Can curiosity solve the e-e dilemma?

Readings for today

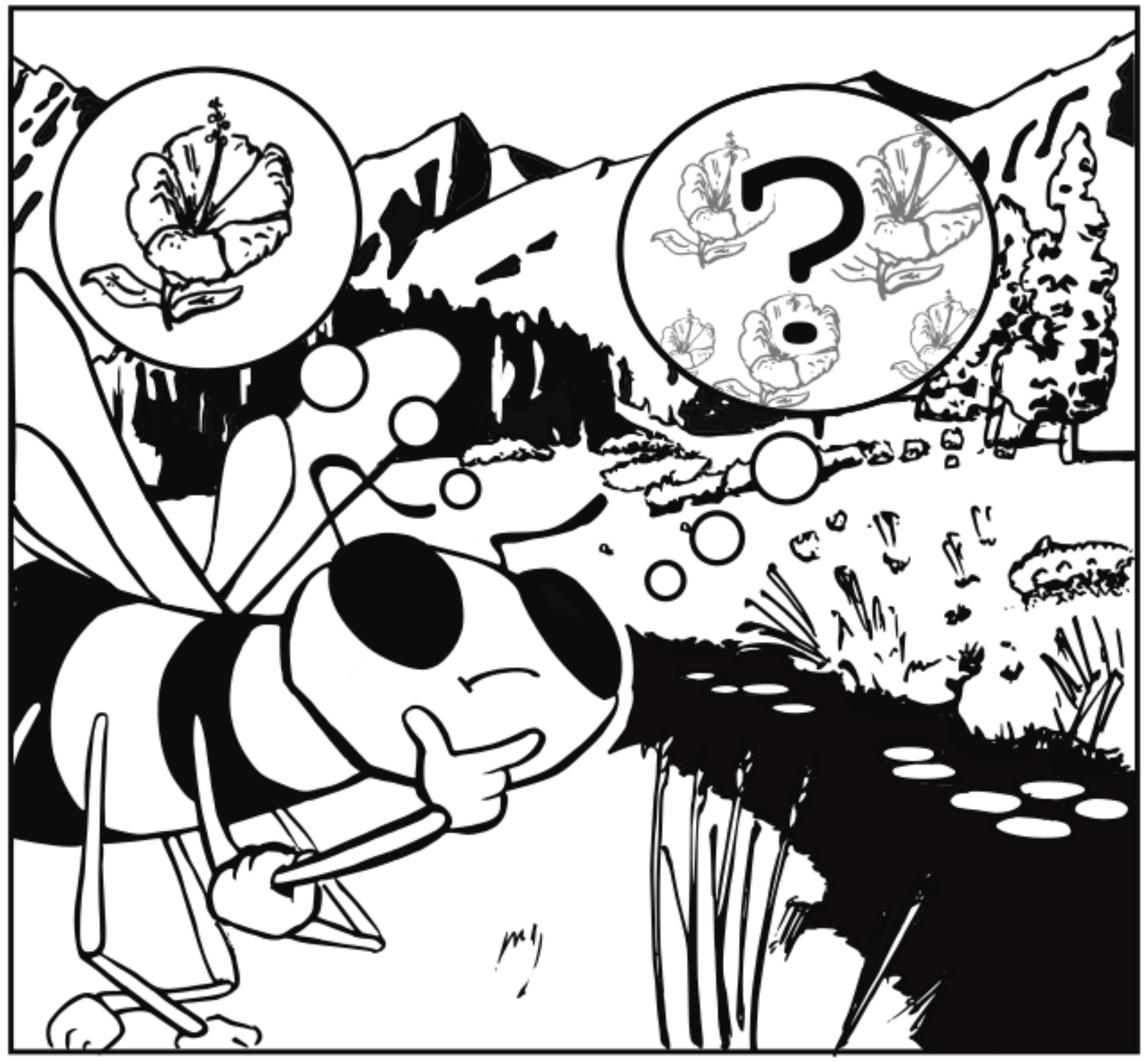
- Peterson, E. J., & Verstynen, T. D. (2022). Embracing curiosity eliminates the exploration-exploitation dilemma. bioRxiv, 671362.

Topics

- Rethinking information value
- A way around the e-e dilemma

Rethinking information value

Rethinking the dilemma



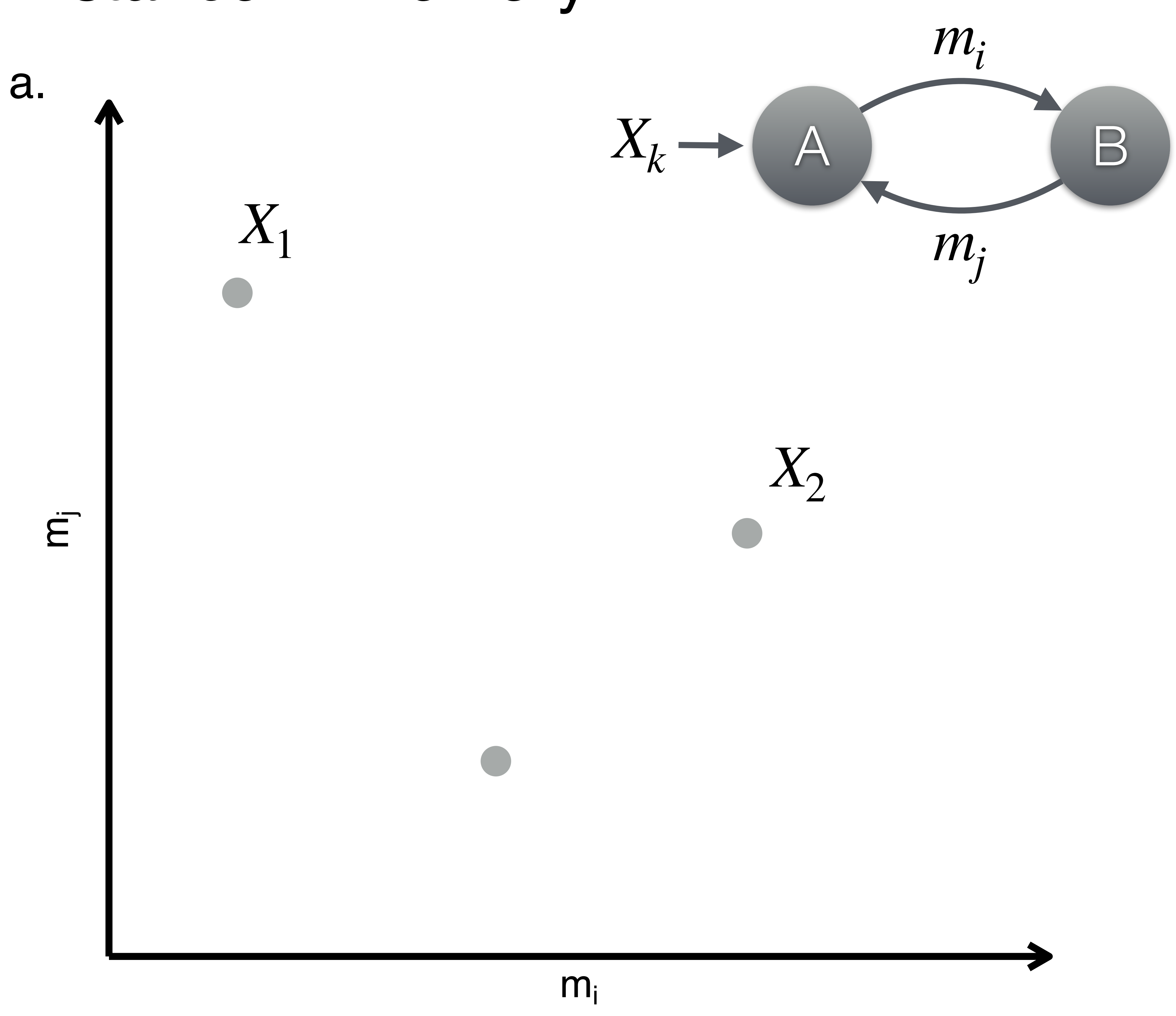
- Value-based decision making is dominated by explore-exploit policy, π.
- Mathematically optimal solution is intractable for reward collection.

$$\max \sum_{s \in S, T} \gamma R$$

(Peterson & Verstynen, 2022)

Information value: E

Distance in memory M



The problem of learning:

Memory

$$M' \leftarrow f(X, M)$$

Updated memory

Observation

Learning function

The problem of forgetting:

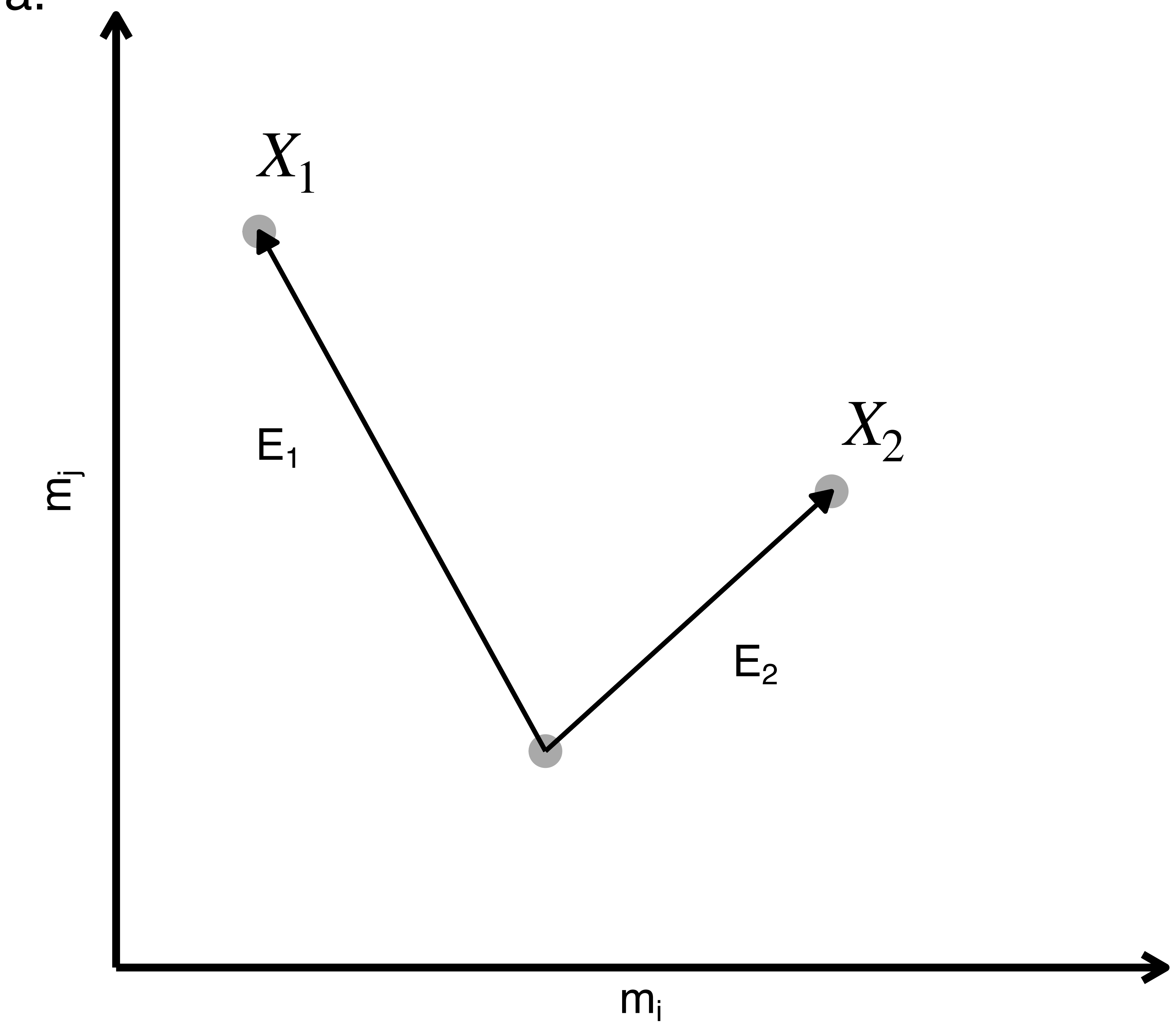
$$f^{-1}(X, M') \rightarrow M$$

(Peterson & Verstynen, 2022)

Information value: E

Distance in memory M

a.



Axiom of Memory:
E depends only on the difference $\Delta\mathbf{M}$ between $\mathbf{M}$ and $\mathbf{M}'$

Axiom of Specificity:
If all $\Delta\mathbf{M}$ are equal, then $E = 0$

Axiom of Scholarship:
$E \geq 0$

Axiom of Equilibrium:
For the same observation E should approach 0 in finite time.

(Peterson & Verstynen, 2022)

Universal information value

This geometric definition encompasses all prior information value measures:

Information value (Howard 1966)

$$V^*_{avg} = V_{avg} - V'_{avg}$$
$$= \sum_i p(x_i)\{\max_j[V(u_j|x_i)]\} - \max_j\{\sum_i p(x_i)V(u_j|x_i)\}$$

- x : event
- i : state
- $p(x_i)$: probability of event i
- u: action
- C : cost per bit

Information value (Sheridan 1995)

$$V^*_{net} = V^*_{avg} - H^*_{avg}$$
$$= \sum_i p(x_i)\{\max_j[V(u_j|x_i)]\}$$
$$- \max_j\{\sum_i p(x_i)V(u_j|x_i)\} + C\sum_i p(x_i)log_2[p(x_i)]$$

KL-divergence

$$D_{KL}(p(X)\,||\,q(X)) = \Sigma_{x\in X}p(x)log\frac{p(x)}{q(x)}dx$$

A way around the e-e dilemma

Curiosity as directed exploration

Directed exploration

$$Q(a) = r(a) + IB(a)$$

How good we expect a to be

Information bonus

Example: Upper confidence bound

variance of the posterior distribution

$$p(a) = Q(a) + 2\sigma(a)$$

(Wilson et al. 2020)

A scheduling problem

An alternative view:

- Turn the dilemma into a *two objective problem*
- Mathematically tractable

$$\max \sum_{s \in S, T} R$$

$$\max \sum_{s \in S, T} E$$

Optimal E learning:

- $\hat{E}$ has optimal substructure
- So the optimal learning policy is the Bellman eq.

$$V^*_{\hat{E}}(\mathbf{S}) = \underset{\mathbf{A} \in \mathbb{A}}{\operatorname{argmax}} \left[\hat{E}_t + V^*_{\hat{E}}(\Lambda(\mathbf{S}, \mathbf{A})) \right]$$

State

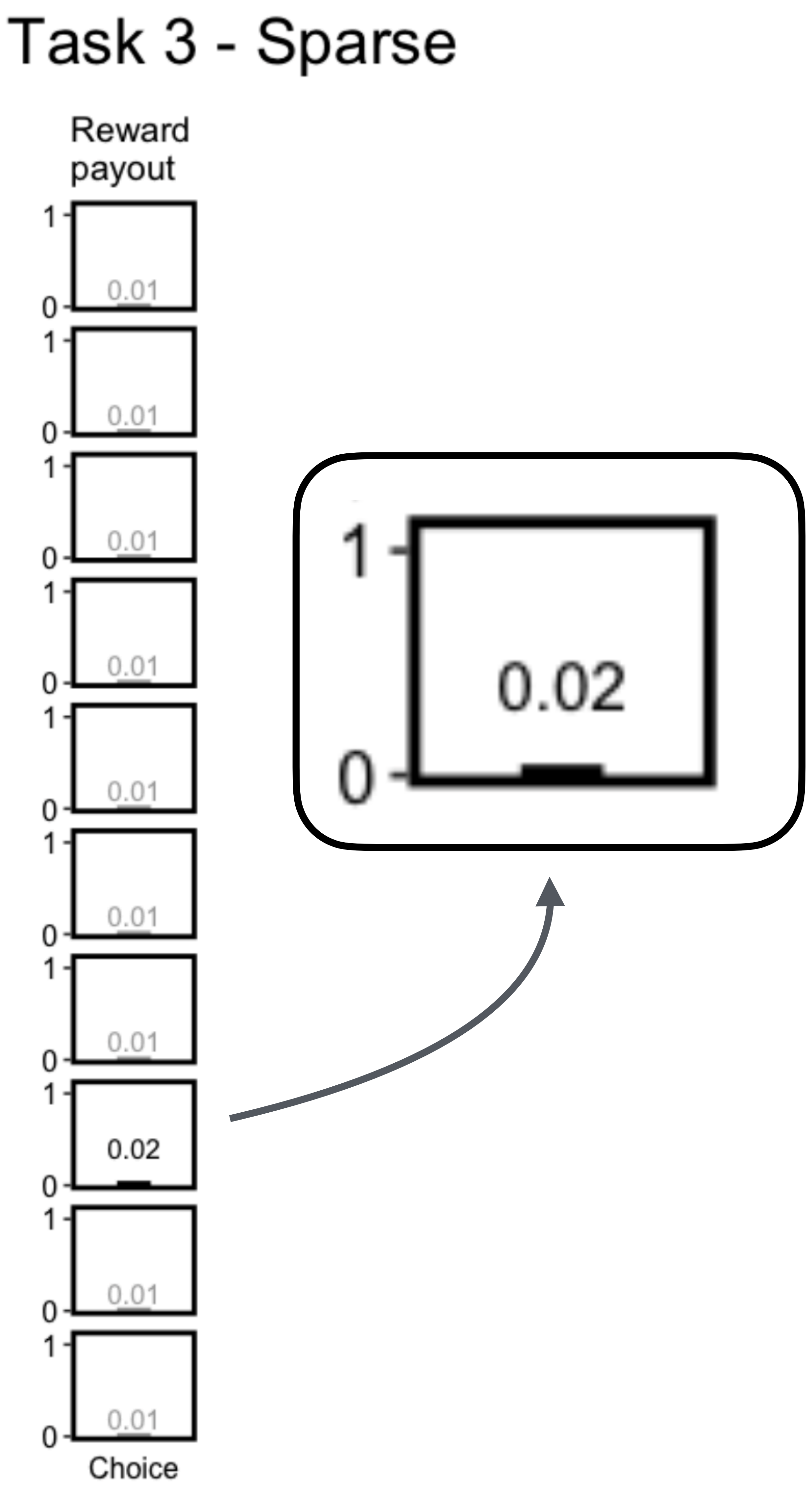
Action

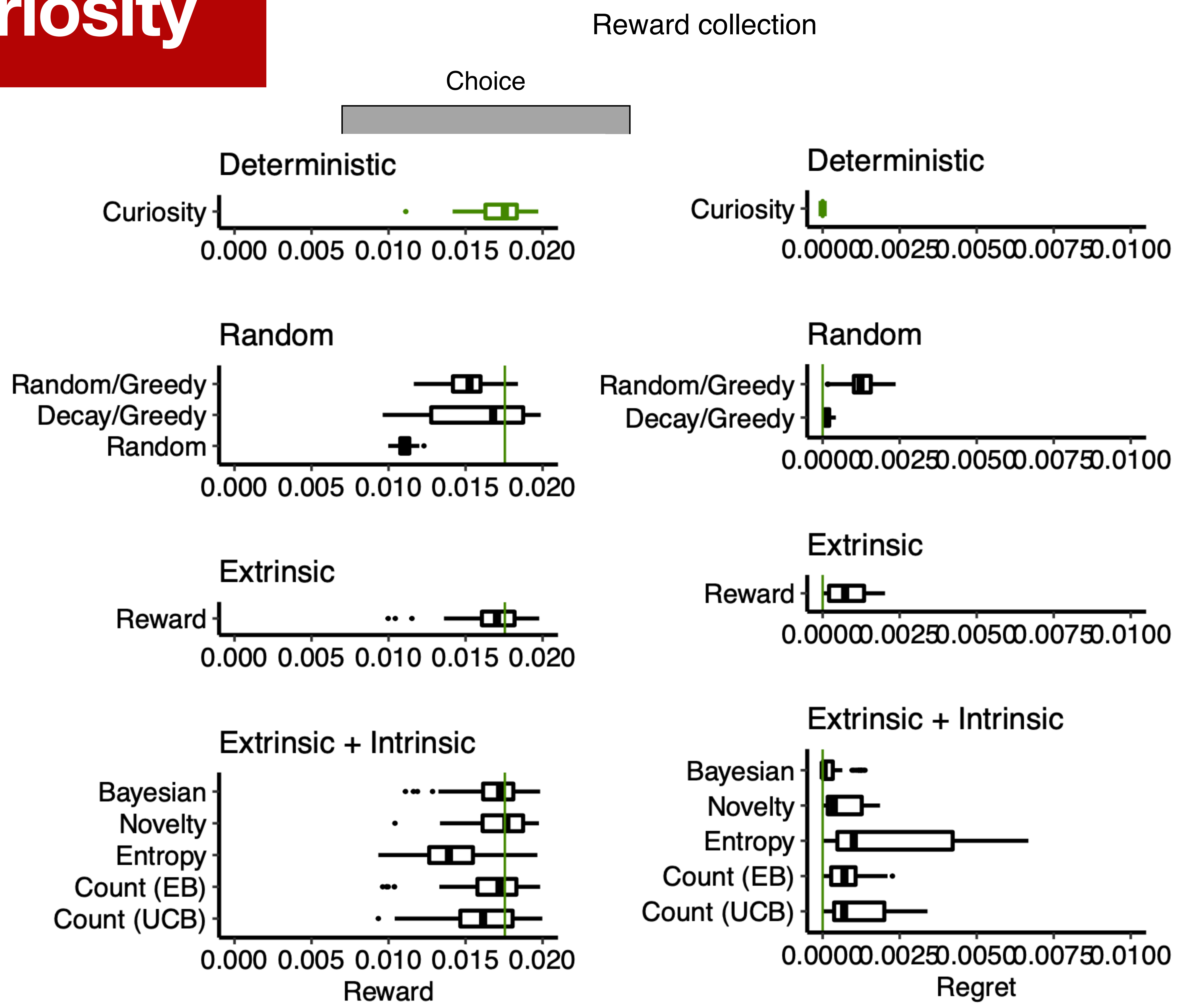
Optimal meta-greedy policy:

$$\Pi_\pi = \begin{cases} \pi^*_{\hat{E}} : E > R \\ \pi_R : E < R \end{cases}$$

(Peterson & Verstynen, 2022)

Evaluating optimal curiosity





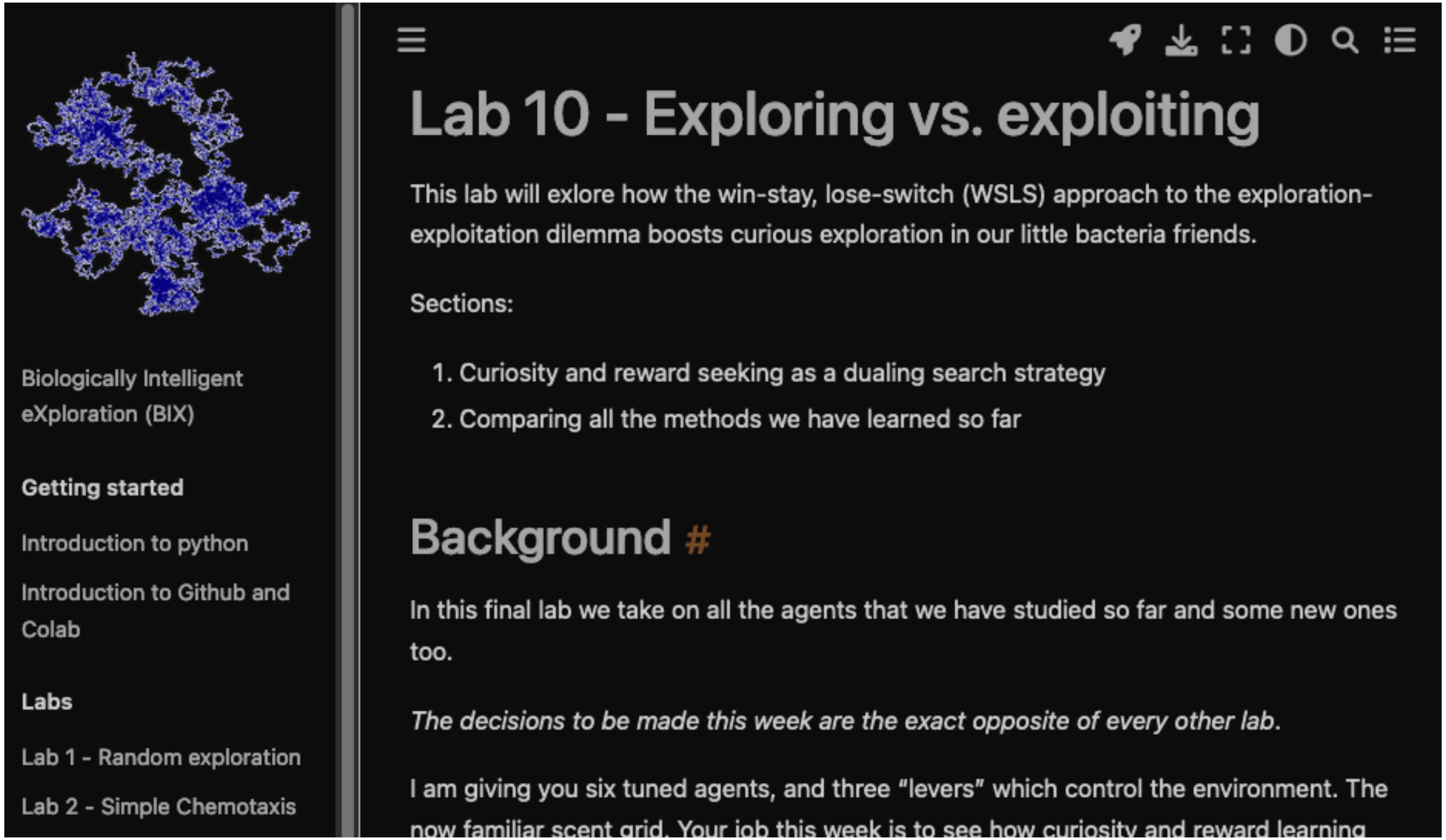
(Peterson & Verstynen, 2022)

Take home message

- The universal definition of information value is about how much a signal can change an underlying memory.
- If you treat maximizing rewards versus maximizing information as separate objectives, the exploration-exploitation dilemma disappears.

Lab 10: Curiosity-driven exploration

URL: https://coaxlab.github.io/BIX-book/notebooks/lab10-exploration_vs_exploitation.html



Biologically Intelligent eXploration (BIX)

Getting started

Introduction to python

Introduction to Github and Colab

Labs

Lab 1 - Random exploration

Lab 2 - Simple Chemotaxis

Lab 10 - Exploring vs. exploiting

This lab will exlore how the win-stay, lose-switch (WSLS) approach to the exploration-exploitation dilemma boosts curious exploration in our little bacteria friends.

Sections:

1. Curiosity and reward seeking as a dualing search strategy
2. Comparing all the methods we have learned so far

Background

In this final lab we take on all the agents that we have studied so far and some new ones too.

The decisions to be made this week are the exact opposite of every other lab.

I am giving you six tuned agents, and three “levers” which control the environment. The now familiar scent grid. Your job this week is to see how curiosity and reward learning