Q1 Reward function 1 Heatmap 0 3 5 6 10 9 1.0 2 8.0 3 4 · 0.6 5 · 6 0.4 8 -0.2 9 .