Report as to 7th September

0. Review

Environment(Data Amount)	Open Loop Control		Closed Loop Control	
	A* + rollout	PPO + rollout	(A* based)LQR	AIP
Real Hand(100% of 150K)	Not yet	Not yet	Not yet	Not yet
Gazebo Hand(0.1% of 1.6M)	Done	Done	Done	Not yet
Reacher-v2(0.1% of 1M)	Done	Done	Done	Not yet
Acrobot-v1(100% of 1M)	Done	Done	Not Available	Not yet

1. Success Rate

Table 1. Gazebo Hand (0.1% Model)

Goal Location	0	2	7	8	15
A*	0%	100%	100%	0%	0%
PPO	0%	100%	100%	100%	0%
LQR	100%	100%	100%	100%	100%

Table 2. Reacher (0.1% Model)

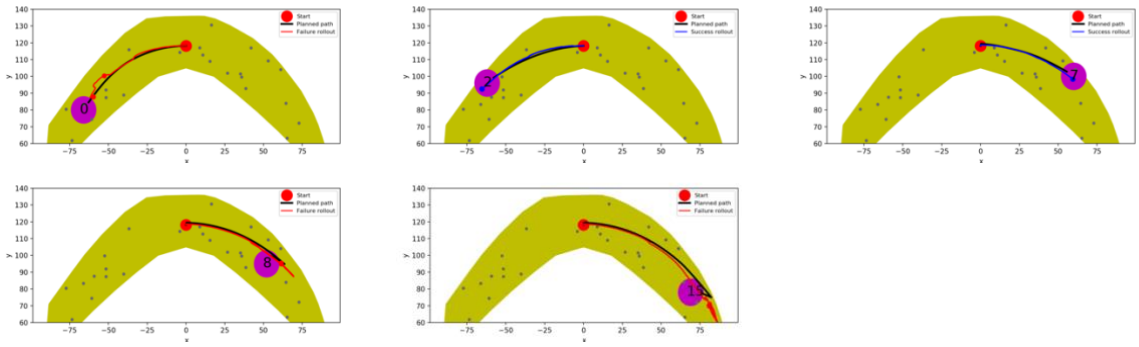
Goal Location	1	2	5
A*	0%	100%	0%
PPO	0%	0%	0%
LQR	100%	0%	100%

Table 3. Acrobot (100% Model)

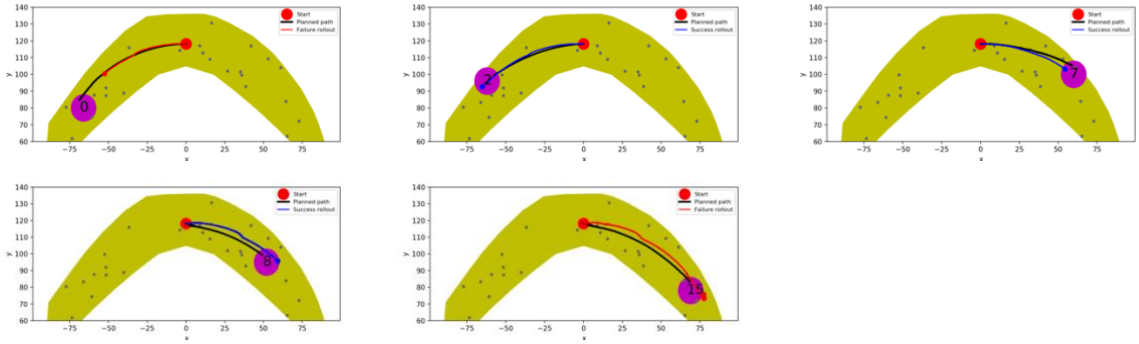
Goal Height	1.0
A*	60%
PPO	30%
LQR	-

2. Gazebo Hand (0.1% Model) Plots

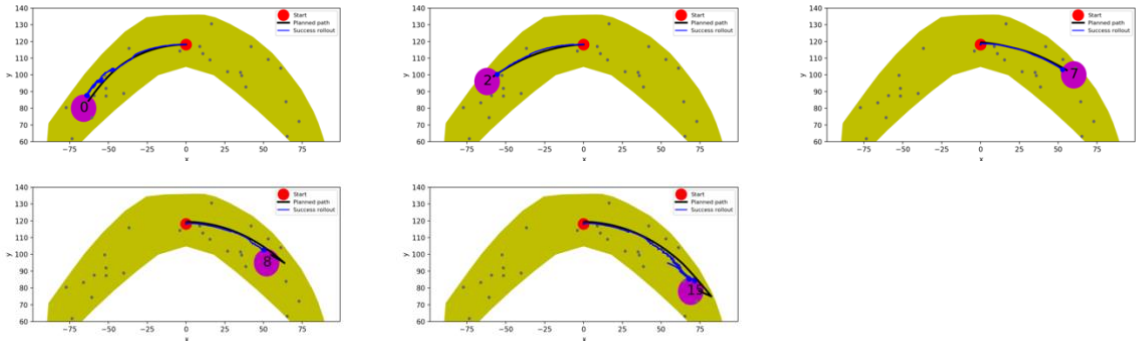
- A*



- PPO(Separate Model: Goal location not as a part of state; Sparse Reward)

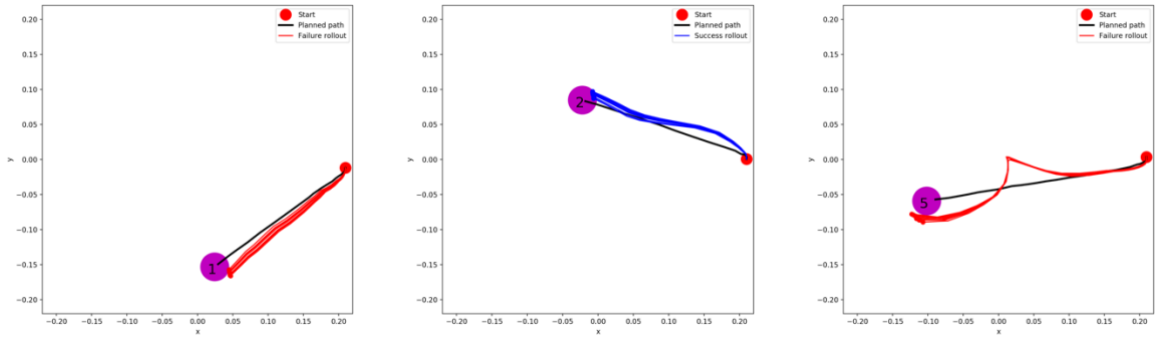


- LQR

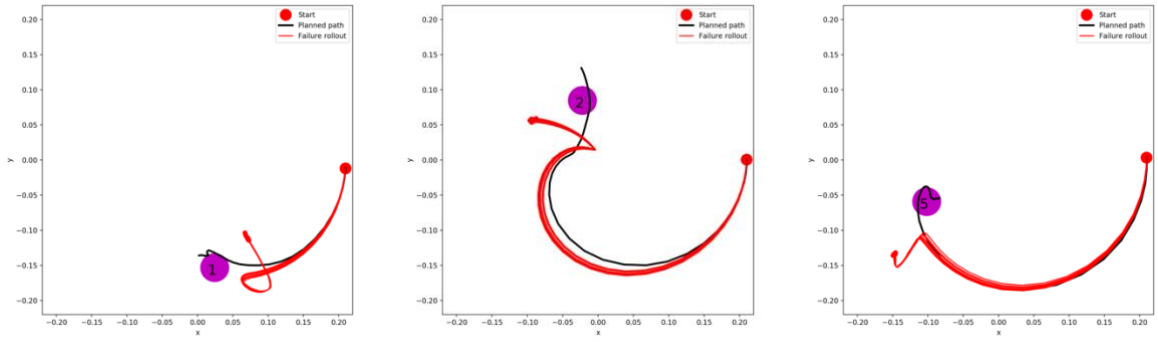


3. Reacher (0.1% Model) Plots

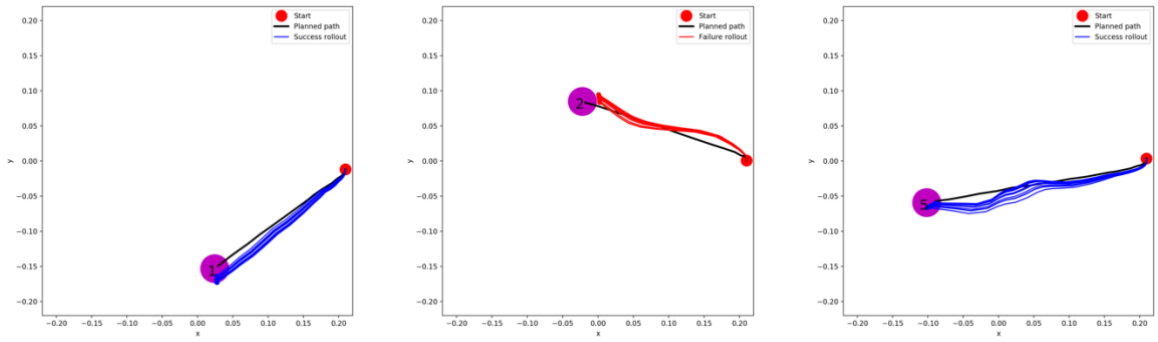
- A*



- PPO(General Model: Goal location as a part of state; Continuous Reward)

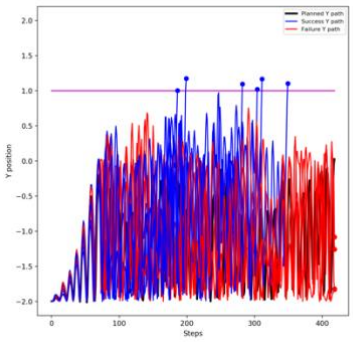


- LQR

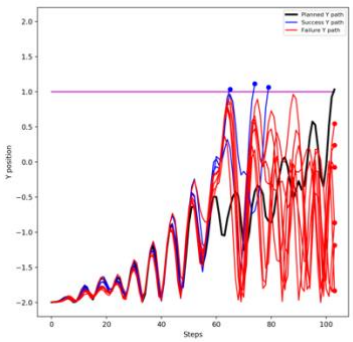


4. Acrobot (100% Model) Plots

- A*



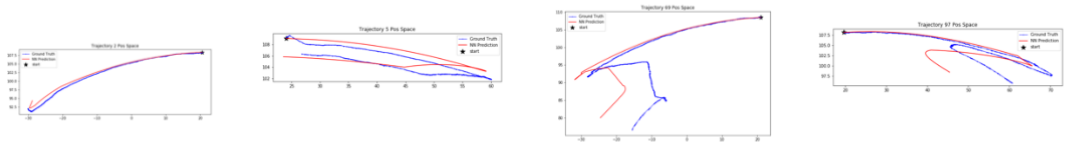
- PPO(Sparse Reward)



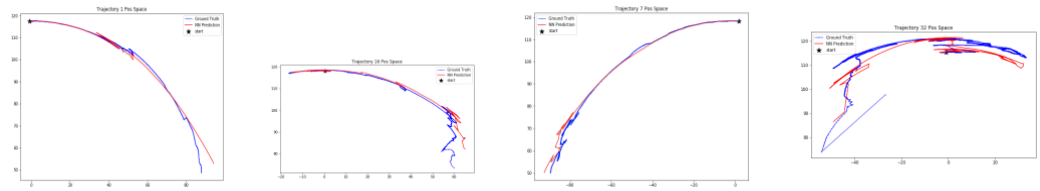
Appendix

1. Transition Model:

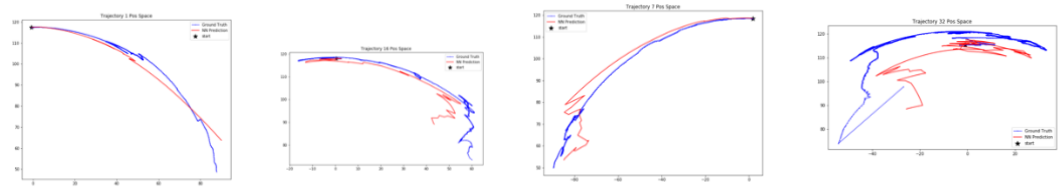
- Real Hand(100% Model):



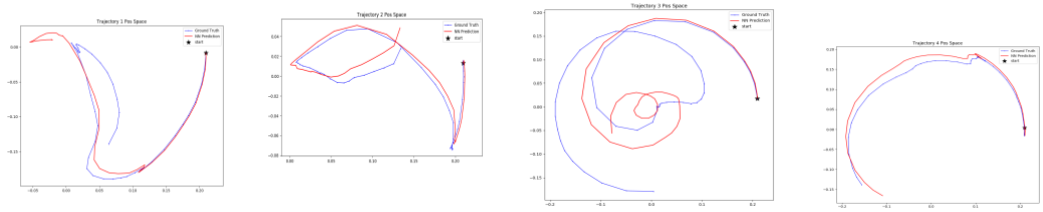
- Gazebo Hand(100% Model):



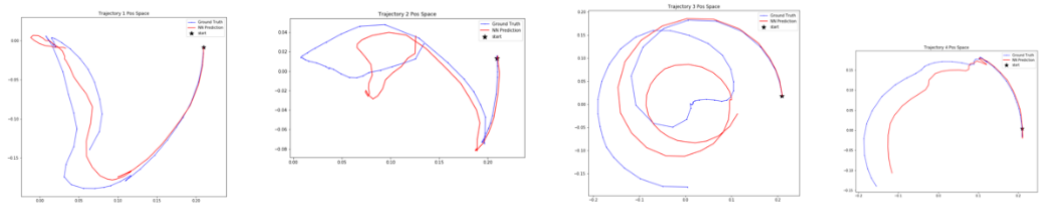
- Gazebo Hand(0.1% Model):



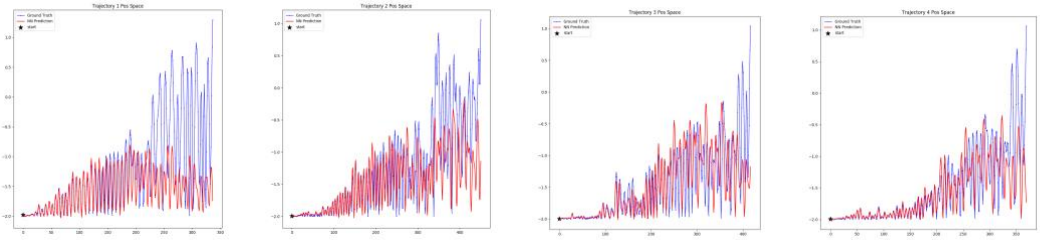
- Reacher(100% Model):



- Reacher(100% Model):



- Acrobot(100% Model):



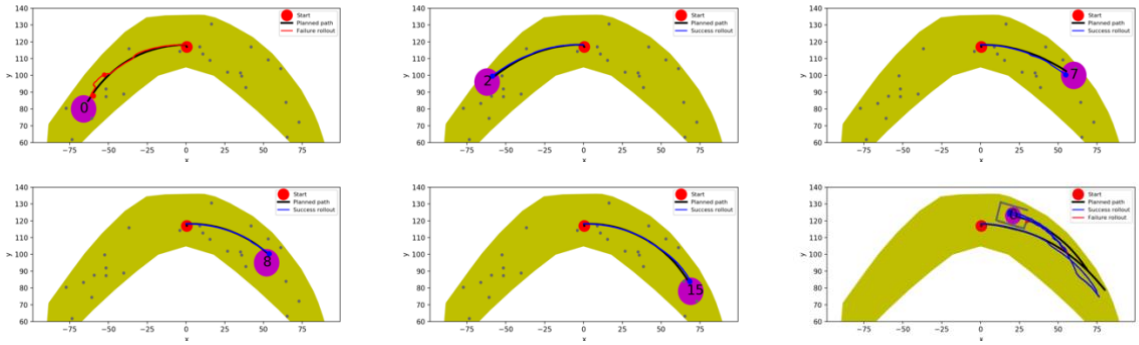
2. Gazebo Hand(100% Model):

- Goal reach rate:

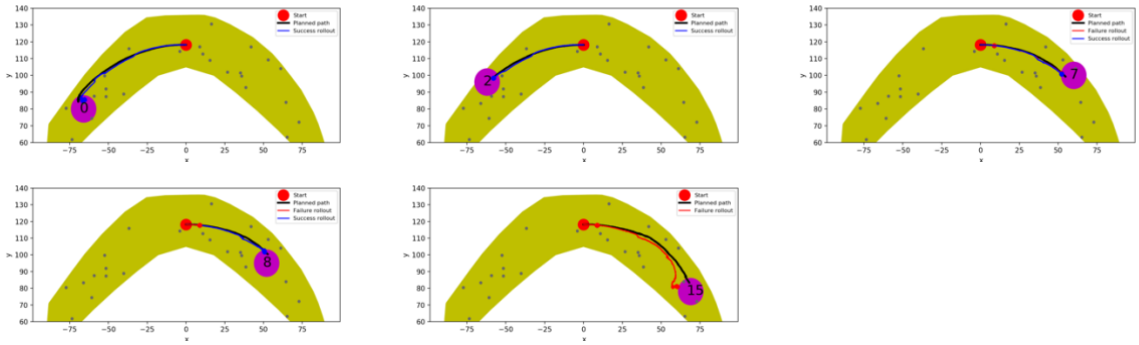
Table 4. Gazebo Hand (100% Model)

Goal Location	0	2	7	8	15	horseshoe
A*	0%	100%	100%	100%	100%	90%
PPO	100%	100%	60%	40%	20%	-

- A*:



- PPO (General Model: Goal location as a part of state; Continuous Reward):



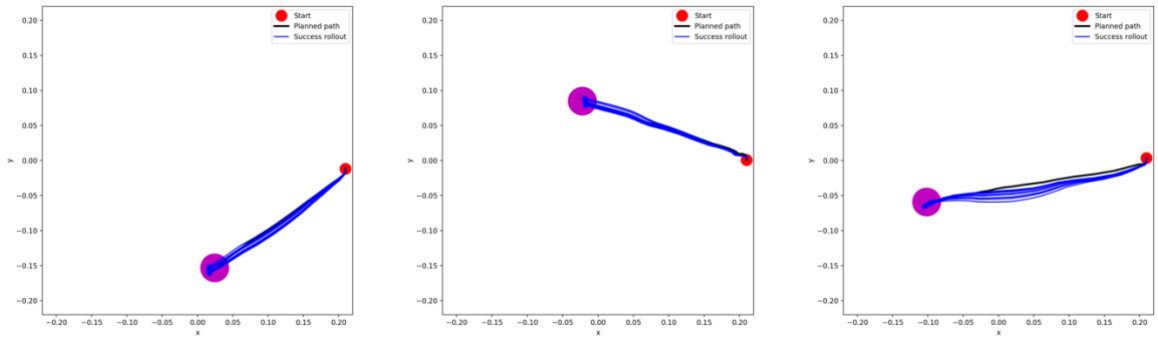
3. Reacher(100% Model):

- Goal reach rate:

Table 2. Reacher (0.1% Model)

Goal Location	1	2	5
A*	100%	100%	100%
PPO	60%	0%	0%

- A*:



- PPO (General Model: Goal location as a part of state; Continuous Reward):