EE209AS Computational Robotics

Prof. Ankur Mehta

Lab1

1. MDP system

General description on MDP code:

There are total 4 classes and one main program written in Matlab (Class: MDP, MDPUI, PolicyIteration, ValueIteration; Main: SolveMDP). Class MDP consists of all functions required in problem 1. MDPUI includes all functions necessary to draw robot motions. Class ValueIteration and PolicyIteration contains corresponding functions to solve each problem. SolveMDP uses those classes and defines the scenarios.

Constructor: MDP(W, L, Pe, IsDirectionalRewardOn)

W, L are size of the grid (x,y)

Pe is error probability in motion

IsDirectionalRewardOn, is true, reward can be gain if the robot is facing down

a) Function: createStateSpace() in MDP class generates state space set. Each column includes three states, x, y positions and robot orientations: (s=[x;y;h]). The number of the columns represents the size Ns = 768 in this case since there are 64 possible combinations of x y positions, and 12 orientations at each cell (8*8*12).

b) Function: createActionSpace() in MDP class generates action space set. Each column includes two motions, {Still, Forwards, Backwards} and {Still, Right, Left}. The number of the columns represents the size Na = 7 in this case since the possible combinations are {(Still, Still), (Forwards, Still), (Forwards, Left), (Forwards, Right), (Backwards, Still), (Backwards, Left), (Backwards, Left)}. *Note that turn without moving forwards or backwards are not allowed.

c) Function Psa = calcActionProb(s,a,sPrime) in MDP class returns probability of getting given s'. Pe is specified when MDP class was instantainous.

d) Function: sPrime = motionSequence(s,a) returns s' given that current state, s, and action to take with error probability, Pe (defined when generating instance of MDP class). If Pe != 0, then there would be three possible s' returned using this function.

Function sPrime = generatePosibleSPrime(s,a) will returns all possible s' states (1 or 3 in this case), instead.

2. Planning Problem

a) Function reward = prob2Reward8By8(s) returns corresponding rewards at each cell for this problem. If IsDirectionalRewardOn is true positive reward can be obtain iff the robot is facing down.

3. Policy Iteration

a) PolicyIteration constructor populate a matrix (8*8*12) with stand still action at all cells.

b) Function: showPolicyTrajectory(s0,Pie) plot the robot trajectory with given initial state (s0=[x,y,h]) with policy matrix Pie. Inside of the function, it's using MDPUI class to formulate grid world and the robot identification.

c) Since the initial policy is stand still at everywhere, the robot stays at initial position(s=[1,6,6])

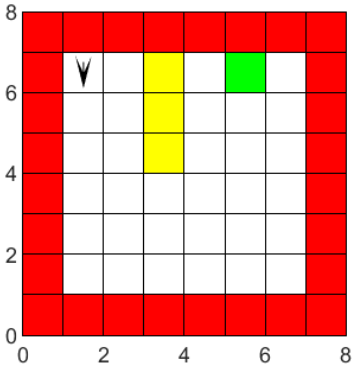


Figure 1: The robot trajectory with initial policy $\pi_0$

d) Function: Vpie = calcValue(Vpre,Pie) generates values of the current policy Pie using following equation:

$$V_{i+1}^{\pi_k}(s) \leftarrow \sum_{s'} T(s, \pi_k(s), s') \left[ R(s, \pi_k(s), s') + \gamma V_i^{\pi_k}(s') \right]$$

Where LHS is value at specified state, s, and T, represent probability of reaching state, s' from original state, s, given a action specified by the policy, $\pi$. R is reward at the new state, s', and the V(s') is value at s' cell at previous iteration. $\gamma$ is a discount factor on the value.

For example,, with initial policy, V(s=[6,7,*]) is equal to 1 if error probability, Pe =0 because the reward is 1 given action stand still (from s=[6,7,*] to s=[6,7,*]) and $V_0 = 0$.

e) The value of the trajectory is zero at s=[1,6,6], since the action is stand still and the cell itself has no reward.

f) Function [Vpie,Pie] = calcValueAndPolicy(Vpre,Pie) returns refined policy and corresponding value matrix with given policy to refine and values from previous step, using following equation:

$$\pi_{k+1}(s) = \arg\max_a \sum_{s'} T(s,a,s') \left[ R(s,a,s') + \gamma V^{\pi_k}(s') \right]$$

Where LHS is refined policy and RHS is calculating value with all possible actions at the cell, and choosing action with maximum value at that step.

For example, after initial policy evaluation, at s=[6,6,0] (one cell down and toward the goal) the actions which maximize the value could be {(Forwards, Still), (Forwards, Right), (Forwards, Left)}, all of which results in reward 1. Due to the behavior of "max" function in Matlab, the first action stored in the action space matrix will be picked if the value is identical for different actions.

g) Function: [V,Pie] = calcOptimalValueAndPolicy() returns optimal value and policy matrix. Initial policy is generated when calling constructor of the class. In order to obtain optimal policy and values, at first the policy is evaluated and refined such that the values are maximized at that time step. At converge, the policy will not change after policy refinement. ( $\pi_{k+1} = \pi_k$)

h) The trajectory of the robot under the optimal policy was found to be:

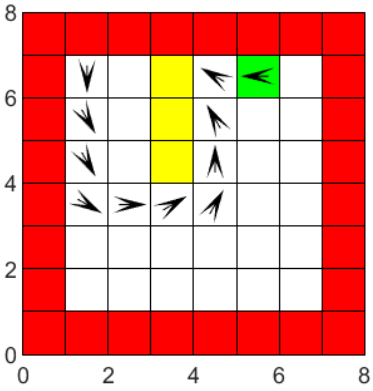


Figure 2: the robot trajectory with optimal policy found by policy iteration.

Where Pe = 0, discount factor of 0.9, initial state, s0=[1,6,6]

Value matrix is listed in Appendix A.

i) The time required to calculate policy iteration (excluding time to draw trajectory) over 5 repeated runs on Ryzen Threadripper 1950x (3.7GHz 32 threads) was 1.16s.

4. Value Iteration

a) Function [V, Pie] = calcOptimalValueAndPolicy(tol) in ValueIteration class returns optimal values and policy matrices with give tolerance. Values are calculated using the following equations:

$$V_{i+1}^{*}(s) \leftarrow \max_{a} \sum_{s'} T(s,a,s') \left[R(s,a,s') + \gamma V_i^{*}(s')\right]$$

$$\pi_{i+1}^{*}(s) \leftarrow \arg\max_{a \in A} \sum_{s'} T(s,a,s') \left[R(s,a,s') + \gamma V_i^{*}(s')\right]$$

At step 0, the values are defined to be 0 since no action is allowed. In order to calculate value at i+1 at a state, s, all possible actions will be tested and the maximum value will be picked. The action which maximize the value at a state is the optimal policy.

For example, at the goal, stand still will results in the maximum value at any time step. At s=[6,5,0] (two cells below and pointing toward the goal), at first step, i =1, the value is zero for any actions since reward is always 0 and there are no values at surrounding cells either. At i = 2, now actions {(Straight, Still), (Straight, Right), (Straight, Left)} have the maximum values since at i = 2 the states, s=[6,6,11 or 0 or 1] have positive values, which originated from the actions going straight to the goal (positive reward). As step goes, the values will spread all over the cells and value iteration will converge where $abs(V_{i+1} - V_i) < tolerance$.

b) The trajectory of the robot under the optimal policy was found to be:

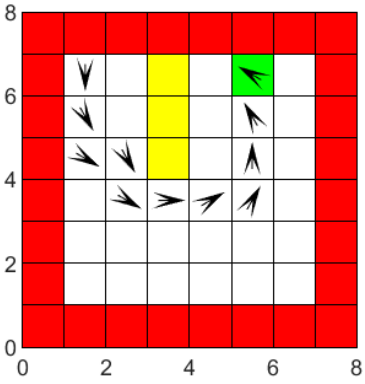


Figure 3: The robot trajectory with optimal policy found by value iteration.

Where Pe = 0, discount factor of 0.9, initial state, s0=[1,6,6]

Comparing to the result of the policy iteration, the robot trajectory was not identical, but both reached the goal with the minimum number of the steps (10 moves). Hence, both reasonable.

Value matrix is listed in Appendix A.

c) The time required to calculate policy iteration (excluding time to draw trajectory) over 5 repeated runs on Ryzen Threadripper 1950x (3.7GHz 32 threads) was 4.16s.

5. Additional Scenarios

a) When Pe =25%, discount factor of 0.9, initial state, s0=[1,6,6], trajectory and value have changed as following:

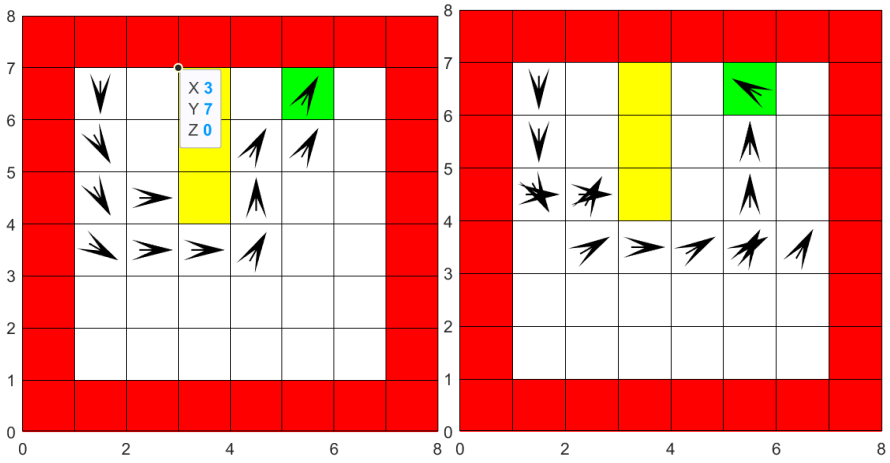


Figure 4: Left Policy iteration (5.19s), Right Value iteration(12.7s). Pe =25%, $\gamma = 0.9$, s0 = [1,6,6], s_goal = [6,7,*].

Note that the trajectory will be different each time we run the robot due to error probability. Another example of the trajectory using the same policy above is following:

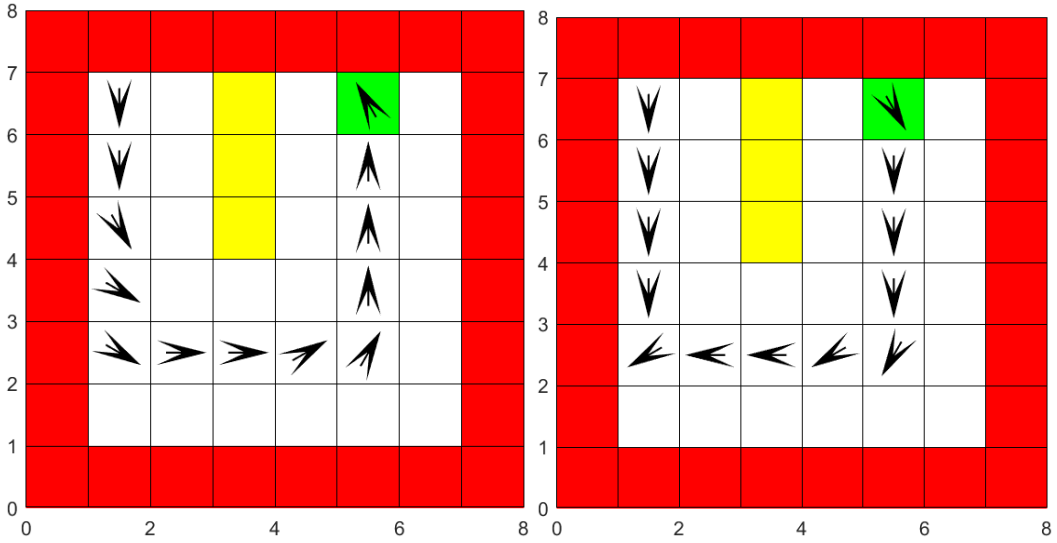

Figure 5: Left Policy iteration (5.19s), Right Value iteration(12.7s). Pe =25%, $\gamma = 0.9$, s0 = [1,6,6], s_goal = [6,7,*].

Value matrix is listed in Appendix B.

b) When Pe =0%, discount factor of 0.9, initial state, s0=[1,6,6], trajectory and value have changed as following and reward is given iff the robot is facing down at the goal:

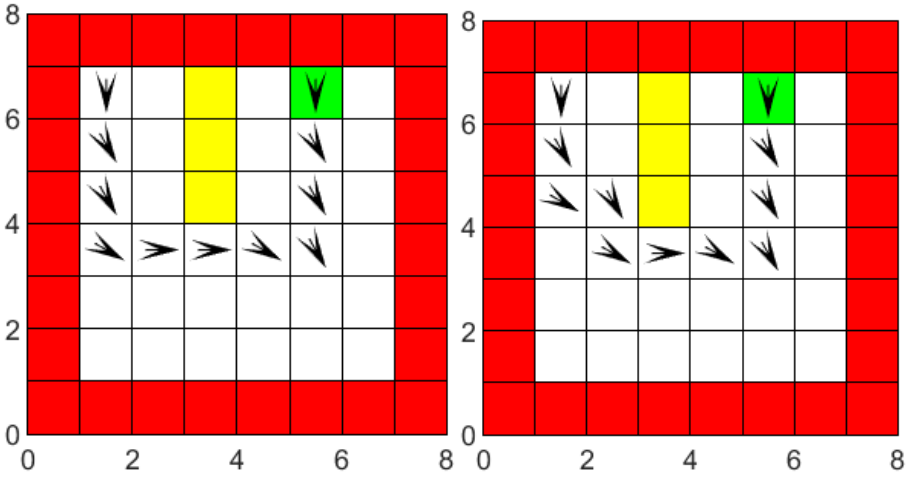

Figure 6: Left Policy iteration (1.18s), Right Value iteration(4.06). Pe =0%, $\gamma = 0.9$, s0 = [1,6,6], s_goal = [6,7,6].

Value matrix is listed in Appendix C.

When Pe =25%, discount factor of 0.9, initial state, s0=[1,6,6], trajectory and value have changed as following and reward is given iff the robot is facing down at the goal:

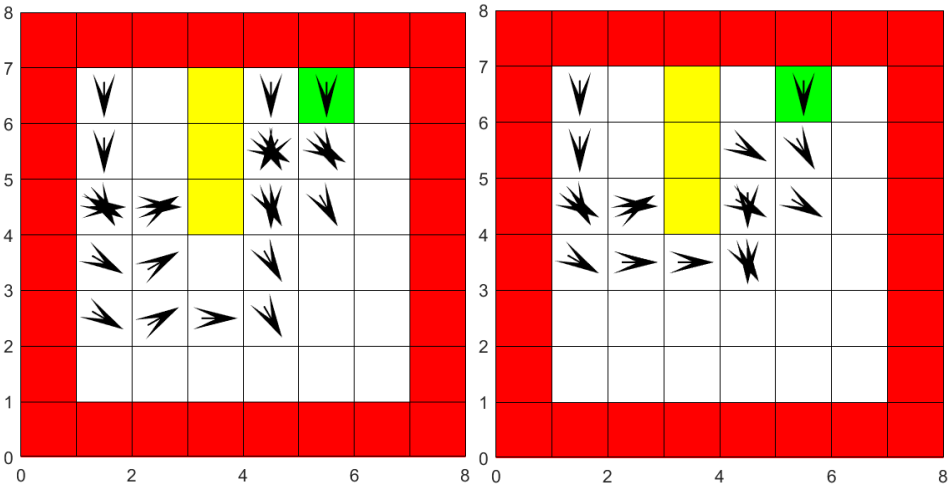

Figure 7: Left Policy iteration (5.85s), Right Value iteration (8.65s). Pe =25%, $\gamma = 0.9$, s0 = [1,6,6], s_goal = [6,7,6].

Value matrix is listed in Appendix D.

c) Even though both methods resulted in different optimal policy, the robot at s=[1,6,6] could reach the goal with minimum number of steps (10 steps). When Pe !=0, the robot trajectory tended to draw a bigger circle (especially for policy iteration Fig.2 vs. Fig. 5). Which is reasonable since if the robot faces the wall (negative reward cells) appropriate action would be go backwards, which does not contribute to reach the goal. Or taking a risk to go into the wall should have less values than any other safer options.

As to calculation time, policy iteration finished about 4 times quicker than value iteration even though each step took about 1.5 times more time to compute. Value iteration required more steps so that values converge, while the policy stopped changing before (Vi+1 - Vi) < tolerance. When Pe = 25%, the computational time by 2 to 4 times. This code skips calculations on state with probability, Psa = 0. Technically, all combinations of going one state to another state with some actions are (Ns*Na*Ns'). But for the sake of time, Ns' was reduced to 1 or 3 for Pe =0 and Pe != 0, respectively. Therefore, having non-zero Pe simply increase number of combinations from Ns*Na*1 to Ns*Na*3. Hence, 3 times more computing time was expected.

When orientation of the robot at the goal was specified, the robot turned around and approached backwards to the goal. When Pe = 0, the robot took 10 steps which was the same number of steps as the orientation was not defined. From Appendix C, the values are now lower even at the goal cell unless the robot is facing down, and only the state, s=[6,7,6] would have stand still action since otherwise there should be neighborhood cells with higher values. When Pe=25%, the robot wondered a lot more than the original scenarios (s_goal = [6,7,*]) simply because the robot won't able to approach the goal with an appropriate orientations easily if there was error in motion and the robot needed to move out of the goal if it failed to direct face down.

One interesting finding was that there was tendency in the action the robot chose which depended on the order of action space (A). If there are actions which results in the same reward/value then the program needs to decide which action to prefer. In this program, the action was distinguished using index in action space matrix. Again, action space, A, has one action in each column:

Table 1: Action space A1

| Index | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Motion | Still | Forwards | Forwards | Forwards | Backwards | Backwards | Backwards |
| Turn | Still | Left | Still | Right | Left | Still | Right |

Table 1: Action space A2

| Index | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Motion | Still | Forwards | Forwards | Forwards | Backwards | Backwards | Backwards |
| Turn | Still | Right | Still | Left | Right | Still | Left |

Index 4 in A means the robot go forwards and then turn right, for instance.

When finding maximum value and its corresponding action, the program takes the lower indexed action if there are multiple actions with the same max value. For example, at s=[6,6,0] (one cell down and toward the goal), the reward is 1 for action index 2 to 4. Then, for the action space, A1, the program selects action index 2 which is Forward and turn left and has the lowest index number out of three. On the other hand, with the action space A2, the index 2 is now go forwards and turn right instead. Therefore, the robot is now likely to make right turn instead of left. Fig 8 and 9 show how the optimal policies differ based on different action spaces.

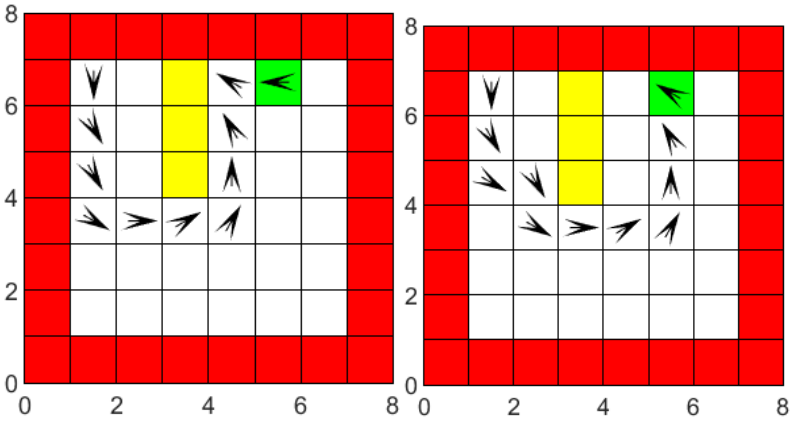


Figure 8: policy tendency due to action space order with action space A1. (left policy iteration, right value iteration) (Pe=0, discount factor =0.9, s0=[1,6,6])

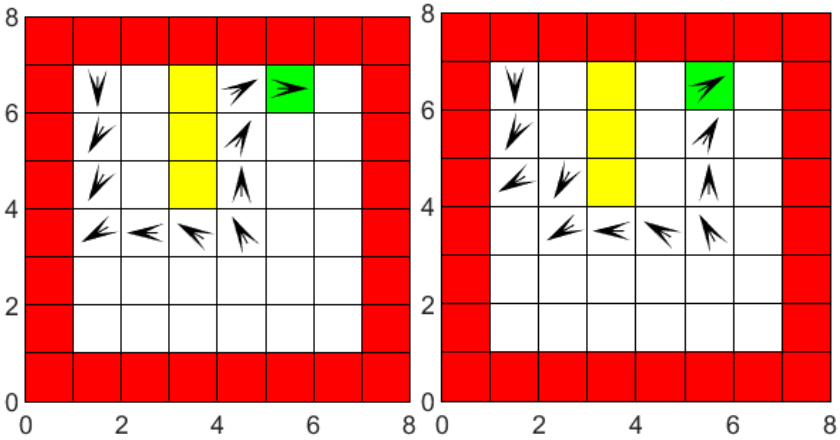


Figure 9: policy tendency due to action space order with action space A2. (left policy iteration, right value iteration) (Pe=0, discount factor =0.9, s0=[1,6,6])

Fig. 9 trajectory tends to turn right, and the robot moved backwards, which looked unreasonable. However, the number of steps was still the minimum, and in that sense the trajectory is appropriate.

In order to choose straight forward actions, it is necessary to formulate the action space in that way that the higher priority actions come in lower indexes, or add some constraints (for instance, optimize energy or so).

# Appendix A

## Pe = 0, $\gamma$ =0.9, s0 = [1,6,6], s_goal=[6,7,*]

Policy Iteration

Pe =0, s_goal = [6,7,*]

val(:,:,1) =

-187.4980 -187.1106 -187.3995 -187.0121 -187.3109 -186.9235 -187.2311 -187.5080
2.8894 3.3199 2.9879 3.4184 3.0765 3.5070 3.1563 2.8407
3.3199 3.7982 3.4184 3.8967 3.5070 3.9853 3.5868 3.2281
3.7982 4.3296 3.8967 4.4281 3.9853 -5.4833 -12.5963 -13.2367
4.3296 4.9201 5.5762 6.3052 7.1152 6.4037 7.2137 6.4923
4.9201 5.5762 6.3052 7.1152 8.0152 9.0152 9.0152 9.1137
4.4281 4.9201 5.6747 6.3052 7.2137 6.4923 7.2137 6.5721
-186.0147 -185.4833 -184.8928 -184.2367 -183.5077 -184.2367 -183.5077 -183.5077

val(:,:,2) =

-97.4980 -97.1106 -96.6801 -97.0121 -96.5816 -96.9235 -97.2311 -97.5080
2.8894 3.3199 3.7982 3.4184 3.8967 3.5070 3.1563 2.8407
3.3199 3.7982 4.3296 3.8967 4.4281 3.9853 3.5868 3.2281
3.7982 4.3296 4.9201 4.4281 5.0186 -2.8848 -3.5963 -2.7863
4.3296 4.9201 5.5762 6.3052 7.1152 8.0152 7.2137 8.1137
4.9201 5.5762 6.3052 7.1152 8.0152 9.0152 9.0152 9.1137
4.4281 4.9201 5.6747 6.3052 7.2137 8.0152 7.3023 8.2023
-96.1033 -95.5719 -94.9814 -94.3253 -93.5963 -92.7863 -93.4279 -92.6179

val(:,:,3) =

-97.4980 2.8894 3.3199 3.7982 3.4184 3.0765 2.7689 -97.5798
-97.1106 3.3199 3.7982 4.3296 3.7982 3.4184 3.0765 -97.2311
-96.6801 3.7982 4.3296 4.9201 3.5070 3.1563 2.8407 -97.4434
-96.2018 4.3296 4.9201 5.5762 6.3052 7.1152 8.0152 -92.8848
-95.6704 4.9201 5.5762 6.3052 7.1152 8.0152 9.0152 -91.9848
-96.1033 4.4281 5.0186 5.6747 6.4037 7.2137 9.0152 -92.6977

-95.5719  5.0186  5.6747  6.4037  7.2137  8.1137  9.1137 -91.7977

-96.0147  4.5167  5.1072  5.7633  6.4923  7.3023  8.2023 -92.6179

val(:,:,4) =

-187.4980  2.8894  3.3199  3.7982  3.4184  3.0765  2.7689 -187.5080

-187.1106  3.3199  3.7982  4.3296  2.9879  2.6891  2.4202 -187.8218

-186.6801  3.7982  4.3296  4.9201  3.4184  3.0765  2.7689 -183.5963

-186.2018  4.3296  4.9201  5.5762  6.3052  7.1152  8.0152 -182.7863

-186.5816  3.8967  4.4281  5.0186  5.5762  6.3052  9.0152 -183.5963

-186.1033  4.4281  5.0186  5.6747  6.4037  7.2137  9.0152 -182.7863

-186.4930  3.9853  4.5167  5.1072  5.7633  6.4923  9.1137 -183.4279

-186.0147  4.5167  5.1072  5.7633  6.4923  7.3023  8.2023 -182.6179

val(:,:,5) =

-97.4980  2.8894  3.3199  3.7982  3.4184  3.0765  2.7689 -97.5798

-97.1106  3.3199  3.7982  4.3296  3.7982  3.4184  3.0765 -97.2311

-96.6801  3.7982  4.3296  4.9201  3.5070  3.1563  2.8407 -97.4434

-96.2018  4.3296  4.9201  5.5762  6.3052  7.1152  8.0152 -92.8848

-95.6704  4.9201  5.5762  6.3052  7.1152  8.0152  9.0152 -91.9848

-96.1033  4.4281  5.0186  5.6747  6.4037  7.2137  9.0152 -92.6977

-95.5719  5.0186  5.6747  6.4037  7.2137  8.1137  9.1137 -91.7977

-96.0147  4.5167  5.1072  5.7633  6.4923  7.3023  8.2023 -92.6179

val(:,:,6) =

-97.4980 -97.1106 -96.6801 -97.0121 -96.5816 -96.9235 -97.2311 -97.5080

  2.8894  3.3199  3.7982  3.4184  3.8967  3.5070  3.1563  2.8407

  3.3199  3.7982  4.3296  3.8967  4.4281  3.9853  3.5868  3.2281

  3.7982  4.3296  4.9201  4.4281  5.0186 -2.8848 -3.5963 -2.7863

  4.3296  4.9201  5.5762  6.3052  7.1152  8.0152  7.2137  8.1137

  4.9201  5.5762  6.3052  7.1152  8.0152  9.0152  9.0152  9.1137

  4.4281  4.9201  5.6747  6.3052  7.2137  8.0152  7.3023  8.2023

-96.1033 -95.5719 -94.9814 -94.3253 -93.5963 -92.7863 -93.4279 -92.6179

val(:,:,7) =

-187.4980 -187.1106 -187.3995 -187.0121 -187.3109 -186.9235 -187.2311 -187.5080
2.8894 3.3199 2.9879 3.4184 3.0765 3.5070 3.1563 2.8407
3.3199 3.7982 3.4184 3.8967 3.5070 3.9853 3.5868 3.2281
3.7982 4.3296 3.8967 4.4281 3.9853 -5.4833 -12.5963 -13.2367
4.3296 4.9201 5.5762 6.3052 7.1152 6.4037 7.2137 6.4923
4.9201 5.5762 6.3052 7.1152 8.0152 9.0152 9.0152 9.1137
4.4281 4.9201 5.6747 6.3052 7.2137 6.4923 7.2137 6.5721
-186.0147 -185.4833 -184.8928 -184.2367 -183.5077 -184.2367 -183.5077 -183.3561

val(:,:,8) =

-97.4980 -97.1106 -96.6801 -97.0121 -96.5816 -96.9235 -97.2311 -97.5080
2.8894 3.3199 3.7982 3.4184 3.8967 3.5070 3.1563 2.8407
3.3199 3.7982 4.3296 3.8967 4.4281 3.9853 3.5868 3.2281
3.7982 4.3296 4.9201 4.4281 5.0186 -2.8848 -3.5963 -2.7863
4.3296 4.9201 5.5762 6.3052 7.1152 8.0152 7.2137 8.1137
4.9201 5.5762 6.3052 7.1152 8.0152 9.0152 9.0152 9.1137
4.4281 4.9201 5.6747 6.3052 7.2137 8.0152 7.3023 8.2023
-96.1033 -95.5719 -94.9814 -94.3253 -93.5963 -92.7863 -93.4279 -92.6179

val(:,:,9) =

-97.4980 2.8894 3.3199 3.7982 3.4184 3.0765 2.7689 -97.5798
-97.1106 3.3199 3.7982 4.3296 3.7982 3.4184 3.0765 -97.2311
-96.6801 3.7982 4.3296 4.9201 3.5070 3.1563 2.8407 -97.4434
-96.2018 4.3296 4.9201 5.5762 6.3052 7.1152 8.0152 -92.8848
-95.6704 4.9201 5.5762 6.3052 7.1152 8.0152 9.0152 -91.9848
-96.1033 4.4281 5.0186 5.6747 6.4037 7.2137 9.0152 -92.6977
-95.5719 5.0186 5.6747 6.4037 7.2137 8.1137 9.1137 -91.7977
-96.0147 4.5167 5.1072 5.7633 6.4923 7.3023 8.2023 -92.6179

val(:,:,10) =

-187.4980 2.8894 3.3199 3.7982 3.4184 3.0765 2.7689 -187.5080

-187.1106 3.3199 3.7982 4.3296 2.9879 2.6891 2.4202 -187.8218

-186.6801 3.7982 4.3296 4.9201 3.4184 3.0765 2.7689 -183.5963

-186.2018 4.3296 4.9201 5.5762 6.3052 7.1152 8.0152 -182.7863

-186.5816 3.8967 4.4281 5.0186 5.5762 6.3052 9.0152 -183.5963

-186.1033 4.4281 5.0186 5.6747 6.4037 7.2137 9.0152 -182.7863

-186.4930 3.9853 4.5167 5.1072 5.7633 6.4923 9.1137 -183.4279

-186.0147 4.5167 5.1072 5.7633 6.4923 7.3023 8.2023 -182.6179

val(:,:,11) =

-97.4980 2.8894 3.3199 3.7982 3.4184 3.0765 2.7689 -97.5798

-97.1106 3.3199 3.7982 4.3296 3.7982 3.4184 3.0765 -97.2311

-96.6801 3.7982 4.3296 4.9201 3.5070 3.1563 2.8407 -97.4434

-96.2018 4.3296 4.9201 5.5762 6.3052 7.1152 8.0152 -92.8848

-95.6704 4.9201 5.5762 6.3052 7.1152 8.0152 9.0152 -91.9848

-96.1033 4.4281 5.0186 5.6747 6.4037 7.2137 9.0152 -92.6977

-95.5719 5.0186 5.6747 6.4037 7.2137 8.1137 9.1137 -91.7977

-96.0147 4.5167 5.1072 5.7633 6.4923 7.3023 8.2023 -92.6179

val(:,:,12) =

-97.4980 -97.1106 -96.6801 -97.0121 -96.5816 -96.9235 -97.2311 -97.5080

2.8894 3.3199 3.7982 3.4184 3.8967 3.5070 3.1563 2.8407

3.3199 3.7982 4.3296 3.8967 4.4281 3.9853 3.5868 3.2281

3.7982 4.3296 4.9201 4.4281 5.0186 -2.8848 -3.5963 -2.7863

4.3296 4.9201 5.5762 6.3052 7.1152 8.0152 7.2137 8.1137

4.9201 5.5762 6.3052 7.1152 8.0152 9.0152 9.0152 9.1137

4.4281 4.9201 5.6747 6.3052 7.2137 8.0152 7.3023 8.2023

-96.1033 -95.5719 -94.9814 -94.3253 -93.5963 -92.7863 -93.4279 -92.6179

Value Iteration

Pe=0, s_goal = [6,7,*]

val(:,:,1) =

-186.6005 -186.2131 -186.6005 -186.2131 -186.6005 -186.2131 -186.6005 -186.9492

3.7869 4.2174 3.7869 4.2174 3.7869 4.2174 3.7869 3.3995

4.2174 4.6957 4.2174 4.6957 4.2174 4.6957 4.2174 3.7869

4.6957 5.2271 4.6957 5.2271 4.6957 -4.7729 -11.7973 -12.5263

5.2271 5.8176 6.4737 7.2027 8.0127 7.2027 8.0127 7.2027

5.8176 6.4737 7.2027 8.0127 8.9127 9.9127 9.9127 9.9127

5.2271 5.8176 6.4737 7.2027 8.0127 7.2027 8.0127 7.2027

-185.3043 -184.7729 -184.1824 -183.5263 -182.7973 -183.5263 -182.7973 -182.7973

val(:,:,2) =

-96.6005 -96.2131 -95.7826 -96.2131 -95.7826 -96.2131 -96.6005 -96.9492

3.7869 4.2174 4.6957 4.2174 4.6957 4.2174 3.7869 3.3995

4.2174 4.6957 5.2271 4.6957 5.2271 4.6957 4.2174 3.7869

4.6957 5.2271 5.8176 5.2271 5.8176 -1.9873 -2.7973 -1.9873

5.2271 5.8176 6.4737 7.2027 8.0127 8.9127 8.0127 8.9127

5.8176 6.4737 7.2027 8.0127 8.9127 9.9127 9.9127 9.9127

5.2271 5.8176 6.4737 7.2027 8.0127 8.9127 8.0127 8.9127

-95.3043 -94.7729 -94.1824 -93.5263 -92.7973 -91.9873 -92.7973 -91.9873

val(:,:,3) =

-96.6005 3.7869 4.2174 4.6957 4.2174 3.7869 3.3995 -96.9492

-96.2131 4.2174 4.6957 5.2271 4.6957 4.2174 3.7869 -96.6005

-95.7826 4.6957 5.2271 5.8176 4.2174 3.7869 3.3995 -96.9492

-95.3043 5.2271 5.8176 6.4737 7.2027 8.0127 8.9127 -91.9873

-94.7729 5.8176 6.4737 7.2027 8.0127 8.9127 9.9127 -91.0873

-95.3043 5.2271 5.8176 6.4737 7.2027 8.0127 9.9127 -91.9873

-94.7729 5.8176 6.4737 7.2027 8.0127 8.9127 9.9127 -91.0873

-95.3043 5.2271 5.8176 6.4737 7.2027 8.0127 8.9127 -91.9873

val(:,:,4) =

-186.6005 3.7869 4.2174 4.6957 4.2174 3.7869 3.3995 -186.9492

-186.2131 4.2174 4.6957 5.2271 3.7869 3.3995 3.0508 -187.2630

-185.7826 4.6957 5.2271 5.8176 4.2174 3.7869 3.3995 -182.7973

-185.3043 5.2271 5.8176 6.4737 7.2027 8.0127 8.9127 -181.9873

-185.7826 4.6957 5.2271 5.8176 6.4737 7.2027 9.9127 -182.7973

-185.3043 5.2271 5.8176 6.4737 7.2027 8.0127 9.9127 -181.9873

-185.7826 4.6957 5.2271 5.8176 6.4737 7.2027 9.9127 -182.7973

-185.3043 5.2271 5.8176 6.4737 7.2027 8.0127 8.9127 -181.9873

val(:,:,5) =

-96.6005 3.7869 4.2174 4.6957 4.2174 3.7869 3.3995 -96.9492

-96.2131 4.2174 4.6957 5.2271 4.6957 4.2174 3.7869 -96.6005

-95.7826 4.6957 5.2271 5.8176 4.2174 3.7869 3.3995 -96.9492

-95.3043 5.2271 5.8176 6.4737 7.2027 8.0127 8.9127 -91.9873

-94.7729 5.8176 6.4737 7.2027 8.0127 8.9127 9.9127 -91.0873

-95.3043 5.2271 5.8176 6.4737 7.2027 8.0127 9.9127 -91.9873

-94.7729 5.8176 6.4737 7.2027 8.0127 8.9127 9.9127 -91.0873

-95.3043 5.2271 5.8176 6.4737 7.2027 8.0127 8.9127 -91.9873

val(:,:,6) =

-96.6005 -96.2131 -95.7826 -96.2131 -95.7826 -96.2131 -96.6005 -96.9492

3.7869 4.2174 4.6957 4.2174 4.6957 4.2174 3.7869 3.3995

4.2174 4.6957 5.2271 4.6957 5.2271 4.6957 4.2174 3.7869

4.6957 5.2271 5.8176 5.2271 5.8176 -1.9873 -2.7973 -1.9873

5.2271 5.8176 6.4737 7.2027 8.0127 8.9127 8.0127 8.9127

5.8176 6.4737 7.2027 8.0127 8.9127 9.9127 9.9127 9.9127

5.2271 5.8176 6.4737 7.2027 8.0127 8.9127 8.0127 8.9127

-95.3043 -94.7729 -94.1824 -93.5263 -92.7973 -91.9873 -92.7973 -91.9873

val(:,:,7) =

-186.6005 -186.2131 -186.6005 -186.2131 -186.6005 -186.2131 -186.6005 -186.9492

3.7869 4.2174 3.7869 4.2174 3.7869 4.2174 3.7869 3.3995

4.2174 4.6957 4.2174 4.6957 4.2174 4.6957 4.2174 3.7869

4.6957 5.2271 4.6957 5.2271 4.6957 -4.7729 -11.7973 -12.5263

5.2271 5.8176 6.4737 7.2027 8.0127 7.2027 8.0127 7.2027

5.8176 6.4737 7.2027 8.0127 8.9127 9.9127 9.9127 9.9127

5.2271 5.8176 6.4737 7.2027 8.0127 7.2027 8.0127 7.2027

-185.3043 -184.7729 -184.1824 -183.5263 -182.7973 -183.5263 -182.7973 -182.7973

val(:,:,8) =

-96.6005 -96.2131 -95.7826 -96.2131 -95.7826 -96.2131 -96.6005 -96.9492

3.7869 4.2174 4.6957 4.2174 4.6957 4.2174 3.7869 3.3995

4.2174 4.6957 5.2271 4.6957 5.2271 4.6957 4.2174 3.7869

4.6957 5.2271 5.8176 5.2271 5.8176 -1.9873 -2.7973 -1.9873

5.2271 5.8176 6.4737 7.2027 8.0127 8.9127 8.0127 8.9127

5.8176 6.4737 7.2027 8.0127 8.9127 9.9127 9.9127 9.9127

5.2271 5.8176 6.4737 7.2027 8.0127 8.9127 8.0127 8.9127

-95.3043 -94.7729 -94.1824 -93.5263 -92.7973 -91.9873 -92.7973 -91.9873

val(:,:,9) =

-96.6005 3.7869 4.2174 4.6957 4.2174 3.7869 3.3995 -96.9492

-96.2131 4.2174 4.6957 5.2271 4.6957 4.2174 3.7869 -96.6005

-95.7826 4.6957 5.2271 5.8176 4.2174 3.7869 3.3995 -96.9492

-95.3043 5.2271 5.8176 6.4737 7.2027 8.0127 8.9127 -91.9873

-94.7729 5.8176 6.4737 7.2027 8.0127 8.9127 9.9127 -91.0873

-95.3043 5.2271 5.8176 6.4737 7.2027 8.0127 9.9127 -91.9873

-94.7729 5.8176 6.4737 7.2027 8.0127 8.9127 9.9127 -91.0873

-95.3043 5.2271 5.8176 6.4737 7.2027 8.0127 8.9127 -91.9873

val(:,:,10) =

-186.6005 3.7869 4.2174 4.6957 4.2174 3.7869 3.3995 -186.9492

-186.2131 4.2174 4.6957 5.2271 3.7869 3.3995 3.0508 -187.2630

-185.7826 4.6957 5.2271 5.8176 4.2174 3.7869 3.3995 -182.7973

-185.3043 5.2271 5.8176 6.4737 7.2027 8.0127 8.9127 -181.9873

-185.7826 4.6957 5.2271 5.8176 6.4737 7.2027 9.9127 -182.7973

-185.3043 5.2271 5.8176 6.4737 7.2027 8.0127 9.9127 -181.9873

-185.7826 4.6957 5.2271 5.8176 6.4737 7.2027 9.9127 -182.7973

-185.3043 5.2271 5.8176 6.4737 7.2027 8.0127 8.9127 -181.9873

val(:,:,11) =

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| -96.6005 | 3.7869 | 4.2174 | 4.6957 | 4.2174 | 3.7869 | 3.3995 | -96.9492 |
| -96.2131 | 4.2174 | 4.6957 | 5.2271 | 4.6957 | 4.2174 | 3.7869 | -96.6005 |
| -95.7826 | 4.6957 | 5.2271 | 5.8176 | 4.2174 | 3.7869 | 3.3995 | -96.9492 |
| -95.3043 | 5.2271 | 5.8176 | 6.4737 | 7.2027 | 8.0127 | 8.9127 | -91.9873 |
| -94.7729 | 5.8176 | 6.4737 | 7.2027 | 8.0127 | 8.9127 | 9.9127 | -91.0873 |
| -95.3043 | 5.2271 | 5.8176 | 6.4737 | 7.2027 | 8.0127 | 9.9127 | -91.9873 |
| -94.7729 | 5.8176 | 6.4737 | 7.2027 | 8.0127 | 8.9127 | 9.9127 | -91.0873 |
| -95.3043 | 5.2271 | 5.8176 | 6.4737 | 7.2027 | 8.0127 | 8.9127 | -91.9873 |

val(:,:,12) =

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| -96.6005 | -96.2131 | -95.7826 | -96.2131 | -95.7826 | -96.2131 | -96.6005 | -96.9492 |
| 3.7869 | 4.2174 | 4.6957 | 4.2174 | 4.6957 | 4.2174 | 3.7869 | 3.3995 |
| 4.2174 | 4.6957 | 5.2271 | 4.6957 | 5.2271 | 4.6957 | 4.2174 | 3.7869 |
| 4.6957 | 5.2271 | 5.8176 | 5.2271 | 5.8176 | -1.9873 | -2.7973 | -1.9873 |
| 5.2271 | 5.8176 | 6.4737 | 7.2027 | 8.0127 | 8.9127 | 8.0127 | 8.9127 |
| 5.8176 | 6.4737 | 7.2027 | 8.0127 | 8.9127 | 9.9127 | 9.9127 | 9.9127 |
| 5.2271 | 5.8176 | 6.4737 | 7.2027 | 8.0127 | 8.9127 | 8.0127 | 8.9127 |
| -95.3043 | -94.7729 | -94.1824 | -93.5263 | -92.7973 | -91.9873 | -92.7973 | -91.9873 |

# Appendix B

## Pe = 0.25, $\gamma$ =0.9, s0 = [1,6,6], s_goal=[6,7,*]

Policy iteration

Pe = 25

val(:,:,1) =

-203.8241 -204.0248 -203.8032 -203.9880 -204.6233 -204.5907 -204.5908 -204.5547
2.2129 2.3158 2.3173 2.3348 2.3617 2.0201 1.7045 1.3412
2.6035 2.7688 2.8070 2.7893 2.8465 2.3336 1.9227 1.5060
2.8616 3.0371 3.1490 3.0590 3.1910 -6.4582 -12.1478 -13.9484
3.6287 4.1708 4.6649 5.2408 5.6530 6.3053 5.6670 6.3434
4.4583 5.2406 6.1154 7.1370 8.3712 9.7863 9.7814 9.8144
3.6780 4.1819 4.6977 5.2538 5.6698 6.3156 5.6840 6.3590
-202.7286 -202.6354 -202.7070 -202.6076 -201.5501 -200.5878 -201.5232 -200.5518

val(:,:,2) =

-185.0177 -113.5767 -114.2515 -116.9023 -123.9680 -139.8889 -149.1278 -211.8433
-30.3514 2.5439 2.5708 2.2110 1.8997 1.5726 1.0204 -70.9907
-29.8948 2.9742 3.0991 2.9509 2.6341 2.2152 1.7151 -39.8528
-29.8673 3.2175 3.1686 3.2636 2.7520 -4.6905 -7.5971 -47.9780
-30.6291 4.0728 4.6765 5.1400 5.8313 6.0409 4.6902 -31.7091
-37.0698 4.5983 5.5636 6.4714 7.5471 8.8575 9.7856 -24.7782
-68.8638 2.7181 4.0544 4.6166 5.2256 5.8604 6.8417 -26.9551
-209.7132 -147.1988 -138.1364 -121.8699 -114.6946 -111.8995 -110.0415 -181.7222

val(:,:,3) =

-185.1378 -30.1696 -29.8627 -30.2475 -32.6789 -39.7652 -70.3964 -211.5533
-113.5719 2.5671 2.9180 2.9249 2.0547 1.7305 0.8589 -149.1733
-113.8363 2.9729 3.3918 3.4396 2.0526 1.7375 1.1775 -140.0772
-115.5303 3.3949 3.8043 3.3338 2.3355 1.5732 -1.4462 -129.6844
-121.7904 3.8860 4.6541 5.3071 6.1175 6.7993 5.8871 -124.4940

-137.8829  3.5000  4.3867  5.0692  5.8529  6.9358  9.7889 -114.8294

-146.8775  2.5485  4.4316  5.1097  5.8952  6.8222  8.8947 -110.7079

-209.6501 -68.8174 -37.5169 -30.0434 -27.8858 -26.8273 -25.2908 -182.3818

val(:,:,4) =

-203.8231  2.2964  2.5984  2.7638  1.8029  1.5144  1.0929 -207.1412

-203.8402  2.6576  3.0256  3.2323  1.8771  1.5856  1.1310 -209.3604

-203.8005  2.9962  3.4745  3.7279  1.8174  1.5364  1.1123 -207.1226

-203.8078  3.4460  4.0623  4.6066  5.2724  5.9304  7.7004 -205.5264

-202.8930  3.5822  4.1859  4.8295  5.5047  6.4358  9.7865 -202.6166

-202.6779  3.5407  4.0884  4.6452  5.3035  5.9578  9.7915 -201.5035

-202.8625  3.6125  4.2273  4.8597  5.5364  6.4684  9.8140 -202.5866

-202.6371  3.5774  4.1281  4.6797  5.3280  5.9820  7.7501 -201.4654

val(:,:,5) =

-210.6473 -69.7793 -38.6992 -31.6190 -30.8073 -30.8288 -31.2401 -186.7553

-148.3014  1.7009  2.6903  3.0440  1.9289  1.6442  1.1422 -117.5239

-139.1430  2.4915  3.2289  3.5567  2.0027  1.6727  0.8857 -122.3367

-123.0910  2.8234  3.6239  4.1662  4.7688  2.9774  3.8808 -119.9152

-116.0091  3.3134  4.4630  5.1389  5.8756  6.8081  8.8696 -118.7760

-113.1326  3.8338  4.5384  5.2084  5.8723  6.9587  9.7934 -134.8601

-112.2170  4.0956  4.7322  5.3824  6.1447  6.8277  5.9071 -143.6071

-183.8358 -29.2124 -28.4591 -28.2235 -29.1204 -35.2459 -65.2595 -207.4233

val(:,:,6) =

-210.6313 -148.2189 -139.0750 -122.9870 -116.4783 -114.7141 -114.4291 -186.9137

 -69.7572  1.6015  2.2760  2.5572  2.3424  1.9546  1.5283 -31.6921

 -38.8014  2.6950  2.9664  3.1507  2.0441  1.6630  1.1160 -34.1406

 -31.6618  3.0490  3.4134  3.4811  3.8898 -3.0382 -4.6471 -46.1521

 -29.3442  3.8183  4.2952  4.8321  5.2092  5.8478  6.8255 -28.7438

 -28.2234  4.8474  5.6226  6.5366  7.5583  8.8732  9.7949 -32.6351

 -28.8685  4.0957  4.6553  5.1742  5.8541  6.0649  4.7046 -66.6216

-184.0816 -112.3125 -112.4336 -114.0578 -120.1152 -135.4717 -144.5980 -207.2311

val(:,:,7) =

-203.8224 -204.0234 -203.8016 -203.9868 -204.6239 -204.5915 -204.5909 -204.5551

2.2140 2.3161 2.3171 2.3347 2.3610 2.0193 1.7039 1.3403

2.6047 2.7689 2.8061 2.7899 2.8458 2.3334 1.9226 1.5022

2.8637 3.0387 3.1515 3.0605 3.1925 -6.4583 -12.1504 -13.9518

3.6318 4.1739 4.6687 5.2432 5.6572 6.3053 5.6711 6.3432

4.4620 5.2448 6.1195 7.1417 8.3759 9.7898 9.7959 9.8169

3.6817 4.1859 4.7016 5.2585 5.6736 6.3210 5.6877 6.3633

-202.7278 -202.6350 -202.7063 -202.6073 -201.5506 -200.5868 -201.5235 -200.5511

val(:,:,8) =

-185.0157 -113.5753 -114.2504 -116.9022 -123.9677 -139.8883 -149.1262 -211.8429

-30.3535 2.5451 2.5704 2.2104 1.8989 1.5714 1.0207 -70.9902

-29.8945 2.9746 3.1004 2.9516 2.6349 2.2153 1.7154 -39.8565

-29.8645 3.2190 3.1717 3.2659 2.7533 -4.6897 -7.5939 -47.9810

-30.6261 4.0759 4.6793 5.1443 5.8328 6.0459 4.6940 -31.7080

-37.0671 4.6019 5.5676 6.4762 7.5500 8.8674 9.7966 -24.7723

-68.8644 2.7196 4.0579 4.6203 5.2292 5.8655 6.8472 -26.9558

-209.7138 -147.2004 -138.1368 -121.8712 -114.6962 -111.9006 -110.0412 -181.7199

val(:,:,9) =

-185.1391 -30.1666 -29.8612 -30.2481 -32.6785 -39.7651 -70.3961 -211.5536

-113.5713 2.5686 2.9188 2.9254 2.0545 1.7302 0.8599 -149.1757

-113.8348 2.9748 3.3925 3.4407 2.0532 1.7367 1.1778 -140.0775

-115.5274 3.3977 3.8073 3.3349 2.3363 1.5777 -1.4410 -129.6867

-121.7878 3.8883 4.6577 5.3115 6.1201 6.8076 5.8908 -124.4950

-137.8816 3.4986 4.3907 5.0736 5.8572 6.9427 9.7970 -114.8259

-146.8752 2.5478 4.4351 5.1137 5.8985 6.8290 8.8995 -110.7036

-209.6496 -68.8167 -37.5182 -30.0458 -27.8877 -26.8279 -25.2876 -182.3812

val(:,:,10) =

-203.8235 2.2978 2.5998 2.7627 1.8038 1.5148 1.0996 -207.1414

-203.8394 2.6592 3.0266 3.2331 1.8775 1.5854 1.1327 -209.3611

-203.8005 2.9983 3.4756 3.7287 1.8181 1.5367 1.1171 -207.1228

-203.8070 3.4473 4.0643 4.6091 5.2748 5.9354 7.7037 -205.5288

-202.8958 3.5793 4.1839 4.8263 5.5081 6.4432 9.7904 -202.6122

-202.6810 3.5373 4.0902 4.6417 5.3058 5.9626 9.7972 -201.5003

-202.8651 3.6099 4.2256 4.8567 5.5352 6.4691 9.8175 -202.5830

-202.6398 3.5745 4.1253 4.6766 5.3249 5.9834 7.7543 -201.4629

val(:,:,11) =

-210.6358 -69.7732 -38.6970 -31.6195 -30.8073 -30.8289 -31.2391 -186.7552

-148.2945 1.7074 2.6923 3.0452 1.9298 1.6446 1.1505 -117.5242

-139.1389 2.4931 3.2310 3.5570 2.0033 1.6731 0.8947 -122.3378

-123.0906 2.8241 3.6216 4.1652 4.7709 2.9835 3.8850 -119.9181

-116.0106 3.3138 4.4595 5.1355 5.8721 6.8066 8.8743 -118.7806

-113.1352 3.8316 4.5357 5.2055 5.8709 6.9595 9.7965 -134.8600

-112.2201 4.0924 4.7289 5.3790 6.1410 6.8255 5.9149 -143.6016

-183.8380 -29.2137 -28.4591 -28.2237 -29.1194 -35.2424 -65.2555 -207.4107

val(:,:,12) =

-210.6264 -148.2162 -139.0714 -122.9856 -116.4781 -114.7141 -114.4288 -186.9134

-69.7513 1.6155 2.2753 2.5582 2.3434 1.9554 1.5314 -31.6919

-38.8006 2.6962 2.9672 3.1524 2.0443 1.6631 1.1329 -34.1408

-31.6636 3.0481 3.4124 3.4790 3.8897 -3.0354 -4.6428 -46.1529

-29.3470 3.8155 4.2934 4.8301 5.2059 5.8451 6.8257 -28.7486

-28.2268 4.8435 5.6190 6.5328 7.5534 8.8701 9.7986 -32.6419

-28.8716 4.0919 4.6513 5.1703 5.8492 6.0606 4.7211 -66.6136

-184.0831 -112.3131 -112.4339 -114.0573 -120.1131 -135.4674 -144.5945 -207.2279

Value Iteration

Pe = 25

val(:,:,1) =

-203.6830 -203.8811 -203.6830 -203.8811 -204.5210 -204.5030 -204.5210 -204.5030
2.3554 2.4461 2.4547 2.4461 2.4547 2.1079 1.7730 1.3931
2.7428 2.9000 2.9395 2.9000 2.9395 2.4135 1.9869 1.5518
2.9999 3.1636 3.2806 3.1636 3.2806 -6.3747 -12.0669 -13.8687
3.7632 4.3067 4.7937 5.3708 5.7729 6.4301 5.7729 6.4301
4.5923 5.3744 6.2447 7.2678 8.4977 9.9127 9.9127 9.9127
3.7979 4.3022 4.8109 5.3706 5.7729 6.4301 5.7729 6.4301
-202.6330 -202.5401 -202.6330 -202.5401 -201.4746 -200.5081 -201.4746 -200.5081

val(:,:,2) =

-184.8747 -113.4341 -114.1230 -116.7800 -123.8723 -139.8141 -149.0625 -211.7840
-30.2126 2.6789 2.7003 2.3355 2.0150 1.6654 1.0880 -70.9384
-29.7577 3.1096 3.2315 3.0607 2.7408 2.2984 1.7767 -39.8090
-29.7317 3.3415 3.3025 3.3760 2.8384 -4.6173 -7.5258 -47.9183
-30.4949 4.2066 4.8037 5.2716 5.9501 6.1705 4.8018 -31.6340
-36.9406 4.7303 5.6944 6.6018 7.6726 8.9905 9.9127 -24.6853
-68.7495 2.8321 4.1556 4.7126 5.3234 5.9507 6.9268 -26.8847
-209.6082 -147.0976 -138.0613 -121.8071 -114.6373 -111.8457 -109.9856 -181.6717

val(:,:,3) =

-184.9997 -30.0254 -29.7261 -30.1119 -32.5695 -39.6858 -70.3240 -211.4952
-113.4316 2.7035 3.0550 3.0485 2.1858 1.8328 0.9242 -149.1170
-113.6989 3.1098 3.5250 3.5728 2.1637 1.8145 1.2321 -140.0350
-115.3930 3.5253 3.9399 3.4431 2.4233 1.6866 -1.3903 -129.6455
-121.6592 4.0131 4.7786 5.4303 6.2358 6.9250 6.0037 -124.4432
-137.7609 3.6204 4.5177 5.1971 5.9782 7.0629 9.9127 -114.7634
-146.7737 2.6594 4.5382 5.2139 5.9974 6.9250 8.9905 -110.6400
-209.5625 -68.7284 -37.4419 -29.9714 -27.8189 -26.7599 -25.2182 -182.3326

val(:,:,4) =

-203.6826  2.4396  2.7365  2.9021  1.9200  1.6032  1.1679 -207.0697

-203.6973  2.7956  3.1623  3.3605  2.0137  1.6858  1.2085 -209.2833

-203.6826  3.1345  3.6076  3.8607  1.9200  1.6032  1.1679 -207.0697

-203.6973  3.5778  4.1946  4.7352  5.4013  6.0574  7.8259 -205.4274

-202.7676  3.7074  4.3206  4.9521  5.6231  6.5553  9.9127 -202.5245

-202.5545  3.6613  4.2073  4.7611  5.4013  6.0574  9.9127 -201.4138

-202.7676  3.7074  4.3206  4.9521  5.6231  6.5553  9.9127 -202.5245

-202.5545  3.6613  4.2073  4.7611  5.4012  6.0574  7.8259 -201.4138

val(:,:,5) =

-210.5018 -69.6421 -38.5678 -31.4883 -30.6956 -30.7442 -31.1741 -186.6886

-148.1591  1.8384  2.8215  3.1673  2.0548  1.7203  1.2121 -117.4546

-139.0206  2.6120  3.3530  3.6803  2.0986  1.7423  0.9472 -122.2564

-122.9773  2.9438  3.7434  4.2838  4.8886  3.0992  4.0002 -119.8192

-115.8998  3.4271  4.5892  5.2610  5.9971  6.9250  8.9905 -118.6755

-113.0256  3.9423  4.6467  5.3156  5.9780  7.0629  9.9127 -134.7672

-112.1166  4.1973  4.8286  5.4813  6.2355  6.9250  6.0037 -143.5277

-183.7476 -29.1265 -28.3752 -28.1415 -29.0414 -35.1665 -65.1816 -207.3619

val(:,:,6) =

-210.4942 -148.0800 -138.9541 -122.8728 -116.3717 -114.6260 -114.3624 -186.8572

 -69.6192  1.7370  2.3906  2.6675  2.4488  2.0376  1.5957 -31.6345

 -38.6684  2.8250  3.0927  3.2614  2.1455  1.7436  1.1849 -34.0826

 -31.5337  3.1770  3.5375  3.5860  3.9910 -2.9359 -4.5443 -46.0788

 -29.2193  3.9441  4.4168  4.9525  5.3204  5.9507  6.9268 -28.6521

 -28.1003  4.9701  5.7410  6.6559  7.6719  8.9905  9.9127 -32.5402

 -28.7581  4.2095  4.7601  5.2844  5.9499  6.1705  4.8018 -66.5447

-183.9880 -112.2207 -112.3435 -113.9699 -120.0303 -135.3858 -144.5346 -207.1788

val(:,:,7) =

-203.6830 -203.8811 -203.6830 -203.8811 -204.5210 -204.5030 -204.5210 -204.5030

   2.3554  2.4461  2.4547  2.4461  2.4547  2.1079  1.7730  1.3931

   2.7428  2.9000  2.9395  2.9000  2.9395  2.4135  1.9869  1.5518

2.9999 3.1636 3.2806 3.1636 3.2806 -6.3747 -12.0669 -13.8687

3.7632 4.3067 4.7937 5.3708 5.7729 6.4301 5.7729 6.4301

4.5923 5.3744 6.2447 7.2678 8.4977 9.9127 9.9127 9.9127

3.7979 4.3022 4.8109 5.3706 5.7729 6.4301 5.7729 6.4301

-202.6330 -202.5401 -202.6330 -202.5401 -201.4746 -200.5081 -201.4746 -200.5081

val(:,:,8) =

-184.8747 -113.4341 -114.1230 -116.7800 -123.8723 -139.8141 -149.0625 -211.7840

-30.2126 2.6789 2.7003 2.3355 2.0150 1.6654 1.0880 -70.9384

-29.7577 3.1096 3.2315 3.0607 2.7408 2.2984 1.7767 -39.8090

-29.7317 3.3415 3.3025 3.3760 2.8384 -4.6173 -7.5258 -47.9183

-30.4949 4.2066 4.8037 5.2716 5.9501 6.1705 4.8018 -31.6340

-36.9406 4.7303 5.6944 6.6018 7.6726 8.9905 9.9127 -24.6853

-68.7495 2.8321 4.1556 4.7126 5.3234 5.9507 6.9268 -26.8847

-209.6082 -147.0976 -138.0613 -121.8071 -114.6373 -111.8457 -109.9856 -181.6717

val(:,:,9) =

-184.9997 -30.0254 -29.7261 -30.1119 -32.5695 -39.6858 -70.3240 -211.4952

-113.4316 2.7035 3.0550 3.0485 2.1858 1.8328 0.9242 -149.1170

-113.6989 3.1098 3.5250 3.5728 2.1637 1.8145 1.2321 -140.0350

-115.3930 3.5253 3.9399 3.4431 2.4233 1.6866 -1.3903 -129.6455

-121.6592 4.0131 4.7786 5.4303 6.2358 6.9250 6.0037 -124.4432

-137.7609 3.6204 4.5177 5.1971 5.9782 7.0629 9.9127 -114.7634

-146.7737 2.6594 4.5382 5.2139 5.9974 6.9250 8.9905 -110.6400

-209.5625 -68.7284 -37.4419 -29.9714 -27.8189 -26.7599 -25.2182 -182.3326

val(:,:,10) =

-203.6826 2.4396 2.7365 2.9021 1.9200 1.6032 1.1679 -207.0697

-203.6973 2.7956 3.1623 3.3605 2.0137 1.6858 1.2085 -209.2833

-203.6826 3.1345 3.6076 3.8607 1.9200 1.6032 1.1679 -207.0697

-203.6973 3.5778 4.1946 4.7352 5.4013 6.0574 7.8259 -205.4274

-202.7676 3.7074 4.3206 4.9521 5.6231 6.5553 9.9127 -202.5245

-202.5545 3.6613 4.2073 4.7611 5.4013 6.0574 9.9127 -201.4138

-202.7676  3.7074  4.3206  4.9521  5.6231  6.5553  9.9127 -202.5245

-202.5545  3.6613  4.2073  4.7611  5.4012  6.0574  7.8259 -201.4138

val(:,:,11) =

-210.5018 -69.6421 -38.5678 -31.4883 -30.6956 -30.7442 -31.1741 -186.6886

-148.1591  1.8384  2.8215  3.1673  2.0548  1.7203  1.2121 -117.4546

-139.0206  2.6120  3.3530  3.6803  2.0986  1.7423  0.9472 -122.2564

-122.9773  2.9438  3.7434  4.2838  4.8886  3.0992  4.0002 -119.8192

-115.8998  3.4271  4.5892  5.2610  5.9971  6.9250  8.9905 -118.6755

-113.0256  3.9423  4.6467  5.3156  5.9780  7.0629  9.9127 -134.7672

-112.1166  4.1973  4.8286  5.4813  6.2355  6.9250  6.0037 -143.5277

-183.7476 -29.1265 -28.3752 -28.1415 -29.0414 -35.1665 -65.1816 -207.3619

val(:,:,12) =

-210.4942 -148.0800 -138.9541 -122.8728 -116.3717 -114.6260 -114.3624 -186.8572

 -69.6192  1.7370  2.3906  2.6675  2.4488  2.0376  1.5957 -31.6345

 -38.6684  2.8250  3.0927  3.2614  2.1455  1.7436  1.1849 -34.0826

 -31.5337  3.1770  3.5375  3.5860  3.9910 -2.9359 -4.5443 -46.0788

 -29.2193  3.9441  4.4168  4.9525  5.3204  5.9507  6.9268 -28.6521

 -28.1003  4.9701  5.7410  6.6559  7.6719  8.9905  9.9127 -32.5402

 -28.7581  4.2095  4.7601  5.2844  5.9499  6.1705  4.8018 -66.5447

-183.9880 -112.2207 -112.3435 -113.9699 -120.0303 -135.3858 -144.5346 -207.1788

# Appendix C

## Pe = 0, $\gamma$ =0.9, s0 = [1,6,6], s_goal=[6,7,6]

Policy Iteration

Pe =0, s_goal=[6,7,6]

val(:,:,1) =

-187.4980 -187.1106 -187.3995 -187.0121 -187.3109 -186.9235 -187.2311 -187.5080
2.8894 3.3199 2.9879 3.4184 3.0765 3.5070 3.1563 2.8407
3.3199 3.7982 3.4184 3.8967 3.5070 3.9853 3.5868 3.2281
2.8894 3.4184 3.0765 3.4184 3.1563 -6.4132 -14.9349 -14.4035
3.4184 3.7982 4.4281 4.9201 4.5167 5.1072 4.5965 5.1072
3.0765 3.4184 3.9853 4.4281 3.9853 4.5965 4.1369 4.5965
3.4184 3.9853 4.4281 5.1072 4.5965 5.1072 4.6683 5.2588
-186.8437 -186.4132 -185.9349 -185.4035 -185.9349 -185.4035 -185.7985 -185.2671

val(:,:,2) =

-97.4980 -97.1106 -96.6801 -97.0121 -96.5816 -96.9235 -97.2311 -97.5080
2.8894 3.3199 3.7982 3.4184 3.8967 3.5070 3.1563 2.8407
3.3199 3.7982 4.3296 3.8967 4.4281 3.9853 3.5868 3.2281
2.8894 3.4184 3.7982 3.5070 3.9853 -5.4833 -4.8928 -5.4833
3.4184 3.7982 4.4281 5.0186 5.6747 5.1072 5.6747 5.1870
3.0765 3.4184 3.9853 4.4281 5.1072 4.5965 5.1072 4.6683
3.4184 3.9853 4.4281 5.1072 5.6747 5.1870 5.8431 5.2588
-96.9235 -96.4930 -96.0147 -95.4833 -94.8928 -95.3317 -94.7412 -95.2671

val(:,:,3) =

-97.4980 2.8894 3.3199 3.7982 3.4184 3.0765 2.7689 -97.5798
-97.1106 3.3199 3.7982 4.3296 3.7982 3.4184 3.0765 -97.2311
-97.3995 3.7982 4.3296 4.9201 3.5070 3.1563 2.8407 -97.4434
-97.0121 3.4184 3.8967 4.4281 5.0186 5.6747 5.0186 -95.4833
-97.3995 3.8967 4.4281 5.0186 5.5762 6.3052 5.7633 -95.9349
-96.9235 3.5070 3.9853 4.5167 5.1072 5.6747 5.1870 -95.3317

-97.2311 3.9853 4.5167 5.1072 5.7633 6.4923 5.7633 -95.7985

-96.9235 3.5868 4.0651 4.5965 5.1870 5.8431 5.2588 -95.2671

val(:,:,4) =

-187.4980 2.8894 3.3199 3.7982 3.4184 3.0765 2.7689 -187.5080

-187.1106 3.3199 3.7982 4.3296 2.9879 2.6891 2.4202 -187.8218

-186.6801 3.7982 4.3296 4.9201 3.4184 3.0765 2.7689 -184.8928

-186.2018 4.3296 4.9201 5.5762 6.3052 7.1152 6.4037 -182.7863

-186.5816 3.8967 4.4281 5.0186 5.5762 6.3052 5.7633 -184.8928

-186.1033 4.4281 5.0186 5.6747 6.4037 7.2137 6.4037 -182.7863

-186.4930 3.9853 4.5167 5.1072 5.7633 6.4923 5.8431 -184.7412

-186.0147 4.5167 5.1072 5.7633 6.4923 7.3023 6.5721 -182.6179

val(:,:,5) =

-97.4980 2.8894 3.3199 3.7982 3.4184 3.0765 2.7689 -97.5798

-97.1106 3.3199 3.7982 4.3296 3.7982 3.4184 3.0765 -97.2311

-96.6801 3.7982 4.3296 4.9201 3.5070 3.1563 2.8407 -97.4434

-96.2018 4.3296 4.9201 5.5762 6.3052 7.1152 6.4037 -94.3253

-95.6704 4.9201 5.5762 6.3052 7.1152 8.0152 7.1152 -91.9848

-96.1033 4.4281 5.0186 5.6747 6.4037 7.2137 6.4923 -94.1569

-95.5719 5.0186 5.6747 6.4037 7.2137 8.1137 7.3023 -91.7977

-96.1033 4.5167 5.1072 5.7633 6.4923 7.3023 6.5721 -94.0851

val(:,:,6) =

-97.4980 -97.1106 -96.6801 -97.0121 -96.5816 -96.9235 -97.2311 -97.5080

2.8894 3.3199 3.7982 3.4184 3.8967 3.5070 3.1563 2.8407

3.3199 3.7982 4.3296 3.8967 4.4281 3.9853 3.5868 3.2281

3.7982 4.3296 4.9201 4.4281 5.0186 -4.3253 -3.5963 -4.2367

4.3296 4.9201 5.5762 6.3052 7.1152 6.4037 7.2137 6.4923

4.9201 5.5762 6.3052 7.1152 8.0152 9.0152 8.1137 9.1137

4.3296 5.0186 5.5762 6.4037 7.1152 6.4923 7.3023 6.5721

-96.1033 -95.5719 -94.9814 -94.3253 -93.5963 -94.1569 -93.4279 -94.0851

val(:,:,7) =

-187.4980 -187.1106 -187.3995 -187.0121 -187.3109 -186.9235 -187.2311 -187.5080
2.8894 3.3199 2.9879 3.4184 3.0765 3.5070 3.1563 2.8407
3.3199 3.7982 3.4184 3.8967 3.5070 3.9853 3.5868 3.2281
3.7982 4.3296 3.8967 4.4281 3.9853 -5.4833 -13.8928 -13.2367
4.3296 4.9201 5.5762 6.3052 5.6747 6.4037 5.7633 6.4923
4.9201 5.5762 6.3052 7.1152 8.0152 9.0152 9.0152 9.1137
4.3296 5.0186 5.5762 6.4037 5.7633 6.4037 5.8431 6.5721
-186.0147 -185.4833 -184.8928 -184.2367 -184.8928 -184.2367 -184.7412 -184.0851

val(:,:,8) =

-97.4980 -97.1106 -96.6801 -97.0121 -96.5816 -96.9235 -97.2311 -97.5080
2.8894 3.3199 3.7982 3.4184 3.8967 3.5070 3.1563 2.8407
3.3199 3.7982 4.3296 3.8967 4.4281 3.9853 3.5868 3.2281
3.7982 4.3296 4.9201 4.4281 5.0186 -4.3253 -3.5963 -4.2367
4.3296 4.9201 5.5762 6.3052 7.1152 6.4037 7.2137 6.4923
4.9201 5.5762 6.3052 7.1152 8.0152 9.0152 8.1137 9.1137
4.3296 5.0186 5.5762 6.4037 7.1152 6.4923 7.3023 6.5721
-96.1033 -95.5719 -94.9814 -94.3253 -93.5963 -94.1569 -93.4279 -94.0851

val(:,:,9) =

-97.4980 2.8894 3.3199 3.7982 3.4184 3.0765 2.7689 -97.5798
-97.1106 3.3199 3.7982 4.3296 3.7982 3.4184 3.0765 -97.2311
-96.6801 3.7982 4.3296 4.9201 3.5070 3.1563 2.8407 -97.4434
-96.2018 4.3296 4.9201 5.5762 6.3052 7.1152 6.4037 -94.3253
-95.6704 4.9201 5.5762 6.3052 7.1152 8.0152 7.1152 -91.9848
-96.1033 4.4281 5.0186 5.6747 6.4037 7.2137 6.4923 -94.1569
-95.5719 5.0186 5.6747 6.4037 7.2137 8.1137 7.3023 -91.7977
-96.1033 4.5167 5.1072 5.7633 6.4923 7.3023 6.5721 -94.0851

val(:,:,10) =

-187.4980 2.8894 3.3199 3.7982 3.4184 3.0765 2.7689 -187.5080

-187.1106 3.3199 3.7982 4.3296 2.9879 2.6891 2.4202 -187.8218

-186.6801 3.7982 4.3296 4.9201 3.4184 3.0765 2.7689 -184.8928

-186.2018 4.3296 4.9201 5.5762 6.3052 7.1152 6.4037 -182.7863

-186.5816 3.8967 4.4281 5.0186 5.5762 6.3052 5.6747 -184.8928

-186.1033 4.4281 5.0186 5.6747 6.4037 7.2137 6.4037 -182.7863

-186.4930 3.9853 4.5167 5.1072 5.7633 6.4923 5.8431 -184.7412

-186.0147 4.5167 5.1072 5.7633 6.4923 7.3023 6.5721 -182.6179

val(:,:,11) =

-97.4980 2.8894 3.3199 3.7982 3.4184 3.0765 2.7689 -97.5798

-97.1106 3.3199 3.7982 4.3296 3.7982 3.4184 3.0765 -97.2311

-97.3995 3.7982 4.3296 4.9201 3.5070 3.1563 2.8407 -97.4434

-97.0121 3.4184 3.8967 4.4281 5.0186 5.6747 5.1072 -95.5719

-97.3995 3.8967 4.4281 5.0186 5.5762 6.3052 5.5762 -96.0147

-96.9235 3.5070 3.9853 4.5167 5.1072 5.6747 5.1072 -95.4035

-97.2311 3.9853 4.5167 5.1072 5.7633 6.4923 5.7633 -95.8631

-96.9235 3.5868 4.0651 4.5965 5.1870 5.8431 5.2588 -95.2671

val(:,:,12) =

-97.4980 -97.1106 -96.6801 -97.0121 -96.5816 -96.9235 -97.2311 -97.5080

2.8894 3.3199 3.7982 3.4184 3.8967 3.5070 3.1563 2.8407

3.3199 3.7982 4.3296 3.8967 4.4281 3.9853 3.5868 3.2281

2.8894 3.4184 3.7982 3.5070 3.9853 -5.4833 -4.8928 -5.4035

3.4184 3.7982 4.4281 5.0186 5.6747 5.1072 5.6747 5.1072

3.0765 3.4184 3.9853 4.4281 4.9201 4.5965 5.1072 4.5965

3.4184 3.9853 4.4281 5.1072 5.6747 5.1870 5.8431 5.2588

-96.9235 -96.4930 -96.0147 -95.4833 -94.8928 -95.3317 -94.7412 -95.2671

Value Iteration

Pe =0, s_goal=[6,7,6]

val(:,:,1) =

-186.6005 -186.2131 -186.6005 -186.2131 -186.6005 -186.2131 -186.6005 -186.9492
3.7869 4.2174 3.7869 4.2174 3.7869 4.2174 3.7869 3.3995
4.2174 4.6957 4.2174 4.6957 4.2174 4.6957 4.2174 3.7869
3.7869 4.2174 3.7869 4.2174 3.7869 -5.7826 -14.3043 -13.7729
4.2174 4.6957 5.2271 5.8176 5.2271 5.8176 5.2271 5.8176
3.7869 4.2174 4.6957 5.2271 4.6957 5.2271 4.6957 5.2271
4.2174 4.6957 5.2271 5.8176 5.2271 5.8176 5.2271 5.8176
-186.2131 -185.7826 -185.3043 -184.7729 -185.3043 -184.7729 -185.3043 -184.7729

val(:,:,2) =

-96.6005 -96.2131 -95.7826 -96.2131 -95.7826 -96.2131 -96.6005 -96.9492
3.7869 4.2174 4.6957 4.2174 4.6957 4.2174 3.7869 3.3995
4.2174 4.6957 5.2271 4.6957 5.2271 4.6957 4.2174 3.7869
3.7869 4.2174 4.6957 4.2174 4.6957 -4.7729 -4.1824 -4.7729
4.2174 4.6957 5.2271 5.8176 6.4737 5.8176 6.4737 5.8176
3.7869 4.2174 4.6957 5.2271 5.8176 5.2271 5.8176 5.2271
4.2174 4.6957 5.2271 5.8176 6.4737 5.8176 6.4737 5.8176
-96.2131 -95.7826 -95.3043 -94.7729 -94.1824 -94.7729 -94.1824 -94.7729

val(:,:,3) =

-96.6005 3.7869 4.2174 4.6957 4.2174 3.7869 3.3995 -96.9492
-96.2131 4.2174 4.6957 5.2271 4.6957 4.2174 3.7869 -96.6005
-96.6005 4.6957 5.2271 5.8176 4.2174 3.7869 3.3995 -96.9492
-96.2131 4.2174 4.6957 5.2271 5.8176 6.4737 5.8176 -94.7729
-96.6005 4.6957 5.2271 5.8176 6.4737 7.2027 6.4737 -95.3043
-96.2131 4.2174 4.6957 5.2271 5.8176 6.4737 5.8176 -94.7729
-96.6005 4.6957 5.2271 5.8176 6.4737 7.2027 6.4737 -95.3043
-96.2131 4.2174 4.6957 5.2271 5.8176 6.4737 5.8176 -94.7729

val(:,:,4) =

-186.6005 3.7869 4.2174 4.6957 4.2174 3.7869 3.3995 -186.9492

-186.2131 4.2174 4.6957 5.2271 3.7869 3.3995 3.0508 -187.2630

-185.7826 4.6957 5.2271 5.8176 4.2174 3.7869 3.3995 -184.1824

-185.3043 5.2271 5.8176 6.4737 7.2027 8.0127 7.2027 -181.9873

-185.7826 4.6957 5.2271 5.8176 6.4737 7.2027 6.4737 -184.1824

-185.3043 5.2271 5.8176 6.4737 7.2027 8.0127 7.2027 -181.9873

-185.7826 4.6957 5.2271 5.8176 6.4737 7.2027 6.4737 -184.1824

-185.3043 5.2271 5.8176 6.4737 7.2027 8.0127 7.2027 -181.9873

val(:,:,5) =

-96.6005 3.7869 4.2174 4.6957 4.2174 3.7869 3.3995 -96.9492

-96.2131 4.2174 4.6957 5.2271 4.6957 4.2174 3.7869 -96.6005

-95.7826 4.6957 5.2271 5.8176 4.2174 3.7869 3.3995 -96.9492

-95.3043 5.2271 5.8176 6.4737 7.2027 8.0127 7.2027 -93.5263

-94.7729 5.8176 6.4737 7.2027 8.0127 8.9127 8.0127 -91.0873

-95.3043 5.2271 5.8176 6.4737 7.2027 8.0127 7.2027 -93.5263

-94.7729 5.8176 6.4737 7.2027 8.0127 8.9127 8.0127 -91.0873

-95.3043 5.2271 5.8176 6.4737 7.2027 8.0127 7.2027 -93.5263

val(:,:,6) =

-96.6005 -96.2131 -95.7826 -96.2131 -95.7826 -96.2131 -96.6005 -96.9492

3.7869 4.2174 4.6957 4.2174 4.6957 4.2174 3.7869 3.3995

4.2174 4.6957 5.2271 4.6957 5.2271 4.6957 4.2174 3.7869

4.6957 5.2271 5.8176 5.2271 5.8176 -3.5263 -2.7973 -3.5263

5.2271 5.8176 6.4737 7.2027 8.0127 7.2027 8.0127 7.2027

5.8176 6.4737 7.2027 8.0127 8.9127 9.9127 8.9127 9.9127

5.2271 5.8176 6.4737 7.2027 8.0127 7.2027 8.0127 7.2027

-95.3043 -94.7729 -94.1824 -93.5263 -92.7973 -93.5263 -92.7973 -93.5263

val(:,:,7) =

-186.6005 -186.2131 -186.6005 -186.2131 -186.6005 -186.2131 -186.6005 -186.9492

3.7869 4.2174 3.7869 4.2174 3.7869 4.2174 3.7869 3.3995

4.2174 4.6957 4.2174 4.6957 4.2174 4.6957 4.2174 3.7869

```
   4.6957   5.2271   4.6957   5.2271   4.6957   -4.7729  -13.1824  -12.5263
   5.2271   5.8176   6.4737   7.2027   6.4737   7.2027   6.4737   7.2027
   5.8176   6.4737   7.2027   8.0127   8.9127   9.9127   9.9127   9.9127
   5.2271   5.8176   6.4737   7.2027   6.4737   7.2027   6.4737   7.2027
-185.3043 -184.7729 -184.1824 -183.5263 -184.1824 -183.5263 -184.1824 -183.5263
```

val(:,:,8) =

```
  -96.6005  -96.2131  -95.7826  -96.2131  -95.7826  -96.2131  -96.6005  -96.9492
   3.7869   4.2174   4.6957   4.2174   4.6957   4.2174   3.7869   3.3995
   4.2174   4.6957   5.2271   4.6957   5.2271   4.6957   4.2174   3.7869
   4.6957   5.2271   5.8176   5.2271   5.8176   -3.5263   -2.7973   -3.5263
   5.2271   5.8176   6.4737   7.2027   8.0127   7.2027   8.0127   7.2027
   5.8176   6.4737   7.2027   8.0127   8.9127   9.9127   8.9127   9.9127
   5.2271   5.8176   6.4737   7.2027   8.0127   7.2027   8.0127   7.2027
  -95.3043  -94.7729  -94.1824  -93.5263  -92.7973  -93.5263  -92.7973  -93.5263
```

val(:,:,9) =

```
  -96.6005   3.7869   4.2174   4.6957   4.2174   3.7869   3.3995  -96.9492
  -96.2131   4.2174   4.6957   5.2271   4.6957   4.2174   3.7869  -96.6005
  -95.7826   4.6957   5.2271   5.8176   4.2174   3.7869   3.3995  -96.9492
  -95.3043   5.2271   5.8176   6.4737   7.2027   8.0127   7.2027  -93.5263
  -94.7729   5.8176   6.4737   7.2027   8.0127   8.9127   8.0127  -91.0873
  -95.3043   5.2271   5.8176   6.4737   7.2027   8.0127   7.2027  -93.5263
  -94.7729   5.8176   6.4737   7.2027   8.0127   8.9127   8.0127  -91.0873
  -95.3043   5.2271   5.8176   6.4737   7.2027   8.0127   7.2027  -93.5263
```

val(:,:,10) =

```
 -186.6005   3.7869   4.2174   4.6957   4.2174   3.7869   3.3995 -186.9492
 -186.2131   4.2174   4.6957   5.2271   3.7869   3.3995   3.0508 -187.2630
 -185.7826   4.6957   5.2271   5.8176   4.2174   3.7869   3.3995 -184.1824
 -185.3043   5.2271   5.8176   6.4737   7.2027   8.0127   7.2027 -181.9873
 -185.7826   4.6957   5.2271   5.8176   6.4737   7.2027   6.4737 -184.1824
 -185.3043   5.2271   5.8176   6.4737   7.2027   8.0127   7.2027 -181.9873
```

-185.7826  4.6957  5.2271  5.8176  6.4737  7.2027  6.4737 -184.1824

-185.3043  5.2271  5.8176  6.4737  7.2027  8.0127  7.2027 -181.9873

val(:,:,11) =

-96.6005  3.7869  4.2174  4.6957  4.2174  3.7869  3.3995 -96.9492

-96.2131  4.2174  4.6957  5.2271  4.6957  4.2174  3.7869 -96.6005

-96.6005  4.6957  5.2271  5.8176  4.2174  3.7869  3.3995 -96.9492

-96.2131  4.2174  4.6957  5.2271  5.8176  6.4737  5.8176 -94.7729

-96.6005  4.6957  5.2271  5.8176  6.4737  7.2027  6.4737 -95.3043

-96.2131  4.2174  4.6957  5.2271  5.8176  6.4737  5.8176 -94.7729

-96.6005  4.6957  5.2271  5.8176  6.4737  7.2027  6.4737 -95.3043

-96.2131  4.2174  4.6957  5.2271  5.8176  6.4737  5.8176 -94.7729

val(:,:,12) =

-96.6005 -96.2131 -95.7826 -96.2131 -95.7826 -96.2131 -96.6005 -96.9492

 3.7869  4.2174  4.6957  4.2174  4.6957  4.2174  3.7869  3.3995

 4.2174  4.6957  5.2271  4.6957  5.2271  4.6957  4.2174  3.7869

 3.7869  4.2174  4.6957  4.2174  4.6957 -4.7729 -4.1824 -4.7729

 4.2174  4.6957  5.2271  5.8176  6.4737  5.8176  6.4737  5.8176

 3.7869  4.2174  4.6957  5.2271  5.8176  5.2271  5.8176  5.2271

 4.2174  4.6957  5.2271  5.8176  6.4737  5.8176  6.4737  5.8176

-96.2131 -95.7826 -95.3043 -94.7729 -94.1824 -94.7729 -94.1824 -94.7729

## Appendix D

## Pe = 0.25, $\gamma$ =0.9, s0 = [1,6,6], s_goal=[6,7,6]

Policy Iteration

Pe =25, s_goal = [6,7,6]

val(:,:,1) =

-205.1842 -205.4140 -205.1814 -205.4097 -205.5686 -205.3773 -205.5646 -205.3728
0.5900 0.6289 0.6345 0.6313 0.6395 0.5510 0.4656 0.3685
0.6831 0.7471 0.7628 0.7494 0.7676 0.6309 0.5214 0.4103
0.7540 0.8191 0.8503 0.8217 0.8547 -9.0460 -15.0832 -17.8820
0.8135 0.9162 1.0010 1.0505 1.0735 1.0612 1.0761 1.0659
0.8540 0.9711 1.0840 1.1775 1.2044 1.1795 1.2077 1.1606
0.7928 0.9103 0.9982 1.0530 1.0744 1.0640 1.0778 1.0682
-204.9055 -205.0527 -204.9030 -205.0493 -204.8003 -204.5888 -204.7973 -204.5849

val(:,:,2) =

-186.4041 -115.1667 -115.8100 -118.1053 -124.8261 -140.5869 -149.8718 -212.5741
-31.7517 0.6739 0.6960 0.6042 0.5226 0.4343 0.2867 -71.7155
-31.6967 0.7774 0.8302 0.7909 0.7100 0.5995 0.4666 -40.8307
-31.9776 0.8515 0.8465 0.8596 0.7388 -6.7443 -9.7968 -50.5081
-33.4539 0.9100 1.0243 1.1204 1.1589 1.0484 0.8056 -35.6086
-40.2567 0.9450 1.0814 1.2151 1.3253 1.3233 1.3085 -32.1743
-70.6591 0.6359 0.8539 0.9515 1.0239 1.1455 1.2091 -31.5722
-211.5056 -149.3644 -140.1493 -124.2792 -117.5197 -115.1714 -114.4709 -186.3806

val(:,:,3) =

-186.4218 -31.7256 -31.6863 -32.0344 -33.8033 -40.6843 -71.0281 -212.2326

-115.1682  0.6600  0.7793  0.7899  0.5638  0.4766  0.2440 -149.9972

-115.7470  0.7488  0.9073  0.9236  0.5598  0.4754  0.3258 -140.9040

-117.8767  0.8439  0.9883  0.8871  0.6303 -1.8858 -2.7296 -130.9709

-124.4602  0.8921  1.0366  1.1679  1.2943  1.3558  0.9028 -126.9550

-140.2156  0.9253  1.0776  1.2257  1.3804  1.5411  1.3335 -119.3170

-149.4856  0.6539  0.9339  1.0892  1.2362  1.4352  1.2516 -116.0823

-211.5346 -70.6704 -40.0950 -33.1025 -31.4097 -30.8342 -31.0826 -186.4054

val(:,:,4) =

-205.1848  0.5982  0.6965  0.7492  0.4951  0.4170  0.3054 -208.3990

-205.3876  0.6782  0.8061  0.8704  0.5176  0.4384  0.3165 -211.0171

-205.1820  0.7534  0.9217  0.9973  0.4968  0.4196  0.3078 -208.3968

-205.3839  0.8221  1.0659  1.2150  1.3636  1.5920  1.3552 -209.6021

-205.1006  0.9708  1.1492  1.3172  1.5043  1.7953  1.4101 -207.2808

-204.8630  0.9585  1.1061  1.2454  1.3935  1.5946  1.3581 -205.9820

-205.0970  0.9743  1.1528  1.3206  1.5077  1.7994  1.4134 -207.2778

-204.8584  0.9629  1.1104  1.2497  1.3973  1.5966  1.1533 -205.9783

val(:,:,5) =

-211.7273 -70.8662 -40.4091 -33.5255 -32.1835 -31.9368 -31.9892 -187.7676

-149.6479  0.4523  0.7187  0.8169  0.5290  0.4486  0.3174 -118.9969

-140.3925  0.6428  0.8444  0.9598  0.5438  0.4550  0.2498 -124.6683

-124.5945  0.7294  0.9908  1.1322  1.2692 -1.0135 -1.8920 -123.4700

-117.8731  0.8338  1.2203  1.3978  1.5971  1.8785  1.5653 -123.4834

-115.5064  1.0218  1.2375  1.4161  1.6182  1.9531  1.4113 -139.3839

-114.7885  1.1042  1.2762  1.4429  1.5997  1.8812  0.9737 -148.4981

-186.0865 -31.4754 -31.3779 -31.5816 -32.9780 -39.7054 -70.3879 -211.2676

val(:,:,6) =

-211.7274 -149.6377 -140.3777 -124.5742 -117.9992 -115.9337 -115.3854 -187.8361

-70.8669 0.4393 0.6201 0.6869 0.6325 0.5293 0.4168 -32.7528

-40.4282 0.7096 0.7977 0.8321 0.5563 0.4546 0.3109 -35.6632

-33.5323 0.8215 0.9276 0.9504 1.0613 -6.2864 -9.0537 -48.8941

-31.7991 1.0136 1.1590 1.2990 1.3749 1.5350 1.4950 -32.8632

-31.2218 1.3279 1.5424 1.7916 2.0710 2.4966 2.1013 -38.7522

-31.3634 1.0968 1.2300 1.3479 1.5031 1.5340 0.9750 -70.6884

-186.1575 -114.8636 -115.3793 -117.4426 -123.9880 -139.8120 -149.0657 -211.5594

val(:,:,7) =

-205.1689 -205.4053 -205.1663 -205.4010 -205.5699 -205.3778 -205.5660 -205.3734

0.6085 0.6337 0.6308 0.6360 0.6360 0.5467 0.4624 0.3660

0.7088 0.7524 0.7556 0.7548 0.7602 0.6281 0.5199 0.4087

0.7744 0.8327 0.8656 0.8351 0.8704 -9.0410 -15.0911 -17.8954

0.9742 1.1150 1.2255 1.3505 1.4140 1.3529 1.4178 1.3440

1.2309 1.4459 1.6839 1.9645 2.3014 2.6233 2.6241 2.6263

0.9897 1.1149 1.2323 1.3520 1.4157 1.3549 1.4196 1.3458

-204.9183 -205.0899 -204.9158 -205.0868 -204.8675 -204.6477 -204.8643 -204.6436

val(:,:,8) =

-186.3818 -115.1507 -115.8040 -118.1030 -124.8254 -140.5874 -149.8715 -212.5824

-31.7494 0.6940 0.7009 0.6031 0.5210 0.4322 0.2851 -71.7241

-31.6889 0.8035 0.8414 0.7984 0.7150 0.5998 0.4660 -40.8340

-31.9279 0.8536 0.8643 0.8824 0.7505 -6.7346 -9.8131 -50.5157

-33.2636 1.0937 1.2413 1.3421 1.5010 1.5322 0.9734 -35.3916

-39.9570  1.2632  1.5296  1.7749  2.0705  2.4959  2.1009 -30.8095

-70.6833  0.7188  1.0926  1.2307  1.3776  1.5368  1.4970 -31.0740

-211.5478 -149.4331 -140.1540 -124.3115 -117.5666 -115.2220 -114.5446 -186.2465

val(:,:,9) =

-186.4151 -31.6991 -31.6747 -32.0300 -33.8006 -40.6843 -71.0290 -212.2373

-115.1513  0.6984  0.7941  0.7930  0.5660  0.4755  0.2432 -150.0005

-115.6929  0.8045  0.9164  0.9291  0.5637  0.4734  0.3243 -140.9090

-117.7402  0.9186  1.0264  0.8988  0.6365 -1.9277 -2.7377 -130.9753

-124.2488  1.0408  1.2604  1.4254  1.5969  1.8788  0.9711 -126.8825

-140.0719  0.8853  1.2012  1.3788  1.6161  1.9507  1.4108 -118.6344

-149.2742  0.6524  1.2094  1.3862  1.5996  1.8806  1.5677 -115.2376

-211.5008 -70.6364 -40.1530 -33.1902 -31.4980 -30.9720 -31.1294 -186.0062

val(:,:,10) =

-205.1702  0.6294  0.7107  0.7537  0.4992  0.4187  0.3075 -208.4342

-205.3581  0.7215  0.8220  0.8741  0.5215  0.4378  0.3162 -211.0909

-205.1673  0.8069  0.9384  1.0046  0.5009  0.4213  0.3096 -208.4320

-205.3544  0.9105  1.0974  1.2315  1.3923  1.5917  1.1464 -209.9085

-205.2931  0.8847  1.1301  1.2956  1.5046  1.7961  1.4100 -206.5174

-205.0497  0.9135  1.1008  1.2349  1.3949  1.5943  1.3553 -205.3183

-205.2899  0.8881  1.1324  1.2977  1.5070  1.7981  1.4125 -206.5142

-205.0453  0.7787  1.0449  1.1927  1.3688  1.5966  1.3601 -205.3143

val(:,:,11) =

-211.7245 -70.8616 -40.3971 -33.5200 -32.1800 -31.9365 -31.9895 -187.7777

-149.6526  0.4731  0.7324  0.8240  0.5333  0.4507  0.3200 -119.0366

-140.3883  0.6778  0.8756  0.9590  0.5479  0.4575  0.2518 -124.8037

-124.5989  0.7594  0.9605  1.1172  1.2856 -0.9821 -1.8701 -123.8704

-117.9221  0.8328  0.9674  1.0808  1.2316  1.4308  1.2484 -124.3042

-115.6840  0.8706  1.0642  1.2136  1.3825  1.5396  1.3310 -140.0974

-115.0165  0.8398  1.0249  1.1566  1.2962  1.3594  0.9052 -149.3039

-186.2113 -31.4987 -31.3074 -31.5026 -32.8780 -39.6230 -70.4648 -211.3982

val(:,:,12) =

-211.7186 -149.6235 -140.3687 -124.5657 -117.9940 -115.9311 -115.3848 -187.8526

 -70.8525  0.4472  0.6151  0.6934  0.6381  0.5331  0.4200 -32.7559

 -40.4334  0.7348  0.8050  0.8517  0.5587  0.4563  0.3139 -35.6852

 -33.5398  0.8186  0.9205  0.9300  1.0498 -6.2820 -9.0388 -48.9956

 -31.9314  0.8627  0.9703  1.0627  1.0380  1.1453  1.2065 -33.3465

 -31.5374  0.9237  1.0699  1.2090  1.3278  1.3242  1.3073 -40.0495

 -31.5451  0.8930  1.0175  1.1196  1.1599  1.0519  0.8088 -70.6073

-186.2067 -114.8420 -115.3378 -117.3954 -123.9550 -139.8289 -148.9858 -211.4267

Value Iteration

Pe = 25%, s_goal = [6,7,6]

val(:,:,1) =

-205.2397 -205.4691 -205.2397 -205.4691 -205.6272 -205.4362 -205.6272 -205.4362

  0.5340  0.5718  0.5781  0.5718  0.5781  0.4894  0.4025  0.3046

  0.6266  0.6898  0.7059  0.6898  0.7059  0.5681  0.4571  0.3453

  0.6969  0.7610  0.7927  0.7610  0.7927 -9.1088 -15.1464 -17.9456

0.7553 0.8579 0.9418 0.9916 1.0134 1.0017 1.0134 1.0017

0.7956 0.9114 1.0250 1.1165 1.1438 1.1165 1.1438 1.0955

0.7329 0.8499 0.9373 0.9911 1.0120 1.0011 1.0120 1.0011

-204.9663 -205.1137 -204.9663 -205.1137 -204.8644 -204.6530 -204.8644 -204.6530

val(:,:,2) =

-186.4597 -115.2225 -115.8668 -118.1616 -124.8844 -140.6461 -149.9329 -212.6353

-31.8076 0.6171 0.6387 0.5466 0.4645 0.3743 0.2268 -71.7770

-31.7532 0.7202 0.7730 0.7311 0.6503 0.5373 0.4033 -40.8952

-32.0348 0.7933 0.7893 0.7995 0.6764 -6.8081 -9.8599 -50.5730

-33.5123 0.8512 0.9649 1.0611 1.0978 0.9892 0.7441 -35.6736

-40.3153 0.8854 1.0219 1.1545 1.2668 1.2620 1.2453 -32.2398

-70.7186 0.5758 0.7915 0.8886 0.9609 1.0809 1.1434 -31.6394

-211.5658 -149.4251 -140.2132 -124.3444 -117.5855 -115.2377 -114.5376 -186.4485

val(:,:,3) =

-186.4776 -31.7814 -31.7428 -32.0906 -33.8615 -40.7449 -71.0868 -212.2936

-115.2243 0.6032 0.7223 0.7317 0.5075 0.4171 0.1851 -150.0596

-115.8039 0.6914 0.8500 0.8666 0.5010 0.4132 0.2641 -140.9676

-117.9342 0.7859 0.9303 0.8269 0.5686 -1.9457 -2.7918 -131.0362

-124.5188 0.8330 0.9773 1.1083 1.2340 1.2965 0.8418 -127.0209

-140.2745 0.8663 1.0185 1.1662 1.3211 1.4795 1.2708 -119.3836

-149.5469 0.5940 0.8720 1.0269 1.1733 1.3711 1.1869 -116.1496

-211.5970 -70.7328 -40.1588 -33.1668 -31.4743 -30.8994 -31.1482 -186.4741

val(:,:,4) =

-205.2407  0.5420  0.6401  0.6928  0.4372  0.3560  0.2458 -208.4619

-205.4435  0.6215  0.7489  0.8125  0.4622  0.3788  0.2575 -211.0802

-205.2407  0.6959  0.8643  0.9398  0.4372  0.3560  0.2458 -208.4619

-205.4435  0.7640  1.0078  1.1565  1.3048  1.5326  1.2952 -209.6647

-205.1583  0.9126  1.0909  1.2585  1.4452  1.7361  1.3500 -207.3452

-204.9210  0.8999  1.0471  1.1862  1.3334  1.5326  1.2952 -206.0471

-205.1583  0.9126  1.0909  1.2585  1.4452  1.7361  1.3500 -207.3452

-204.9210  0.8999  1.0471  1.1862  1.3334  1.5322  1.0870 -206.0471

val(:,:,5) =

-211.7820 -70.9220 -40.4661 -33.5827 -32.2421 -31.9981 -32.0496 -187.8303

-149.7032  0.3962  0.6612  0.7585  0.4722  0.3859  0.2571 -119.0607

-140.4502  0.5841  0.7860  0.9015  0.4829  0.3914  0.1899 -124.7319

-124.6533  0.6702  0.9320  1.0732  1.2093 -1.0741 -1.9529 -123.5327

-117.9327  0.7740  1.1620  1.3391  1.5379  1.8188  1.5046 -123.5454

-115.5663  0.9615  1.1772  1.3556  1.5573  1.8914  1.3500 -139.4469

-114.8492  1.0432  1.2147  1.3813  1.5379  1.8188  0.9105 -148.5623

-186.1486 -31.5382 -31.4412 -31.6453 -33.0421 -39.7704 -70.4529 -211.3339

val(:,:,6) =

-211.7833 -149.6940 -140.4357 -124.6331 -118.0588 -115.9949 -115.4480 -187.8987

 -70.9227  0.3832  0.5612  0.6271  0.5727  0.4668  0.3542 -32.8165

 -40.4849  0.6519  0.7397  0.7722  0.4960  0.3927  0.2505 -35.7281

 -33.5897  0.7634  0.8695  0.8900  1.0002 -6.3483 -9.1161 -48.9589

 -31.8570  0.9551  1.1001  1.2399  1.3143  1.4742  1.4334 -32.9265

 -31.2801  1.2692  1.4832  1.7323  2.0113  2.4366  2.0410 -38.8147

 -31.4228  1.0371  1.1697  1.2873  1.4416  1.4724  0.9120 -70.7538

-186.2192 -114.9258 -115.4418 -117.5056 -124.0518 -139.8762 -149.1324 -211.6273

val(:,:,7) =

-205.2243 -205.4603 -205.2243 -205.4603 -205.6286 -205.4367 -205.6286 -205.4367

0.5528 0.5766 0.5745 0.5766 0.5745 0.4850 0.3991 0.3020

0.6526 0.6952 0.6985 0.6952 0.6985 0.5653 0.4555 0.3442

0.7179 0.7746 0.8081 0.7746 0.8081 -9.1042 -15.1550 -17.9597

0.9171 1.0573 1.1674 1.2915 1.3550 1.2915 1.3550 1.2803

1.1733 1.3881 1.6255 1.9059 2.2422 2.5641 2.5641 2.5641

0.9306 1.0554 1.1725 1.2915 1.3550 1.2915 1.3550 1.2803

-204.9800 -205.1518 -204.9800 -205.1518 -204.9310 -204.7113 -204.9310 -204.7113

val(:,:,8) =

-186.4372 -115.2061 -115.8607 -118.1593 -124.8836 -140.6466 -149.9326 -212.6436

-31.8051 0.6373 0.6436 0.5453 0.4627 0.3721 0.2250 -71.7857

-31.7453 0.7469 0.7843 0.7386 0.6552 0.5376 0.4026 -40.8980

-31.9850 0.7955 0.8069 0.8224 0.6881 -6.7985 -9.8782 -50.5810

-33.3209 1.0359 1.1830 1.2835 1.4416 1.4724 0.9120 -35.4565

-40.0149 1.2055 1.4713 1.7164 2.0113 2.4366 2.0410 -30.8728

-70.7420 0.6598 1.0311 1.1688 1.3152 1.4742 1.4334 -31.1396

-211.6078 -149.4934 -140.2173 -124.3763 -117.6320 -115.2878 -114.6107 -186.3137

val(:,:,9) =

-186.4707 -31.7545 -31.7314 -32.0864 -33.8587 -40.7447 -71.0879 -212.2983

-115.2074 0.6420 0.7374 0.7351 0.5096 0.4162 0.1841 -150.0626

-115.7494 0.7478 0.8588 0.8718 0.5049 0.4120 0.2624 -140.9726

-117.7971 0.8610 0.9689 0.8384 0.5747 -1.9917 -2.8014 -131.0406

-124.3063 0.9832 1.2017 1.3662 1.5379 1.8188 0.9105 -126.9486

-140.1301  0.8275  1.1431  1.3205  1.5573  1.8914  1.3500 -118.7000

-149.3347  0.5936  1.1486  1.3251  1.5379  1.8188  1.5046 -115.3035

-211.5629 -70.6982 -40.2165 -33.2542 -31.5624 -31.0368 -31.1946 -186.0737

val(:,:,10) =

-205.2258  0.5738  0.6540  0.6971  0.4413  0.3579  0.2471 -208.4969

-205.4138  0.6652  0.7653  0.8166  0.4660  0.3787  0.2569 -211.1539

-205.2258  0.7504  0.8807  0.9468  0.4413  0.3579  0.2471 -208.4969

-205.4138  0.8535  1.0398  1.1736  1.3334  1.5322  1.0871 -209.9718

-205.3519  0.8275  1.0712  1.2363  1.4452  1.7361  1.3500 -206.5803

-205.1086  0.8535  1.0398  1.1736  1.3334  1.5326  1.2952 -205.3817

-205.3519  0.8275  1.0712  1.2363  1.4452  1.7361  1.3500 -206.5803

-205.1086  0.7158  0.9817  1.1293  1.3050  1.5326  1.2952 -205.3817

val(:,:,11) =

-211.7805 -70.9179 -40.4544 -33.5775 -32.2388 -31.9975 -32.0499 -187.8402

-149.7090  0.4168  0.6752  0.7659  0.4766  0.3881  0.2588 -119.1002

-140.4464  0.6195  0.8171  0.9004  0.4872  0.3941  0.1908 -124.8672

-124.6579  0.7007  0.9020  1.0585  1.2264 -1.0418 -1.9302 -123.9337

-117.9821  0.7739  0.9075  1.0213  1.1720  1.3708  1.1868 -124.3682

-115.7446  0.8113  1.0037  1.1528  1.3211  1.4795  1.2708 -140.1622

-115.0780  0.7789  0.9632  1.0947  1.2340  1.2964  0.8415 -149.3710

-186.2737 -31.5604 -31.3699 -31.5655 -32.9414 -39.6872 -70.5312 -211.4664

val(:,:,12) =

-211.7749 -149.6796 -140.4269 -124.6249 -118.0538 -115.9922 -115.4473 -187.9152

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| -70.9087 | 0.3897 | 0.5562 | 0.6338 | 0.5785 | 0.4709 | 0.3572 | -32.8195 |
| -40.4902 | 0.6774 | 0.7473 | 0.7918 | 0.4985 | 0.3945 | 0.2515 | -35.7499 |
| -33.5975 | 0.7605 | 0.8620 | 0.8698 | 0.9891 | -6.3432 | -9.1004 | -49.0598 |
| -31.9901 | 0.8039 | 0.9111 | 1.0032 | 0.9771 | 1.0840 | 1.1440 | -33.4105 |
| -31.5965 | 0.8638 | 1.0103 | 1.1495 | 1.2668 | 1.2620 | 1.2453 | -40.1147 |
| -31.6054 | 0.8326 | 0.9565 | 1.0582 | 1.0976 | 0.9885 | 0.7434 | -70.6731 |
| -186.2678 | -114.9034 | -115.3996 | -117.4577 | -124.0182 | -139.8933 | -149.0520 | -211.4948 |