**Adversarial Attack and Defense:**

There are three major questions to figure out.

1. Does the Fast Graident Sign Method only attacks the first frame, or multiple frames?
2. Does it allow environment rollback or not when the policy is under attack?
3. Does the attack and defense strategies used for single agent applied to multi-agents as well. (When playing a chess game, my opponent’s policy is my environment and vice versa, can I figure out the way he plays and try to attack?)

**TRPO Proof:**

Try to understand the proof and core idea of TRPO mainly focused on the policy advantage part. Refer back to Approximately Optimal Approximate Reinforcement Learning and Neuro-Dynamic Programming.

The core idea of this proof might be related to adversarial attack and defense

**What to do first?**

* Tackle the first two problems of Adversarial Attack and Defense and implement code for the experiments in Adversarial Attacks on Neural Network Policies.
* Figure out TRPO proof and we can have a one-hour meeting to discuss about it in the future