1. The number of states depends on the size and configuration of the board, the position of the agent, the position of the dots, and the position of the ghosts (if any). The number of states can be reduced by using some techniques such as state aggregation, where similar states are grouped together, or feature extraction, where only relevant features of the state are considered.

2. Action: The action is the direction that the agent can move in each state, such as up, down, left, or right. State: The state is the representation of the board at a given time, including the position of the agent, the dots, and the ghosts (if any). Rewards: The rewards are the numerical values that the agent receives after taking an action in a state. For example, the agent may receive a positive reward for collecting a dot, a negative reward for hitting a wall or a ghost, and a zero reward for moving to an empty cell. Goal State: The goal state is the state where the agent has collected all the dots and reached the end point without touching any ghosts.

3. **γ (Gamma) Analysis:**

   Gamma (γ) Values: 0.25, 0.5, 1

   The gamma parameter in Q-learning determines the discount factor for future rewards. Let's analyze the impact of different gamma values on the Pacman agent's behavior:

   γ = 0.25: A low gamma value makes the agent focus more on immediate rewards. The agent may prioritize short-term gains and might not consider the long-term consequences of its actions. This could result in suboptimal decision-making.

   γ = 0.5: A moderate gamma value balances the consideration of immediate and future rewards. The agent considers both short-term and long-term consequences, leading to more strategic decision-making.

   γ = 1: A gamma value of 1 implies that the agent values all future rewards equally. The agent takes into account the entire future trajectory, aiming for the maximum cumulative reward. This can lead to more forward-looking and potentially conservative decision-making.

   α (Alpha) Analysis:

   Alpha (α) Values: Varying Values

The alpha parameter in Q-learning represents the learning rate, determining how much the agent updates its Q-values based on new information. Let's analyze the impact of different alpha values:

Low α: For a low alpha, the agent gives less weight to the most recent information, relying more on past experiences. This might result in slow adaptation to changes in the environment.

Moderate α: A moderate alpha balances the importance of recent experiences and past knowledge. It allows the agent to adapt reasonably quickly to changes without completely discarding its previous learnings.

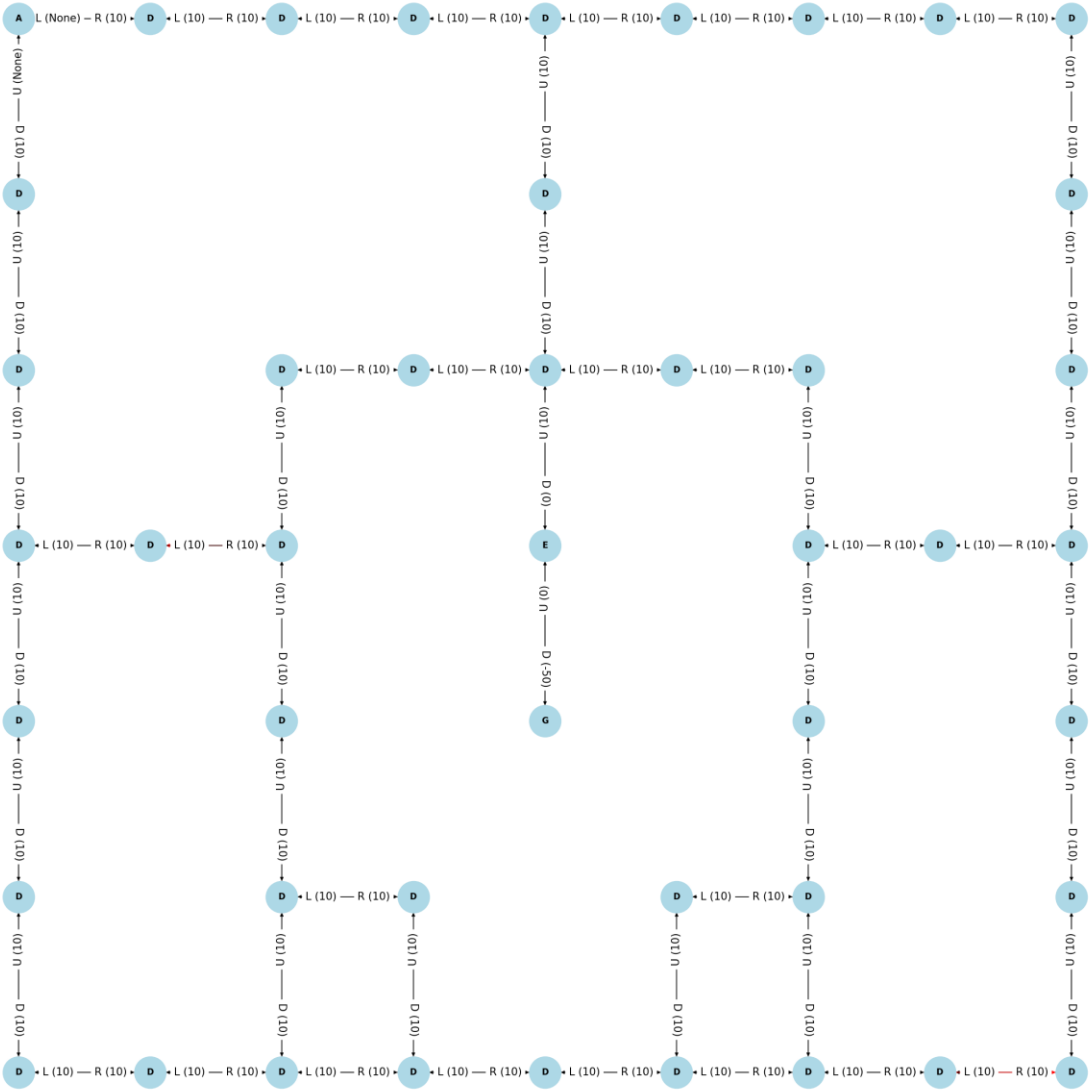
High α: A high alpha makes the agent give more weight to the most recent experiences. The agent adapts quickly to changes in the environment but might be more prone to overfitting and reacting impulsively.

Impact Analysis:

The choice of gamma and alpha values is crucial for the performance of the Q-learning algorithm.

Finding the right balance is essential, and it often involves experimenting with different values.

4.



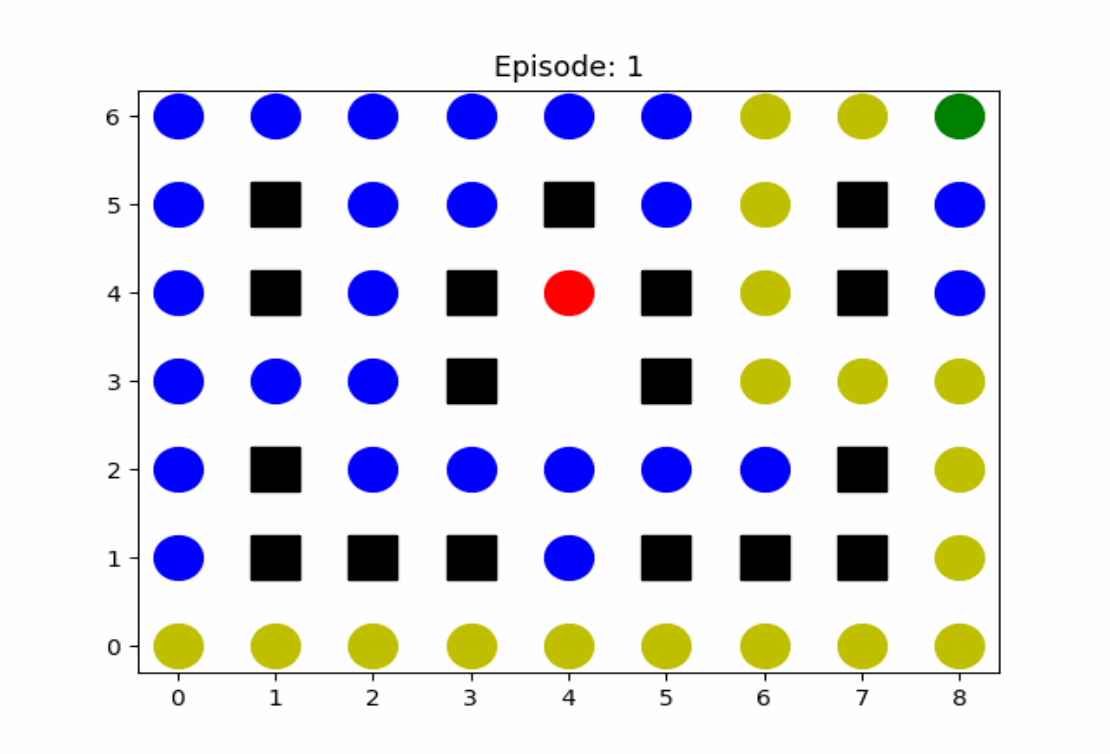
5.

```
Episode 1, Total Reward: 18
[['E' 'U' 'U' 'U' 'U' 'U' 'U' 'U' 'U']
 ['D' 'W' 'W' 'W' 'D' 'W' 'W' 'W' 'U']
 ['D' 'W' 'D' 'D' 'D' 'D' 'D' 'W' 'U']
 ['D' 'D' 'D' 'W' 'E' 'W' 'U' 'U' 'U']
 ['D' 'W' 'D' 'W' 'G' 'W' 'U' 'W' 'D']
 ['D' 'W' 'D' 'D' 'W' 'D' 'U' 'W' 'D']
 ['D' 'D' 'D' 'D' 'D' 'D' 'U' 'U' 'A']]

+-------+------+------+------+-------+
| state |  up  | down | left | right |
+-------+------+------+------+-------+
| (0,0) | 0.01 | 0.07 | 0.00 | 0.19  |
| (0,1) | 0.00 | 0.00 | 0.02 | 0.16  |
| (0,2) | 0.02 | 0.02 | 0.04 | 0.16  |
| (0,3) | 0.01 | 0.02 | 0.04 | 0.19  |
| (0,4) | 0.01 | 0.07 | 0.02 | 0.18  |
| (0,5) | 0.01 | 0.00 | 0.02 | 0.19  |
| (0,6) | 0.00 | 0.01 | 0.00 | 0.18  |
| (0,7) | 0.02 | 0.02 | 0.02 | 0.19  |
| (0,8) | 0.01 | 0.19 | 0.02 | 0.02  |
| (1,0) | 0.04 | 0.00 | 0.00 | 0.00  |
| (1,1) | 0.00 | 0.00 | 0.00 | 0.00  |
| (1,2) | 0.00 | 0.00 | 0.00 | 0.00  |
| (1,3) | 0.00 | 0.00 | 0.00 | 0.00  |
| (1,4) | 0.05 | 0.01 | 0.00 | 0.01  |
| (1,5) | 0.00 | 0.00 | 0.00 | 0.00  |
| (1,6) | 0.00 | 0.00 | 0.00 | 0.00  |
| (1,7) | 0.00 | 0.00 | 0.00 | 0.00  |
| (1,8) | 0.02 | 0.18 | 0.00 | 0.01  |
| (2,0) | 0.00 | 0.00 | 0.00 | 0.00  |
| (2,1) | 0.00 | 0.00 | 0.00 | 0.00  |
| (2,2) | 0.03 | 0.03 | 0.03 | 0.05  |
| (2,3) | 0.02 | 0.02 | 0.05 | 0.05  |
| (2,4) | 0.01 | 0.00 | 0.04 | 0.05  |
| (2,5) | 0.02 | 0.01 | 0.04 | 0.05  |
| (2,6) | 0.00 | 0.04 | 0.02 | 0.00  |
| (2,7) | 0.00 | 0.00 | 0.00 | 0.00  |
| (2,8) | 0.00 | 0.17 | 0.00 | 0.01  |
| (3,0) | 0.10 | 0.00 | 0.00 | 0.01  |
| (3,1) | 0.00 | 0.00 | 0.00 | 0.05  |
| (3,2) | 0.06 | 0.03 | 0.06 | 0.03  |
| (3,3) | 0.00 | 0.00 | 0.00 | 0.00  |
| (3,4) | 0.02 | 0.00 | 0.00 | 0.00  |
| (3,5) | 0.00 | 0.00 | 0.00 | 0.00  |
| (3,6) | 0.00 | 0.17 | 0.02 | 0.04  |
| (3,7) | 0.02 | 0.01 | 0.17 | 0.02  |
| (3,8) | 0.04 | 0.08 | 0.18 | 0.03  |
| (4,0) | 0.08 | 0.00 | 0.00 | 0.00  |
| (4,1) | 0.00 | 0.00 | 0.00 | 0.00  |
| (4,2) | 0.07 | 0.01 | 0.03 | 0.04  |
| (4,3) | 0.00 | 0.00 | 0.00 | 0.00  |
| (4,4) | 0.00 | 0.00 | 0.00 | 0.00  |
| (4,5) | 0.00 | 0.00 | 0.00 | 0.00  |
| (4,6) | 0.02 | 0.17 | 0.01 | 0.01  |
| (4,7) | 0.00 | 0.00 | 0.00 | 0.00  |
| (4,8) | 0.04 | 0.05 | 0.03 | 0.03  |
| (5,0) | 0.10 | 0.00 | 0.00 | 0.00  |
```

```
| (5,1) | 0.00 | 0.00 | 0.00 | 0.00  |
| (5,2) | 0.09 | 0.08 | 0.03 | 0.03  |
| (5,3) | 0.01 | 0.03 | 0.09 | 0.04  |
| (5,4) | 0.00 | 0.00 | 0.00 | 0.00  |
| (5,5) | 0.04 | 0.08 | 0.02 | 0.05  |
| (5,6) | 0.03 | 0.11 | 0.07 | 0.02  |
| (5,7) | 0.00 | 0.00 | 0.00 | 0.00  |
| (5,8) | 0.04 | 0.00 | 0.01 | 0.01  |
| (6,0) | 0.10 | 0.01 | 0.00 | 0.00  |
| (6,1) | 0.00 | 0.00 | 0.10 | 0.04  |
| (6,2) | 0.01 | 0.01 | 0.06 | 0.03  |
| (6,3) | 0.08 | 0.02 | 0.01 | 0.02  |
| (6,4) | 0.01 | 0.04 | 0.08 | 0.01  |
| (6,5) | 0.03 | 0.02 | 0.09 | 0.06  |
| (6,6) | 0.03 | 0.00 | 0.02 | 0.20  |
| (6,7) | 0.00 | 0.00 | 0.00 | 0.19  |
| (6,8) | 0.00 | 0.00 | 0.00 | 0.00  |
+-------+------+------+------+-------+

Episode 2, Total Reward: 18
[['E' 'U' 'U' 'U' 'U' 'U' 'U' 'U' 'U']
 ['D' 'W' 'W' 'W' 'D' 'W' 'W' 'W' 'U']
 ['D' 'W' 'D' 'D' 'D' 'D' 'D' 'W' 'U']
 ['D' 'D' 'D' 'W' 'E' 'W' 'U' 'U' 'U']
 ['D' 'W' 'D' 'W' 'G' 'W' 'U' 'W' 'D']
 ['D' 'W' 'D' 'D' 'W' 'D' 'U' 'W' 'D']
 ['D' 'D' 'D' 'D' 'D' 'D' 'U' 'U' 'A']]
```

6. محیط را کمی عوض کردم و به جای خانه (4,0) دیوار گذاشتم و خانه (5,1) را Empty کردم. همچنین در هر دو محیط از Ghost استفاده کردم. اما در محیط تغییر یافته که تصاویر آن به صورت انیمیشنی در فایل modified_pacman_animation.gif هست، Ghost متحرک است و در نقشه حرکت می کند.

7. با ران کردن کد تعداد reward ها در هر اپیزود بعلاوه Q-table در کنسول چاپ میشود. همچنین 2 فایل گیف موجود است که board بازی را در هر اپیزود نشان می دهند. در این تصاویر رنگ سبز مربوط به agent هست. قرمز ghost است. Dot هایی که استفاده نشده اند آبی هستند و Dot های استفاده شده زرد هستند. Wall ها با مربع مشکی نمایش داده شده اند.



8. در محیط modified روح قابلیت حرکت در بازی را دارد.