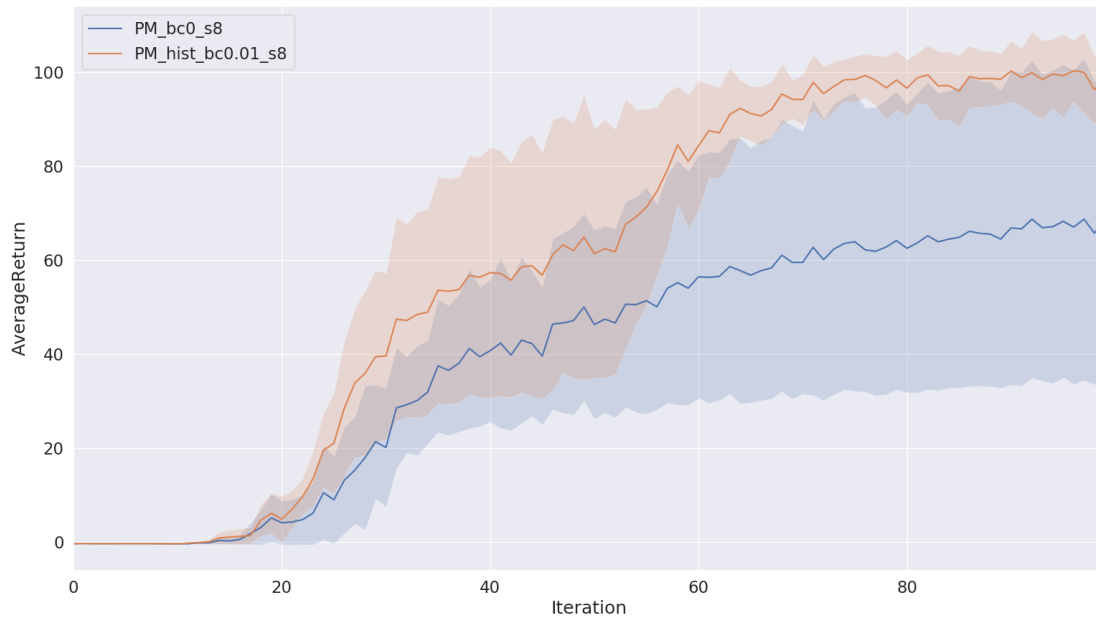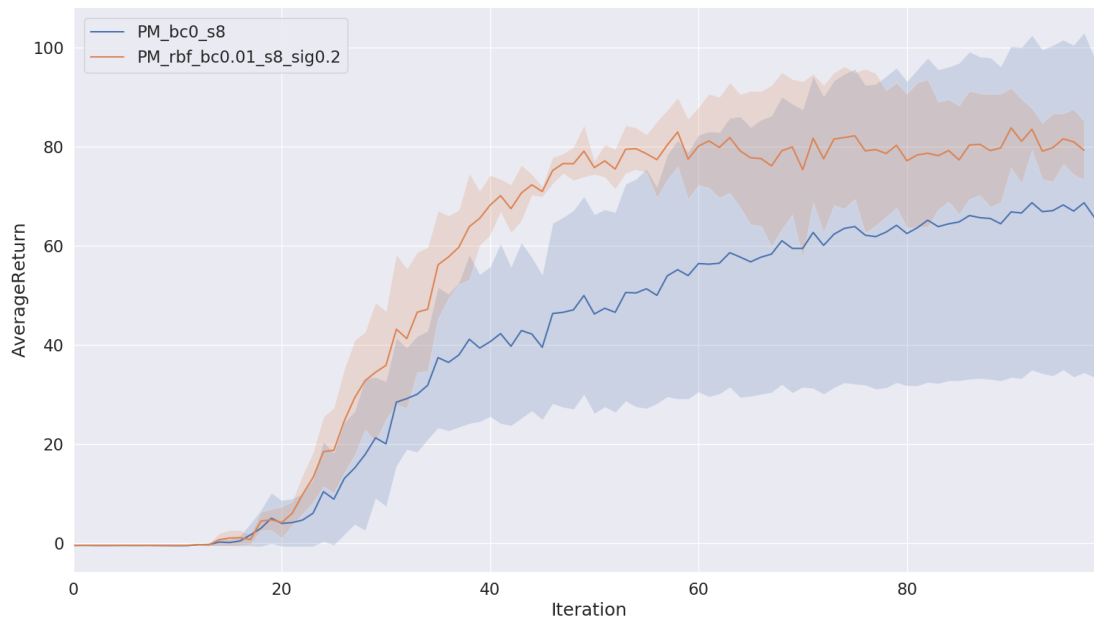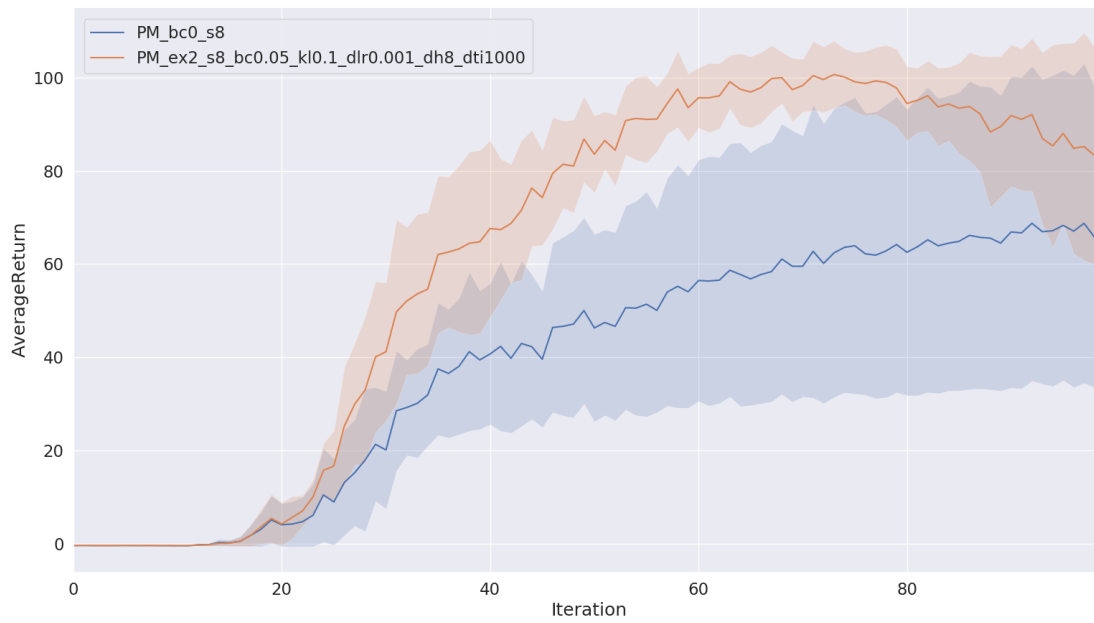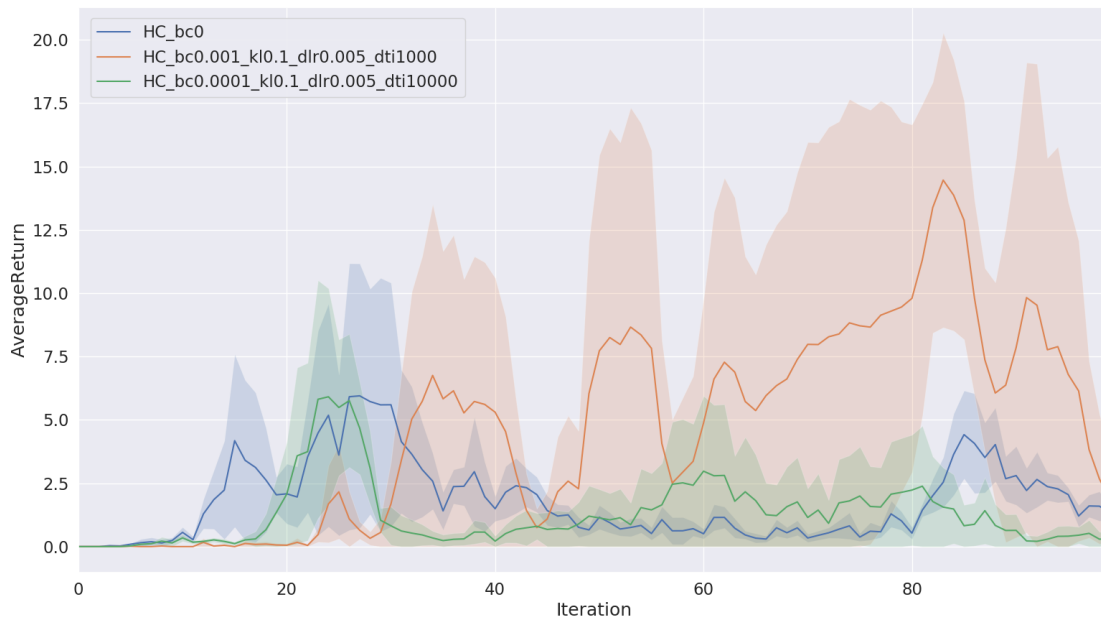# CS 294-112 – Homework#5a

Problem 1

Figure 1. Discrete exploration vs No exploration

Problem 2

Figure 2. RBF exploration vs No exploration

Problem 3

Figure 3. Examplar exploration vs No exploration

Problem 4

Figure 4. Cheetah

The curve with -bc 0.0001 and -dti 10000 performs much better than the one with -bc 0.001 and -dti 1000. Reasons:

- Larger density train iterations fit the density model better, since there is more training for the discriminator.
- Smaller bonus coefficient encourages the exploration in a slower but more stable way. As we can see, the improvement of the red curve comes later than that of the green curve but it is more consistent.