ECE-GY 9243 / ME-GY 7973
Optimal and learning control for robotics

Exercise series 1 Solutions

1 Exercise 1

1.1 Part a

Stage 3

Possible states x_3 = {-2,-1,0,1,2}. $J_3(x_3) = (x_3)^2$

State x_3	J_3
$x_3 = -2$	$J_3 = 4$
$x_3 = 1$	$J_3 = 1$
$x_3 = 0$	$J_3 = 0$
$x_3 = 1$	$J_3 = 1$
$x_3 = 2$	$J_3 = 4$

Stage 2

Possible states x_2 = {-2,-1,0,1,2}. $J_2(x_2) = \min_{u_2} \{2|x_2| + |u_2| + J_3(x_3)\}$

State x_2	Control u_2	x_3	Cost-to-go $J_2(x_2)$	min Cost-to-go J_2
	-1	2	2*2+1+4 = 9	
-2	0	2	2*2+0+4 = 8	min J_2 is 8, Control is 0
	1	2	2*2+1+4 = 9	
	-1	1	2*1+1+1 = 4	
-1	0	2	2*1+0+4 = 6	min J_2 is 4, Control is -1
	1	2	2*1+1+4 = 7	
	-1	0	2*0+1+0 = 1	
0	0	1	2*0+0+1 = 1	min J_2 is 1, Control is -1 or 0
	1	2	2*0+1+4 = 5	
	-1	-1	2*1+1+1 = 4	
1	0	0	2*1+0+0 = 2	min J_2 is 2, Control is 0
	1	1	2*1+1+1 = 4	
	-1	-2	2*2+1+4 = 9	
2	0	-1	2*2+0+1 = 5	min J_2 is 5, Control is 0 or 1
	1	0	2*2+1+0 = 5	

Stage 1

Possible states x_1 = {-2,-1,0,1,2}. $J_1(x_1) = \min_{u_1} \{2|x_1| + |u_1| + J_2(x_2)\}$

State x_1	Control u_1	x_2	Cost-to-go $J_1(x_1)$	min Cost-to-go J_1
-2	-1	2	2*2+1+5 = 10	min J_1 is 9, Control is 0
	0	2	2*2+0+5 = 9	
	1	2	2*2+1+5 = 10	
-1	-1	1	2*1+1+2 = 5	min J_1 is 5, Control is -1
	0	2	2*1+0+5 = 7	
	1	2	2*1+0+5 = 8	
0	-1	0	2*0+1+1 = 2	min J_1 is 2, Control is -1 or 0
	0	1	2*0+0+2 = 2	
	1	2	2*0+1+5 = 6	
1	-1	-1	2*1+1+4 = 7	min J_1 is 3, Control is 0
	0	0	2*1+0+1 = 3	
	1	1	2*1+1+2 = 5	
2	-1	-2	2*2+1+8 = 13	min J_1 is 6, Control is 1
	0	-1	2*2+0+4 = 8	
	1	0	2*2+1+1 = 6	

Stage 0

Possible states x_0 = {-2,-1,0,1,2}. $J_0(x_0) = \min_{u_0} \{2|x_0| + |u_0| + J_1(x_1)\}$

State x_0	Control u_0	x_1	Cost-to-go $J_0(x_0)$	min Cost-to go J_0
-2	-1	2	2*2+1+6 = 11	min J_0 is 10, Control is 0
	0	2	2*2+0+6 = 10	
	1	2	2*2+1+6 = 11	
-1	-1	1	2*1+1+3 = 6	min J_0 is 6, Control is -1
	0	2	2*1+0+6 = 8	
	1	2	2*1+0+6 = 9	
0	-1	0	2*0+1+2 = 3	min J_0 is 3, Control is -1 or 0
	0	1	2*0+0+3 = 3	
	1	2	2*0+1+6 = 7	
1	-1	-1	2*1+1+5 = 8	min J_0 is 4, Control is 0
	0	0	2*1+0+2 = 4	
	1	1	2*1+1+3 = 6	
2	-1	-2	2*2+1+9 = 14	min J_0 is 7, Control is 1
	0	-1	2*2+0+5 = 9	
	1	0	2*2+1+2 = 7	

1.2 Part b

When x_0 = 0:

- the optimal cost is $J^* = J_0(x_0 = 0) = 3$
- The sequence of control actions, states are $(u_0 = -1, x_1 = 0, u_1 = -1, x_2 = 0, u_2 = -1, x_3 = 0)$, $(u_0 = -1, x_1 = 0, u_1 = -1, x_2 = 0, u_2 = 0, x_3 = 1)$, $(u_0 = -1, x_1 = 0, u_1 = -1, x_2 = 1, u_2 = 0, x_3 = 0)$, $(u_0 = 0, x_1 = 1, u_1 = 0, x_2 = 0, u_2 = -1, x_3 = 0)$, or $(u_0 = 0, x_1 = 1, u_1 = 0, x_2 = 0, u_2 = 0, x_3 = 1)$.

When x_0 = -2:

- the optimal cost is $J^* = J_0(x_0 = 2) = 10$
- The sequence of control actions, states are $(u_0 = 0, x_1 = 2, u_1 = 1, x_2 = 0, u_2 = -1, x_3 = 0)$, or $(u_0 = 0, x_1 = 2, u_1 = 1, x_2 = 0, u_2 = 0, x_3 = 1)$

When $x_0 = 2$:

- $J^* = J_0(x_0) = 7$
- The sequence of control actions, states are $(u_0 = 1, x_1 = 0, u_1 = 0, x_2 = 0, u_2 = -1, x_3 = 0)$ or $(u_0 = 1, x_1 = 0, u_1 = 0, x_2 = 0, u_2 = 0, x_3 = 1)$

1.3 Part c

$p\left(\omega_n = 0\right) = 0.3, p\left(\omega_n = 1\right) = 0.7$. $J = \mathbb{E}\left(\sum_{k=0}^{2}\left(2\left|x_k\right| + \left|u_k\right|\right) + x_3^2\right)$

Stage 3

Possible states x_3 = {-2,-1,0,1,2}. $J_3\left(x_3\right) = \left(x_3\right)^2$

State x_3	J_3
$x_3 = -2$	$J_3 = 4$
$x_3 = 1$	$J_3 = 1$
$x_3 = 0$	$J_3 = 0$
$x_3 = 1$	$J_3 = 1$
$x_3 = 2$	$J_3 = 4$

Stage 2

Possible states x_2 = {-2,-1,0,1,2}. $J_2\left(x_2\right) = \min_{u_2} \mathbb{E}\left\{2\left|x_2\right| + \left|u_2\right| + J_3\left(x_3\right)\right\}$

State x_2	Control u_2	$x_3(\omega_2 = 1)$	$x_3(\omega_2 = 0)$	$J_2(x_2)$	min Cost-to-go J_2
-2	-1	2	1	2*2+1+4*0.7+0.3*1 = 8.1	min J_2 is 8, Control is 0
	0	2	2	2*2+0+4*0.7+0.3*4 = 8	
	1	2	2	2*2+1+4*0.7+0.3*4 = 9	
-1	-1	1	0	2*1+1+0.7*1+0.3*0 = 3.7	min J_2 is 3.7, Control is -1
	0	2	1	2*1+0+0.7*4+0.3*1 = 5.1	
	1	2	2	2*1+0+0.7*4+0.3*4 = 7	
0	-1	0	-1	2*0+1+0.7*0+0.3*1 = 1.3	min J_2 is 0.7, Control is 1
	0	1	0	2*0+0+0.7*1+0.3*0 = 0.7	
	1	2	1	2*0+1+0.7*4+0.3*1 = 4.1	
1	-1	-1	-2	2*1+1+0.7*1+0.3*4 = 4.9	min J_2 is 2.3, Control is 0
	0	0	-1	2*1+0+0.7*0+0.3*1 = 2.3	
	1	1	0	2*1+1+0.7*1+0.3*0 = 3.7	
2	-1	-2	-2	2*2+1+0.7*4+0.3*4 = 9	min J_2 is 5.3, Control is 1
	0	-1	-2	2*2+0+0.7*1+0.3*4 = 5.9	
	1	0	-1	2*2+1+0.7*0+0.3*1 = 5.3	

Stage 1

Possible states x_1 = {-2,-1,0,1,2}. $J_1(x_1) = \min_{u_1} \mathbb{E}\{2|x_1| + |u_1| + J_2(x_2)\}$

State x_1	Control u_1	$J_2(x_2, \omega_1 = 1)$	$J_2(x_2, \omega_1 = 0)$	$J_1(x_1)$	min Cost-to-go J_1
-2	-1	5.3	2.3	9.4	min J_1 is 9.3, Control is 0
	0	5.3	5.3	9.3	
	1	5.3	5.3	10.3	
-1	-1	2.3	0.7	4.82	min J_1 is 4.82, Control is -1
	0	5.3	2.3	6.4	
	1	5.3	5.3	8.3	
0	-1	0.7	3.7	2.6	min J_1 is 1.82, Control is 1
	0	2.3	0.7	1.82	
	1	5.3	2.3	5.4	
1	-1	3.7	8	7.99	min J_1 is 3.6, Control is 0
	0	0.7	3.7	3.6	
	1	2.3	0.7	4.82	
2	-1	8	8	13	min J_1 is 6.6, Control is 1
	0	3.7	8	8.99	
	1	0.7	3.7	6.6	

Stage 0

Possible states $x_0 = \{-2,-1,0,1,2\}$. $J_0(x_0) = \min_{u_0} \mathbb{E}\{2|x_0| + |u_0| + J_1(x_1)\}$

State x_0	Control u_0	$J_1(x_1, \omega_0 = 1)$	$J_1(x_1, \omega_0 = 0)$	$J_0(x_0)$	min Cost-to-go J_0
-2	-1	6.6	3.6	10.7	min J_0 is 10.6, Control is 0
	0	6.6	6.6	10.6	
	1	6.6	5.3	11.21	
-1	-1	3.6	1.82	6.066	min J_0 is 6.066, Control is -1
	0	6.6	3.6	7.7	
	1	6.6	6.6	9.6	
0	-1	1.82	4.82	3.72	min J_0 is 3.066, Control is 1
	0	3.6	1.82	3.066	
	1	6.6	3.6	6.7	
1	-1	4.82	9.3	9.164	min J_0 is 4.72, Control is 0
	0	1.82	4.82	4.72	
	1	3.6	1.82	6.066	
2	-1	9.3	9.3	14.3	min J_0 is 7.72, Control is 1
	0	4.82	9.3	10.164	
	1	1.82	4.82	7.72	

1.4 Part d

The optimal controls are different in the two models because their dynamics are different. Indeed, due to the uncertainty in the stochastic model, the optimal controller needs to mitigate for potential events and therefore decisions will be different. One can also see that all the costs are higher for the stochastic case, which makes sense because the outcomes are uncertain, so it average the cost will be higher.

2 Exercise 2

2.1 Part a

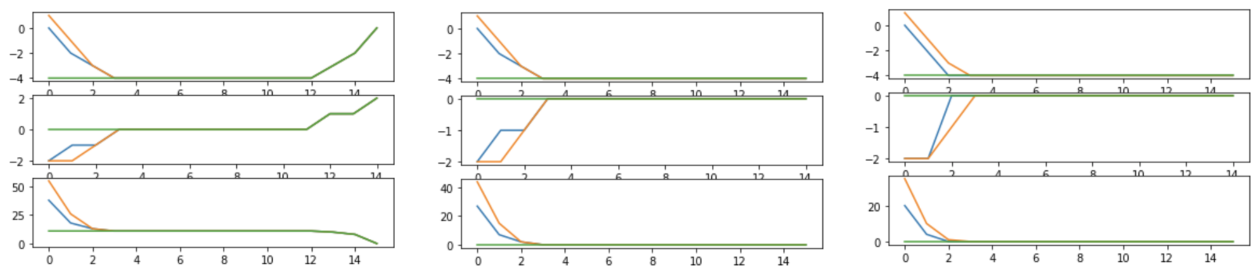
```
optimal cost to go
[[ 11.  11.  11.  11.  11.  11.  11.  11.  11.  11.  11.  11.  11.  12.
   inf  inf]
 [ 13.  13.  13.  13.  13.  13.  13.  13.  13.  13.  13.  12.  11.  10.
   inf  inf]
 [ 18.  18.  18.  18.  18.  18.  18.  18.  18.  18.  17.  16.  15.  12.
    8.  inf]
 [ 26.  26.  26.  26.  26.  26.  26.  26.  26.  26.  25.  24.  22.  18.
   10.  inf]
 [ 38.  38.  38.  38.  38.  38.  38.  38.  38.  37.  36.  35.  32.  27.
   16.   0.]
 [ 55.  55.  55.  55.  55.  55.  55.  55.  55.  54.  53.  51.  47.  39.
   26.  inf]
 [ 78.  78.  78.  78.  78.  78.  78.  78.  77.  76.  75.  72.  67.  56.
   40.  inf]
 [108. 108. 108. 108. 108. 108. 108. 108. 107. 106. 104. 100.  92.  79.
   inf  inf]
 [146. 146. 146. 146. 146. 146. 146. 145. 144. 143. 140. 135. 124. 108.
   inf  inf]]
optimal control
[[ 0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  1.  2. inf]
 [-1. -1. -1. -1. -1. -1. -1. -1. -1. -1. -1.  0.  0.  1. inf]
 [-1. -1. -1. -1. -1. -1. -1. -1. -1. -1. -1. -1. -1.  0.  2.]
 [-2. -2. -2. -2. -2. -2. -2. -2. -2. -2. -2. -2. -1. -1.  1.]
 [-2. -2. -2. -2. -2. -2. -2. -2. -2. -2. -2. -2. -2. -1.  0.]
 [-2. -2. -2. -2. -2. -2. -2. -2. -2. -2. -2. -2. -2. -2. -1.]
 [-2. -2. -2. -2. -2. -2. -2. -2. -2. -2. -2. -2. -2. -2. -2.]
 [-2. -2. -2. -2. -2. -2. -2. -2. -2. -2. -2. -2. -2. -2. inf]
 [-2. -2. -2. -2. -2. -2. -2. -2. -2. -2. -2. -2. -2. -2. inf]]
```

2.2 Part b

The cost to go for $x_4 = 4$ is 146.0.

2.3 Part c-e

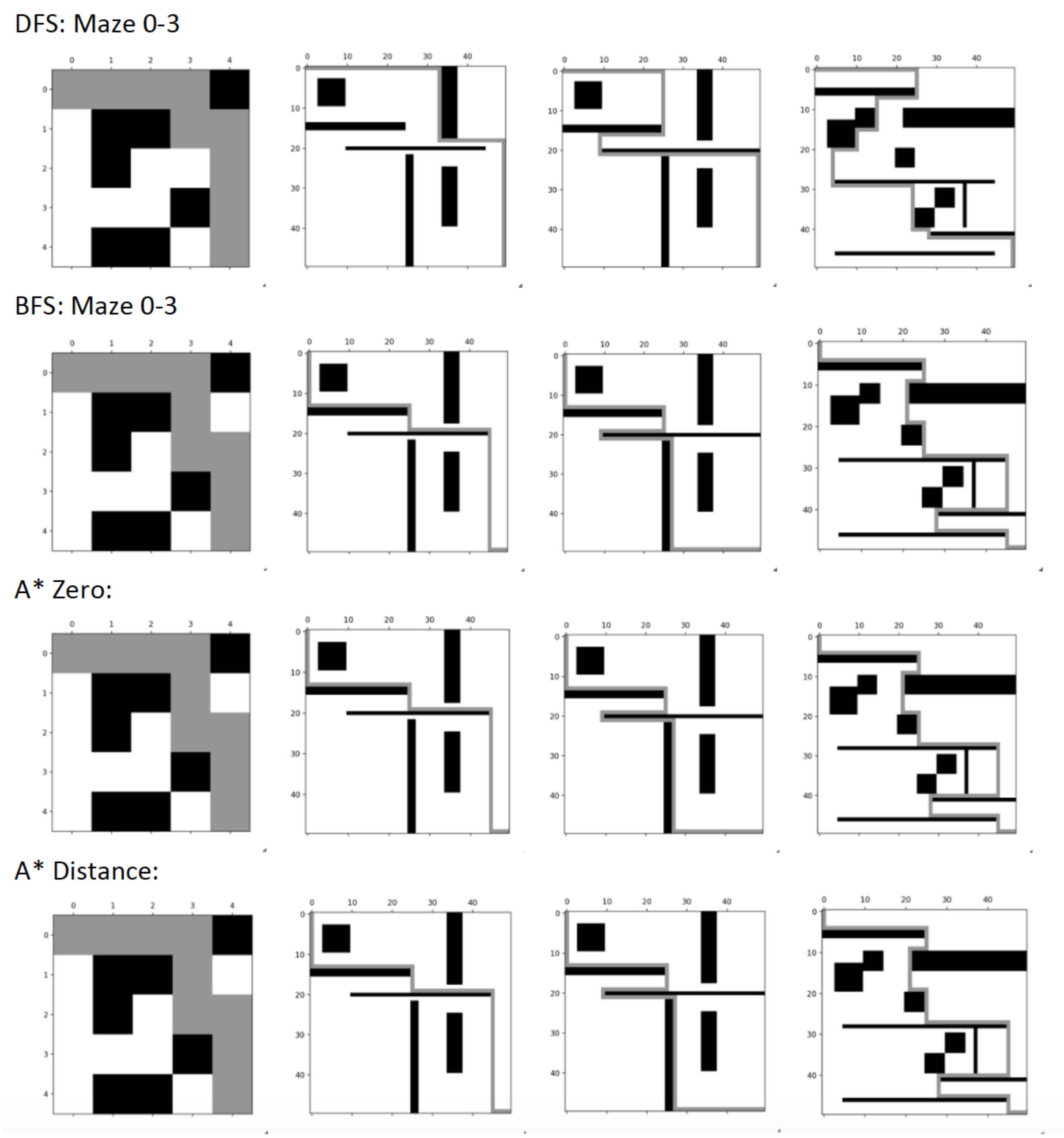
"State sequences", "optimal control" And "associated optimal cost" plots for Part c to e:



c) Associated optimal costs are: 38,55,11.
d) Associated optimal costs are: 27,44,0. Optimal Cost-to-go drops to 0 at early stages. Since state x = -4 give lowest stage cost and also 0 terminal cost, while for previous cost function optimal state has to transfer from state -4 to state 0 at last stages.
e)Associated optimal costs are: 20,35,0. Cost for control is removed, thus optimal cost-to-go can converge faster to 0 (at state -4) with higher control at each stage(if constraint allows).

3 Exercise 3

Path Count/ Node Count	DFS	BFS	A* Zero	A* Distance
Maze0	9/9	9/17	9/17	9/16
Maze1	99/223241	99/2025	99/2025	99/663
Maze2	131/176721	131/2173	131/2173	131/1843
Maze3	141/332734	141/2049	141/2049	141/1703



c) For the A* algorithm, both the A* distance and A* zero are under-estimators of the actual cost. Since A* zero estimates the cost to be 0 and A* distance estimates the distance without any obstacles.
d) **Pros and Cons:** While they all find the optimal path: DFS needs to go through a lot of nodes to find the optimal path; BFS finds the path much faster than DFS; A* Zero finds the path at a similar number of nodes comparing to BFS; A* Distance finds the path faster than all above methods.