Assignment 5 Report for Part A

CSE 415: Introduction to Artificial Intelligence
Winter 2019

Stephan Hagel
Student Number: 1870562
sthagel@uw.edu

Question 1a:

It took 4 iterations to turn 9 of the 27 states green.

Question 1b:

It took 8 total iterations to turn all 27 steps green. Running a 9th iteration does not change the values anymore.

Question 1c:

The policy shown suggests to "go right" in every single state. This happens, because every state has a value of 100.0 and therefore the algorithm gives each state the same value as the goal. Therefore it doesn't bother to go to the goal, since the values are the same.

Question 2a:

Just as in question 1b, it took 8 steps for the initial state to have a non zero value.

Question 2b:

This question does not exist in the assignment.

Question 2c:

The policy shown results in finding the quickest path from each state. Therefore it is already the optimal policy. This happens, because the state values, and also their Q-values, are increased move by move, so the first policy, in which every state has been updated at least once, will already know the fastest path from each state.

Question 2d:

After 56 steps, delta_max was lower than 10^{-8}, so the MDP was considered as converged.

Question 2e:

The policy stays the same after the additional VI steps, since the old policy was already optimal, as explained in question 2c.

Question 3a:

The value of the start state is 0.82. The policy implies to go to the goal, which is worth more points (100 instead of 10), as long as the path to it is not longer by more than one step.

Question 3b:

The start state now has a value of 36.9. This policy implies to always go to the goal, which is worth more points, except from directly at the other goal, where there is no other choice.

Question 4a:

The agent went off the plan in 6 out of 10 cases.

Question 4b:

The agent still reached the goal in 8 out of 10 cases, this means two times it went off the plan but still arrived at the better goal.

Question 4c:

One time the agent ended up one step away from the goal, the other time it was three steps away from the goal.

Question 4d:

The top path of the state space (as drawn in the GUI) and the two states, drawn closest to the worse goal have never been visited by the agent.

Question 5a:

No, as I have seen in question 2, the policy converged much faster than the actual state values.

Question 5b:

Revisiting states would only be needed mostly on states on and around the golden path. States, that rarely, if ever, get visited, like the top part of the state space, would not need many visits.