|  |
| --- |
|  |
| Assignment 2 |
| Robots, Agents, and Humans |
|  |
| **Jim Ihrig and Frank Bruno** |
| **3/28/2011** |

# Grid World

Stuff about the Grid World used from her example.

## Q Learning

Stuff about the Q Learning in the Grid World.

### Design

The transition and reward functions used in your world.

### Exploration Policy

A description of the exploration policy.

### Rule Implementation

A description of the Q-learning rule implementation.

### Final Q-Table

The learning rate, discount factor, and number of episodes used for this Q-Table are 150, 0.7, and 0.7 respectively. It is important to note that the maximum number of steps allowed per episode is 500 for this example shown in . Given that the number of times the episodes were cut short was 45 times.

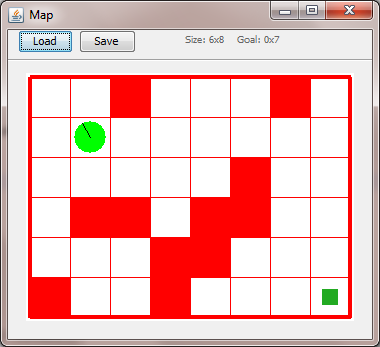


Figure 1 - Example Map used for Q-Learning

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Starting position | (4, 1) | Goal | (0,7) |  |
| Q | up | down | left | right |
| (5, 0) | Outside Grid | 3.86 | Outside Grid | 3.86 |
| (5, 1) | Outside Grid | 5.52 | 2.7 | Wall |
| (5, 2) | Is Wall | Is Wall | Is Wall | Is Wall |
| (5, 3) | Outside Grid | 11.26 | Wall | 11.26 |
| (5, 4) | Outside Grid | 16.09 | 7.88 | 16.09 |
| (5, 5) | Outside Grid | 22.99 | 11.26 | Wall |
| (5, 6) | Is Wall | Is Wall | Is Wall | Is Wall |
| (5, 7) | Outside Grid | 46.96 | Wall | Outside Grid |
| (4, 0) | 2.7 | 2.7 | Outside Grid | 5.52 |
| (4, 1) | 3.86 | 3.86 | 3.86 | 7.88 |
| (4, 2) | Wall | 5.52 | 5.52 | 11.26 |
| (4, 3) | 7.88 | 7.88 | 7.88 | 16.09 |
| (4, 4) | 11.26 | 11.26 | 11.26 | 22.98 |
| (4, 5) | 16.1 | Wall | 16.09 | 32.87 |
| (4, 6) | Wall | 46.95 | 23 | 46.96 |
| (4, 7) | 32.87 | 67.09 | 32.87 | Outside Grid |
| (3, 0) | 3.86 | 1.89 | Outside Grid | 3.86 |
| (3, 1) | 5.52 | Wall | 2.7 | 5.52 |
| (3, 2) | 7.88 | Wall | 3.86 | 7.88 |
| (3, 3) | 11.26 | 5.52 | 5.52 | 11.26 |
| (3, 4) | 16.09 | Wall | 7.88 | Wall |
| (3, 5) | Is Wall | Is Wall | Is Wall | Is Wall |
| (3, 6) | 32.86 | 67.04 | Wall | 67.1 |
| (3, 7) | 46.96 | 95.86 | 46.96 | Outside Grid |
| (0, 0) | 2.7 | 1.32 | Outside Grid | Wall |
| (2, 1) | Is Wall | Is Wall | Is Wall | Is Wall |
| (2, 2) | Is Wall | Is Wall | Is Wall | Is Wall |
| (2, 3) | 7.88 | Wall | Wall | Wall |
| (2, 4) | Is Wall | Is Wall | Is Wall | Is Wall |
| (2, 5) | Is Wall | Is Wall | Is Wall | Is Wall |
| (2, 6) | 46.9 | 95.83 | Wall | 95.84 |
| (2, 7) | 67.09 | 136.94 | 67.06 | Outside Grid |
| (1, 0) | 1.89 | Wall | Outside Grid | 0.93 |
| (1, 1) | Wall | 0.65 | 1.32 | 0.65 |
| (1, 2) | Wall | 0.45 | 0.93 | Wall |
| (1, 3) | Is Wall | Is Wall | Is Wall | Is Wall |
| (1, 4) | Is Wall | Is Wall | Is Wall | Is Wall |
| (1, 5) | Wall | 95.79 | Wall | 95.85 |
| (1, 6) | 67.04 | 136.89 | 67.02 | 136.94 |
| (1, 7) | 95.85 | 195.67 | 95.83 | Outside Grid |
| (0, 0) | Is Wall | Is Wall | Is Wall | Is Wall |
| (0, 1) | 0.93 | Outside Grid | Wall | 0.45 |
| (0, 2) | 0.65 | Outside Grid | 0.65 | Wall |
| (0, 3) | Is Wall | Is Wall | Is Wall | Is Wall |
| (0, 4) | Wall | Wall | Wall | 95.79 |
| (0, 5) | 67.07 | Outside Grid | 67.03 | 136.9 |
| (0, 6) | 95.81 | Outside Grid | 95.67 | 195.65 |
| (0, 7) | 135.98 | Outside Grid | 136.69 | Outside Grid |

### Tests using different parameters

Talk about the tests done using the different parameters. The tests done were

|  |  |  |  |
| --- | --- | --- | --- |
| Episodes | Learning Rate | Discount Factor | Max steps taken per episode |
| 300 | 0.5 | 0.5 | 1000 |
| 150 | 1 | 0.9 | 500 |
| 75 | 0.75 | .7 | 250 |

### Action Sequences

Two example action sequences generated using your policy.

## A\* Path Planning

• pseudocode of your planner

• f, g, and h values for one example map

# Conclusion