

MDP Questions

1. Synchronous value iteration required 71 iterations, and asynchronous value iteration required 50.
2. Command line arguments:

a.

```
discountFactor:      0.99
maxStateUtilityError: 1.0E-6
keyLossProbability:   0.5
positiveTerminalReward: 1.0
negativeTerminalReward: -1.0
stepCost:            -0.04
solutionTechnique:   av
```

These are the default arguments provided, and result in a path that goes straight North to get the key, returns straight South, then goes East to the goal.

b.

```
discountFactor:      0.99
maxStateUtilityError: 1.0E-6
keyLossProbability:   0.5
positiveTerminalReward: 1.0
negativeTerminalReward: -1.0
stepCost:            -0.02
solutionTechnique:   av
```

If all parameters are kept the same except the step cost, which is halved, the specified policy will be to go West twice, North twice, West once (into the Lose-Key square), then North and East to acquire the key. Once the key is taken, the path is as in (a).

c.

```
discountFactor:      0.99
maxStateUtilityError: 1.0E-6
keyLossProbability:   1.0
positiveTerminalReward: 1.5
negativeTerminalReward: -1.5
stepCost:            -0.02
solutionTechnique:   av
```

These parameters cause the policy to be to return directly between the lose-key and the -1 (now -1.5) square on the return from fetching the key.

3. discountFactor: 0.9
 maxStateUtilityError: 1.0E-6
 keyLossProbability: 1.0
 positiveTerminalReward: 1.5
 negativeTerminalReward: -1.5
 stepCost: -0.02
 solutionTechnique: av

These are the parameters from (2c), but with a discount factor of 0.9 rather than 0.99. This causes the policy to be to go east no matter whether the agent has the key or not. With this policy, the maze is not solvable, which is a fairly drastic change.