Homework Assignment

Course: Deep Reinforcement Learning

Sophie Haas, Johanna Linkemeyer, Tjorven Weber (Group 9)

July 12, 2022

1 Review for Group 12

1.1 Task 1

We are unsure if your buffer works as intended. It looks like you start replacing the elements from the start of the buffer once the max-size is reached. However, instead, the first element should be 'deleted' and the new element appended to the end. So basically, you should have first 'moved' the entire array one to the left and then fill the last item.

1.2 Task 2

You created a nice, small functional model, looks good!

1.3 Task 3

The graph shows convergence after some epochs which is really nice, we like that you have included not only the separate values but also the blue average curve. You have a delay in your network and fulfil all requirements of the homework. Well done!

1.4 General

On a more general note, your code could need some more comments to ensure readability for others. You chose good variable names and since we worked on the same task, it was still understandable, but for a generally good coding practice, docstrings and comments where appropriate would be great.

2 Review for Group 29

Dear group 29, you did a great job with your homework! :)
Replay Buffer The replay buffer is nicely done and even includes multiprocessing for the sampling process. Impressive!
Network Training The organization of the training in the agent class is also very good, especially for code readability. Also, it seems to work quite well, as the control over the

lunar lander visibly improves with increasing epochs.
On top of that, the overall setup and organization is really good, as you included a nice readme, good comments, visualizations of the training etc.
Overall, the homework is very nicely done!
Best wishes, group 9

3 Review for Group 46

Hi everyone,
great job with the homework! The overall setup and organization of your code is very good. Your coding style is nice and pretty clear - a few more comments here and there might be helpful to others though.
Your replay buffer fulfills all the requirements and is implemented in a nice way.
Interestingly, you implemented two separate models. We just used a copy of the actual q-net as target-net but I think both is fine :)
Also, you managed to achieve great visualizations, both with the graph and the videos (which really wasn't that easy to get to work for us, so props to everyone who managed). Plus, the summary of the models and the usual prints are nice.
All in all, very nice work! It seems to work, so that's a clear success :)
BR, group 9