Programming Assignment 4

Implementation of a First Visit Monte Carlo (FVMC) based control algorithm

In this assignment, you have to modify the FVMC based policy evaluation code that was already given in order to make a FVMC based control algorithm

1. Modify the FVMC code to first estimate state action values rather than state values
2. Introduce a policy improvement step
3. Make sure that the policy which is used at every step chooses to explore all actions uniformly at random (with some probability that diminishes with the number of episodes)

Comparison of FVMC and EVMC based control algorithms

1. In this section, we will compare the "learning performance" of FVMC and EVMC algorithms.
2. Consider the "Forest managment" problem from our previous assignment
3. Implement a class to model the Forest Management MDP - this should especially have a function that simulates how the forest's state evolves from one time step to the next under the influence of an action (either cut or wait). The class should be written in a way similar to how the "Machine repair" example was done in the code discussed in class.
4. We will apply both FVMC and EVMC based control algorithms for this Forest management problem.
5. For both algorithms, for each episode keep track of the total discounted reward from time 0. Then we have $G_{FVMC}[e]$ and $G_{EVMC}[e]$ where e is the episode index. Plot and compare $G_{FVMC}[e]$ and $G_{EVMC}[e]$ as a function of e.

Submission of this assignment will be through Moodle. For this assignment you have to submit a Jupyter notebook. The submission file should have the format "[SCCODE]_[NAME]_PA4.ipynb" (where you have to insert your SCCODE and NAME).