Group U Frozen Lake

Cong Peng 210064410, Qinlin Xie 210171822, Qinjian Shangguan 210273126

School of Electronic Engineering and Computer Science, Queen Mary University of London, UK

.peng@hss20.qmul.ac.uk, q.xie@hss20.qmul.ac.uk,

q.shangguan@hss20.qmul.ac.uk

email[@qmul.ac.uk](mailto:%7d@Springer.com)

1. Explain how your code for this assignment is organized. Did you make implementation decisions that deviate significantly from what we suggested?
2. How many iterations did policy iteration require to find an optimal policy for the big frozen lake? How many iterations did value iteration require? Which algorithm was faster?
3. How many episodes did Sarsa control require to find an optimal policy for the small frozen lake? How many episodes did Q-learning control require? **Hint:** you may use policy evaluation to compare the value of each policy obtained by these algorithms to the value of an optimal policy.
4. In linear action-value function approximation, how can each element of the parameter vector ***θ*** be interpreted when each possible pair of state s and action a is represented by a different feature vector ***φ***(s, a) where all elements except one are zero? Explain why the tabular model-free reinforcement learning algorithms that you implemented are a special case of the non-tabular model-free reinforcement learning algorithms that you implemented.
5. Try to find an optimal policy for the big frozen lake by tweaking the parameters for Sarsa control and Q-learning control (maximum number of episodes, learning rate, and exploration factor). You must use policy evaluation to confirm that the resulting policy is optimal. Even if you fail, describe your experience.

References