

General comments- important!

1. In theoretical question 3, many people did or overestimated the update for the reward function. Note that since we are only changing a single state-action reward and we can keep that information inside a matrix, this is a $O(1)$ operation.

2. In the same question, Many did not consider the cost of the optimal policy, for calculating the cost of an iteration, the correct cost is $|S|+|A|$ for calculating a BFS where we change the direction of the edges, and connect all the states with unknown state-action to a single new node, and run a BFS over that graph, where the resulting directed tree will be the optimal policy. So the final cost per iteration is $(|S|+|A|)$. This implies that the final cost is $|A|*|S|*(|S|+|A|)$.

3. In question 4: "What is the mean recurrence time to a corner?"

some of you have interpreted it as mean time between corners.

In this question, every square is a (different) state, and therefore the answer should be 168 and not 168/4.

We have decided not to reduce points in this case.

36626208	T.Q3.1 - Did not consider the cost for calculating the optimal policy -5p,	95
300472826	T.Q3.1 - Did not consider the cost for calculating the optimal policy -5p,	95
38113585	T.Q3.1 - Did not consider the cost for calculating the optimal policy -5p,	95
200397354	V	100
308402163	V	100
203486550	V	100
201375656	T.Q3.1 - Did not consider the cost for calculating the optimal policy -5p	95
203569264	T.Q3.1 - Did not consider the cost for calculating the optimal policy -5p	95
301538674	T.Q3.1 - Did not consider the cost for calculating the optimal policy -5p	95
203308192	V	100
312527302	V	100
203269931	V	100
203455126	T.Q3.1 - Good work, note that a proof by contradiction is a bit shorter,	100
311460125	T.Q3.1 - Good work, note that a proof by contradiction is a bit shorter,	100
203521984	T.Q3.1 - Insufficient explanation -3p T.Q3.4 - Insufficient explanation for why the All-Pairs solves this problem -4p, P.Q2 - Where is the plot? -5p	88
203774849	T.Q3.1 - Insufficient explanation -3p T.Q3.4 - Insufficient explanation for why the All-Pairs solves this problem -4p, P.Q2 - Where is the plot? -5p	88
312433279	T.Q3.1 - Insufficient explanation -3p T.Q3.4 - Insufficient explanation for why the All-Pairs solves this problem -4p, P.Q2 - Where is the plot? -5p	88
203571278	P.Q2 - Where is the explanation of the result? -2p	98
203628482	P.Q2 - Where is the explanation of the result? -2p	98
204037840	T.Q3.2 - This is not an optimal policy -5p,	95
308352525	T.Q3.2 - This is not an optimal policy -5p,	95
204257299	T.Q2 - No need to run Dijkstra, you can run BFS -1p,	99
203119433	T.Q2 - No need to run Dijkstra, you can run BFS -1p,	99
203285366	T.Q2 - No need to run Dijkstra, you can run BFS -1p,	99
204389043	T.Q2 - No need to run full policy iteration, you can run BFS -1p,	99
300602448	T.Q2 - No need to run full policy iteration, you can run BFS -1p,	99
204738736	T.Q3.2 - Did not consider the cost for calculating the optimal policy -5p,	95
205781693	T.Q3.2 - Did not consider the cost for calculating the optimal policy -5p,	95
205925894	T.Q3.2 - Did not consider the cost for calculating the optimal policy -5p,	95
313288615	T.Q3.2 - Did not consider the cost for calculating the optimal policy -5p,	95
208143552	T.Q3.2 - Did not consider the cost for calculating the optimal policy -5p,	95
209044791	T.Q3.2 - Did not consider the cost for calculating the optimal policy -5p,	95
208414466	T.Q3.2 - Did not consider the cost for calculating the optimal policy -5p,	95
305751000	T.Q3.2 - Did not consider the cost for calculating the optimal policy -5p,	95
212648406	T.Q3.2 - Did not consider the cost for calculating the optimal policy -5p,	95
201271251	T.Q3.2 - Did not consider the cost for calculating the optimal policy -5p,	95
301250650	T.Q3.2 - Did not consider the cost for calculating the optimal policy -5p,	95
301558086	T.Q3.2 - Did not consider the cost for calculating the optimal policy -5p,	95
300746930	T.Q3.2 - Did not consider the cost for calculating the optimal policy -5p,	95
301794376	T.Q3.2 - Did not consider the cost for calculating the optimal policy -5p,	95
203626775	T.Q3.2 - Did not consider the cost for calculating the optimal policy -5p,	95
305584690	T.Q3.2 - Did not consider the cost for calculating the optimal policy -5p,	95
301803862	V	100
201243573	V	100

305546590	V	100
308223726	T.Q3.2 - Did not consider the cost for calculating the optimal policy -5p,	95
311233761	T.Q3.2 - Did not consider the cost for calculating the optimal policy -5p,	95
308315431	V	100
312497977	V	100
308428572	V	100
204219646	V	100
316879899	V	100
302461215	V	100
203703327	V	100
319124244	T.Q3 - No need to run full policy iteration, you can run BFS -1p, T.Q4.3 - final answer should be 166 moves -5p	94
308419290	T.Q3 - No need to run full policy iteration, you can run BFS -1p, T.Q4.3 - final answer should be 166 moves -5p	94
		96.66071429