

Maximum final average reward over 10 initializations

Regularisation strength

1e-4.0  
1e-3.5  
1e-3.0  
1e-2.5  
1e-2.0  
1e-1.5  
1e-1.0  
1e-0.5

1e-7.00 1e-6.74 1e-6.47 1e-6.21 1e-5.95 1e-5.68 1e-5.42 1e-5.16 1e-4.89 1e-4.63 1e-4.37 1e-4.11 1e-3.84 1e-3.58 1e-3.32 1e-3.05 1e-2.79 1e-2.53 1e-2.26 1e-2.00

Learning rate

