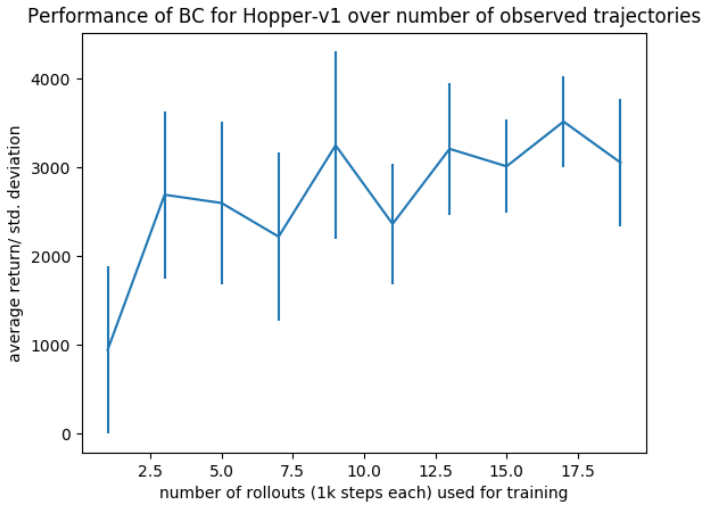
Homework 1

Results (Sections 3 and 4)

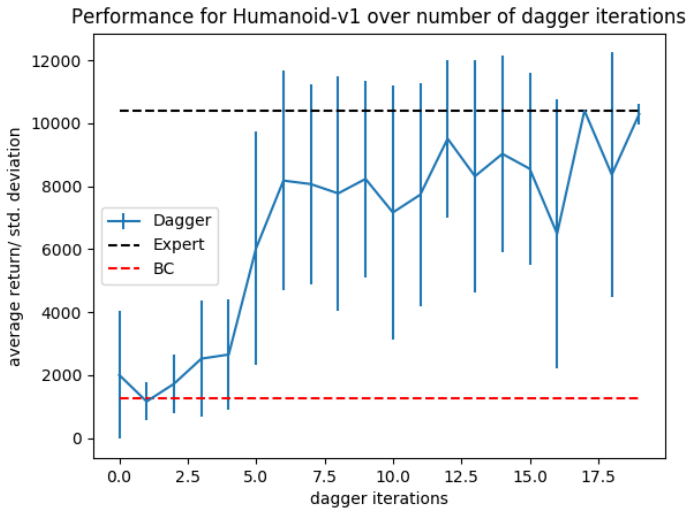
	Expert	BC	Dagger
Hopper	3777 +- 3.1	3496.8 +- 450	
Humanoid	10414 +- 58	1261.5 +- 846.5	9258 +- 2571
Ant	4830 +- 89	4804 +- 112	

Table1: Results show average returns and standard deviation over 20 runs. All BC-agents are trained on 100 trajectories (of 1000 steps each) for 95000 iterations. For all examples of BC and Dagger the function approximator is chosen to be a neural network with 3 layers and 400 neurons each. For BC and Dagger the learning rates are piecewise constant with boundaries at 1e4, 2e4 and 5e4 and lr: [1e-3,1e-4,1e-5,1e-6].

Section 3



Section 4



In every dagger iteration 10 new trajectories (with 1k steps each) are collected and the network is trained for 5000 steps. Overall 20 iterations of Dagger are used.