Session 3

1. Blackbox drone problem -> id decision making process of enemy drone based on partial observability

2. Drone captured by us. Info -> dmp is q-learning -> Questions: What does this tell us about the behavior of other enemy drone's of the same class? # static environment # online-v-offline learning # counter measures # more complex approaches

3. Reverse engineer the drone and determine different optimal path we conclude that enemy cannot have same intelligence / training data as us. Based on the drone's optimal path, can we determine what the enemy knows about our environment? i.e. assumes that base is at different location or does not know that another base exists along the way.

4. What is a more realistic learning architecture?

Segway to reward shaping etc. in session 4.

Session 4

Convoy

Human in the loop RL as improvement on vanilla learning strategies

But: how do humans actually learn the knowledge/rules/behavior that they communicate to the agent? Adaptive/generative learning in groups…

Trade-offs for convoy case