

1 Behavioral Cloning

1.2

Behavioral Cloning	
Ant	Hopper
Mean	Mean
4550.343	938.227
STD	STD
236.455	104.239

Table 1: We are collecting Ant and Hopper, Ant being the one which was amenable to behavior cloning, Hopper not. Settings were about 5 trajectories, with 2 hidden layers of size 64. Batch size was 1000, training batches were 100. Each were run for 1 epoch. Ant expert policy mean was 4713.653. Hopper expert policy mean was 3772.670

1.3

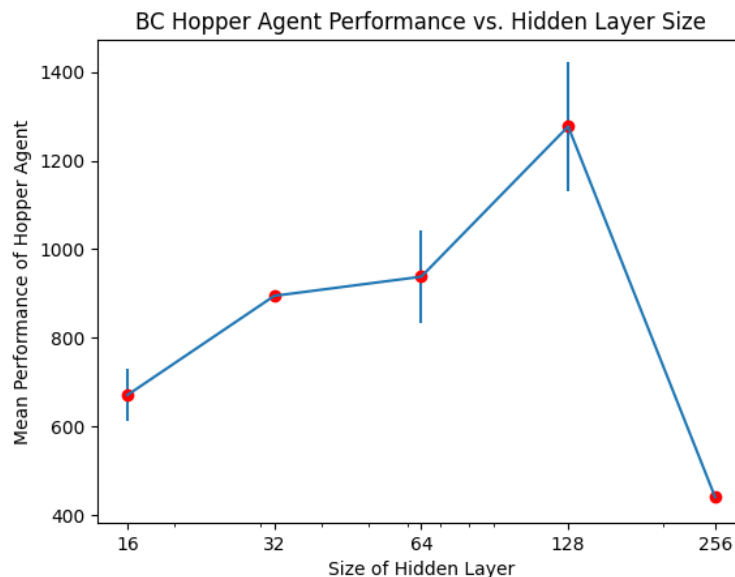


Figure 1: Plotted agent performance against size of the Neural Network. I wanted to see whether or not the ability for the agent to imitate was a matter of network complexity, but it probably is not.

2 DAgger

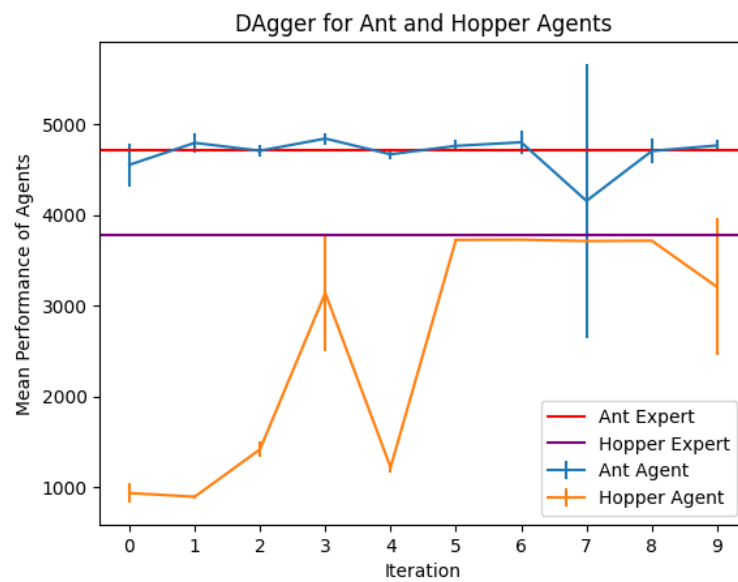


Figure 2: Ant and Hopper agents with DAgger. The settings are the same as in 1.2. But it is run with DAgger for 10 epochs