Homework 4 Submission

Use this template to record your answers for Homework 4. Add your answers using LaTeXand then save your document as a PDF to upload to Gradescope. You are required to use this template to submit your answers. **You should not alter this template in any way** other than to insert your solutions. You must submit all 9 pages of this template to Gradescope. Do not remove the instructions page(s). Altering this template or including your solutions outside of the provided boxes can result in your assignment being graded incorrectly.

You should also export your code as a .py file and upload it to the **separate** Gradescope coding assignment. Remember to mark all teammates on **both** assignment uploads through Gradescope.

Instructions for Specific Problem Types

On this homework, you must fill in blanks for each problem. Please make sure your final answer is fully included in the given space. **Do not change the size of the box provided.** For short answer questions you should **not** include your work in your solution. Only provide an explanation or proof if specifically asked.

Fill in the blank: What is the course number?

10-703

Problem 0: Collaborators

Enter your team members' names and Andrew IDs in the boxes below. If you worked in a team with fewer than three people, leave the extra boxes blank.

Name 1:		Andrew ID 1:	
Name 2:		Andrew ID 2:	
Name 3:		Andrew ID 3:	

Problem 1: CMA-ES (24 pts)

1.1 Plot of CMA-ES on simple objective function (10 pts)

Solution

1.2 RL reward of fixed policies (4 pts)

$x = (-1, -1, -1, -1, -1)$: 15.6

$x = (1, 0, 1, 0, 1)$: SOL

$x = (0, 1, 2, 3, 4)$: SOL

1.3 Plot of CMA-ES on Cartpole (10 pts)

Solution

Problem 2: Imitation Learning (62 pts)

Problem 2.1: BC (14 pts)

2.1.1 Loss Plot + Final Loss Value BC (6 pts)

Solution

2.1.2 Rewards Plot BC (6 pts)

Solution

2.1.3 GIF link BC (2 pts)

Solution

Problem 2.2: DAgger (18 pts)

2.2.1 Loss Plot DAgger (6 pts)

Solution

2.2.2 Rewards Plot DAgger (6 pts)

Solution

2.2.3 GIF link DAgger (2 pts)

Solution

2.2.4 Compare DAgger training with BC (written, 4 pts)

Solution

Problem 2.3: Diffusion Policy (30 pts)

2.3.1 Loss Plot + Final Loss value Diffusion Policy (6 pts)

Solution

2.3.2 Rewards Diffusion Policy (15 pts)

2.3.2.1; 3 actions evaluated in a row (5 pts)

avg trajectory time: SOL mean: SOL median: SOL max: SOL

2.3.2.2; 2 actions evaluated in a row (5 pts)

avg trajectory time: SOL mean: SOL median: SOL max: SOL

2.3.2.3; 1 action evaluated in a row (5 pts)

avg trajectory time: SOL mean: SOL median: SOL max: SOL

2.3.3 GIF link Diffusion Policy (2 pts)

Solution

2.3.4 Compare diffusion policy and simple model runtime (written, 4 pts)

Solution

2.3.5 Compare diffusion policy with different actions in a row runtime (written, 3 pts)

Solution