

In tactical scenarios, there is an increased demand for the use of cooperative, distributed systems of UAVs which enable more complex and larger objectives. Oftentimes, these systems require the use of centralized, human-in-the-loop control approaches to accomplish their goals which may be slow, inaccurate, and computationally intensive. As a result, autonomy and decentralization are desirable traits for any such system of UAVs.

The goal of this project is the implementation of an algorithm by which a system of multiple UAVs can to learn how to navigate in a contested environment that is initially unknown. Our implementation will use modified Q-Learning techniques to allow UAVs to navigate complex dynamic environments. Missions will contain different surveillance target layouts and the UAVs will learn from their observations. The UAVs' path planning will be rewarded for minimizing risk to the UAVs while maximizing information acquired. The UAVs will work cooperatively to explore the unknown environment by sharing information about the risk environment seen by each agent in the system.

The simulation environment for this project is built in MATLAB. In this environment, risks and targets can be programmatically generated. This simulation will be able to replicate entire missions, including all the relevant jamming, observation, and performance information of the UAVs and targets.

Methods

From our initial research, reinforcement learning appeared to be the most common method of generating path planning for multiple agents in the complex dynamic environments these UAVs would operate in. Most common implementations of reinforcement learning for this centered on some form of Q-Learning. This method was investigated, and an initial simulation was created but some flaws were identified. The main flaw dealt with the discretation of states. Unlike ground vehicles, aircraft paths have significantly more maneuvering freedom. The discretation of states for Q-Learning can cause issues because this requires the UAV navigate a series of discrete points. A more representative system would predict the reward of the aircraft along more realistic paths. Additionally, because aircraft have many state variables, deep reinforcement learning was investigated, but was deemed too complex at this time. This decision was made because current literature suggests that successful implementation of deep reinforcement learning requires additional pre/post processing of the data to get stable implementations [1].

Ultimately, a modified Q-Learning algorithm known as the Cooperative and Geometric Learning Algorithm (CGLA) was chosen for the implementation of this system. This algorithm was proposed by a group of researchers primarily based out of Beihang University in China in 2013 [3]. It is an online reinforcement learning approach which was specifically designed for helping multi-agent UAV systems navigate through uncertain environments, so it is unsurprising that it was deemed as a useful starting point for this project.

The algorithm assumes that all agents know their initial start positions, as well as a set of target locations on the map. All agents share an underlying "G matrix" which encodes the expected risk of all states on the map. This matrix is effectively an equivalent to the "Q matrix" of Q-Learning, which encodes expected reward associated with a set of actions. As agents navigate toward the targets and encounter previously unknown risks, the G matrix is updated and all agents are able to modify their paths accordingly [4].

It should be noted that a primary distinction between CGLA and Q-Learning is that all states on the map are considered valid next states. This is meant to emulate realistic motion in UAVs which, as previously mentioned, are not constrained to discretized step-wise motion. This has a tendency to make updates to the G matrix somewhat computationally expensive.

Preliminary Experiments

The current state of the project is a basic implementation of CGLA within a simplified MATLAB environment. Currently, we are only simulating a single UAV with a single target. It is also assumed that all threats are stationary and known by the agent from the beginning. This is in contrast with the real scenario of a full multi-agent system with multiple targets and multiple dynamic, unknown threats. These simplifications made were meant to reduce the time and complexity involved in getting the CGLA implementation up and running.

In its current state, the CGLA implementation is effectively a single agent path planner. The optimal path is generated through the path of minimum cost according to the G matrix described above, and relies on a trade-off between time to reach the target and risk of the traversed path. In the algorithm's current implementation we are able to visualize the algorithm's performance and subsequently tune any relevant parameters to achieve what may be considered a desirable path. The figures on the next page illustrate the trade off between risk and time of flight in a single agent scenario with multiple threats.

It should be noted that in its current state, the algorithm scales very poorly as the number of states on the map increases. As such, improvements in performance are a high priority going forward, perhaps through increased vectorization or other means.

Future Work

Given the results of our first experiments, there a few avenues we could pursue to improve our algorithm. First, we plan to introduce unknown threats to the environment. These will add improve other results because it is more realistic for the UAVs to discover their environment, rather than be omniscient from the start. For this to occur, an observation radius needs to be added to the UAVs within which the agents can sense threats. Then, we will add multiple cooperative UAVs to the simulation. These UAVs will cooperate by sharing risk information about the environment they observe. Using this update infor-

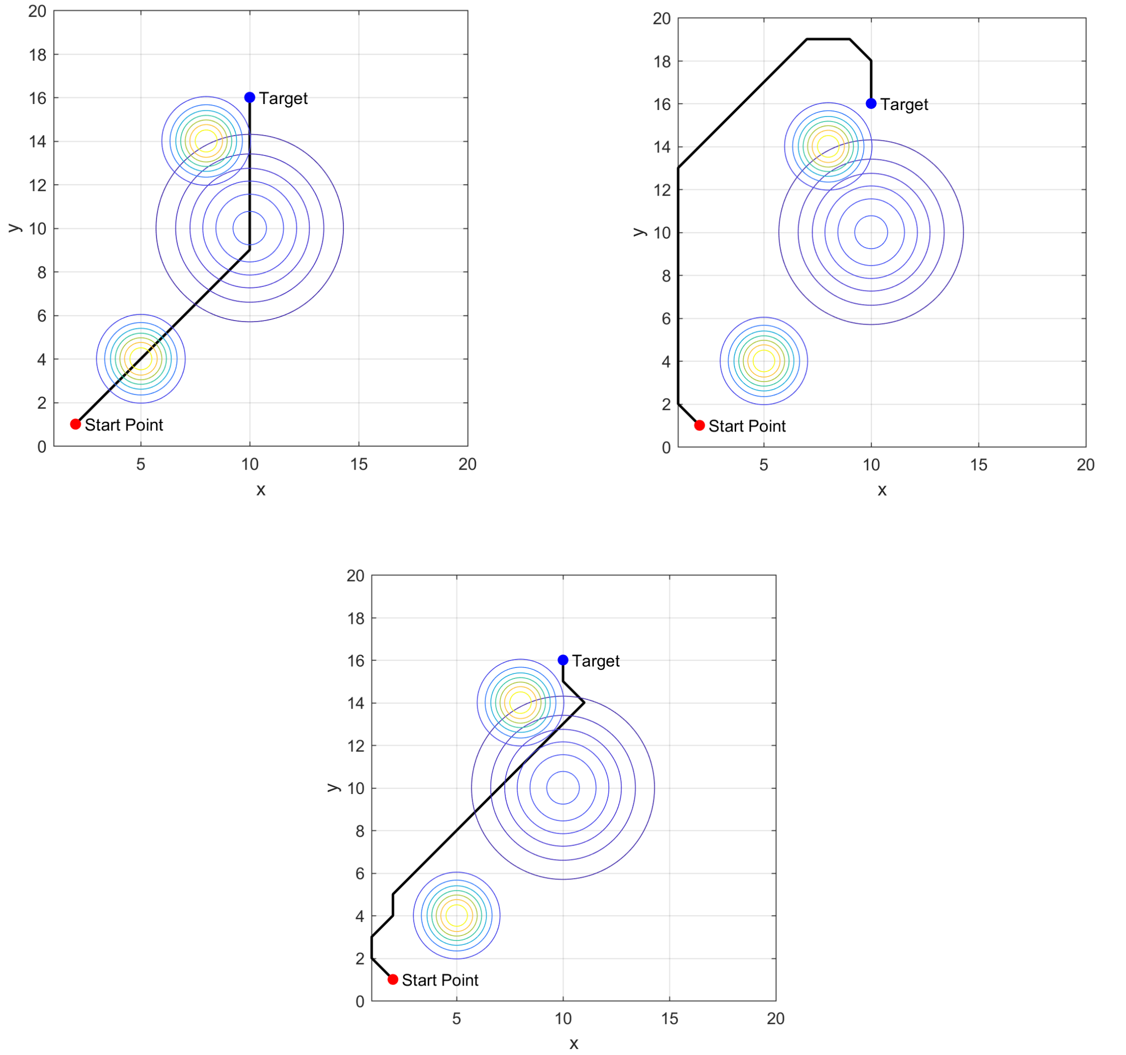

Figure 1: Sample CGLA paths associated with low risk aversion (top left), high risk aversion (top right), and moderate risk aversion (bottom center); circular contours indicate areas of concentrated risk

mation, they will be able to more quickly discover the unknown threats and move to the targets. Comparing this algorithm to other methods of path planning, such as Q-Learning or deterministic algorithms could provide interesting information regarding its suitability for real world applications. Finally, we plan to add fuel consumption to our UAVs paired with additional reward for returning to base. This feature would build a basis for very realistic simulations because most UAVs aren't disposable, so they need to return to some safe point as well.

Contributions

The work on this project has been an even split. Cameron McMillan worked on preliminary literature review, explored the preliminary implementation of the Q-Learning algorithm, and assisted in the implementation of CGLA. Lloyd Maza also worked on preliminary literature review and the implementation of CGLA.

References

[1] Matiisen, T., (2015, December 22). Guest Post (Part I): Demystifying Deep Reinforcement Learning [Web log post] Retrieved October 10, 2017, from https://www.intelnervana.com

[2] U.S. Air Force. Air Force Materiel Command. (2017). Resilient Autonomous Systems (RAS) (Solicitation No. BAA-AFRL-RIK-2016-0005). Rome, NY

[3] Zhang, B., Mao, Z., Liu, W., et. al. "Geometric Reinforcement Learning for Path Planning of UAVs," Journal of Intelligent and Robotic Systems, Volume 77, No. 2, 2015, pp 391–409.

[4] Zhang, B., Mao, Z., Liu, W., et. al. "Cooperative and Geometric Learning Algorithm (CGLA) for path planning of UAVs with limited information," Automatica, Volume 50, No. 3, 2014, pp 809-820.