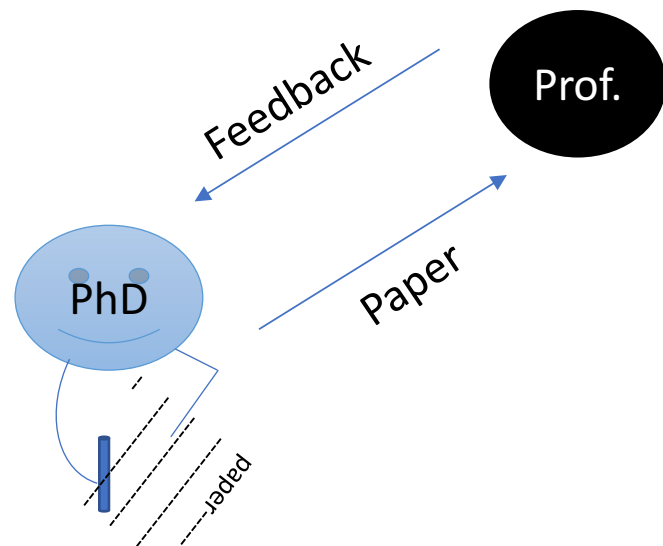


Learning Objective:

$$\text{Max}_{\pi} \sum r_t$$

Value Based/ Policy Based/ Actor-Critic



Dense Reward: Once you have a commitment, you receive a feedback

Sparse Reward: Your professor is busy

Model Based/ Model Free