

# Mechanical Engineering Doctoral Defense

## Reinforcement Learning for Planning and Scheduling in Aviation

School for Engineering of Matter, Transport and Energy

**Jueming Hu**

Advisor: Yongming Liu

### Abstract

Aviation is a complicated field that involves a wide range of operations, from commercial airline flights to Unmanned Aerial Systems (UAS). Planning and scheduling are essential components in the aviation industry that play a significant role in ensuring safe and efficient operations. Reinforcement Learning (RL) has received increasing attention in recent years due to its capability to enable autonomous decision-making. To investigate the potential advantages and effectiveness of RL in aviation planning and scheduling, three topics are explored in-depth, including obstacle avoidance, task-oriented path planning, and maintenance scheduling.

A dynamic and probabilistic airspace reservation concept, called Dynamic Anisotropic (DA) bound, is first developed for UAS, which can be added around the UAS as the separation requirement. A model based on Q-learning is proposed to integrate DA bound with path planning for obstacle avoidance. Moreover, A deep reinforcement learning algorithm based on Proximal Policy Optimization (PPO) is proposed to guide the UAS to destinations while avoiding obstacles through continuous control. Results from case studies demonstrate that the proposed model can provide accurate and robust guidance and resolve conflict with a success rate of over 99%.

Next, the single-UAS path planning problem is extended to a multi-agent system where agents aim to accomplish their own complex tasks. These tasks involve non-Markovian reward functions and can be specified using reward machines. Both cooperative and competitive environments are explored. Decentralized Graph-based reinforcement learning using Reward Machines (DGRM) is proposed to improve computational efficiency for maximizing the global reward in a graph-based Markov Decision Process (MDP). Q-learning with Reward Machines for Stochastic Games (QRM-SG) is developed to learn the best-response strategy for each agent in a competitive environment.

Furthermore, maintenance scheduling is investigated. The purpose is to minimize the system maintenance cost while ensuring compliance with reliability requirements. Maintenance scheduling is formulated as an MDP and determines when and what maintenance operations to conduct. A Linear Programming-enhanced Rollout (LPRT) method is developed to solve both constrained deterministic and stochastic maintenance scheduling with an infinite horizon. LPRT categorizes components according to their health condition and makes decisions for each category.



April 5, 2023; 10 AM; ERC 490;

Zoom Link: <https://asu.zoom.us/j/87867072856?pwd=V090Z0l4SzNkM0FOQTZ3VHBhL25lQT09>