Homework 1 (Due 09/28/2023, 10PM EST)

Starter code repos/resources:

Stable baseline 3 (SB3) Github repo:

https://github.com/DLR-RM/stable-baselines3

SB3 documentation:

https://stable-baselines3.readthedocs.io/en/master/guide/rl.html

SB3 RL Baseline 3 Zoo (pretrained SB3 models):

https://github.com/DLR-RM/rl-baselines3-zoo

To submit HW1, please send the link(s) to the two Google Colab notebooks that are required in the homework

1. **Review the tutorial notebooks in SB3:**

   https://github.com/Stable-Baselines-Team/rl-colab-notebooks/tree/sb3

   [Requirements]:

   - Getting Started Colab Notebook
   - Saving, loading Colab Notebook
   - Multiprocessing Colab Notebook
   - Monitor Training Colab Notebook
   - Atari games Colab Notebook
   - PyBullet: Normalizing input features Colab Notebook
   - Pre-training using Behavior Cloning Colab Notebook
   - RL Baselines3 Zoo Colab Notebook
   - Advanced Saving and Loading Colab Notebook
   - Getting Started Colab Notebook
   - Gym Wrappers, saving and loading models Colab Notebook
   - Multiprocessing Colab Notebook
   - Callbacks and hyperparameter tuning Colab Notebook
   - Creating a custom gym environment Colab Notebook

2. **Train an agent to play any game/task in SB3 using A2C with Colab:**

   Hint: in this task, you don't need to modify the source code and just need to use the APIs provided in SB3 to run the Google Colab experiments and to train/test the RL agent.

   [Requirements]:

   - Pick a task/game in SB3.
   - Add a TensorBoard to visualize the training curves.
   - Include/record the final evaluation video.
   - Saving/loading the policy/model.

   [Submission]:

   Submit a link to the Colab notebook including the training/testing experiments and results.

3. **Compare the n-step advantage with n-step return (mentioned in the class), vanilla advantage, GAE, as well as MC advantage for A2C algorithm:**

   Hint: SB3 implements Generalized Advantage Estimation (GAE) for A2C. In particular, you can find the implementation of the advantage in the method *def compute_returns_and_advantage method* in *buffer.py* (*stable-baselines3/stable_baselines3/common/buffers.py*) (https://github.com/DLR-RM/stable-baselines3/blob/master/stable_baselines3/common/buffers.py). You can also play with the hyper-parameter (`gae_lambda`) to get different advantages without making model/algo implementation code changes.

   [Requirements]:

   - Compare the n-step advantage with the (vanilla) advantage, MC advantage, as well as GAE. Note that MC advantage is just optional for this assignment.

   [Software Setup for Development]:

   1. git clone the SB3 source code repo (development version) to your local machine (no need to install it on your local machine). https://stable-baselines3.readthedocs.io/en/master/guide/install.html
   2. Download the Google Drive Desktop (to sync with your Google Drive).

https://www.google.com/drive/download/

3. Sync the cloned SB3 code repo (on your local machine) with the one in your Google Drive using Google Drive Desktop.
4. Connect your cloned SB3 code repo stored in your Google Drive to Google Colab.
5. Install the cloned SB3 code repo in your Google Drive using Google Colab (install it on Colab).
6. Use any editor (e.g. Sublime, PyCharm, etc.) for the code implementation (modify the code in the cloned SB3 on your local machine).
7. Test your implementation and running experiments using Colab.
8. Push your code in the cloned SB3 on your local machine to your repo in your GitHub account.

[Submission]:

(We don't need a link to your Google Drive. We only need the link to your GitHub using the Colab notebook)

1. Create a new Colab notebook.
2. *!pip install git+"your github URL"*
3. Training/testing experiments using the Colab notebook.