A Multi-Region Brain Model to Elucidate the Role of Hippocampus in Spatially Embedded Decision-Making

Eva Yi Xie[1,2], Jaedong Hwang[1], Carlos Brody[2,3], David Tank[2], Ila Fiete[1]

1 MIT

2 Princeton Neuroscience Institute

3 Howard Hughes Medical Institute

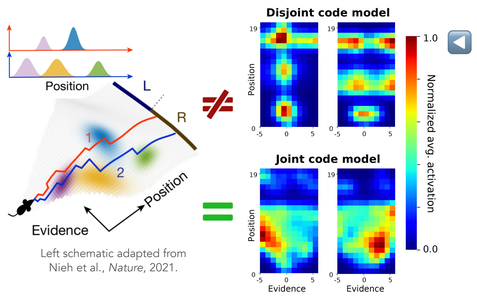
Project Page

Introduction & Motivations

1. How do **multiple** brain regions coordinate to support cognition?
2. Why does the **hippocampus** form a **map** of both position (physical variable) and evidence (cognitive variable) in spatially embedded decision-making tasks (Nieh et al.)?
3. Can **brain-like inductive bias** help machine learning?

We built a "virtual brain" inside a reinforcement-learning (RL) agent, with variants spanning all possible counterfactuals of how the multi-region (entorhinal-hippocampal-cortical) decision-making circuit **wires** and **communicates** internally.

We Use Tower Task: A Neuroscience Benchmark for Studying Spatially Embedded Decision-Making



Left schematic adapted from Nieh et al., *Nature*, 2021.

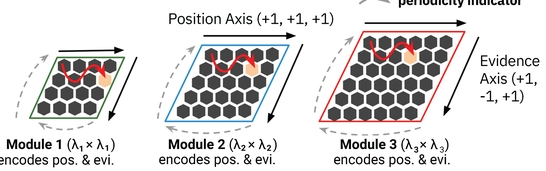
Fig 1. The agent learns to turn to the side with more towers.

Nieh et al. found that the hippocampal activities form a conjunctive map of both position and evidence, reproduced by our joint code model M5.

Rethinking the Neural Code of Grid Cells

Example: An agent moves 3 positions forward, encounters evi. value +1, -1, +1.

Joint grid representation:



Disjoint grid representation:

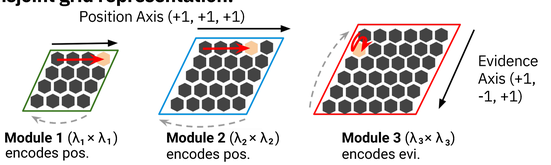
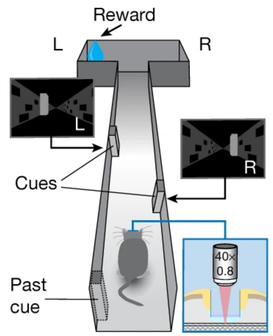
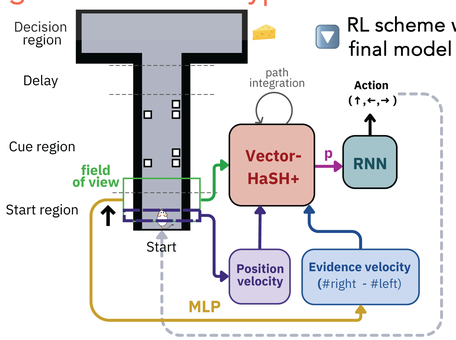


Fig 2. While we simulate the traditional setting where grid cells **only encode position** (M1, M2), we also test what happens when grid cells **encode both** position and evidence *jointly* (M3, M5) or *disjointly* (M4).

Alternative Multi-Region Interaction Hypotheses





RL scheme with final model M5

The hippocampus-grid cell relation is pre-defined & fixed. The grid module states are shifted based on velocities *(just position: M1-M2; both position and evidence: M3-M5)*. Sensory-hippocampal connections are bidirectionally modifiable with Hebbian rules (Chandra et al.). Prefrontal cortex (RNN) maps the hippocampal vector to actions. Numbers correspond to the order of computation.

Model	Grid cell code	Place cell code	MLP input	RNN input
M0	-	-	-	s
M0+	-	-	s	s & v_{pos} & v_{evi}
M1	pos.	g	-	p
M2	pos.	g & s	-	p
M3	joint pos. & evi.	g	s	p
M4	disjoint pos. & evi.	g & s	s	p
M5*	joint pos. & evi.	g & s	s	p

Table 1. Summary of the neural coding and information flow in each model variant. Our final model, M5, is marked with ∗.

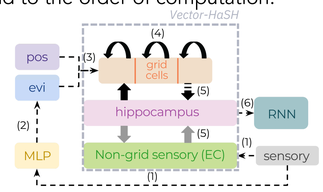


Fig 3. Final model M5

Joint Integration Model Induces Efficient Learning

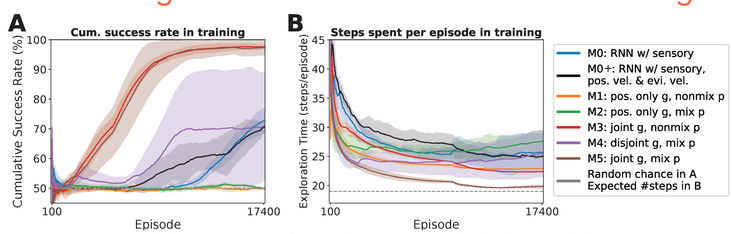


Fig 4. Joint integration model (M5) **(A)** performs the tower task the best, while **(B)** navigates the fastest.

Only Joint Grid Models Reproduce Experimental Hippocampal Maps

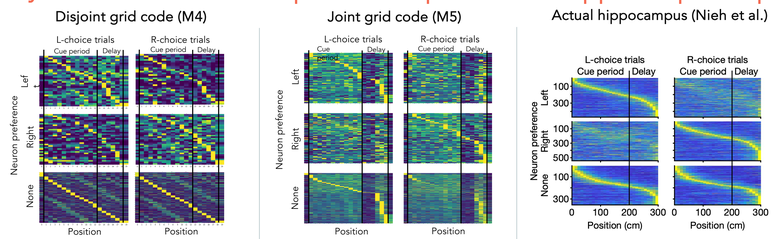
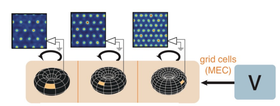


Fig 5. Only joint grid code models (M3, M5) exhibit **choice-specific** HPC neurons

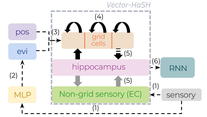
Multi-Region Interactions (M5)

Models of hypotheses (Source of Evidence)	Not grid cells	grid cells
Not sensory	M1	M3 Final model
sensory	M2	M4, M5*

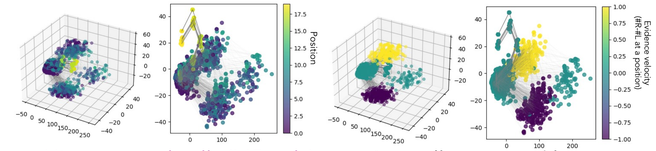
Table 2. Why does the hippocampus form a conjunctive map of position and evidence? Overview of how model variants correspond to alternative hypotheses of neural coding and information flow based on the evidence source. Our final model M5 is marked with ∗.



Grid cells jointly encode velocity of evidence and position.



Some regions (MLP) predict evidence velocity from sensory.



Hippocampal cells (M5 only) contains well-separated low-dimensional representation of positional & evidence velocity (local #right - #left towers). This seems to be a consequence of projection from the joint grid code & sensory input.

Takeaways/Predictions

1. Our results predict that grid cell firing combines position (physical variable) with evidence (cognitive variable) in spatially embedded decision-making, just like the hippocampus. Our collaborators are directly testing this falsifiable hypothesis in experiments.

   → Neural algorithms that support spatial navigation may be *repurposed* for abstract cognitive functions

2. Adding structured memory maps (grid cells) to RL can potentially reduce training demands in agents.

References

1. Nieh, E.H., *et al.* Geometry of abstract learned knowledge in the hippocampus. *Nature* (2021).
2. Chandra, S., *et al.* Episodic and associative memory from spatial scaffolds in the hippocampus. *Nature* (2025).

Twitter/X: @EvaYiXie; evayixie@princeton.edu