**Dear all students,**

**Follow the instructions carefully, please:**

* Kindly you can find the following 3 questions you have whole **140 minutes** to answer the questions and **5 minutes of upload** time (not extendable more).
* For the questions that you need to answer in the text, add your answer in this world file and upload it into Moodle. For the questions that you need to write code, upload your python file into Moodle. Rename both files to your English name or student ID.
* You can use Slides and class code examples.
* Must not compress as zip or use WinRAR, only Attah one python code and one word files answer.
* No need to mention any communication during the exam is acceptable.
  + phones only in the pocket all the time, no open apps or browser in background and using the internet is completely prohibited, (you must disconnect your personal laptop or computer from the network during the exam).

**Good Luck.**

--------------------------------------------------------------------------------------------------

English Name:

Chinese name:

Student ID:

--------------------------------------------------------------------------------------------------

**Questions:**

1. A) In Value Function Approximation when we should use state-value function and when state-action function? Explain clearly with one agent example. B) explain why we have Multiple forward pass problem in Deep neural network part of Deep Q-learning algorithm and how we can solve it?

Answer:

Gives us the ability to compute Q-values for all possible actions in a given state with only a single forward pass through the network

1. A) In both DQL and Actor-critic we saw how to define two deep networks. What are the differences of the target and predicted networks in DQL with actor and critic networks in Actor-critic approaches? Explain the 2 main differences.  
   B) Consider a 2 DoFs robotics arm and explain what is the exactly output of our DRL algorithm for continues action space (output of actor network)? Shortly, you need to explain what is the output by stating numerical range and it’s meaning first, then how agent executes that output clearly. You can show example with values or draw in word file if needed.

Answer:

1. Considering the Q-Learning code provided for OpenAI Gym environment and the 'FrozenLake-v1' environment from Gym. Your objective is to enhance specific elements of this script and conduct an analysis of the agent's performance.
2. **Apply Variable Learning Rate (20%):** Adapt the Q-Learning algorithm with a variable learning rate (α). Rather than a fixed α value, create a schedule where α starts at a higher value and decreases progressively as the number of episodes goes up. Provide a justification for your chosen schedule and plot comparisons.
3. **Alter the Reward Dynamics (30%):** The current reward system is basic and depends on the environment. Implement a novel reward or penalty mechanism within the algorithm (for instance, extra rewards for achieving specific states or penalties for certain actions). Clarify how your modifications could enhance the learning efficiency of the agent.
4. **Implement a Different Exploration Technique (30%):** Substitute the current epsilon-greedy strategy with a different exploration approach. Explain this new method and discuss its potential advantages over epsilon-greedy in certain situations.
5. **Discussion on the Algorithm (20%):** Reflect on the functionality of the Q-Learning algorithm in the 'FrozenLake-v1' setup. Address any challenges or limitations and propose possible enhancements or alternative strategies that might be utilized to overcome these issues. (answer this part in the word file)

Answer:

Upload your python code (don’t compress) and comment in front of all changed lines that is related your answers like #A#, or #B#, …. You better to add short comments for your motivations in the code.