ECEN 689-605/700 Dependable Learning Systems Fall 2018 Homework 2

Due 1pm U.S. central time, September 19, Wednesday
Submission must be through ecampus
No late submission accepted!

Please implement a reinforcement learning system consisting of an agent and an environment. The environment is specified by the model in homework 1, including the state transition probabilities and reward functions. Please note the agent does not know the probabilities and reward functions in advance. When your agent take an action at certain state, you can simulate the environment to obtain the reward and the next state. Please implement Sarsa (10 points) and Q-learning (10 points) for the agent in C/C++, java or Python. Finally, please report the Q(s,a) values for all (s,a) pairs. The parameters are:

- ε=0.2
- learning rate α=0.5
- discount factor γ=0.5
- Each episode start with s_0, and end with s_H, which is the terminal state 0 reward, and the subscript here indicates time step. So the agent will take H actions in each episode.
- For each method of Sarsa and Q-learning, please run K episodes. The first episode initializes with s_1 defined in homework 1, the 2nd episode starts with s_2, … the fourth episode starts with s_1, and so on.
- For each method, please generate results for the following two settings
    - H=9, K = 9
    - H=27, K = 27