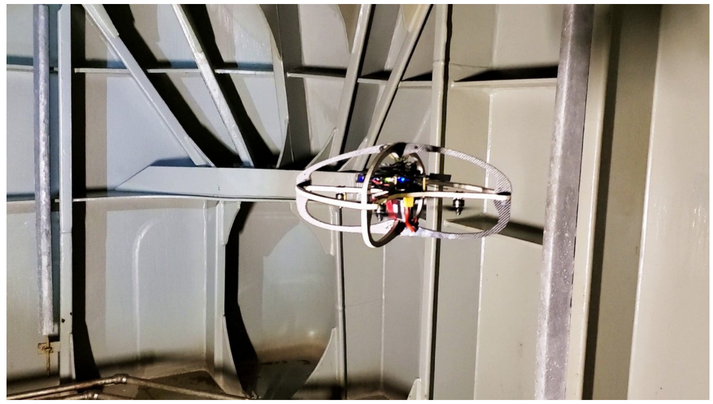
Assured Learning-based Safe Navigation within Degraded and Constrained Environments



Overview: This thesis aims to develop the new methods and technological solutions required to enable learning-based verifiably safe navigation of aerial robotic systems operating inside sensor-degraded and constrained/cluttered environments. In particular, a novel navigation policy, derived through assured reinforcement learning techniques, is proposed that will guide an autonomous flying system across reference waypoints, while autonomously ensuring a) the avoidance of 3D obstacles inside dense and narrow environments, b) sampling of data from the subsets of the environment that present interest to the robot's mission, and c) accounting for the effects of localization uncertainty to the problem of safe navigation. To offer guarantees for field deployment, the envisioned work aims to introduce reachability analysis inside the training process and a verification tool to certify the safety and performance of the learned policy. The results will be demonstrated in the context of challenging field experiments involving the autonomous navigation inside ballast water tanks. Ballast tanks represent particularly confined and often cluttered settings that are simultaneously dark and textureless, alongside being GPS-denied.

Tasks and Sub-objectives

- Literature review - understanding of Q-Learning/Q-Learning in Continuous State/Action Spaces, study of the Deep Deterministic Policy Gradients method.
- Modeling of the Navigation and Information Sampling Planning Problem - understanding of the application of Q-Learning for Partially Observable Markov Decision Processes - understanding of the Recurrent Deterministic Policy Gradients method.
- Modeling of the RMF collision-tolerant robot in simulation (prior model in Gazebo available).
- Research to propose Policy, Objective function approximation and Network design for Assured Learning-based Safe Navigation within Degraded and Constrained Environments.
- Derivation of Combined Simulation- and Experiments-based Dataset for the Assured Learning-based Safe Navigation.
- Implementation on the RMF (photo) Aerial Robot and Evaluation in real-life Field Experiments.

Starting Literature

[1] Hu, H., Fazlyab, M., Morari, M. and Pappas, G.J., 2020. Reach-SDP: Reachability Analysis of Closed-Loop Systems with Neural Network Controllers via Semidefinite Programming. arXiv preprint arXiv:2004.07876.

[2] Bajcsy, A., Bansal, S., Bronstein, E., Tolani, V. and Tomlin, C.J., 2019, December. An efficient reachability- based framework for provably safe autonomous navigation in unknown environments. In 2019 IEEE 58th Conference on Decision and Control (CDC) (pp. 1758-1765). IEEE.

[3] Achiam, J., Held, D., Tamar, A. and Abbeel, P., 2017. Constrained policy optimization. arXiv preprint arXiv:1705.10528.

[4] Wierstra, D., Foerster, A., Peters, J. and Schmidhuber, J., 2007, September. Solving deep memory POMDPs with recurrent policy gradients. In International Conference on Artificial Neural Networks (pp. 697-706). Springer, Berlin, Heidelberg.

[5] Sutton, R.S. and Barto, A.G., 2018. Reinforcement learning: An introduction. MIT press.

[6] Bertsekas, D.P., 2019. Reinforcement learning and optimal control. Athena Scientific.