RL homework: Markov Decision Processes

2-3 bonus points

Problem:
Choose one of the following environments from OpenAI Gym Atari [https://gym.openai.com/envs/#atari] with the proposed simplifications:
● Breakout-v0
● MsPacman-v0 (with static ghosts)
● SpaceInvaders-v0 (with static invaders)

Assignment:
Implement the chosen environment (using the same interface as the one in Gym [https://gym.openai.com/docs/#environments]) in Python according to the following specifications:

2 points:
● The environment is FULLY observable
● No graphical rendering is needed
● Use ONLY discrete states (i.e., instead of images, you will return the real environment state as observation)
● Use a reduced state space (e.g., the state of the grid should be small - at most 10x10). Adapt the number of walls, ghosts or invaders accordingly
● Use ONLY discrete actions
● Except the agent (i.e., the player) everything is STATIC
● Write a simple test program where you create an instance of the environment and you execute a sequence of at least 5 random actions, printing (or showing in any form you prefer) the observation you get out of it.

1 more point: expose the transition model from the environment to the agent, implement the policy iteration algorithm and run it on your environment.

FAQ:
● No groups are allowed, each homework is for 1 individual.
● Plagiarism will be verified among students, previous students and internet archives. In case of plagiarism, the homework will be evaluated as **-3 points (they will be subtracted from the final score)**
● What is an interface?
https://stackoverflow.com/questions/2866987/what-is-the-definition-of-interface-in-object-oriented-programming