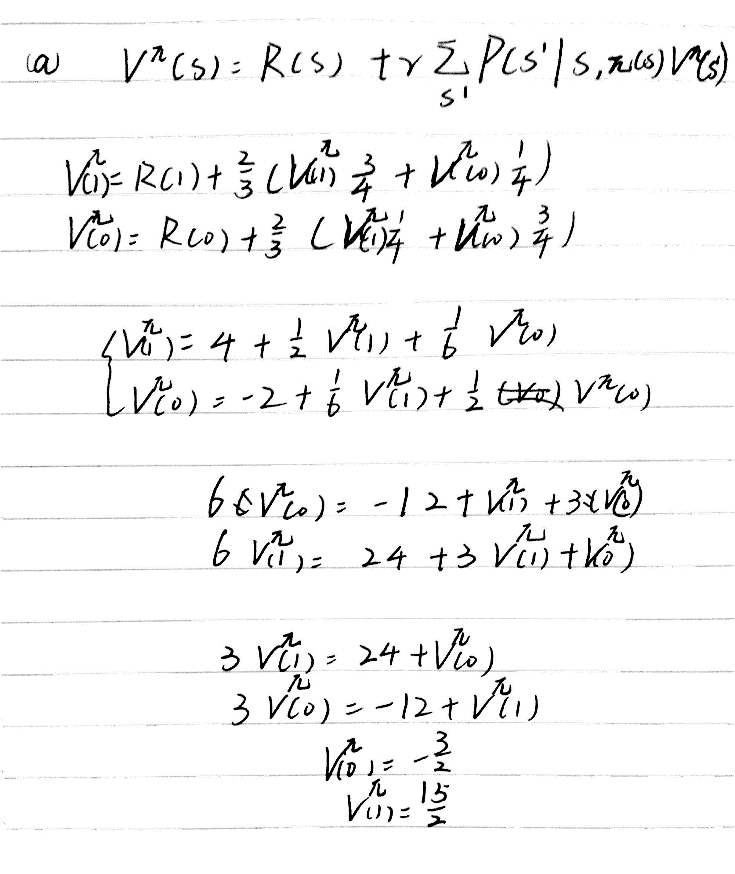
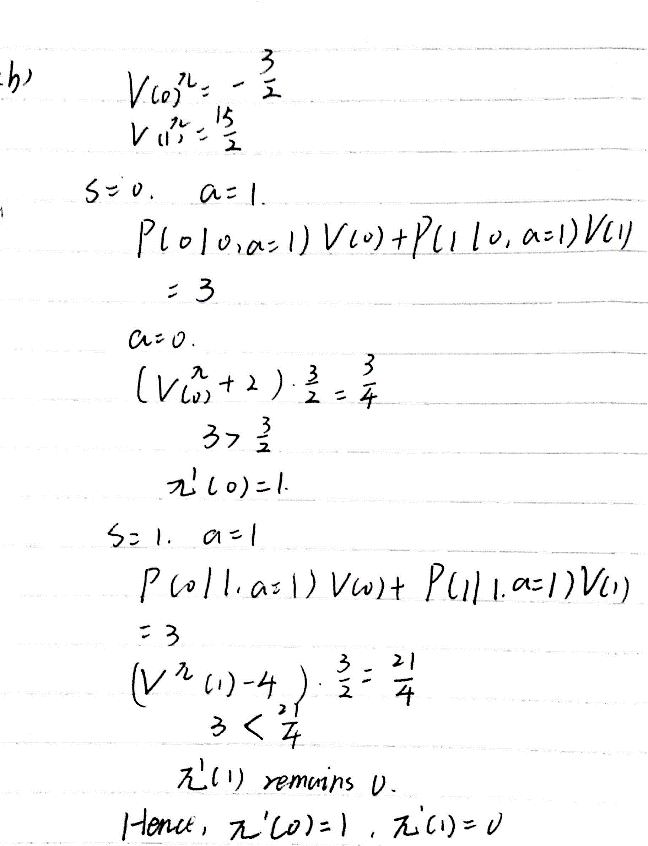
CSE150 PA5 Report

Yunfan Kang A14282934

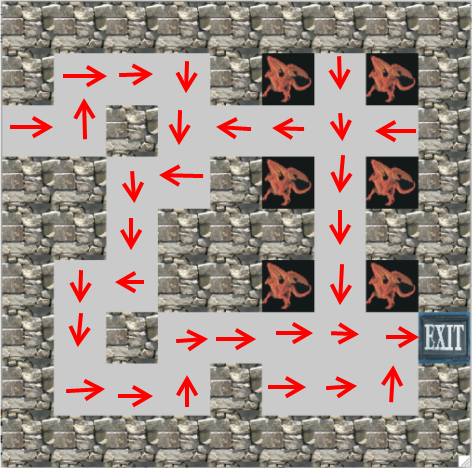
1(a)



b.



2.



For problem 2, I wrote a file named MDP.py to give the definition of MDP class. I set the transitionFunction as a defaultdict(list) with keys being (s, action) and values being lists of (s’, action). Object mdp is initialized in both value\_iteration.py and policy\_iteration.py with the data in the files provided. The following description will only focus on the functions.

1. Value iteration:

Create two lists U and U2, both of size 81, and initialize all their entries to be 0.

Set maxChange = 0

While True:

Copy the entries of U2 to U.

Set maxChange = 0

For every state s of the mdp:

For every action:

Get all (s2, prob) pairs through transitionFunction[(s, action)], calculate prob\*U[s2] and put the sum into a list. Finally take the max of that list to decide