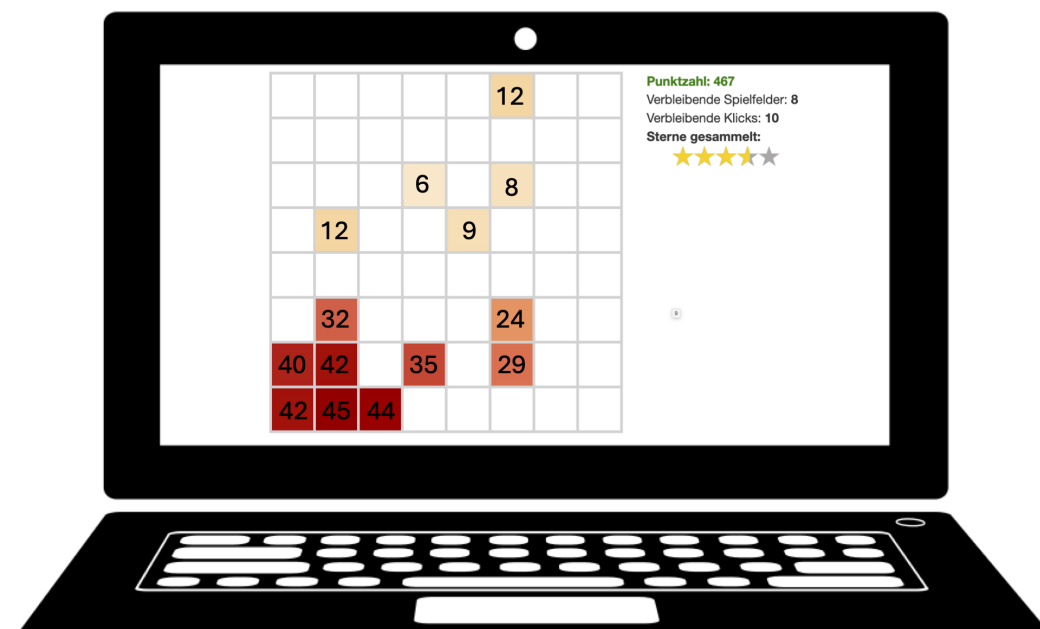
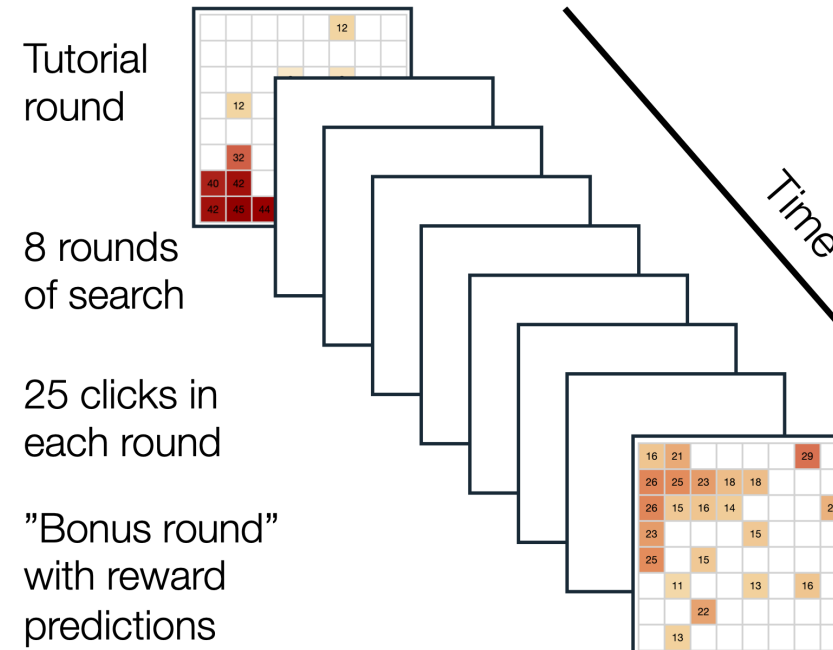


a Experiment and design

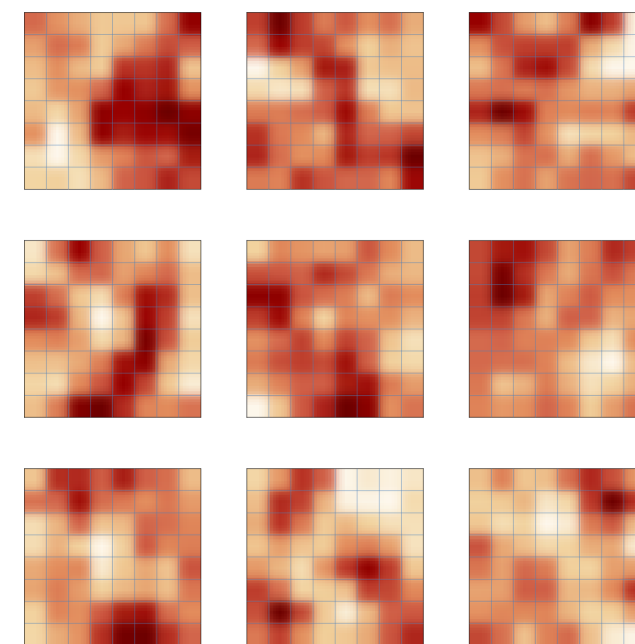
Task



Procedure

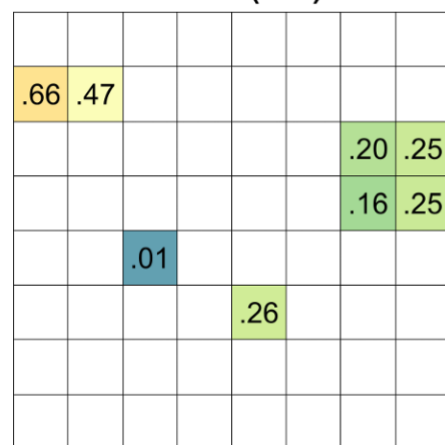


Environments

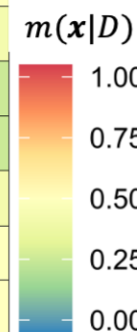
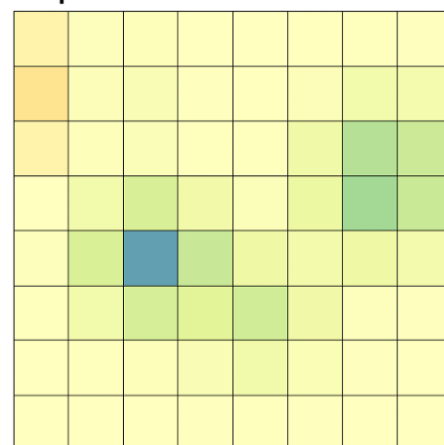


b Computational GP-UCB model

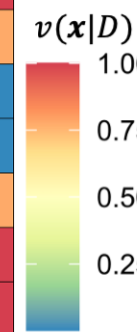
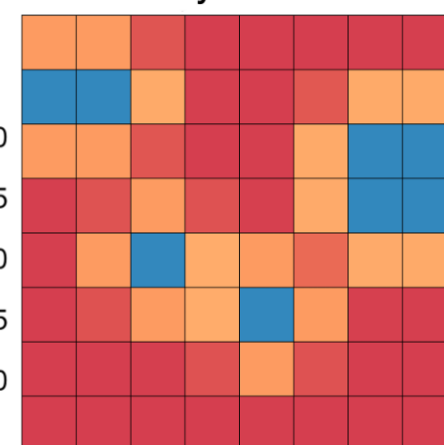
Observations (t=8)



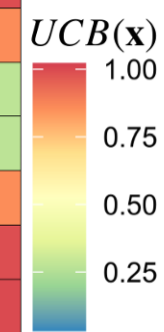
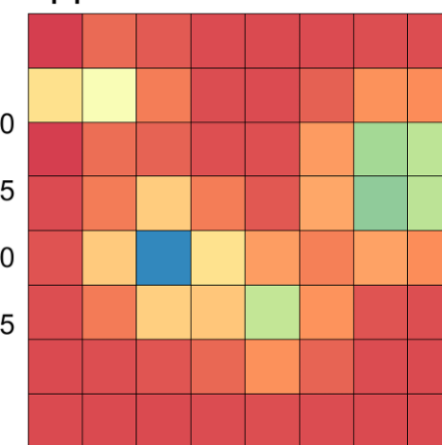
Expected rewards



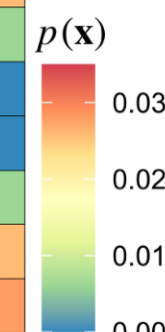
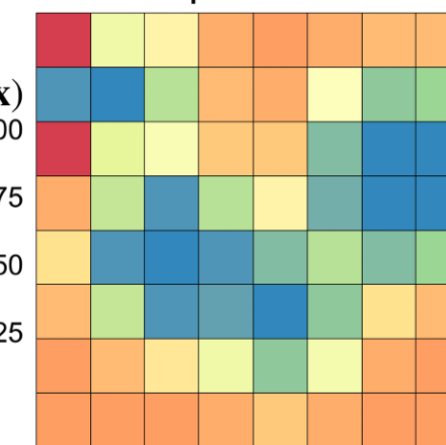
Uncertainty



Upper Confidence Bound (UCB)



Softmax predictions



Generalization λ

Exploration bonus β

Random temperature τ