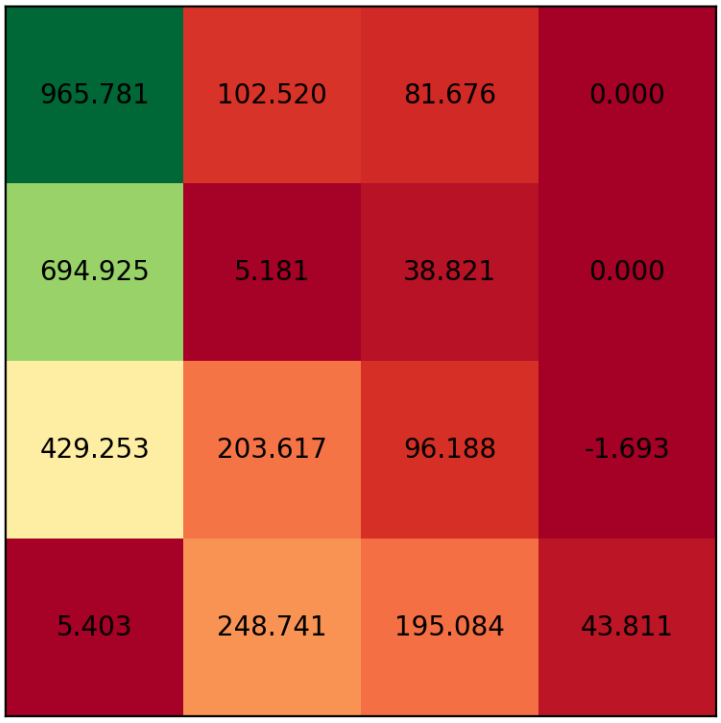
Inverse RL

Exercise 1

a [0 3 0 3 0 0 0 0 3 1 0 0 0 2 1 0]

b [0.00279781, 0.00154556, 0.0021857, 0.00267707,
0.00183428, 0.28992156, 0.00154973, 0.17667529,
0.0022512, 0.00329131, 0.0022425, 0.0944733,
0.2712387, 0.0049202, 0.00237741, 0.14001836]



c

d Maximum Entropy IRL models can learn well more sophisticated environments where simple counting of is not sufficient or a lot of states are not visited and thus the model can handle them with different policies at random. ME IRL forces the model to learn hoe to behave in every state, even if it is not visited at the demonstrations.