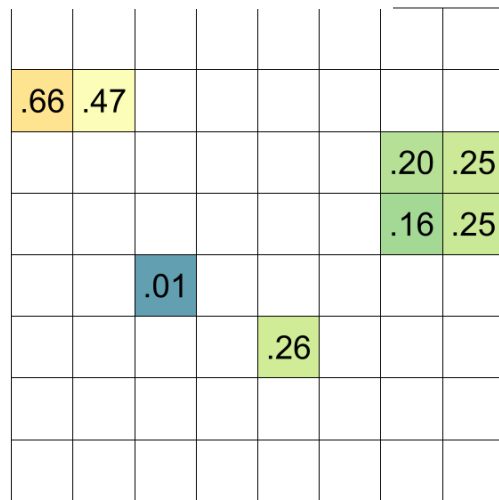
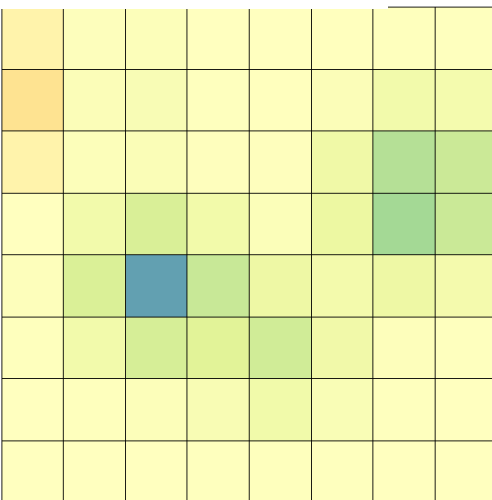


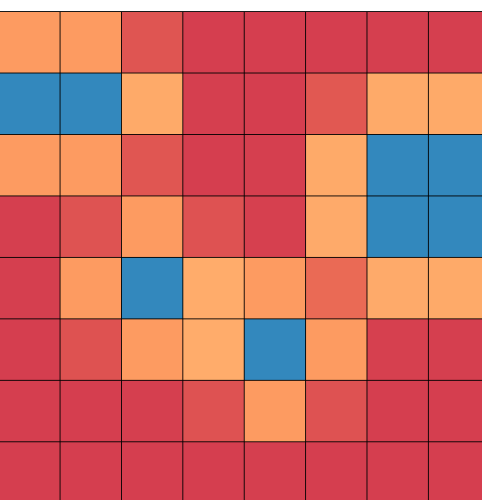
Observations (t=8)



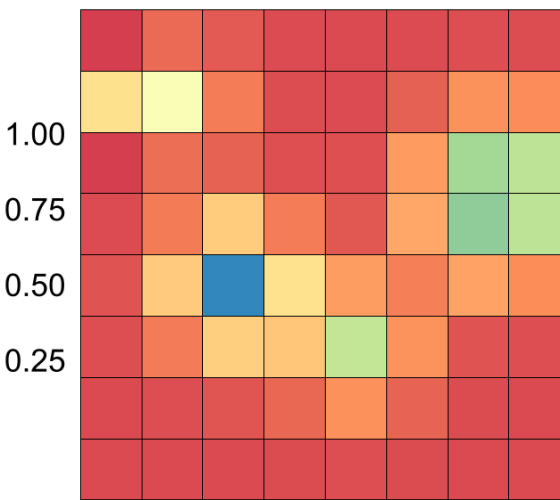
Expected rewards

 $\mu(\mathbf{x})$ 1.00
0.75
0.50
0.25
0.00

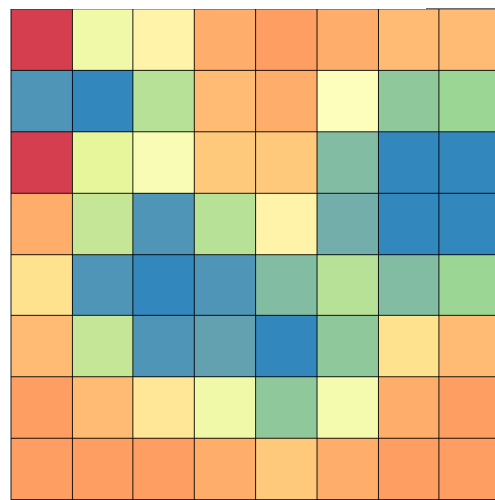
Uncertainty

 $\sigma(\mathbf{x})$ 1.00
0.75
0.50
0.25

Upper Confidence Bound (UCB)

 $UCB(\mathbf{x})$ 1.00
0.75
0.50
0.25

Softmax predictions

 $p(\mathbf{x})$ 0.03
0.02
0.01
0.00Generalization λ Exploration bonus β Random temperature τ