

Qn No.2

1. Based on the question provided in learning the optimal actions to be taken in a specific state, the following observation has been made:

```
"D:\Abeesh\Assignment\machine learning\venv\Scripts\python.exe" "D:/Abeesh/Assignment/machine learning/Assignemt 2/Q2.py"
Q Table
[[0.00000000e+000 0.00000000e+000 0.00000000e+000 0.00000000e+000
 1.00000000e+000 0.00000000e+000]
 [0.00000000e+000 0.00000000e+000 0.00000000e+000 7.41976889e-297
 0.00000000e+000 1.00000000e+000]
 [0.00000000e+000 0.00000000e+000 0.00000000e+000 1.00000000e+000
 0.00000000e+000 0.00000000e+000]
 [0.00000000e+000 7.56550285e-001 3.65809180e-003 0.00000000e+000
 2.39791623e-001 0.00000000e+000]
 [0.00000000e+000 0.00000000e+000 0.00000000e+000 2.45056560e-320
 0.00000000e+000 1.00000000e+000]
 [0.00000000e+000 0.00000000e+000 0.00000000e+000 0.00000000e+000
 4.21778013e-048 1.00000000e+000]]
Optimal Path for starting state 0
0 -> 4 -> 5

Optimal Path for starting state 1
1 -> 5

Optimal Path for starting state 2
2 -> 3 -> 1 -> 5

Optimal Path for starting state 3
3 -> 1 -> 5

Optimal Path for starting state 4
4 -> 5

Optimal Path for starting state 5
5
```

Start Room	Destination Room	Path Covered
0	5	0 -> 4 ->5
1	5	1 -> 5
2	5	2 -> 3 -> 1 -> 5
3	5	3 -> 1 -> 5
4	5	4 -> 5
5	5	5

Hyperparameters used

```
# hyperparameters
gamma = 0.8 #vary this between 0 and 1 i.e., (0,1)
alpha = 0.9 #vary this between 0 and 1 inclusive, i.e., [0,1]
n_episodes = 1E3 #try different values, e.g., 10, 500, 10000 and so forth
epsilon = 0.05 #you can experiment with this as well
```

Now based on the question, by changing the hyperparameter, a new optimal path has been find out, which is as shown below.

```
# hyperparameters
gamma = 0.3 #vary this between 0 and 1 i.e., (0,1)
alpha = 0.9 #vary this between 0 and 1 inclusive, i.e., [0,1]
n_episodes = 400 #try different values, e.g., 10, 500, 10000 and so forth
epsilon = 0.05 #you can experiment with this as well
```

The hyperparameter value of **gamma** is changed to **0.3** and **n_episodes** to **400**

Output

```
Q Table
[[0.00000000e+000 0.00000000e+000 0.00000000e+000 0.00000000e+000
 1.00000000e+000 0.00000000e+000]
 [0.00000000e+000 0.00000000e+000 0.00000000e+000 1.05142911e-005
 0.00000000e+000 9.99989486e-001]
 [0.00000000e+000 0.00000000e+000 0.00000000e+000 1.00000000e+000
 0.00000000e+000 0.00000000e+000]
 [0.00000000e+000 3.18707031e-001 3.30146330e-001 0.00000000e+000
 3.51146639e-001 0.00000000e+000]
 [2.56426235e-140 0.00000000e+000 0.00000000e+000 1.81457827e-060
 0.00000000e+000 1.00000000e+000]
 [0.00000000e+000 0.00000000e+000 0.00000000e+000 0.00000000e+000
 0.00000000e+000 1.00000000e+000]]
Optimal Path for starting state 0
0 -> 4 -> 5

Optimal Path for starting state 1
1 -> 5

Optimal Path for starting state 2
2 -> 3 -> 4 -> 5

Optimal Path for starting state 3
3 -> 4 -> 5

Optimal Path for starting state 4
4 -> 5

Optimal Path for starting state 5
5
```

Start Room	Destination Room	Path Covered
0	5	0 -> 4 -> 5
1	5	1 -> 5
2	5	2 -> 3 -> 4 -> 5
3	5	3 -> 4 -> 5
4	5	4 -> 5
5	5	5

Based on the table, we can do the following conclusion:

- a. Initially to reach the destination from room 2 and 3, it use the room no.1.
- b. Now after changing the hyperparameter value, another optimal route has been find out to reach the destination from room 2 and 3 is via room 4.