Assignment 6 Report for Part A
Junhao Zeng

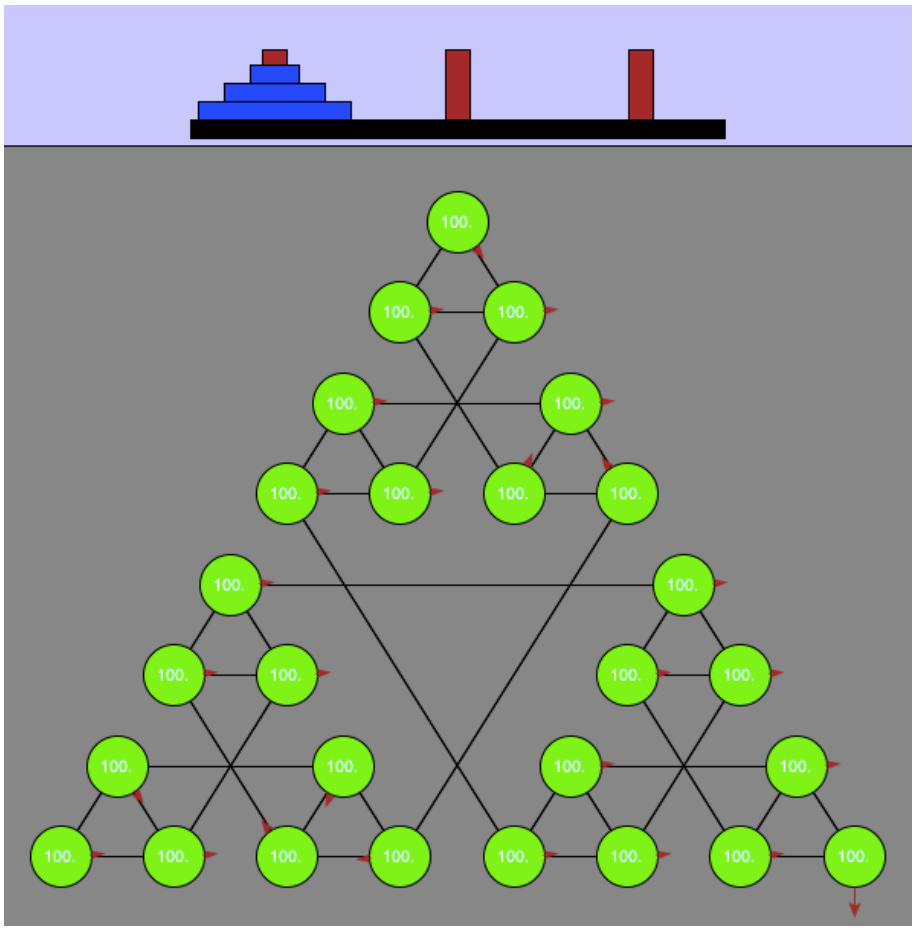
1

1a
answer : 4 iterations are needed to turn 1/3 of the states green

1b
answer : 8 iterations are needed to get all the states to 100

1c

answer:

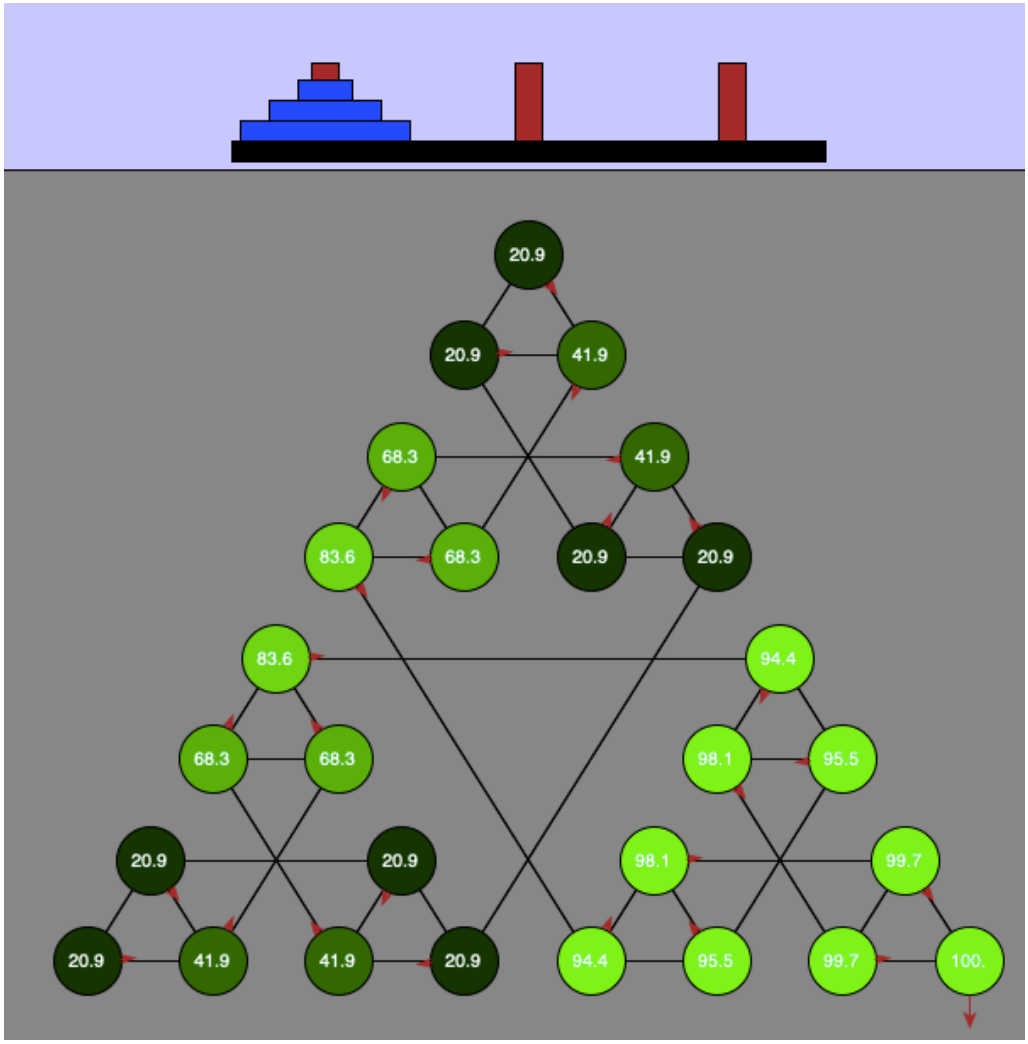


This policy is a bad policy.
Because this policy has a lot of suggested actions which are illegal, which means this is not a good policy. Because the noise is zero, we could never arrive goal state by following the this policy.

2.

2a
answer : 8 iteration are needed

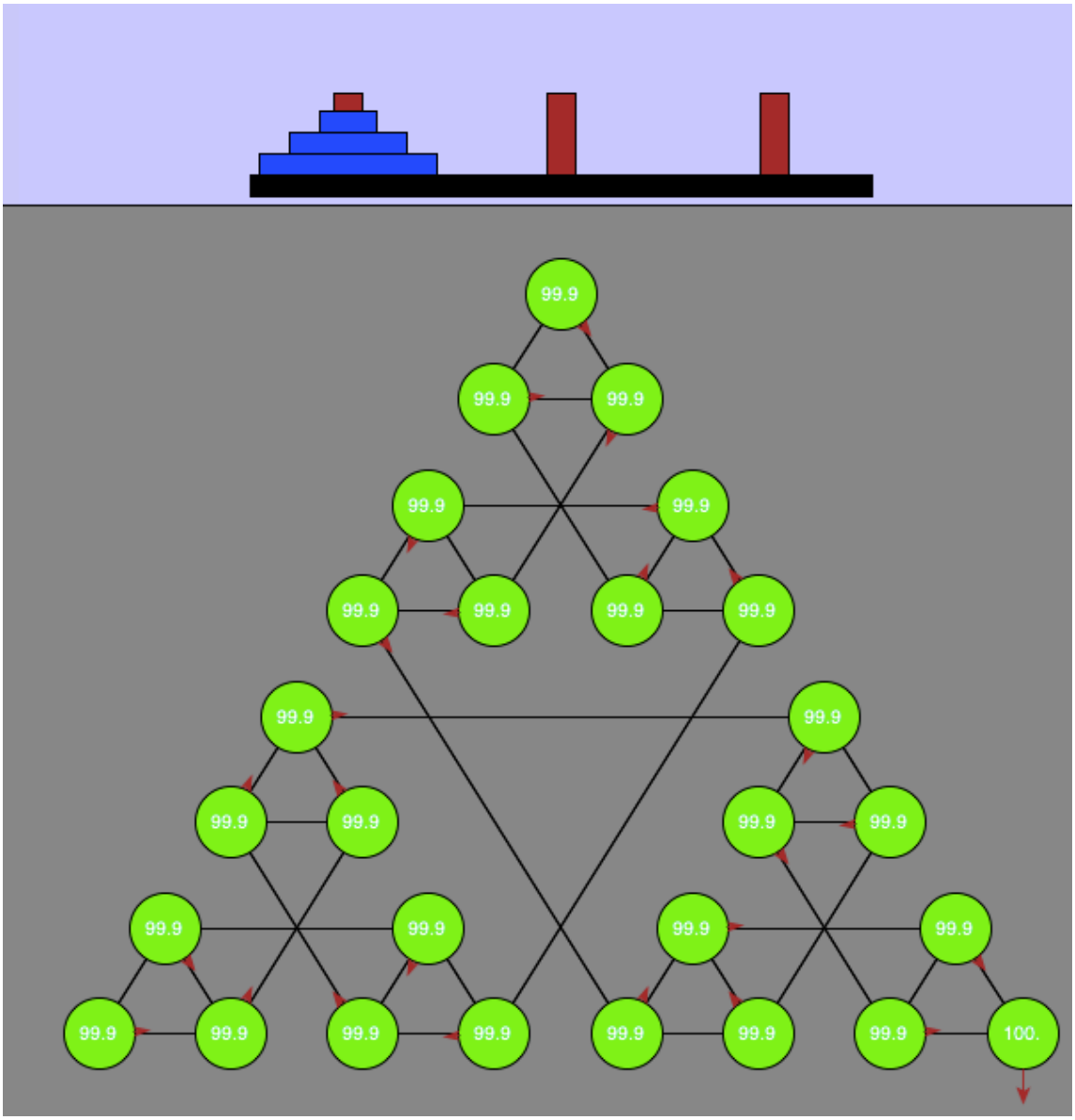
2c



It is a good policy, because gramma is 1, so it has already found the best policy to get the goal state.

2d
answer : 56 iterations are needed

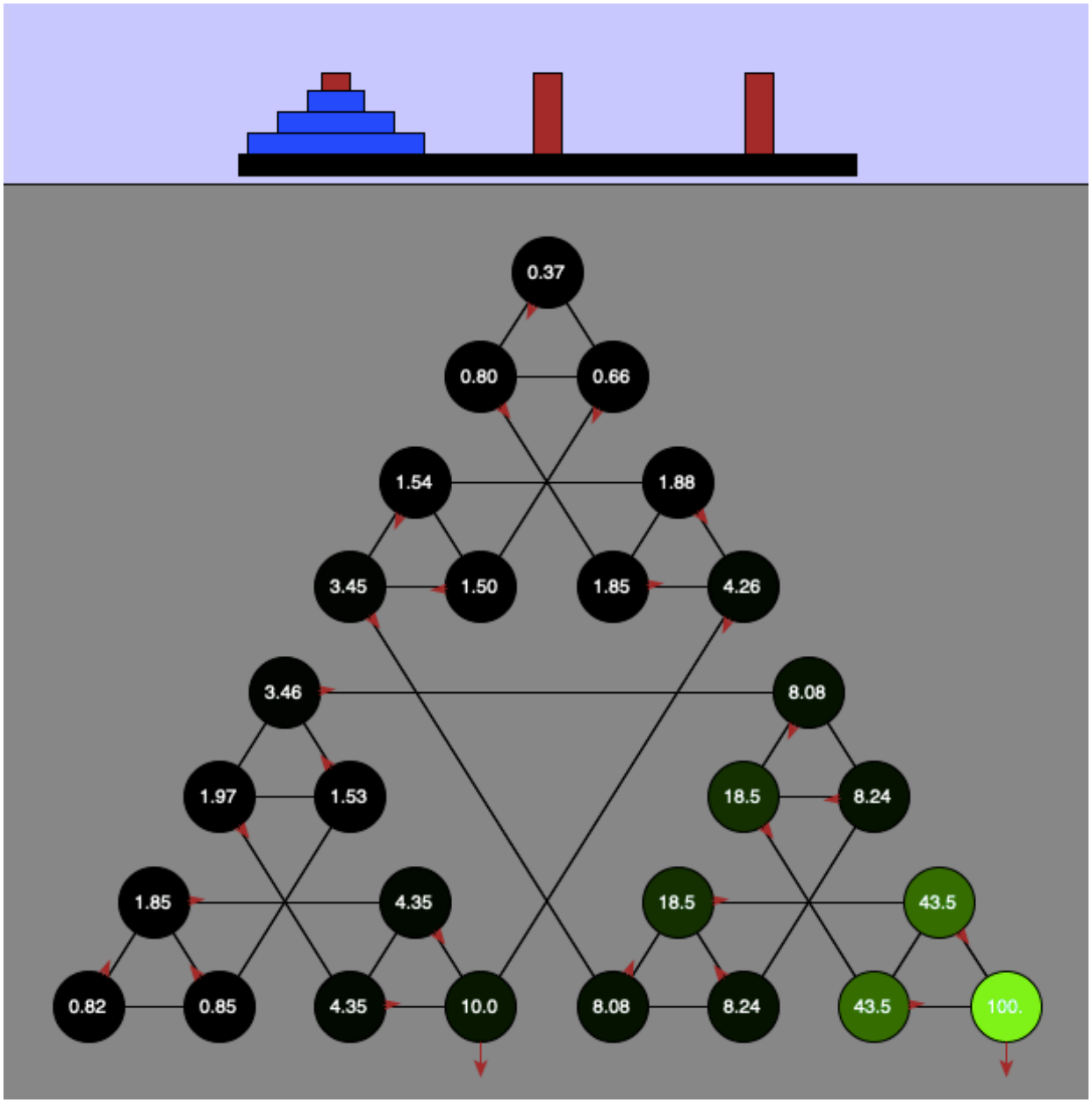
2e



answer : There is no change in policy, because the gramma is 1.

3.

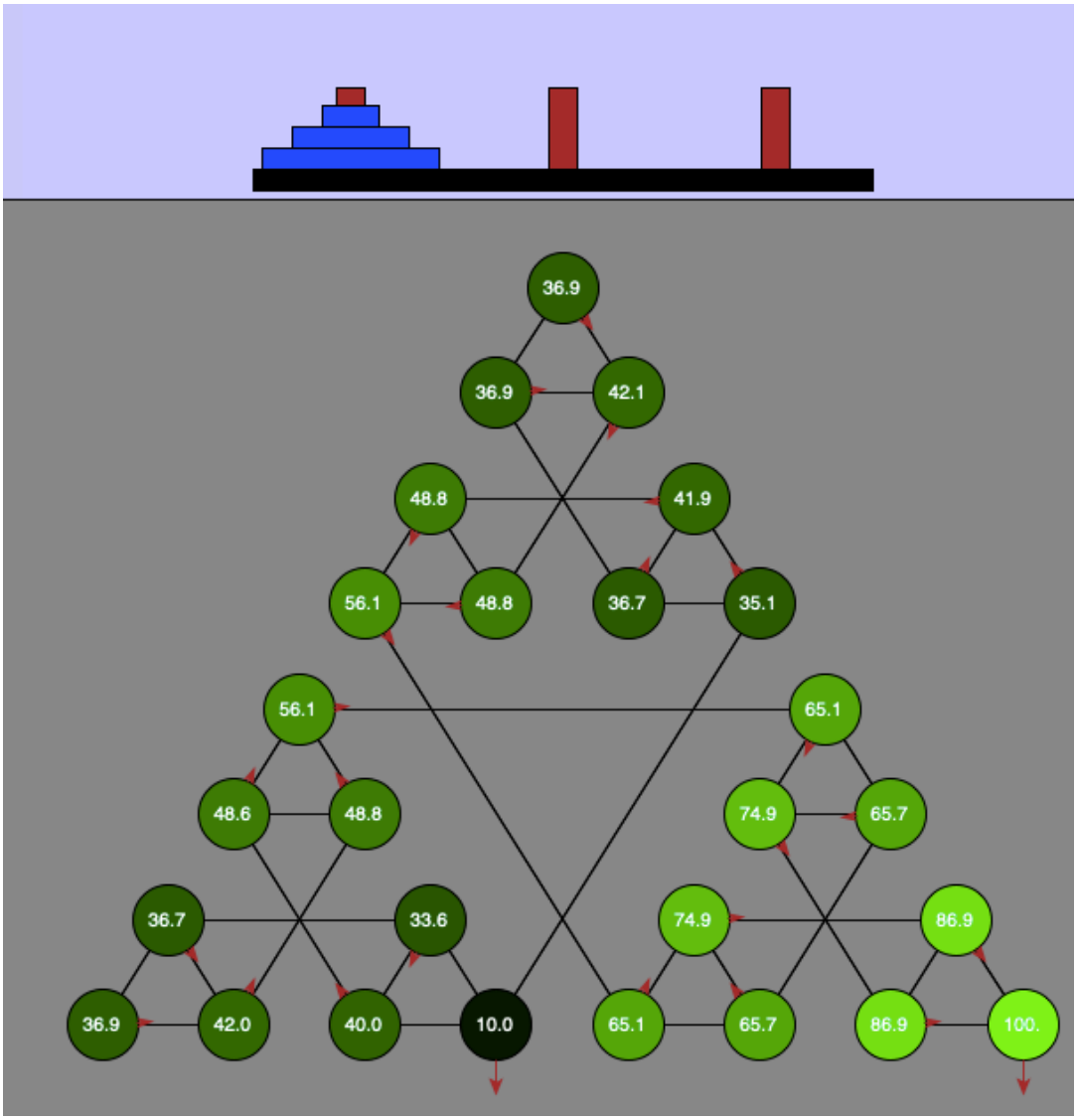
3a



answer : The policy indicates agent should get out as soon as possible instead of getting higher reward from goal state, since the gamma is too small.

the start state value is 0.82

3b.



answer : The policy indicates that we need to avoid to get the goal state 10, and try to get goal state 100, because gramma is 0.9, which is not too small.

The start state value is 36.9

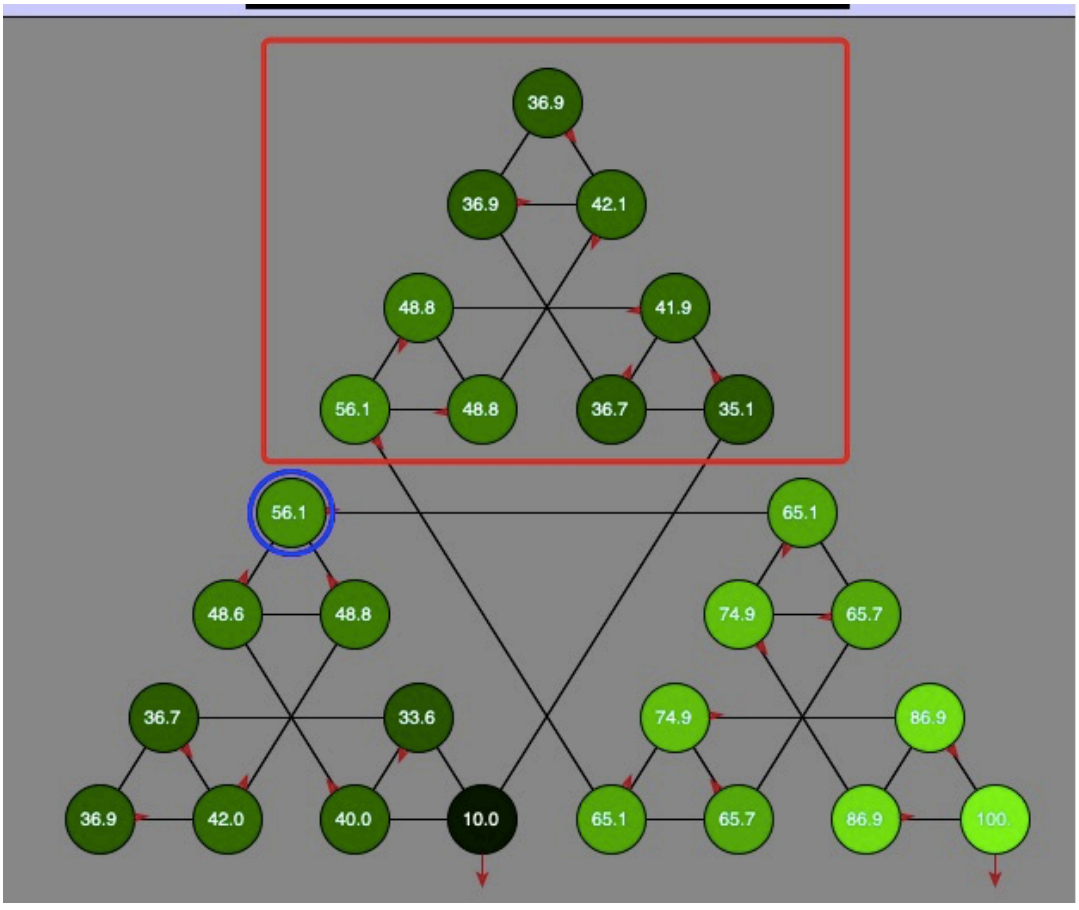
4

4a
four runs is off the plan in ten runs

4b
four runs which the agent arrive in the goal state with golden path in ten runs
eight runs which the agent arrive in the goal state in ten runs

4c
one run is 4 steps away
one run is 1step away

4d



the part in the red rectangle seems never to be visited by the agent

5

5a
answer : no, if the gamma is one, we only need every state have value to get the good policy. However, if gramma is less than 1, then it is essential to converge.

5b
answer : It is very important, because the agent does not know T and R at all. The agent doesn't even know S, but it can recognize states it has visited before, which means agent need to explore all states, even though some states are far from goal state. Besides, our agent need to visit every state a lot to get sample and update the Q value. Therefore, the agent needs to visit all states a lot.