CS685 Homework 4

Yue Hao, yhao3@gmu.edu

1. The problem can be modeled as a MDP, where each vertex is a state and an action is moving from one vertex to its neighboring vertices. The transition model is simple, as the probability of moving to the neighboring vertices is 1 and non-neighboring is 0. The rewards is specified as follow, only the goal state (i.e. v_g) is 1, all other states are 0. Because we have a discount parameter $0 < \gamma < 1$, long paths will be discounted more times than short paths. Therefore the shortest path is the optimal solution of MDP.