Emeka Ezike, Dominic Garrity, and Victor Goncalves

Professor Stephanie Gil

COMPSCI 286

13 April 2022

Programming Assignment 4

Group Information

Group number: 6

Group members: Dominic Garrity (HUID: 61367713), Victor Goncalves (HUID: 11482961), and Emeka Ezike (HUID: 11464432)

Problem 1

We have included a visualization of the pickup probability distribution in Figure 1 below.

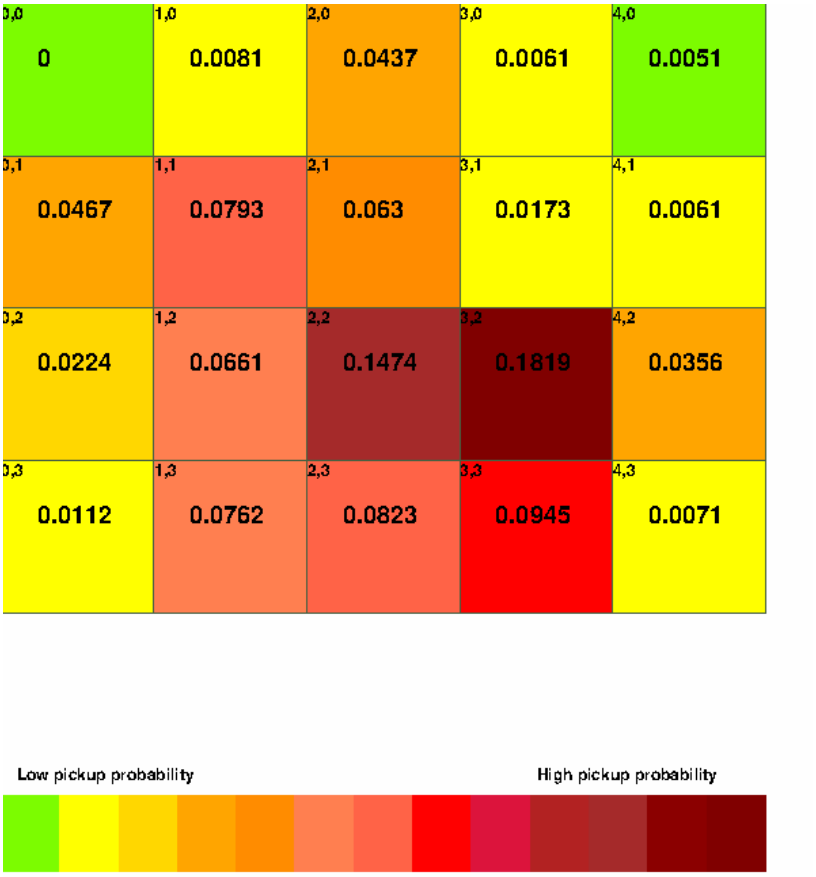


Figure 1

Problem 2

We have included a visualization of the dropoff probability distribution in Figure 2 below.

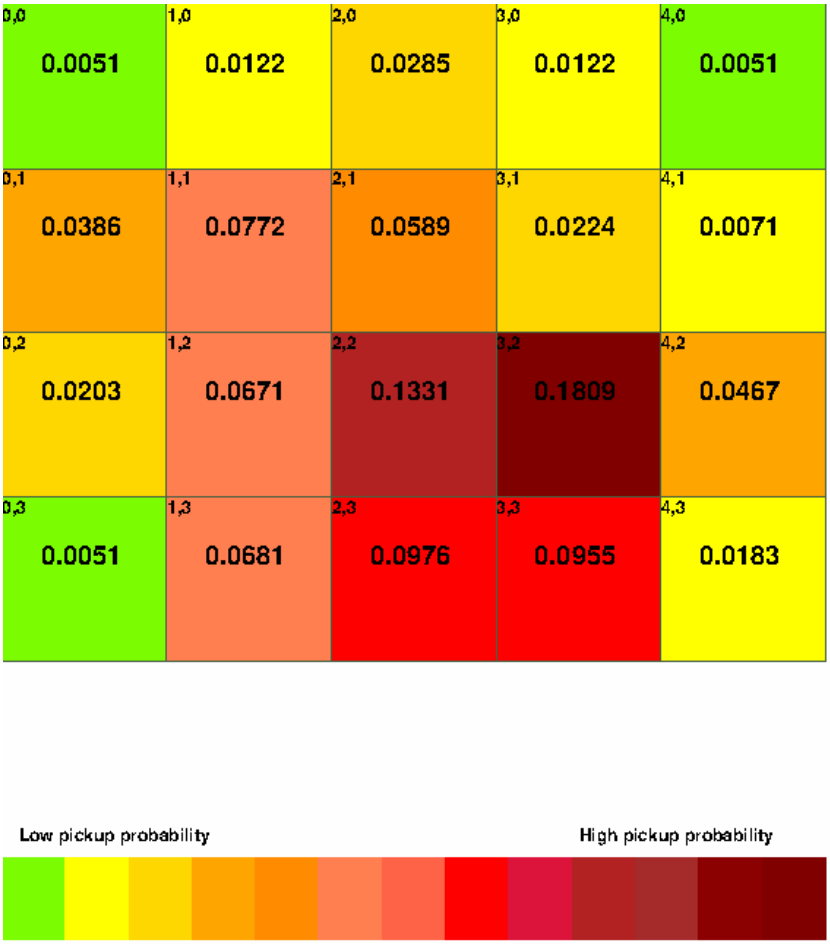
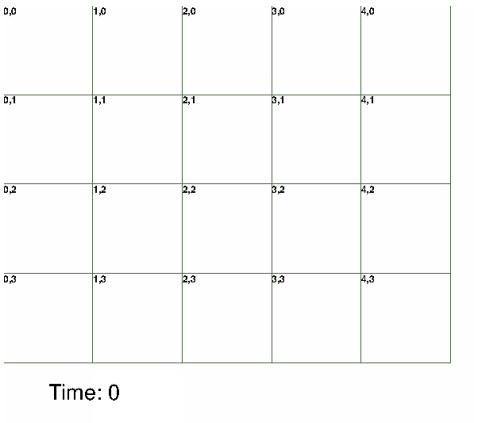


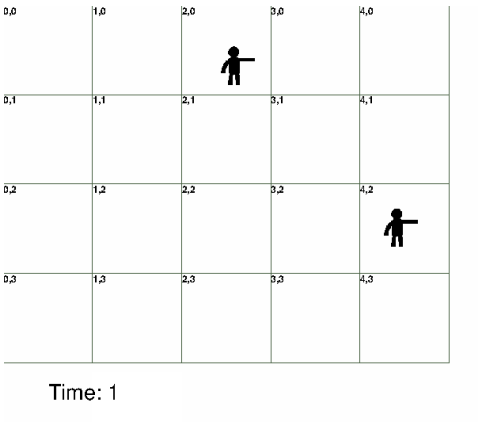
Figure 2

Problem 3

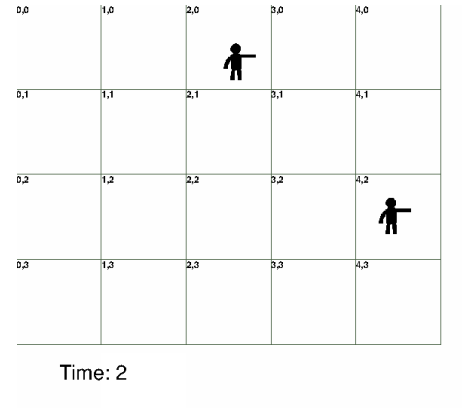
We have included the frames of request.mov in Figure 3 below. These frames show requests that are generated over a horizon of 10. Although it is difficult to determine from just one trial, the locations of the requests seem to align with the pickup probability distribution shown in Figure 1.


0,0
1,0
2,0
3,0
4,0
0,1
1,1
2,1
3,1
4,1
0,2
1,2
2,2
3,2
4,2
0,3
1,3
2,3
3,3
4,3
Time: 0

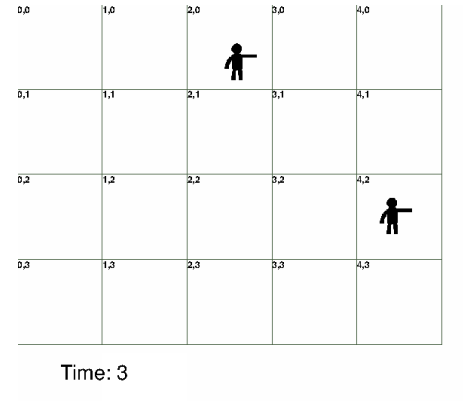
(a)


0,0
1,0
2,0
3,0
4,0
0,1
1,1
2,1
3,1
4,1
0,2
1,2
2,2
3,2
4,2
0,3
1,3
2,3
3,3
4,3
Time: 1

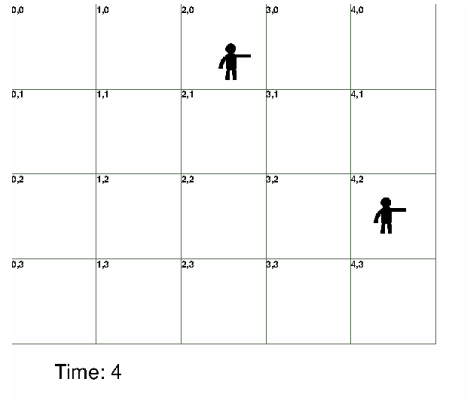
(b)


0,0
1,0
2,0
3,0
4,0
0,1
1,1
2,1
3,1
4,1
0,2
1,2
2,2
3,2
4,2
0,3
1,3
2,3
3,3
4,3
Time: 2

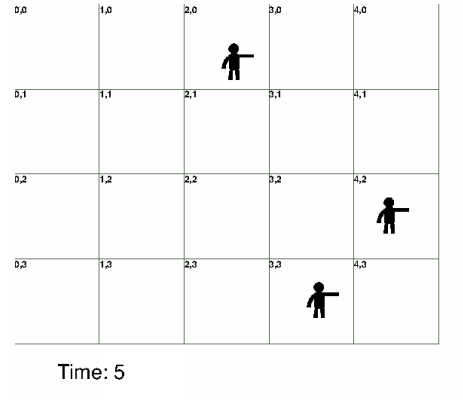
(c)


0,0
1,0
2,0
3,0
4,0
0,1
1,1
2,1
3,1
4,1
0,2
1,2
2,2
3,2
4,2
0,3
1,3
2,3
3,3
4,3
Time: 3

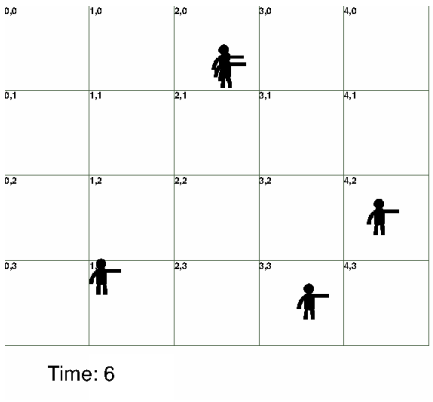
(d)


0,0
1,0
2,0
3,0
4,0
0,1
1,1
2,1
3,1
4,1
0,2
1,2
2,2
3,2
4,2
0,3
1,3
2,3
3,3
4,3
Time: 4

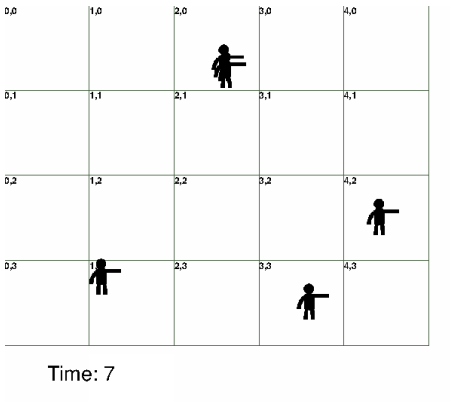
(e)


0,0
1,0
2,0
3,0
4,0
0,1
1,1
2,1
3,1
4,1
0,2
1,2
2,2
3,2
4,2
0,3
1,3
2,3
3,3
4,3
Time: 5

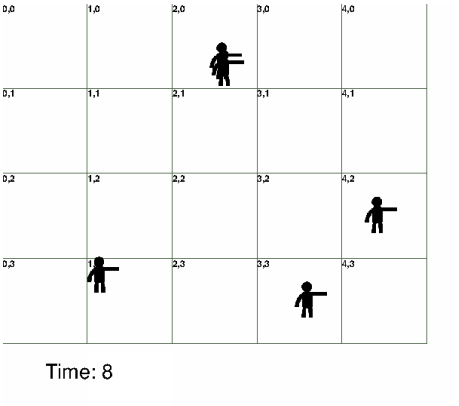
(f)



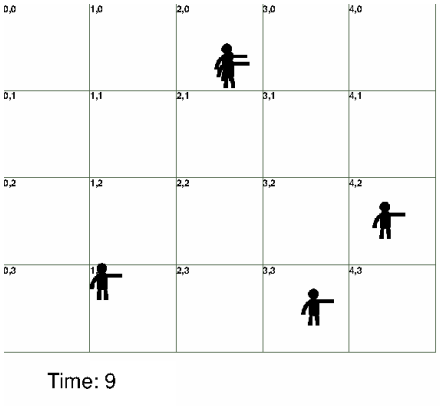
(g)



(h)



(i)



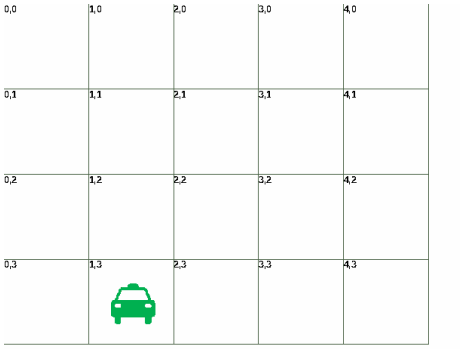
(j)

Figure 3

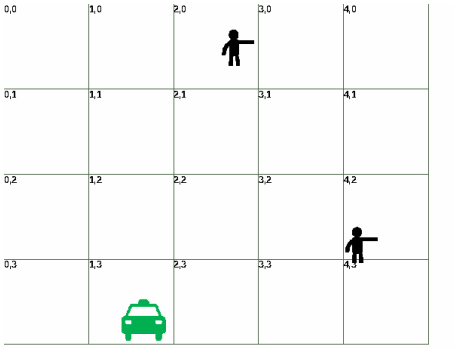
Problem 4

Below, in Figure 5, we have included the frames of base_policy_1.mov in Figure 4 and the frames of base_policy_2.mov. Evidently, the base policy does a fair, but suboptimal, job; the final cost was 31 in the case with one taxi and 27 in the case with two. The policy's relatively mediocre performance makes sense, because it is greedy and, as a result, taxis do not explore all possible outcomes. Similarly, the reduction of cost that came with the introduction of a second taxi makes sense, because additional taxis allow the entire system to more quickly/effectively service requests.

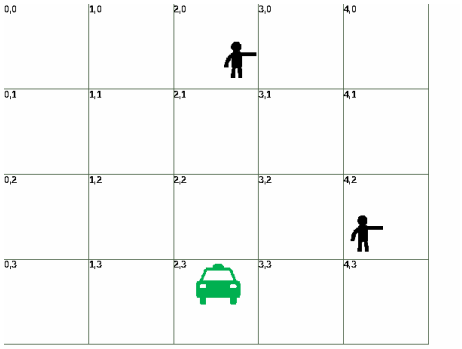
Base Policy (One Taxi)



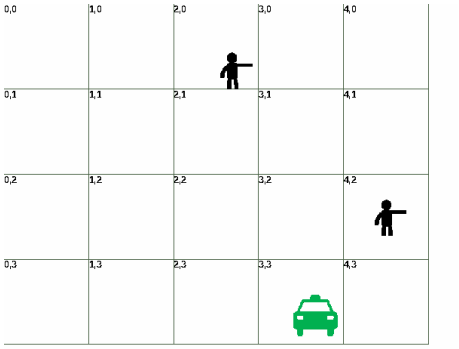
(a)



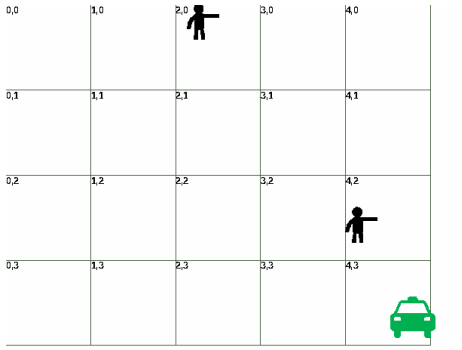
(b)



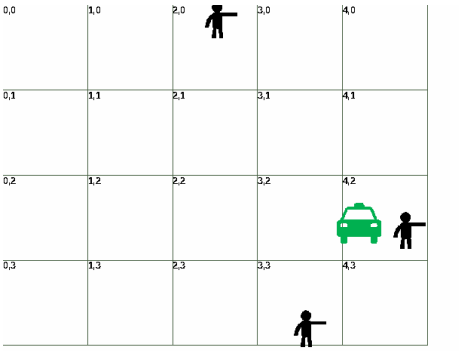
(c)



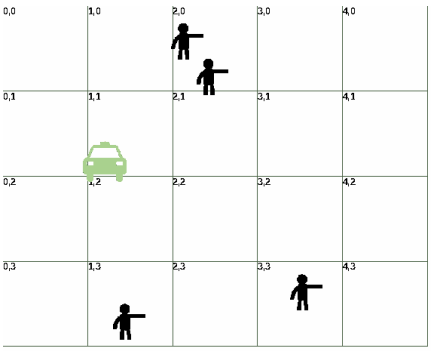
(d)



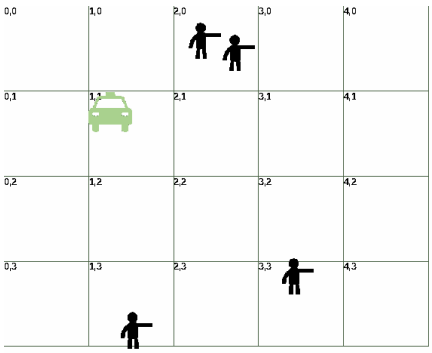
(e)



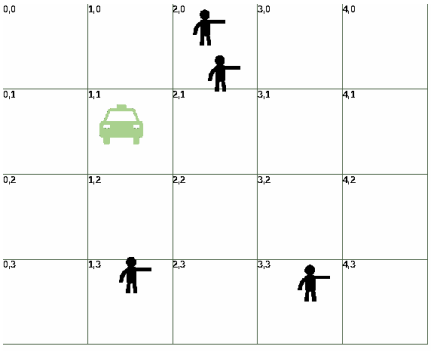
(f)



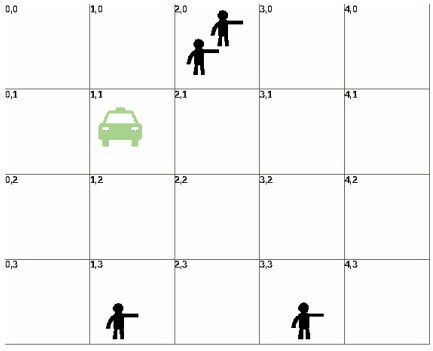
(g)



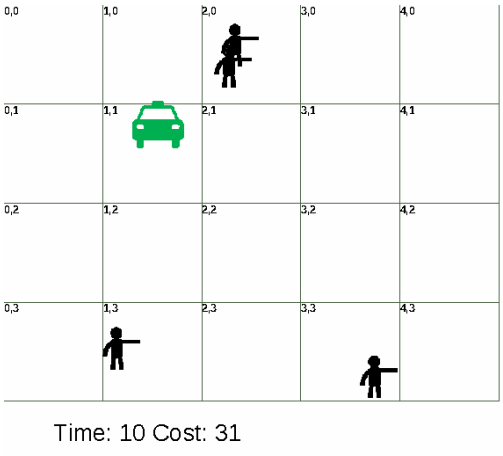
(h)



(i)



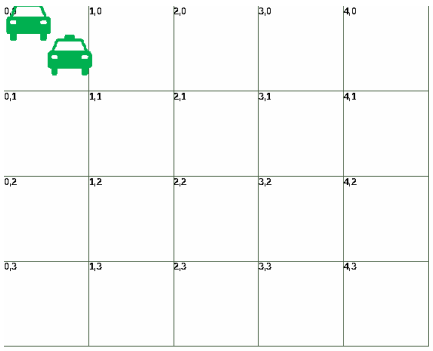
(j)



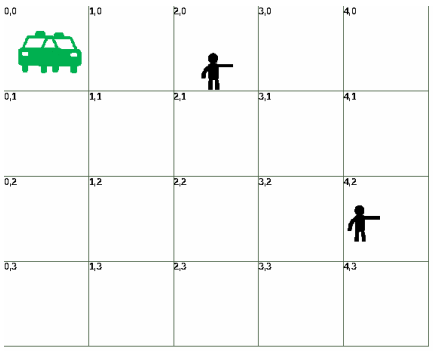
(k)

Figure 4

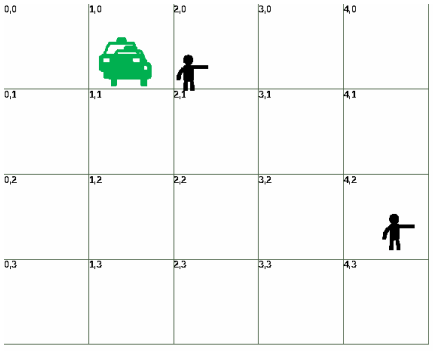
Base Policy (Two Taxis)


Time: 0 Cost: 0

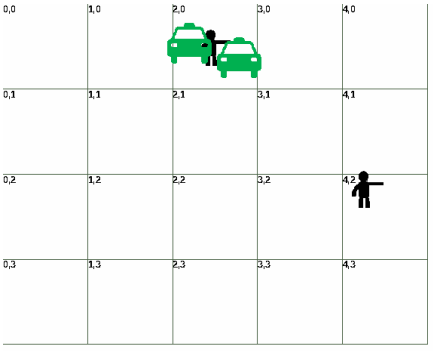
(a)


Time: 1 Cost: 2

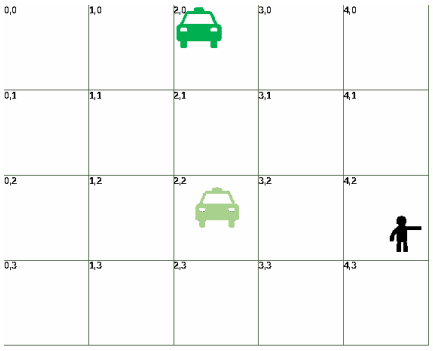
(b)


Time: 2 Cost: 4

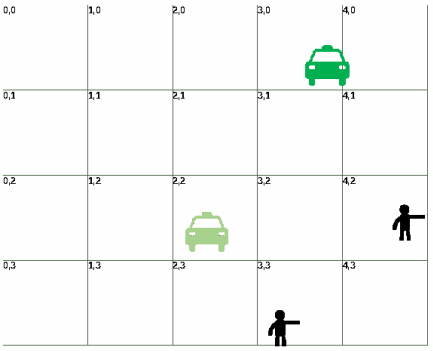
(c)


Time: 3 Cost: 6

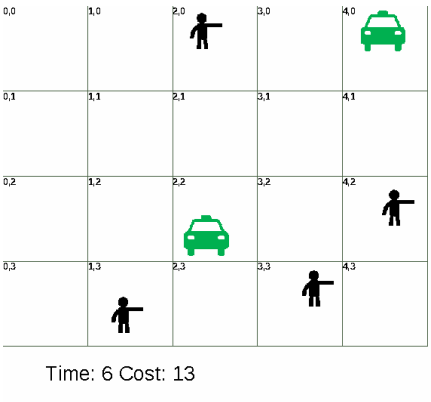
(d)


Time: 4 Cost: 7

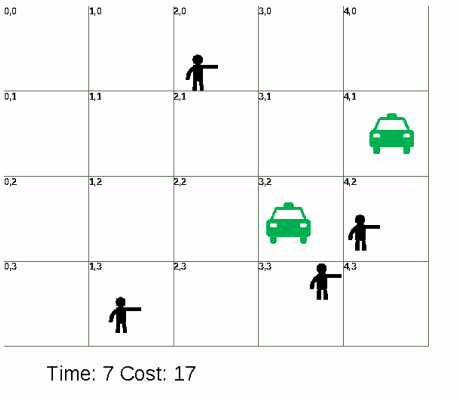
(e)


Time: 5 Cost: 9

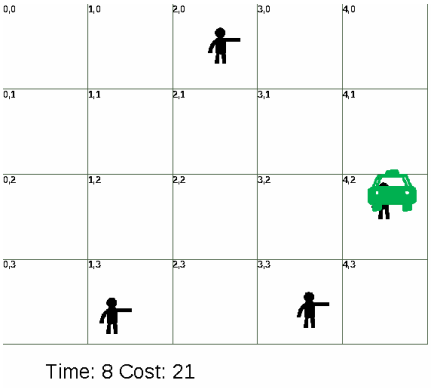
(f)



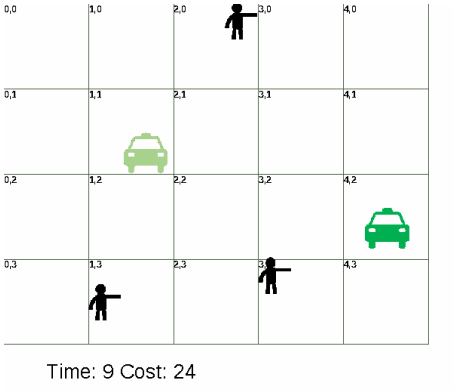
(g)



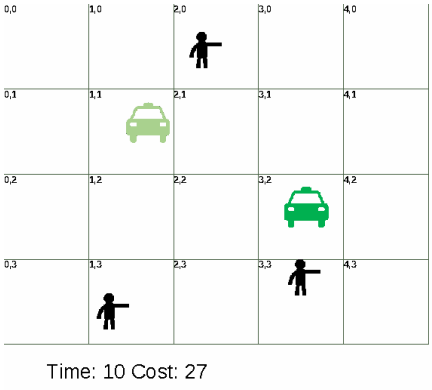
(h)



(i)



(j)



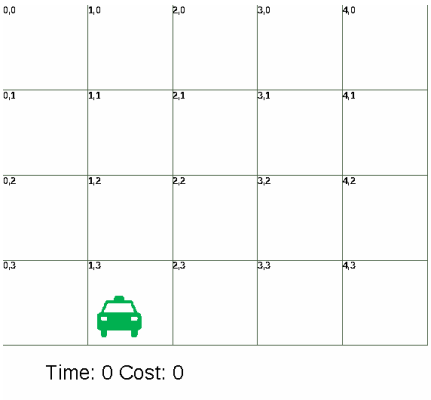
(k)

Figure 5

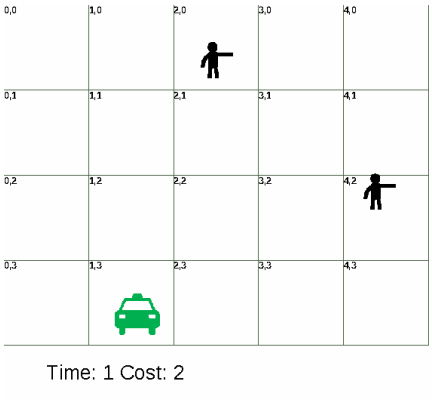
Problem 5

Below, in Figure 7, we have included the frames of standard_rollout_1.mov in Figure 6 and the frames of standard_rollout_2.mov. In both the case with one taxi and the case with two, the standard rollout policy initially relied on 100 Monte Carlo simulations. It resulted in a final cost of 31 with one taxi and 22 with two taxis. Importantly, this means that the standard rollout policy was at least as good as, if not better than, both the base policy and the one-at-a-time rollout policy (as described in Problem 6). This also aligns with Bertsekas' description; by the standard rollout policy, each taxi can fully observe other taxis (full observability) whereas, by the one-at-a-time rollout policy, each taxi can only partially observe other taxis (partial observability). It is worth noting, however, that the standard rollout policy may not be wholly optimal with only 100 Monte Carlo simulations, because it estimates future costs through Monte Carlo simulations, which are probabilistic and therefore susceptible to error/noise. In Figure 6, this is highlighted by the lack of change in the taxi's position between time steps 9 and 10. This sort of suboptimality could be improved through more Monte Carlo simulations. (Note: To evaluate our implementation with more simulations, simply change the three instances of the "simulation" variable in algorithms.py!)

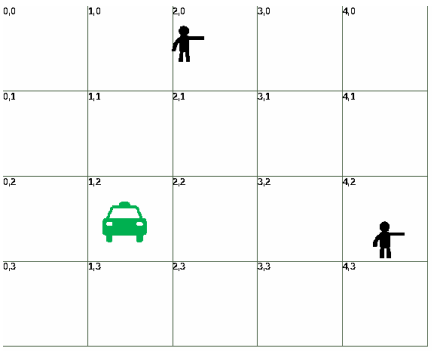
Standard Rollout (One Taxi)



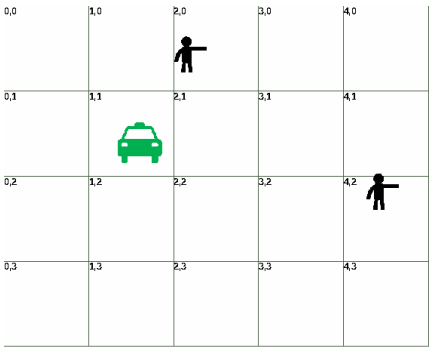
(a)



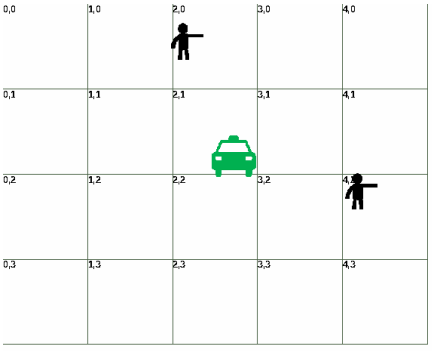
(b)


Time: 2 Cost: 4

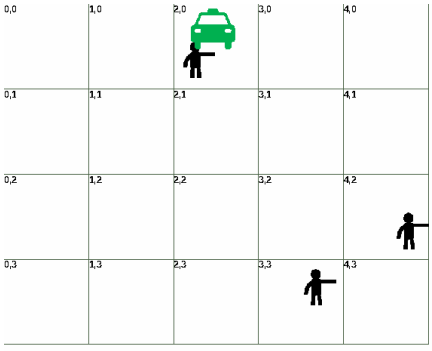
(c)


Time: 3 Cost: 6

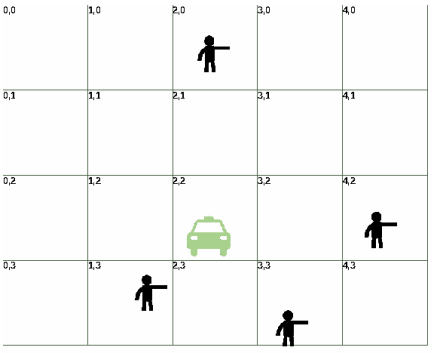
(d)


Time: 4 Cost: 8

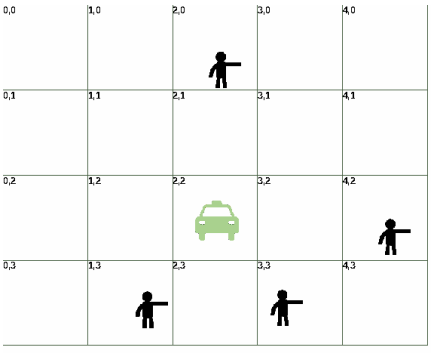
(e)


Time: 5 Cost: 11

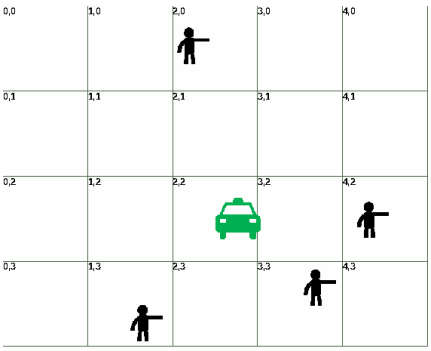
(f)


Time: 6 Cost: 15

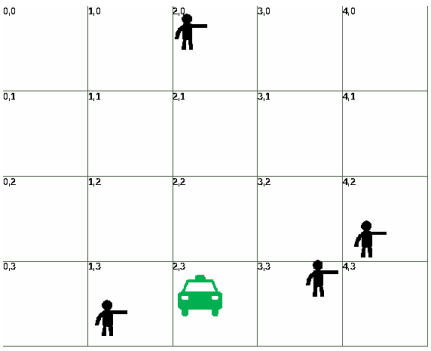
(g)


Time: 7 Cost: 19

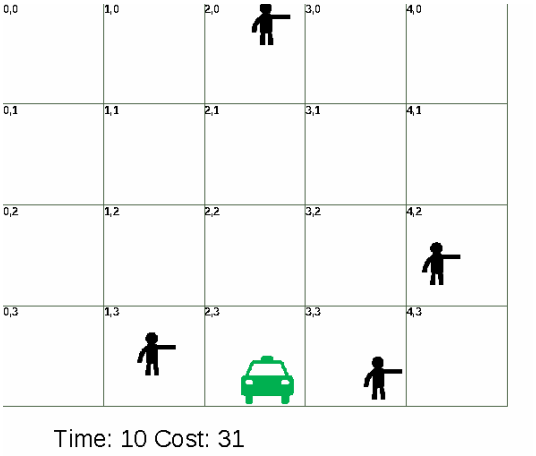
(h)



(i)



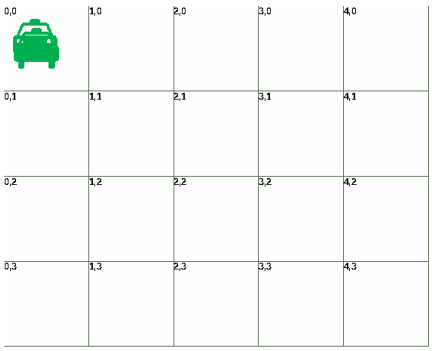
(j)



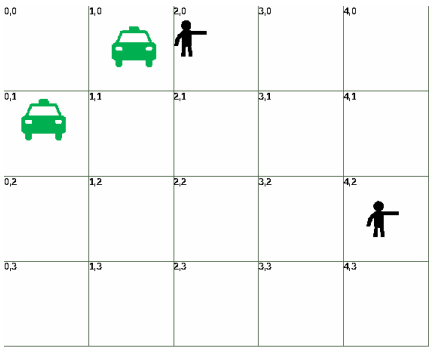
(k)

Figure 6

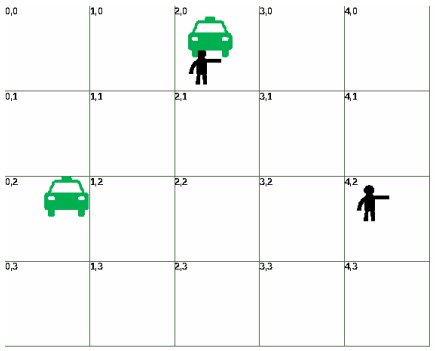
Standard Rollout (Two Taxis)



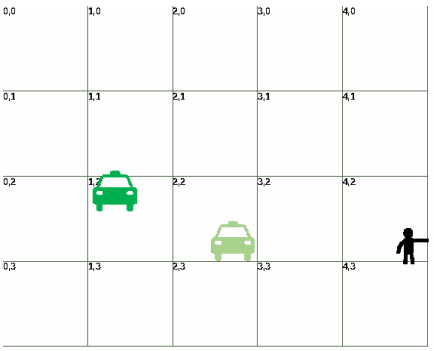
(a)



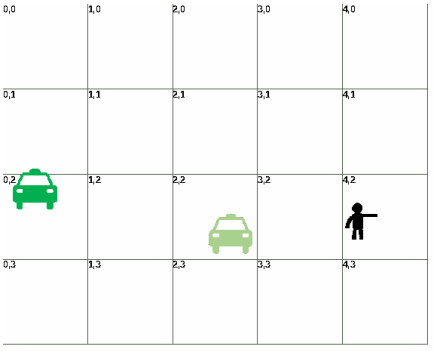
(b)


Time: 2 Cost: 4

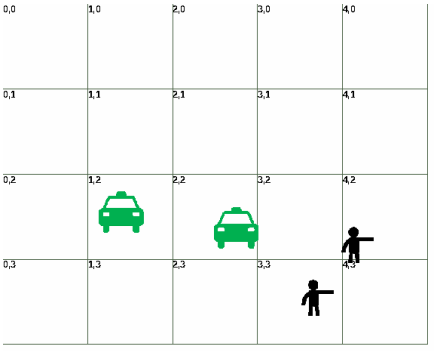
(c)


Time: 3 Cost: 5

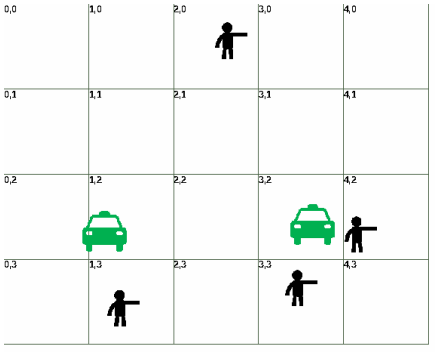
(d)


Time: 4 Cost: 6

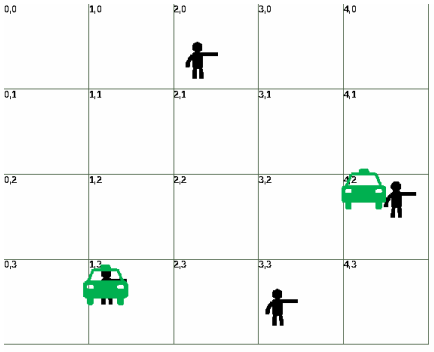
(e)


Time: 5 Cost: 8

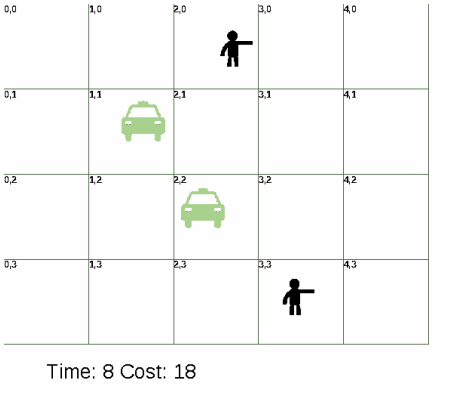
(f)


Time: 6 Cost: 12

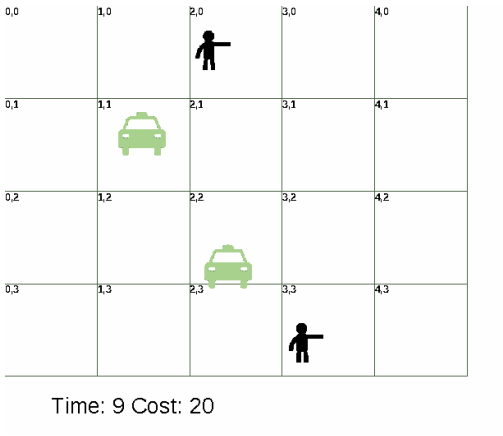
(g)


Time: 7 Cost: 16

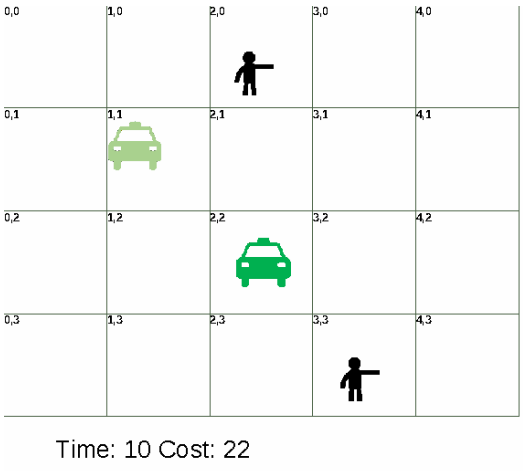
(h)



(i)



(j)



(k)

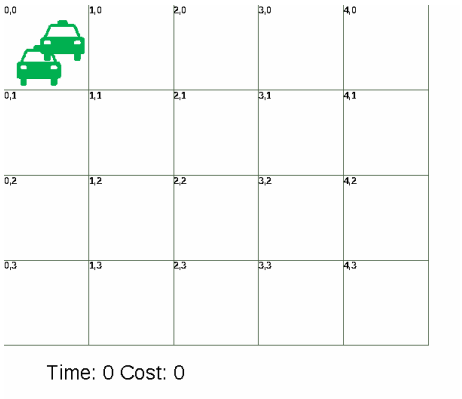
Figure 7

Problem 6

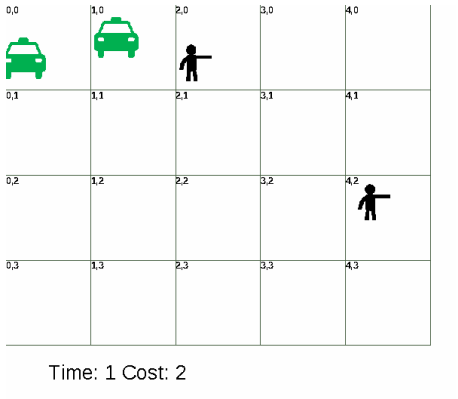
We have included the frames of one_agent_at_a_time_rollout_2.mov in Figure 8 below. As expected, when relying on 100 Monte Carlo simulations, the one-at-a-time rollout policy performed better than the base policy, but worse than the standard rollout policy; the final cost was 23 in the case with two taxis. This means that–as they should–both rollout policies performed better than the base policy, exhibiting the "fundamental cost improvement property of rollout" to which Bertsekas referred in "Multiagent Rollout Algorithms and Reinforcement

Learning." Like the standard rollout policy, though, the one-at-a-time rollout policy may be sub-optimal because of its Monte Carlo simulations, which are probabilistic and therefore susceptible to error/noise. In addition to partial observability, this explains suboptimality like that exhibited between time steps 2 and 3 below. Such optimality could be improved through more Monte Carlo simulations.

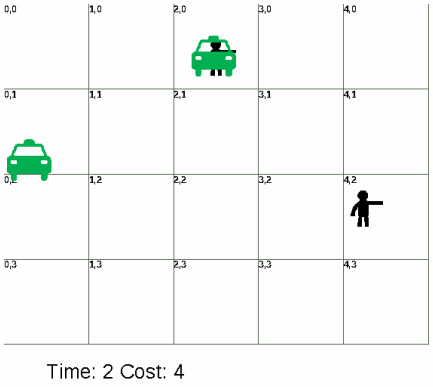
One-at-a-Time Rollout Policy



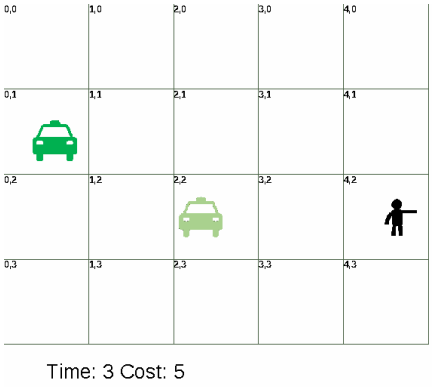
(a)



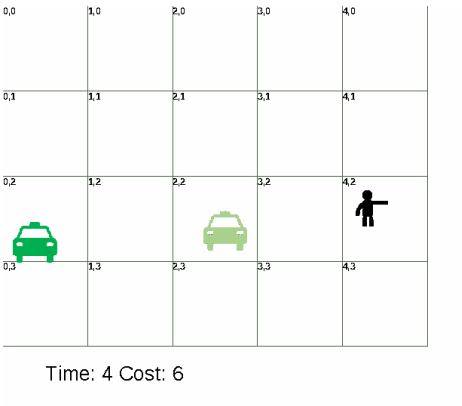
(b)



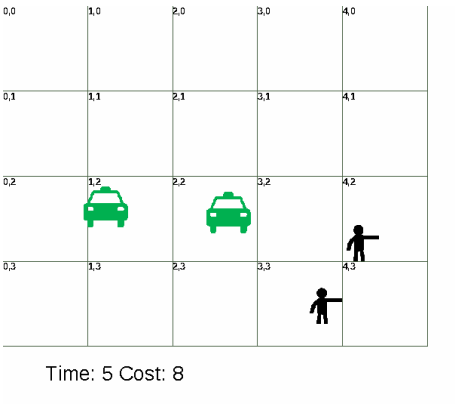
(c)



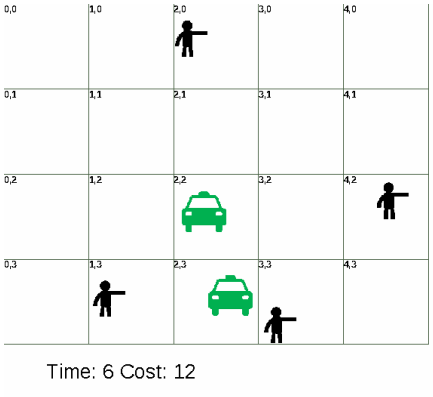
(d)


Time: 4 Cost: 6

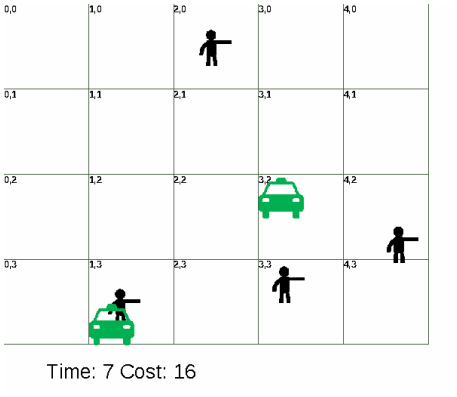
(e)


Time: 5 Cost: 8

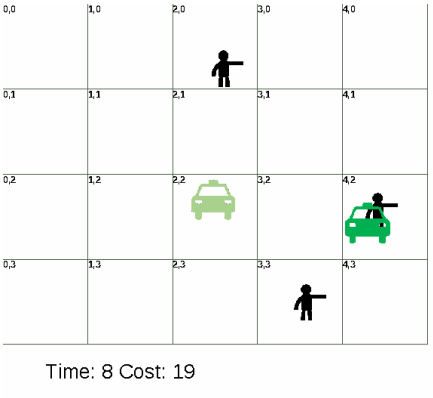
(f)


Time: 6 Cost: 12

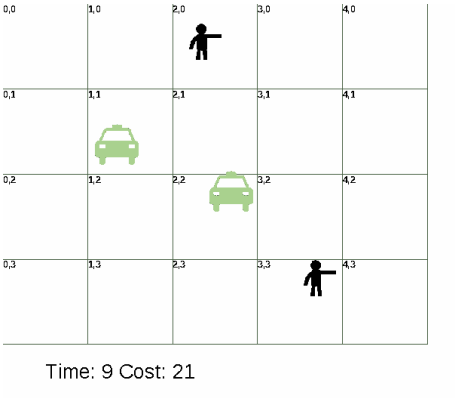
(g)


Time: 7 Cost: 16

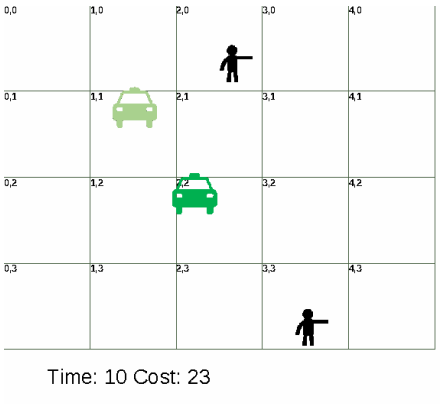
(h)


Time: 8 Cost: 19

(i)


Time: 9 Cost: 21

(j)


0,0
1,0
2,0
3,0
4,0
0,1
1,1
2,1
3,1
4,1
0,2
1,2
2,2
3,2
4,2
0,3
1,3
2,3
3,3
4,3
Time: 10 Cost: 23

(k)

Figure 8