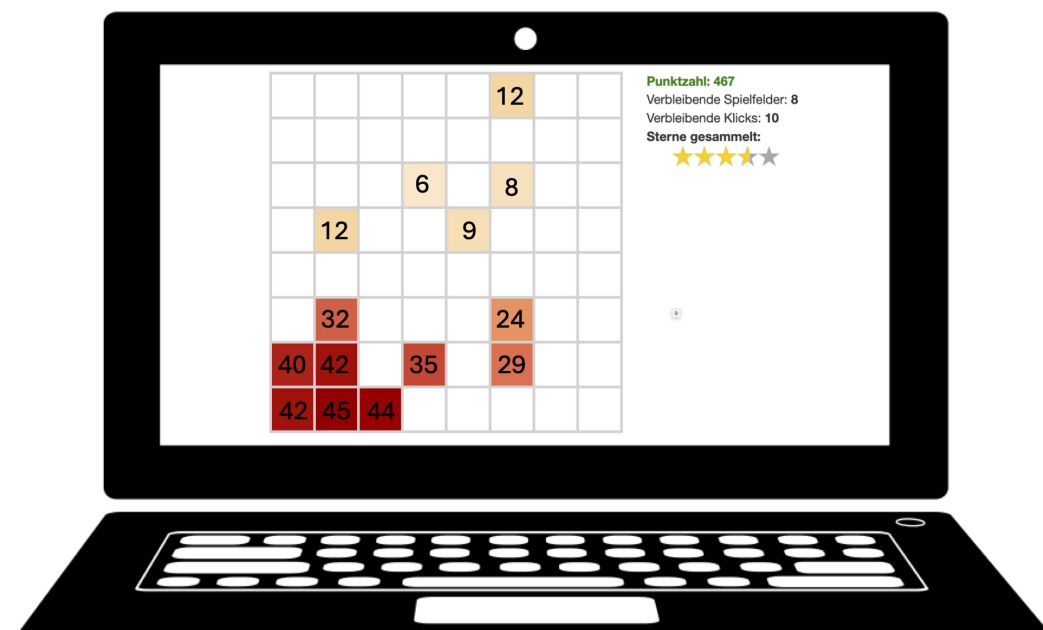
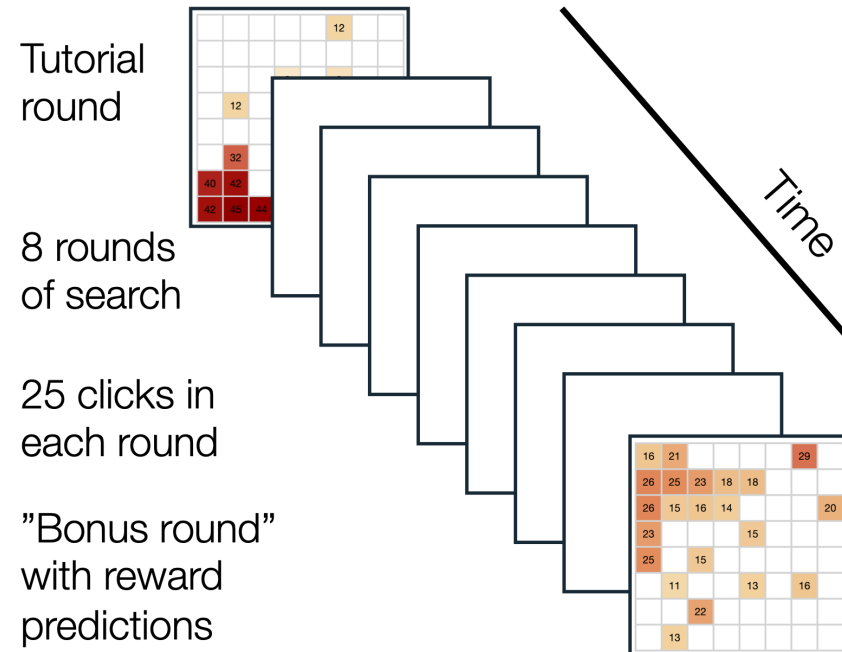


a Experiment and design

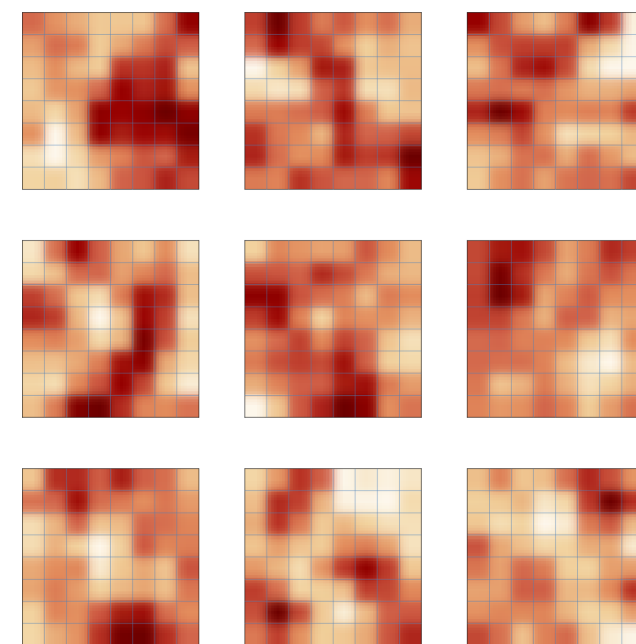
Task



Procedure

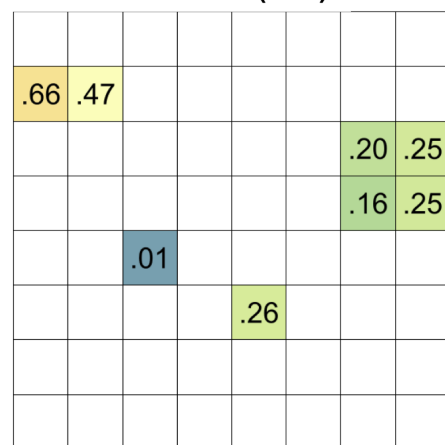


Environments

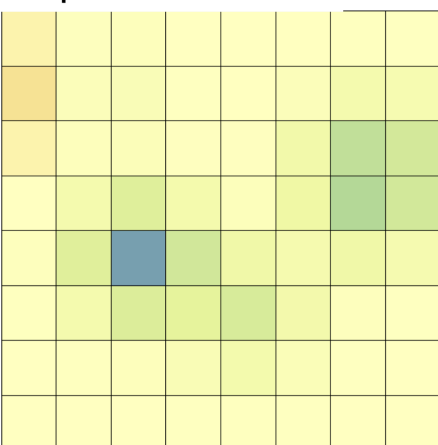


b Computational GP-UCB model

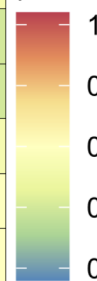
Observations (t=8)



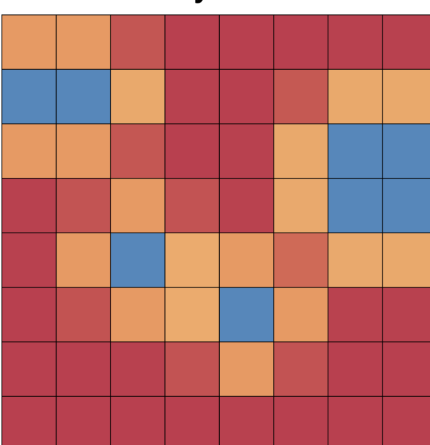
Expected rewards



$\mu(\mathbf{x})$



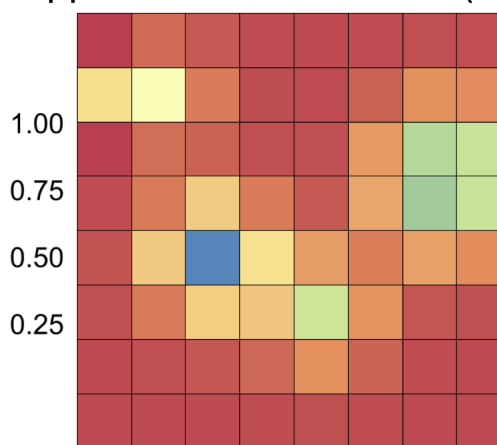
Uncertainty



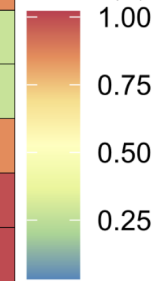
$\sigma(\mathbf{x})$



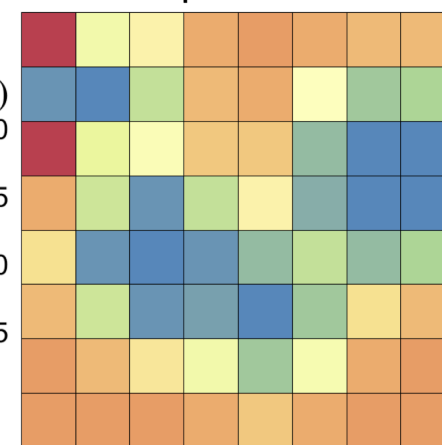
Upper Confidence Bound (UCB)



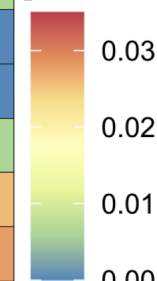
$UCB(\mathbf{x})$



Softmax predictions



$p(\mathbf{x})$



Generalization λ

Exploration bonus β

Random temperature τ