CS294-112 Deep Reinforcement Learning HW4

Heri Zhao

Fall 2018

Problem 1

(a)

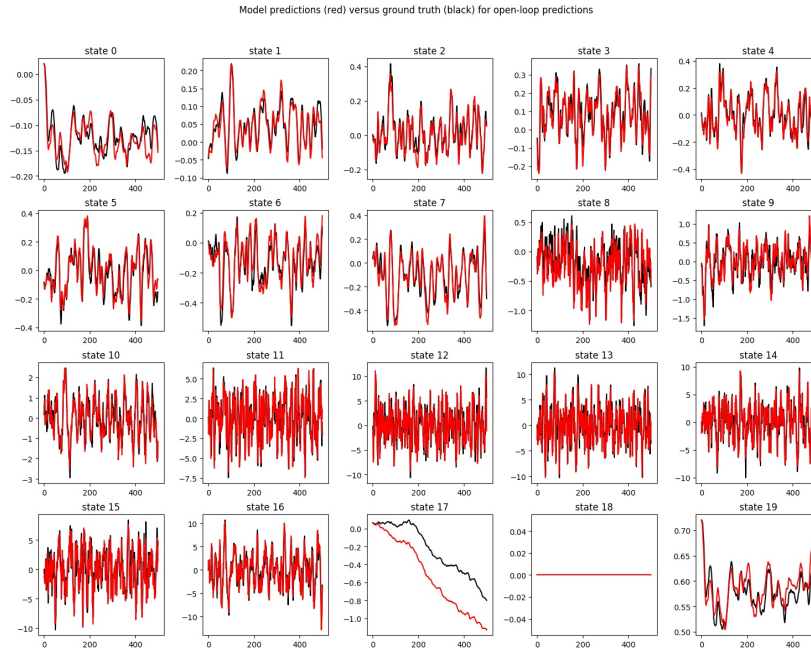Model predictions (red) versus ground truth (black) for open-loop predictions

Figure 1: Mostly accurate dynamics model predictions

(b)

State 17 is the most inaccurate prediction. This is the distribution mismatch problem since we have fixed dataset here. Intuitively, the earlier prediction error will accumulate and we finally end up with large prediction error.

1

Problem 2

	Random policy	Model-based policy
ReturnAvg	-152.333	29.7827
ReturnStd	29.1592	29.6192

Table 1: ReturnAvg & ReturnStd for random policy and model-based policy
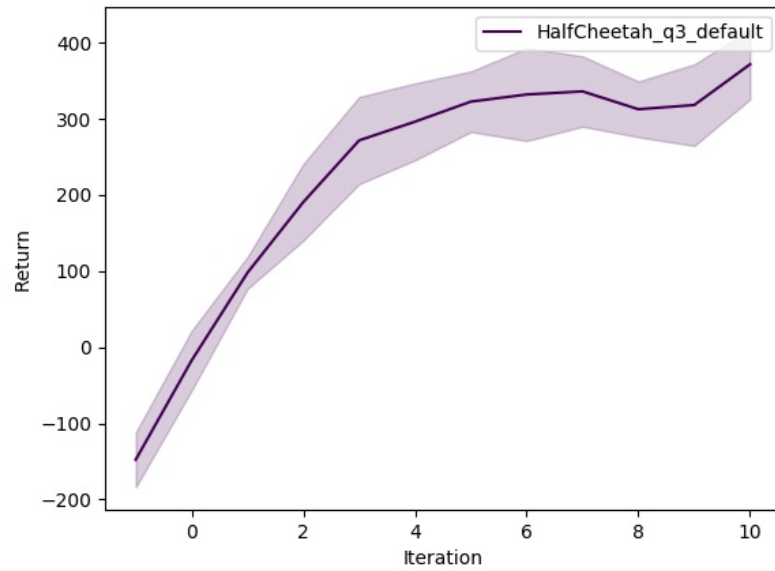
Problem 3a

Figure 2: Returns versus iteration for default setting
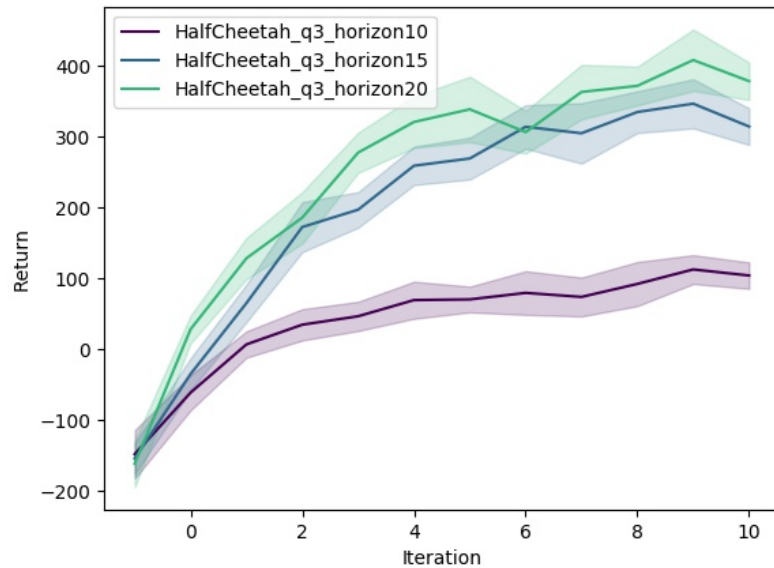
Problem 3b

(a)

Figure 3: Comparison performance when varying the MPC horizon

We can see that the higher horizon for the policy the better performance it generates. This makes sense intuitively since if the policy could plan further in the future, it should perform better.
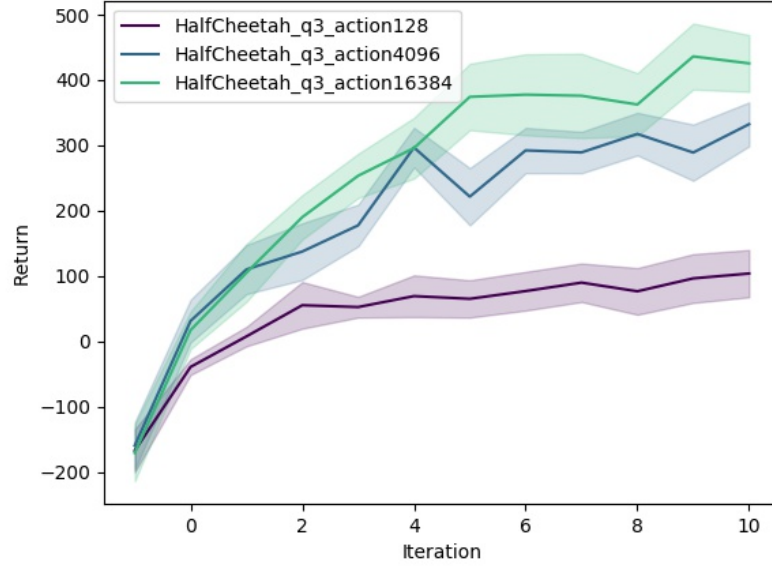
(b)

Figure 4: Comparison performance when varying the number of randomly sampled action sequences used for planning

We can see that the more actions we sampled, the better performance the policy generates, since by sampling more trajectories, we have larger chances to get optimal action.
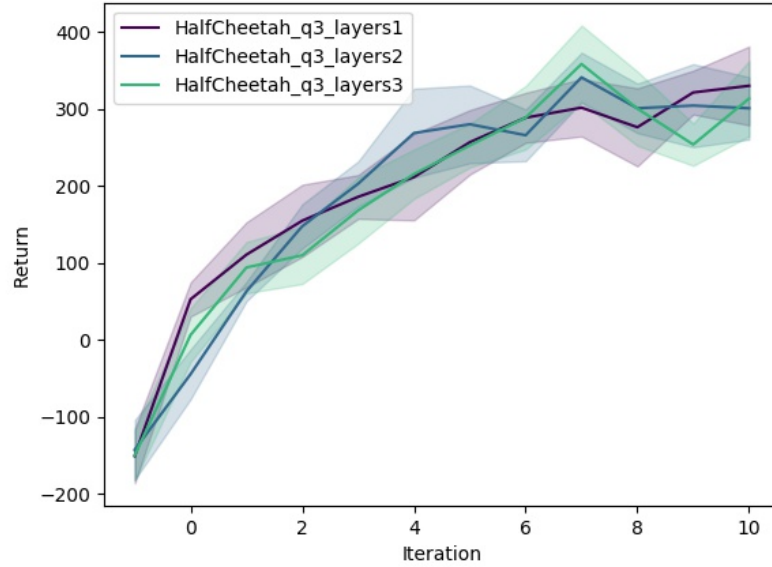
(c)

Figure 5: Comparison performance when varying the number of neural network layers for the learned dynamics model

There was a bug in the code... and that's why they are almost the same. The guess would be that the more layers we have, the better performance we should be able to get.
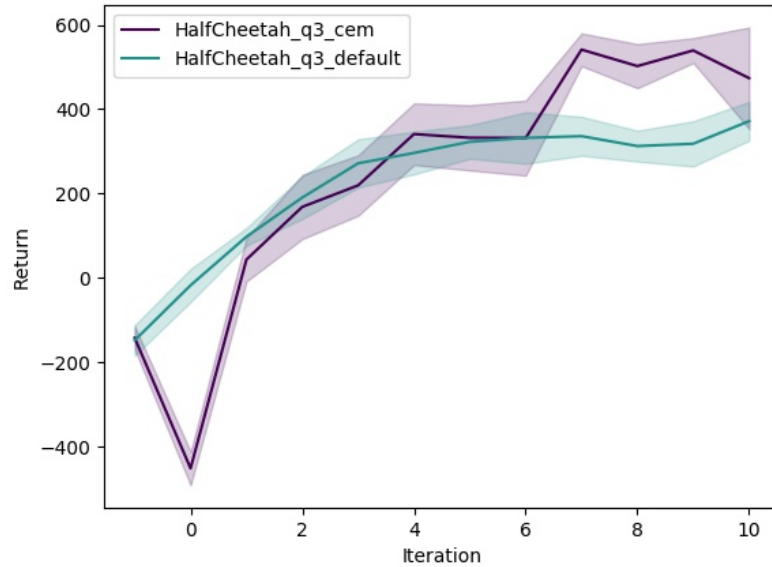
Bonus - CEM

Figure 6: Comparison performance CEM vs uniform random action selection

We can see that at the beginning, the return dropped a lot, this might be caused by our distribution initialization. Later on, the performance increased a lot to around 600 compared with the default 300-350. I think the performance could be better if we tune the parameter *maxits* and *Ne*.

For implementation, please check **_setup_action_selection_cem(self, state_ph)** in model_based_policy.py.

To run:

```
python main.py q3 --cem
```