Programming Assignment 1 - Policy and Value Iteration

Policy on Academic Integrity and Plagiarism

In this assignment, we place a strong emphasis on academic integrity, individual effort, and the authentic learning experience. It is crucial that students understand and adhere to the following guidelines to maintain the academic integrity of the assignment:

1. **Independent Work:** All students are expected to complete the assignment independently. This means that you should not collaborate with others on the assignment unless explicitly allowed by the instructor.
2. **No Code Copying:** Looking at online code repositories or any form of external code is strictly prohibited. While you may refer to online resources to understand the algorithmic concepts, you should not directly copy or replicate code from any online source. The purpose of this assignment is for you to implement the policy and value iteration algorithm yourself, which is an essential part of the learning process.
3. **Original Solutions:** Your code and solutions should be your own original work. Avoid any form of plagiarism, including direct copying, paraphrasing, or reusing code or solutions from online or offline sources.
4. **ChatGPT Assistance:** While you can seek guidance and clarification from ChatGPT regarding algorithmic concepts and problem-solving strategies, you are not allowed to request or receive direct code solutions from ChatGPT. The use of ChatGPT should be for educational guidance and not for generating code solutions.
5. **Proper Citation:** If you do refer to external sources or tutorials to understand the algorithm, be sure to provide proper citations in your code comments or accompanying documentation.

Consequences of Academic Misconduct:

Engaging in academic misconduct, including plagiarism or code copying, is a violation of academic integrity and will result in appropriate consequences, as outlined in the course syllabus or your institution's academic integrity policy. These consequences may include grade penalties, course failure, or disciplinary actions.

Seeking Help:

If you find yourself struggling with the assignment, feel free to seek help from the instructor, teaching assistant - Muhammad Mahen Mughal, or fellow students through appropriate channels. We are here to support your learning journey and are more than willing to provide guidance and assistance in a manner that adheres to the principles of academic integrity.

Remember, the objective of this assignment is not just to complete it but to gain a deep understanding of policy and value iteration and reinforcement learning. Your effort and learning experience are highly valued, and we expect your work to reflect your individual understanding and implementation of the concepts covered in the assignment.

Please read and understand this policy carefully. If you have any questions or concerns regarding academic integrity or the assignment itself, do not hesitate to reach out for clarification.

By submitting your assignment, you acknowledge that you have read and agreed to abide by this policy on academic integrity and plagiarism.

Assignment Description: Policy and Value Iteration in OpenAI Gym's Frozen Lake Environment

Objective:

In this assignment, you will dive into the world of reinforcement learning by implementing policy and value iteration algorithms on the OpenAI Gym's Frozen Lake environment. The primary goal is to understand and apply these fundamental concepts to find the optimal policy in two different grid-world environments, 4x4 and 8x8.

Assignment Components:

1. **Policy Iteration (4x4 and 8x8):**

- You will implement the policy iteration algorithm to find the optimal policy for both the standard 4x4 and larger 8x8 Frozen Lake environments.
- The optimal policy should be calculated and displayed in a tabular format.

2. **Value Iteration (4x4 and 8x8):**
    - The assignment also requires students to implement the value iteration algorithm to derive the optimal policy.
    - Similar to policy iteration, students should display the optimal policy in a table for both the 4x4 and 8x8 environments.
3. **State-Value and Action-Value Functions:**
    - You are tasked with computing the state-value function and action-value function for every state within the Frozen Lake environments.
    - These functions should be presented in a clear and tabular format for easy reference.
4. **Test the policy on massive iterations and answer questions based on it.**
    - The question is listed on the Starter Jupyter Notebok.

Assignment Expectations:

- **Optimal Policy:** Students should demonstrate a clear understanding of how policy and value iteration work to determine the optimal policy. The computed optimal policies should reflect this understanding.
- **Tabular Representation:** Tables displaying the optimal policies, state-value functions, and action-value functions should be neatly organized and comprehensible.

Some Starter code has been provided to you so that you don't have to reinvent the wheel.

Learning Outcomes:

This assignment aims to achieve the following learning outcomes:

- Mastery of the policy and value iteration algorithms.

- Understanding the impact of environment complexity on algorithm performance.
- Proficiency in presenting results in a structured and informative manner.

By completing this assignment, students will gain hands-on experience in reinforcement learning, policy optimization, and the practical application of these concepts in real-world scenarios.