Rapid spatial learning via efficient exploration and inference

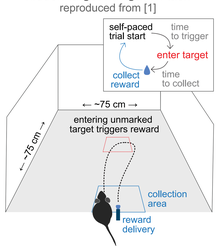
Nada AbdelRahman[1], Wanchen Jiang[2], Joshua T Dudman[1*], & Ann M Hermundstad[1*]

[1]HHMI Janelia Research Campus; [2]ShanghaiTech University; *correspondence

motivation & setup

- **existing models** of learning **fail to capture** many salient aspects of **animal learning**
- we highlight these limitations in the context of navigational learning, using a **hidden-target foraging task for mice [1]**
- we show how agents equipped with **controllable behavioral structure & basic knowledge** about their surroundings can **overcome these limitations**

hidden target navigation task
reproduced from [1]



existing models are unstructured and slow:

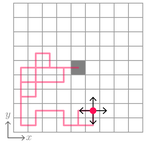
typical agent

states s: □□□□

actions a: ↑ ← ↓ →

value: $Q(a,s)$

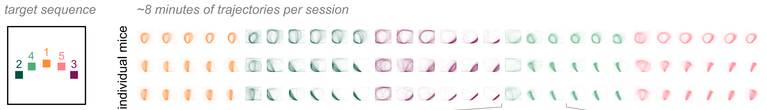
policy: $\begin{cases} (1-\epsilon) & \text{greedy action } \max_a Q(a,s) \\ \epsilon & \text{random action} \end{cases}$



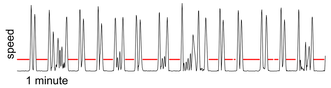
well-known limitations

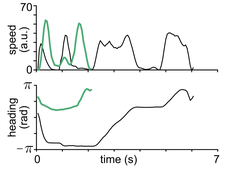
slow to learn:	requires hundreds of thousands of trials to learn even instinctive behaviors [2]
doesn't generalize:	no *a priori* knowledge about the structure of its surroundings
doesn't capture behavior:	exploration is assumed to be random, and has no intrinsic structure

animal learning is structured and fast:



trajectories are characterized by **structured speed profiles delineated by pauses**

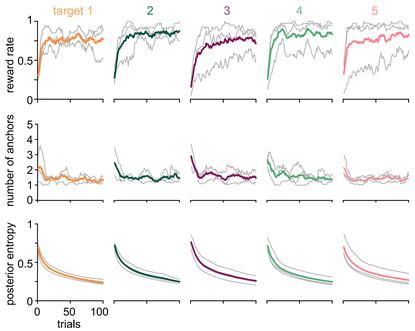




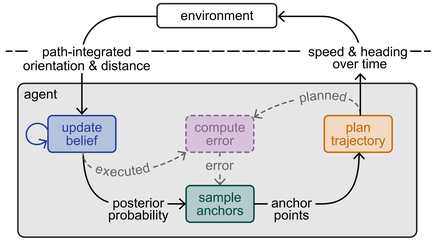
mice learn target switches equally fast, regardless of their spatial separation

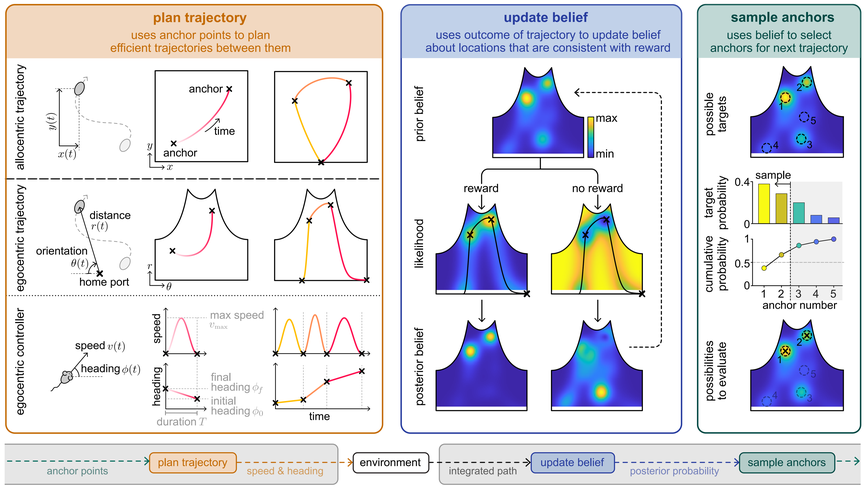
initial search exhibits more pauses (putative anchor points in trajectories)

trajectories contain **sufficient information to localize a belief** about target locations

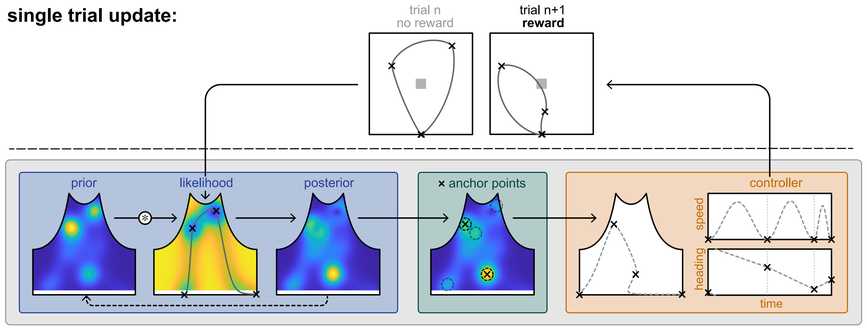


agent architecture:



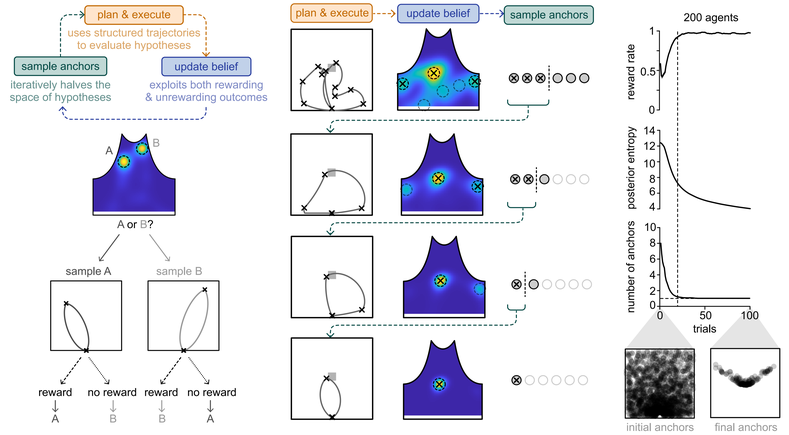


single trial update:

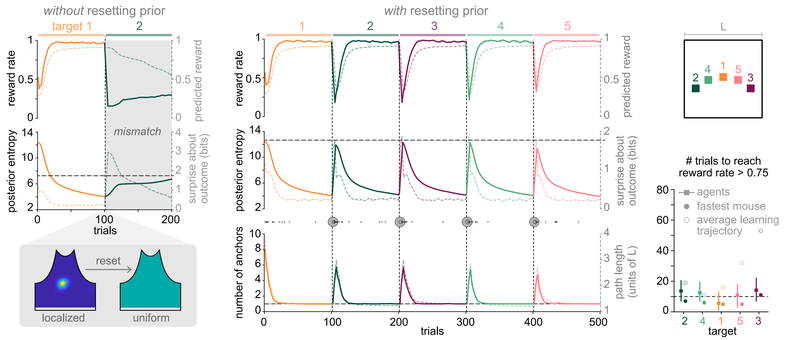


belief, sampler, & planner modules interact to rapidly narrow the hypothesis space,

leading agents to learn compact trajectories that are consistent with reward



agents learn target switches equally fast by exploiting surprise about new outcomes



summary & outlook:

- agents **learn within a handful of trials** to reliably intercept hidden targets, qualitatively capturing both the speed of learning and the temporal evolution of behavioral changes observed in mice
- learning is fast because agents **use structured behavioral trajectories to iteratively evaluate sets of hypotheses** about the consequences of their actions
- naturally enables **memory caching, obstacle avoidance, & coordinate transformations;** to hear more, come to the workshop on "Reimagining RL in the Brain"

references:

[1] Jiang, Xu, Dudman (2022), PMID: 36446934; [2] Shamash et al. (2023), PMID: 37119818