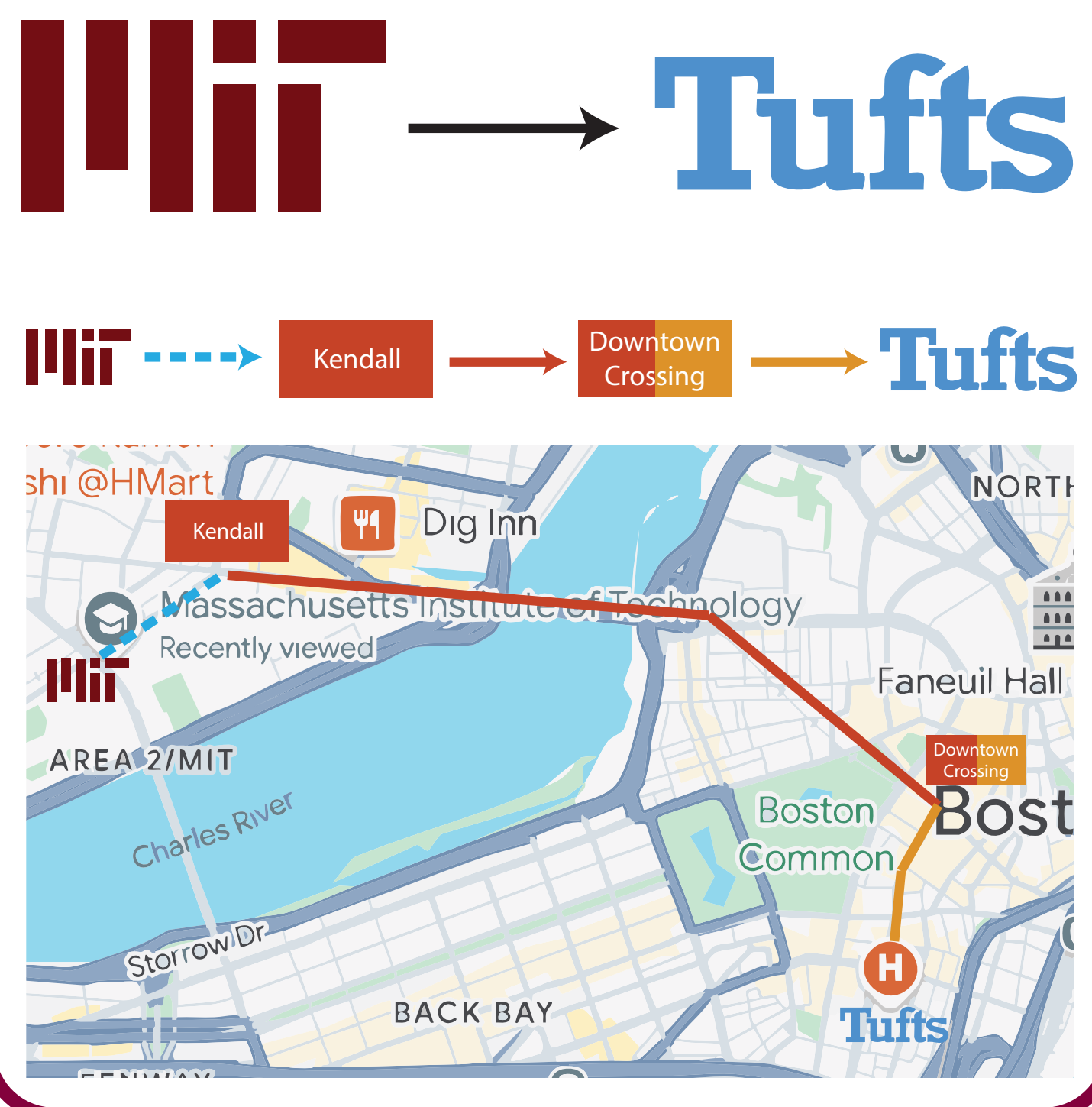


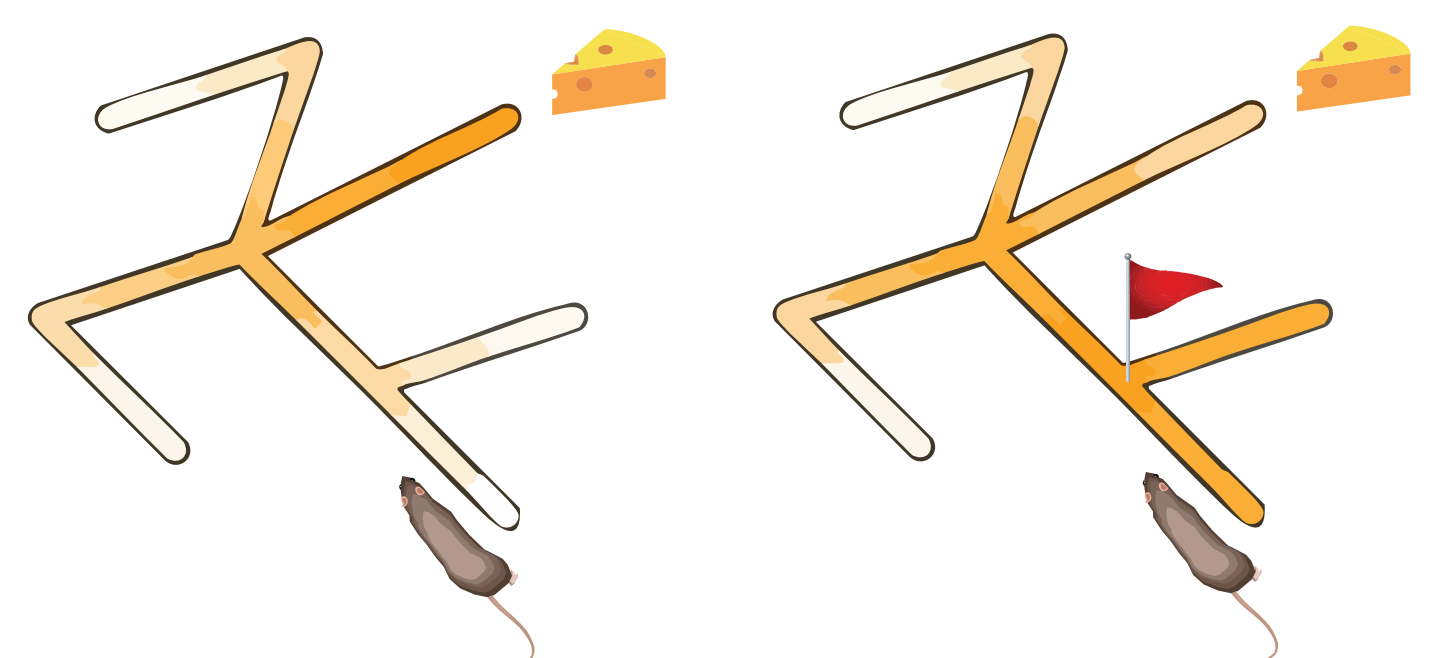
Introduction

In this work, we propose a neural mechanistic theory on how animals plan hierarchically for decision making.

- How to learn hierarchy from experience?
- How to learn hierarchical subgoal values?
- How to plan over learned hierarchy and subgoal values?

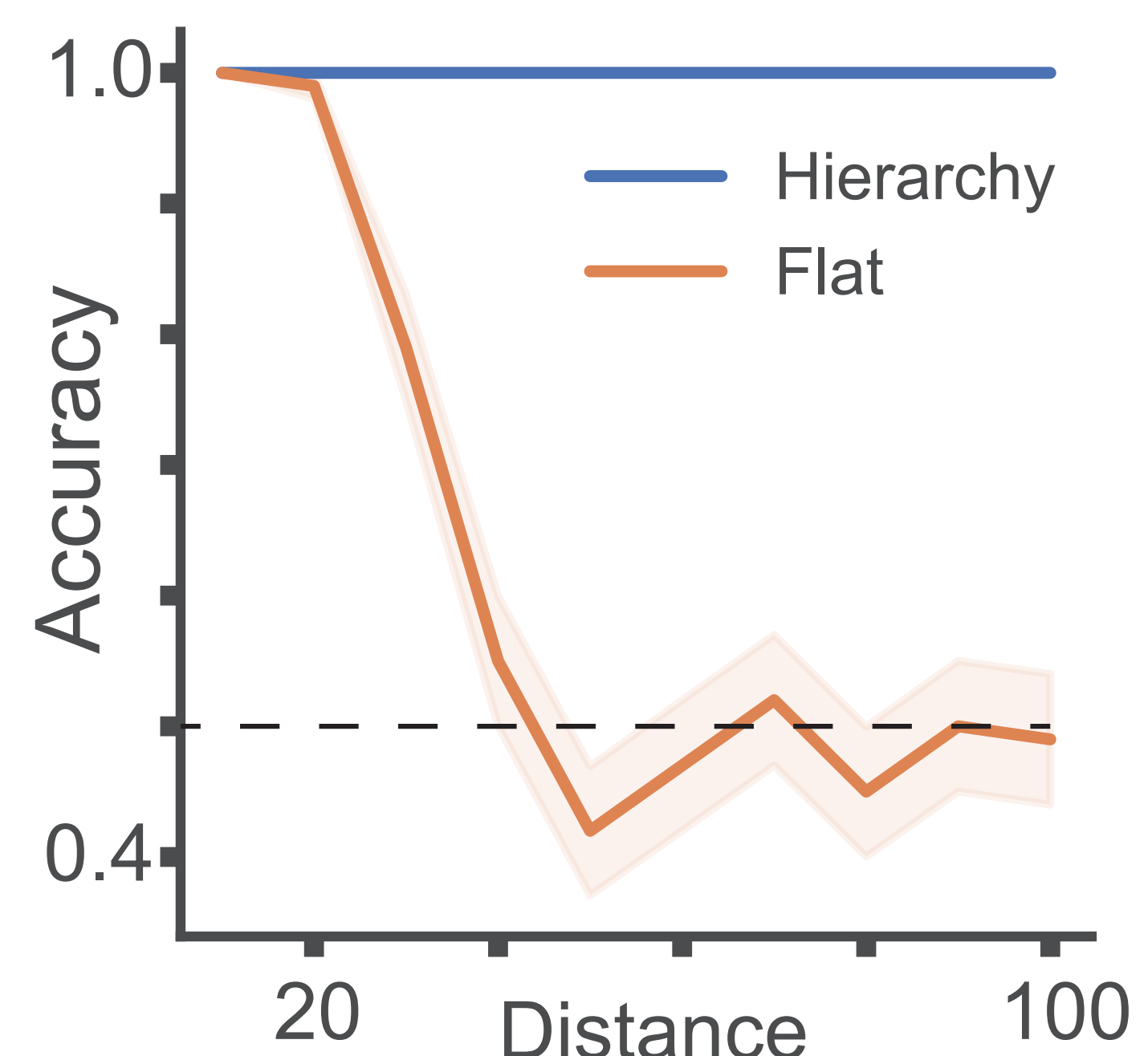


Why flat planning fails

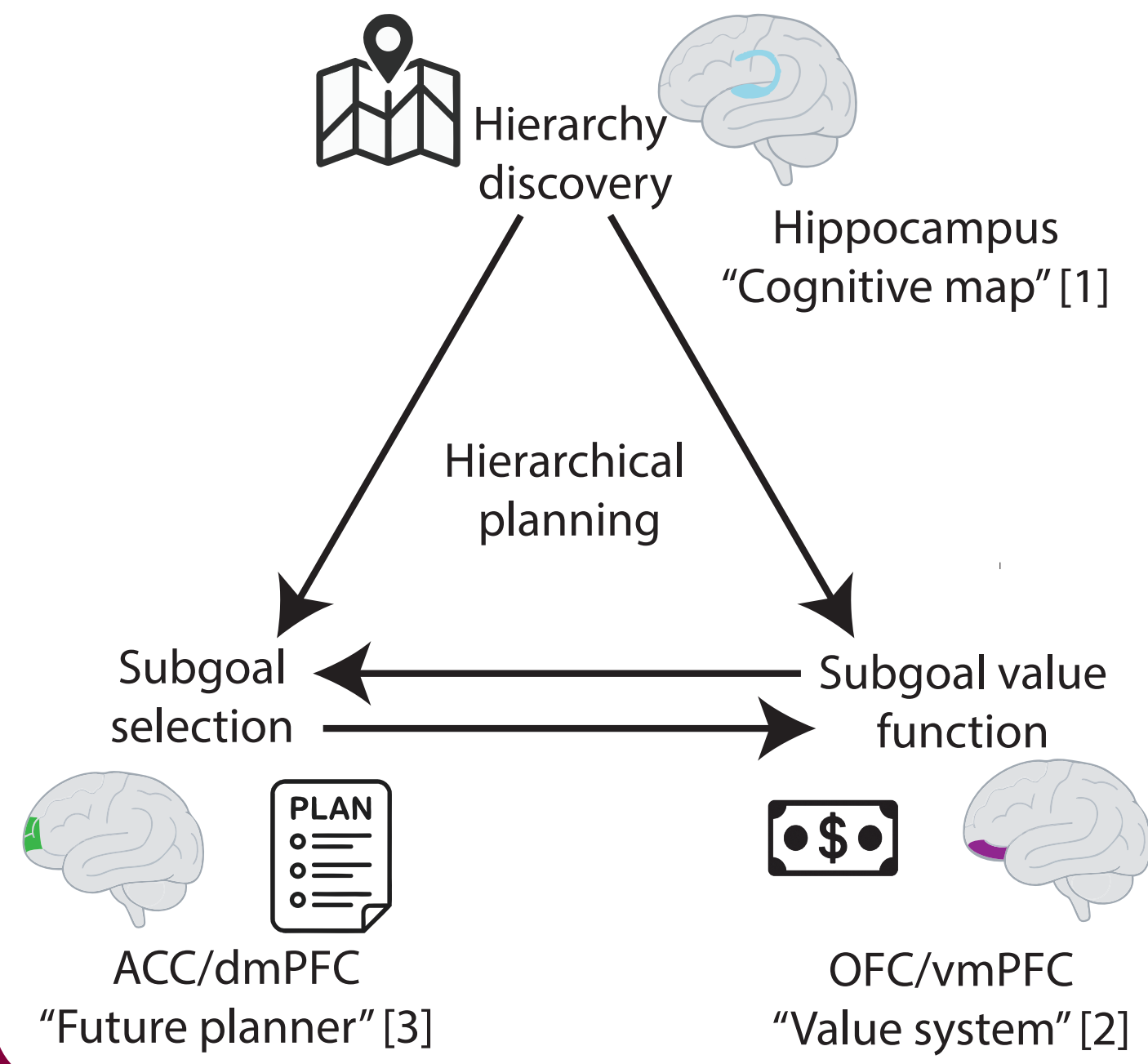


Standard RL learns a value function $V(s) = \gamma^{d(s,g)}$ and chooses actions leading to the highest values.

However, when the goal is far, the value function is close to 0. With neural noise, the decision is unreliable. Hierarchical planning overcomes this by moving value signals to proximal subgoals.



hierarchical planning



Hierarchy discovery via hippocampal replay

Replay compresses behaviors into neural timescales and contains transition information. Clustering replay gives geometry-aware hierarchy.

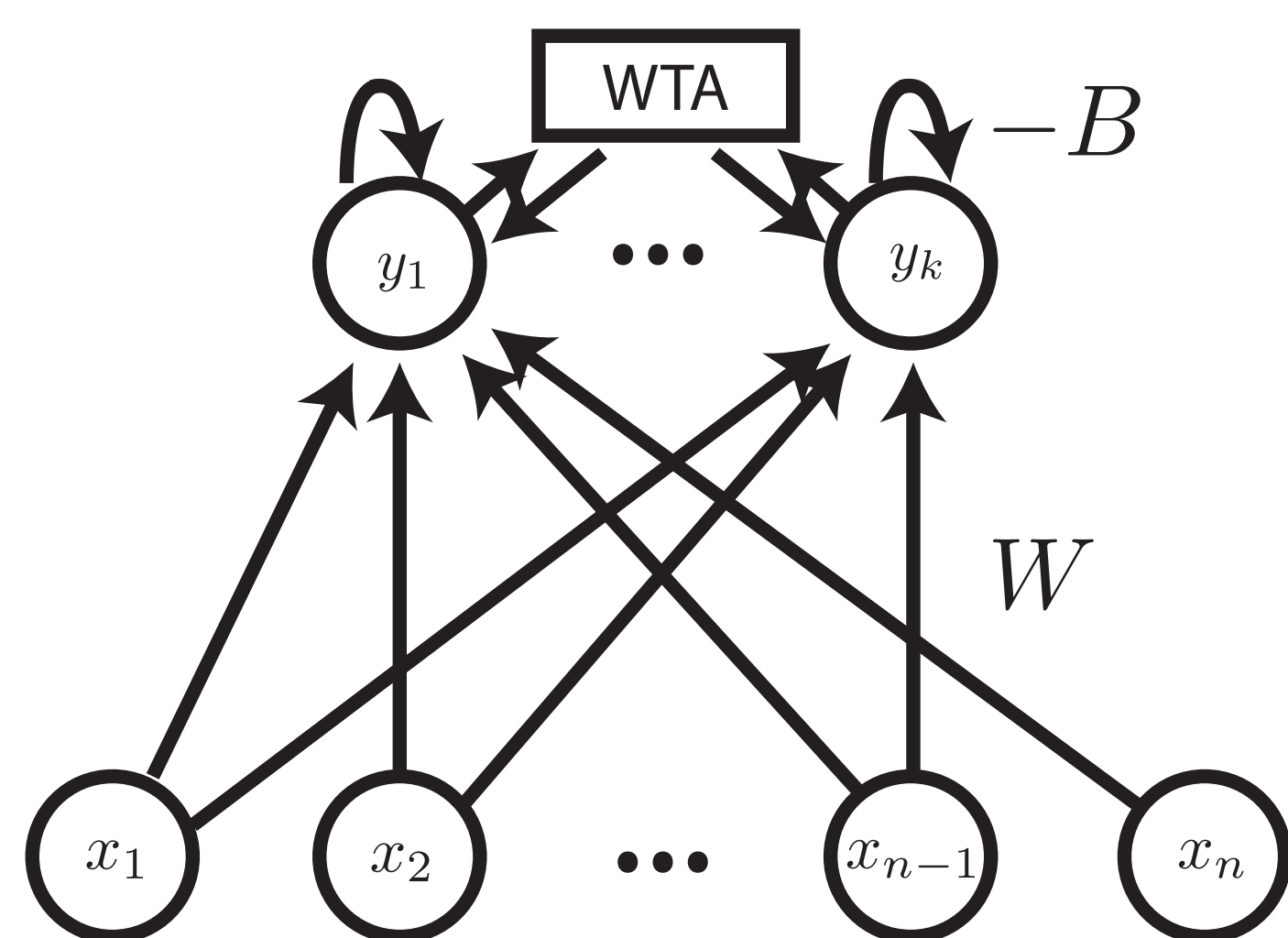
$$s = (x_1, \dots, x_n) \sim \mathcal{D}_{\text{hipp}},$$

$$\phi(s) = \sum_{i=1}^n \gamma^{i-1} x_i.$$

Clustering with k -mean fails to group states by common cues since it does not care about the shape of the clusters.

Therefore, we optimize the sum of within-cluster principal values:

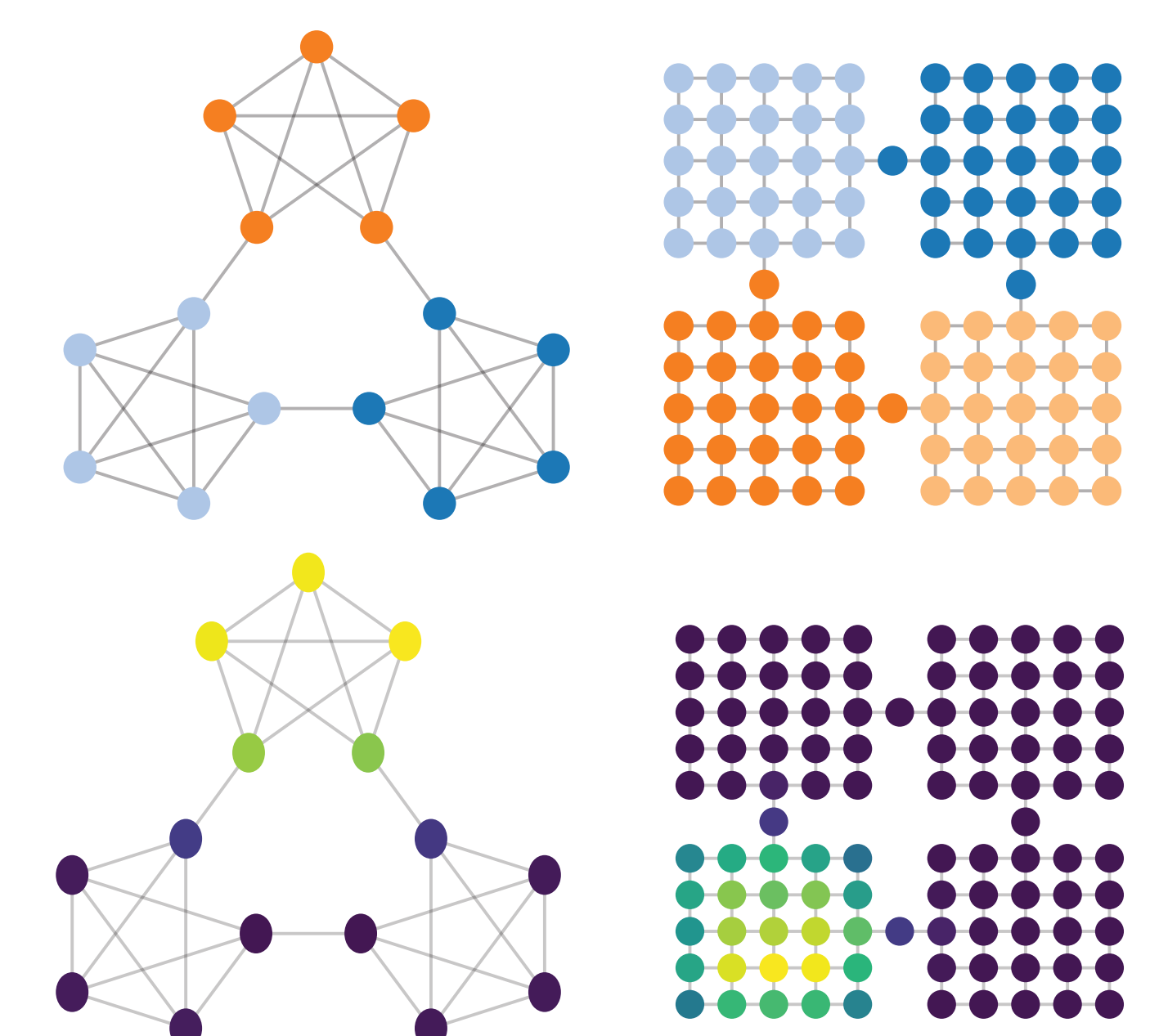
$$\max_{\mathcal{C}_k} \sum_{k=1}^K \lambda_1(\mathbb{E}[xx^\top \mathbf{1}_{x \in \mathcal{C}_k}]).$$



$$\tau_n \dot{r} = Wx - br, \quad y = \text{WTA}(r),$$

$$\tau_w \dot{W} = yx^\top - W, \quad \tau_w \dot{b}_i = y_i^2 - 1,$$

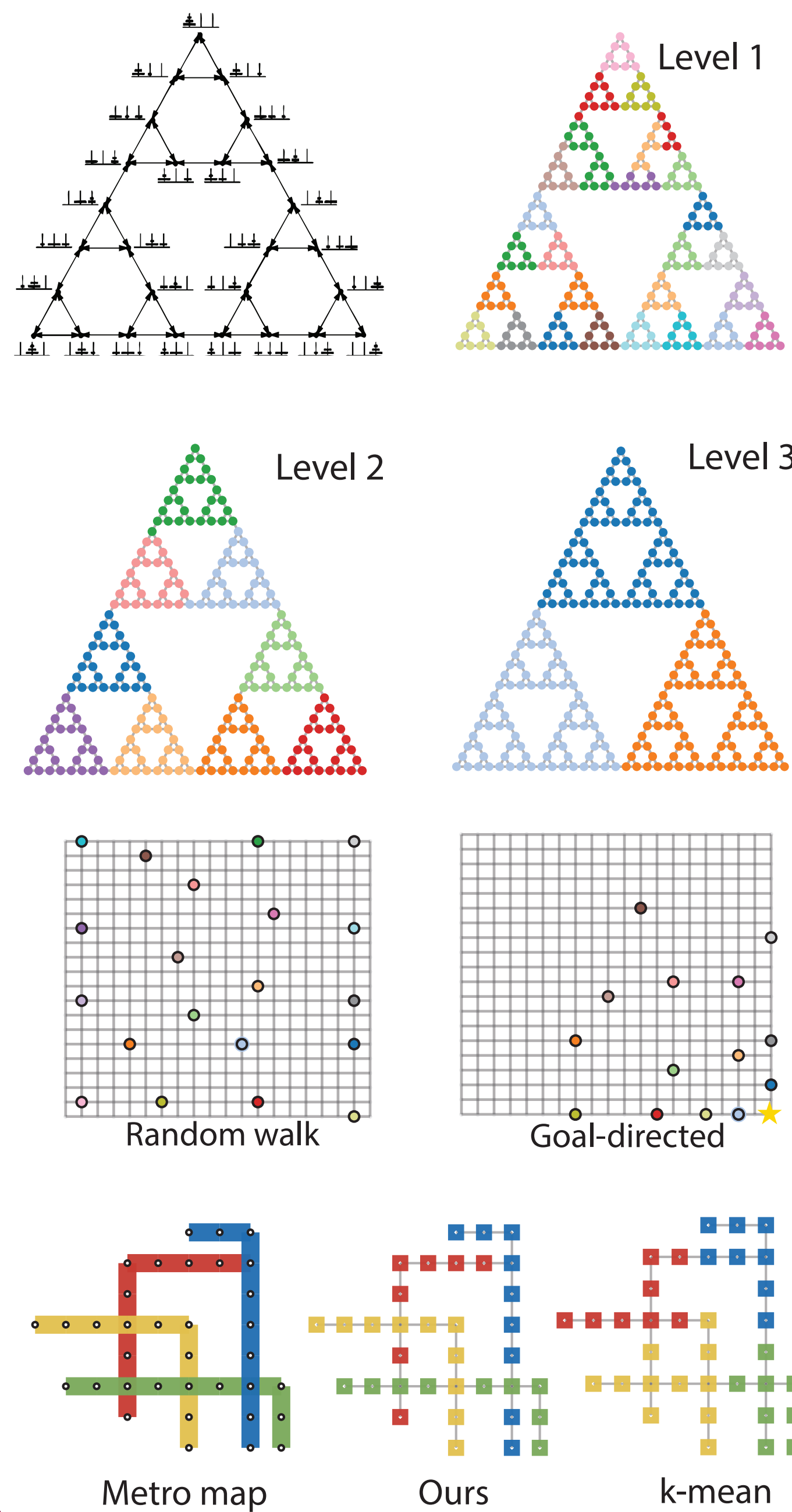
optimizes this objective.



Desired properties

Needs to reflect:

- Geometry
- Task demands
- Common cues

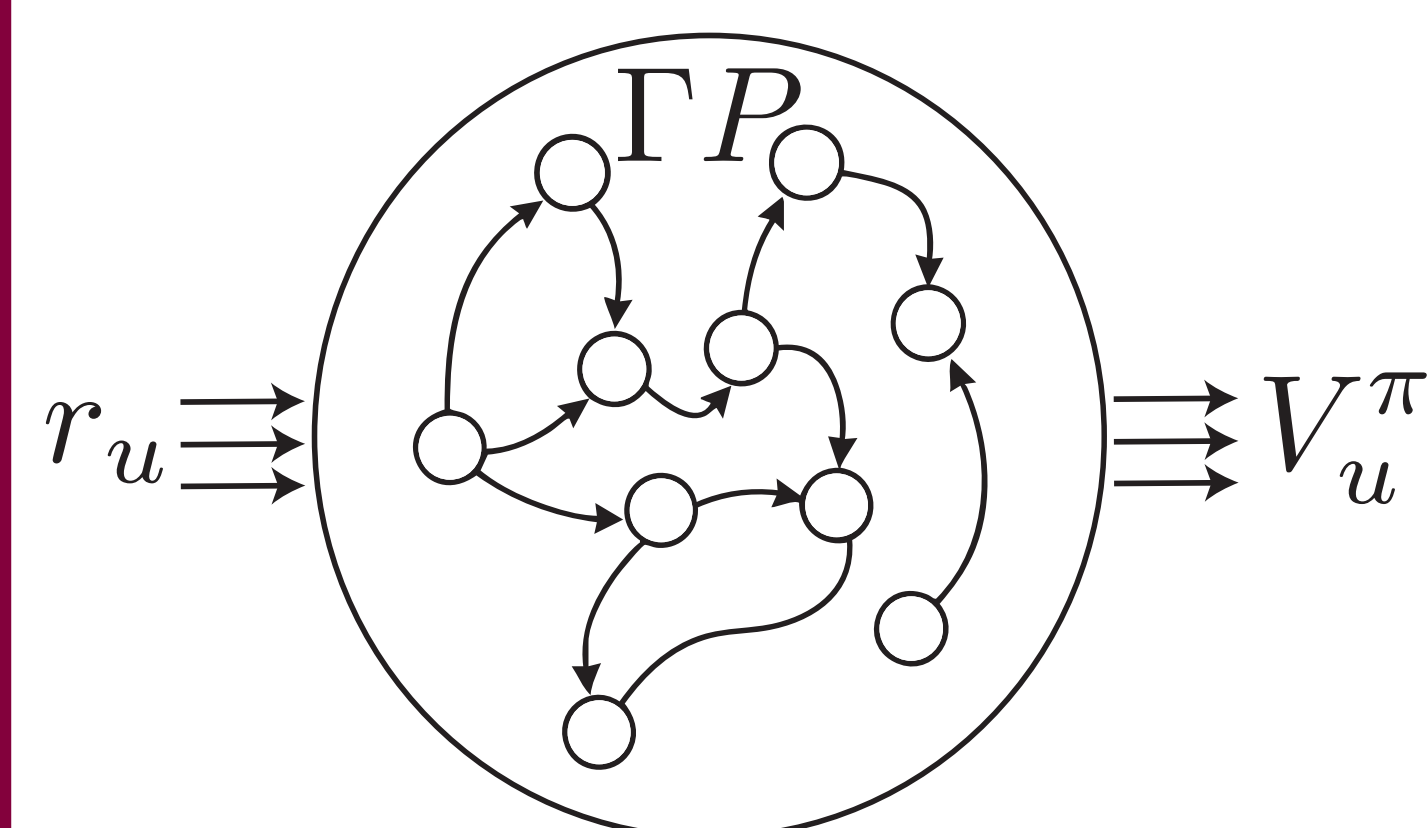


OFC subgoal values via hierarchical SR

In hierarchical planning, the circuit needs to recompute subgoal values rapidly because subgoals can change frequently. Successor representation (SR) does exactly this and can be implemented in a recurrent circuit.

$$\tau_n \dot{x} = -x + Wx + r_u, \quad W = \Gamma P,$$

$$x^* = (I - \Gamma P)^{-1} r_u = V_u^\pi.$$



Hierarchical SR allows OFC to recompute subgoal values quickly when the subgoal changes without relearning the environments.

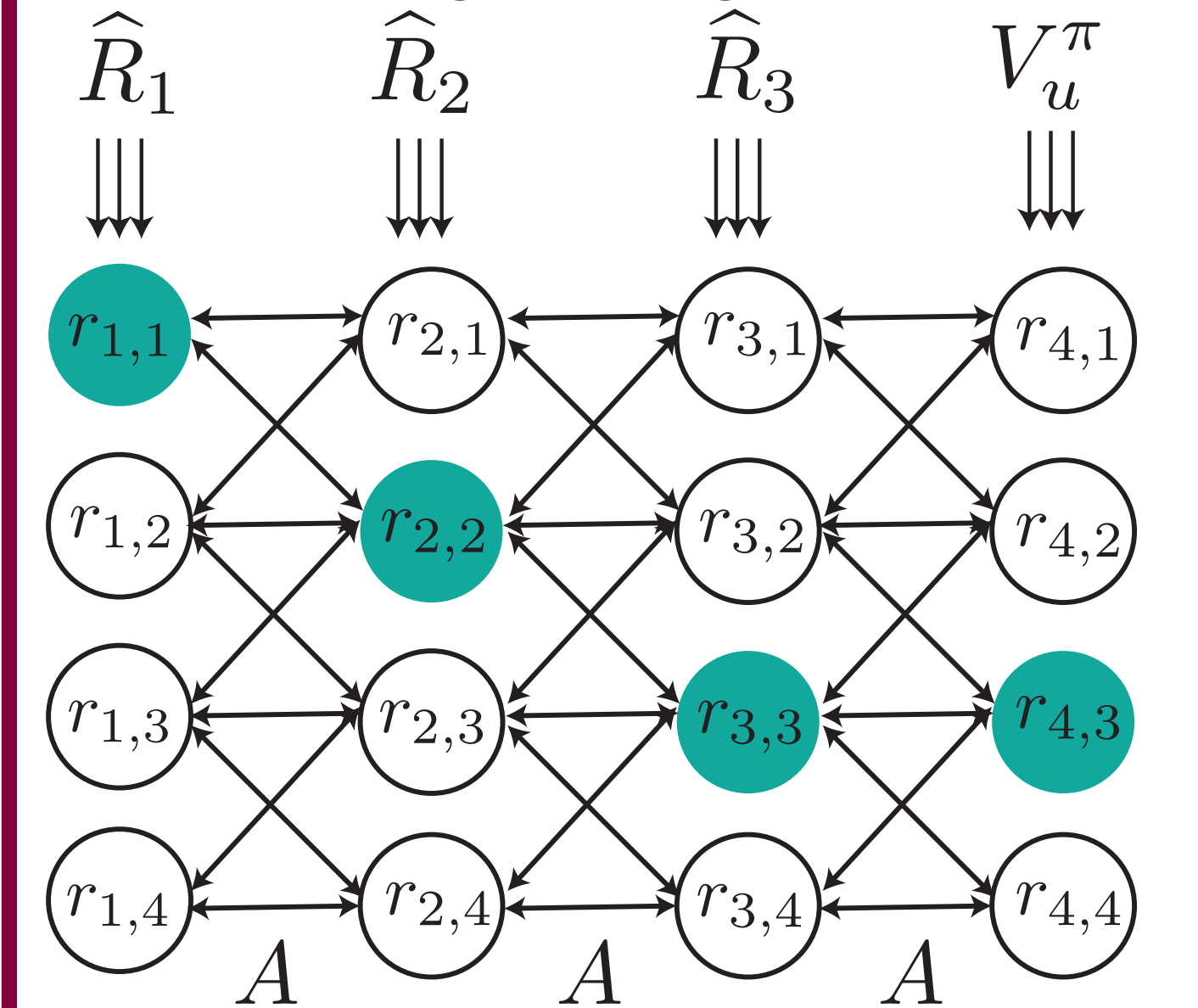
Hebbian learning can learn the effective transition matrix.

$$e_i \leftarrow \gamma e_i + \mathbf{1}[c_t = i],$$

$$\Delta W_{ij} \propto e_j \mathbf{1}[c_{t+1} = i] - W_{ij}.$$

ACC plans via path-level attractors

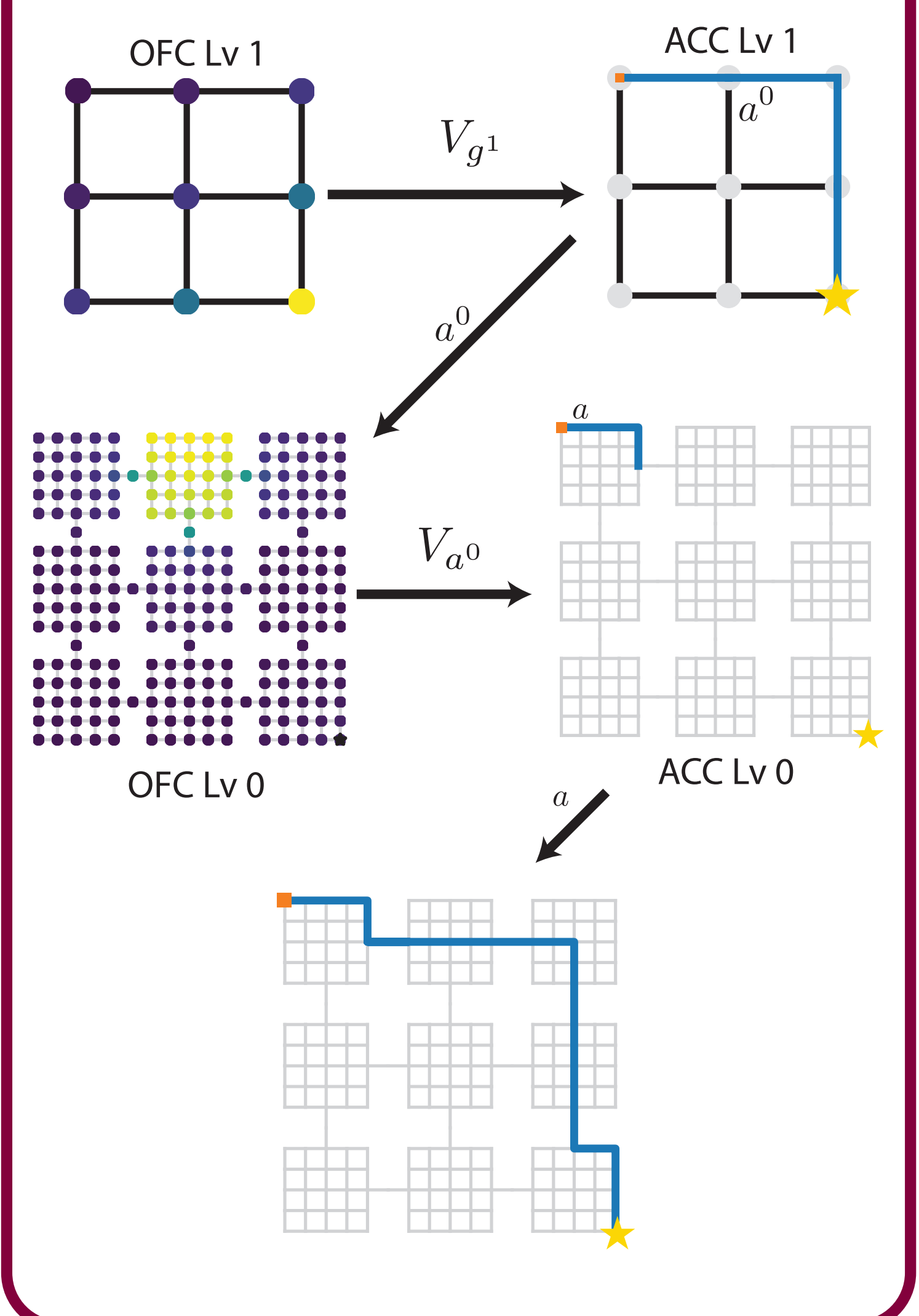
Given the subgoal value functions from OFC, ACC performs path-level winner-take-all to select option sequences maximizing subgoal values.



The attractor computes the marginals of a Gibbs distribution over path values.

The selected high-level action then projects to lower-level OFC as a subgoal reward.

Hierarchical planning via OFC-ACC circuits



Summary

- We propose a multi-area circuit that both discovers hierarchy and plans over it.
- Derive a plasticity rule to optimize a normative clustering objective over replay.
- OFC computes subgoal values via hierarchical SR.
- ACC performs path-level WTA to select plans maximizing OFC subgoal values.
- Build a neural theory for hierarchical model-based RL.

REFERENCES CONTACTS

- [1] Eliav et al., 2021 Email:
[2] Grossman et al., brabeeba@vt.edu 2026 Open to collaboration
[3] Akam et al., 2021 and discussion!