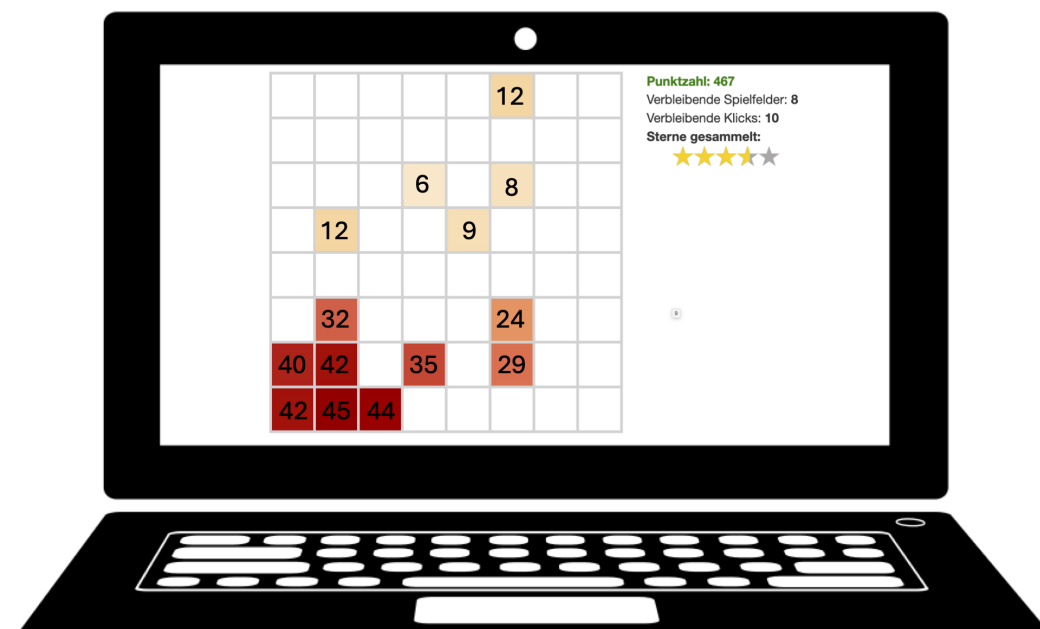
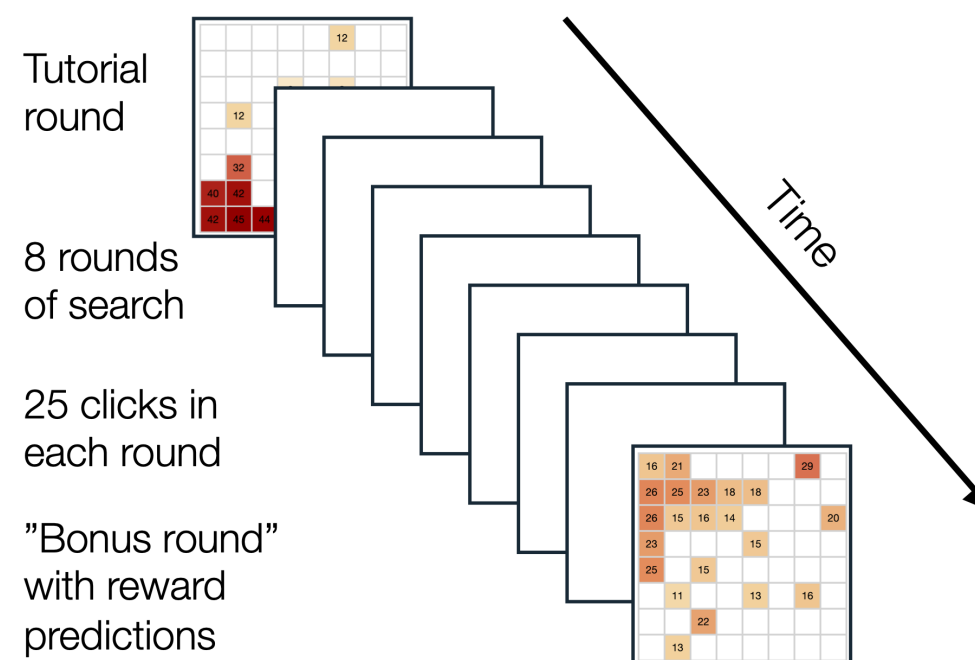


# a Experiment and design

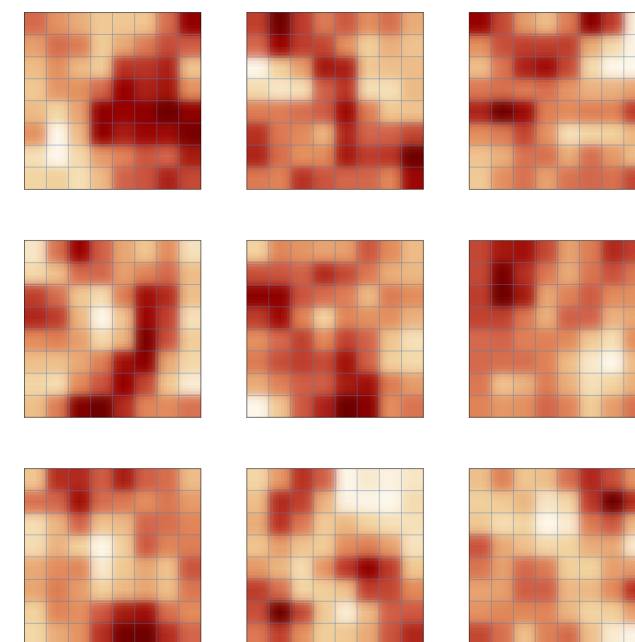
Task



Procedure

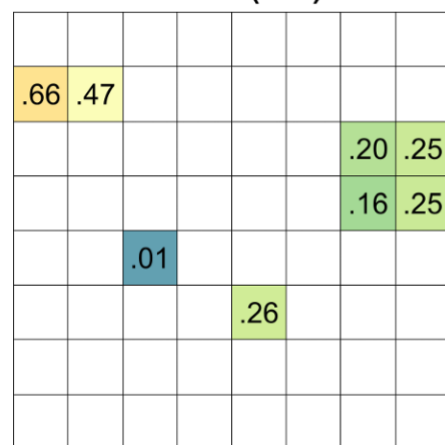


Environments

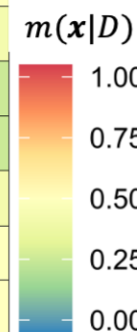
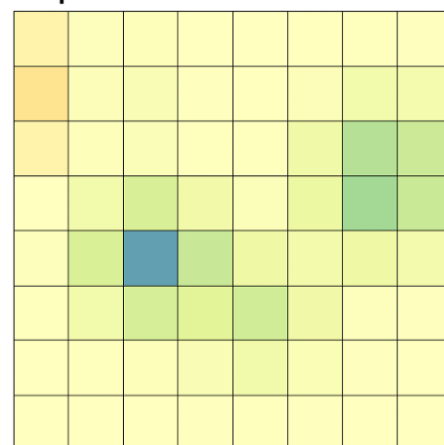


# b Computational GP-UCB model

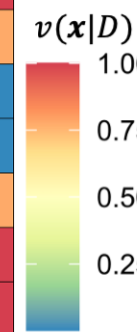
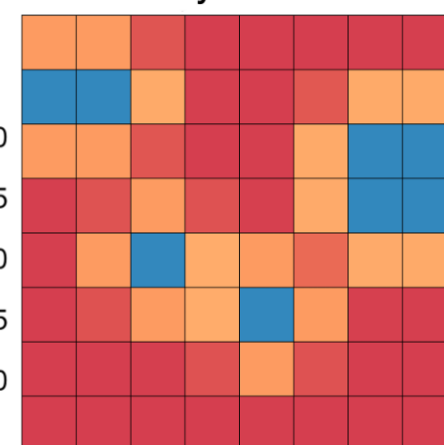
Observations (t=8)



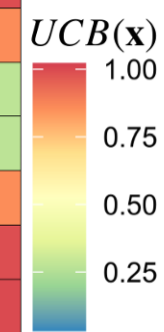
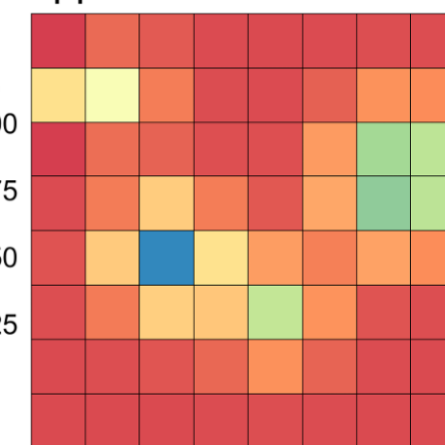
Expected rewards



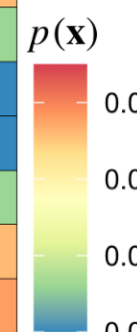
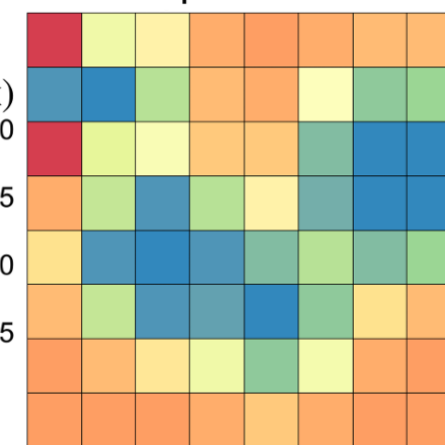
Uncertainty



Upper Confidence Bound (UCB)



Softmax predictions



Generalization  $\lambda$       Exploration bonus  $\beta$       Random temperature  $\tau$