Assignment 1
Reinforcement Learning
Instructor: Ashis Kumer Biswas, PhD
Spring 2023

Title: Conversion of the Non-gym "Goal-vs-Hole-v1" game to gym

You may remember the 2D non-gym game Dr. B introduced in this course named "Goal vs Hole" (v1, not v0), which in fact was a 4x4 board where an agent starts from the top-left cell and choose one of the 4 possible actions: going {up, down, left, right} based on "some" policy that requires training to find optimal path to maximize reward. The specifications of this non-gym agent-environment are detailed below:

```
---------------------------------------------
|           |           |           |           |
|  Start    |           |           |           |
|           |           |           |           |
---------------------------------------------
|           |           |           |           |
|           |           |  Hole     |           |
|           |           |           |           |
---------------------------------------------
|           |           |           |           |
|           |  Hole     |  Goal     |           |
|           |           |           |           |
---------------------------------------------
|           |           |           |           |
|           |           |           |  Hole     |
|           |           |           |           |
---------------------------------------------
```

- The environment is a 4x4 grid.
- It has 4 terminal states: one Goal and three Hole states; if a player moves into any of the four terminal states, the game is over.
- 12 non-terminal states.
- The agent always begins from the "Start" state which is located at the top-left grid cell.
- Rewards are defined as below:
  - 0 if on a non-terminal state,
  - -100 if on Hole state,
  - +100 if on Goal state.
- Full implementation of this non-gym version of the game can be retrieved at https://github.com/ashiskb/RL-workspace/blob/master/openai-gym/notebooks/10_non-gym-RL-intro-2.ipynb (in the https://github.com/ashiskb/RL-workspace.git repository)

Your tasks for this assignment
TLDR: Convert the non-gym implementation into gym.

Now, the long specifications:

1. OpenAI gym (now known as gymnasium) provides ways to create custom environments. (Reference 1: https://gymnasium.farama.org/tutorials/gymnasium_basics/environment_creation/ , Reference 2: https://towardsdatascience.com/creating-a-custom-openai-gym-environment-for-stock-trading-be532be3910e , Reference 3: https://towardsdatascience.com/beginners-guide-to-custom-environments-in-openai-s-gym-989371673952). First task is to review these so that you become familiar with creating any kind of environment. Also review the openAI-gym agent-environment implementations which are presented in this course already and can be retrieved at the GIT repository (https://github.com/ashiskb/RL-workspace/tree/master/openai-gym)

2. Now that you are familiar with the intricacies, think about creating the custom environment class for the "Goal vs Hole v1" environment. In the constructor define the action space and observation (i.e., state) spaces. Define the step, reset, render member methods so that your environment can be instantiated and be used similar to other openAI-gym environments.

3. Make sure to make your agent intelligent by training using any of the tabular approach you learned so far (i.e., Dynamic programming to optimize the Bellman's equation), and do not forget about adjusting the exploration-exploitation rates. You may begin with a discount factor, $\gamma=0.9$. Please check the non-gym implementation to find any other specific details if you would like to bring in to your own gym implementation.

4. Make sure to render multiple episodes of your intelligent agent as gif (animations) files.

5. Please use the name "my_gym_demo.py" as the driver program, and try to make it executable from a shell terminal. If you are unable to do that, no worries. Include a README file explaining your processes, and how to execute your program.

6. Submit the entire package as a zip file into Canvas.