Rapid mapping of abstract domains through extraction and projection of generalized velocity signals via a cognitive foundation model with grid cells

Abhiram Iyer (abiyer@mit.edu)
MIT
Cambridge, MA

Sarthak Chandra (sarthakc@mit.edu)
MIT
Cambridge, MA

Sugandha Sharma (susharma@mit.edu)
MIT
Cambridge, MA

Ila Fiete (fiete@mit.edu)
MIT
Cambridge, MA

Abstract

Grid cells in the medial entorhinal cortex create remarkable spatial maps during navigation, but recent studies show that they also extend to mapping and organizing abstract cognitive spaces. Examples of abstract environments include images with deformable features, like a cartoon bird with stretching legs and neck, or auditory inputs varying in frequency and amplitude. While it is understood how grid cells map physical spaces using velocity estimates, how they map abstract cognitive spaces remains unknown. We hypothesize that the brain maps abstract spaces by extracting low-dimensional velocity signals using the path integration capability of grid cells, which are then error-corrected by the same circuit. We propose the first model neural circuit that explains how grid cells can represent any abstract space. The model processes abstract, time-varying inputs across modalities and identifies minimal velocity representations to capture state transition dynamics. It enforces a self-supervised geometric consistency constraint where movements in closed loops produce velocity estimates summing to zero, a computation itself performed by the grid cell circuit. Our model suggests a way for grid cells to use velocity signals to map high-dimensional abstract environments, explaining how animals perceive velocities in diverse non-spatial contexts and encode cognitive spaces.

Keywords: self-supervised learning; deep learning; grid cells; path integration

Introduction

Grid cells in the medial entorhinal cortex are known for their role in spatial navigation and representation. However, experimental work has shown evidence for grid cells representing abstract, non-spatial domains as well, including visual spaces, auditory spaces, conceptual spaces and several others (Park, Miller, Nili, Ranganath, & Boorman, 2020; Park, Miller, & Boorman, 2021; Nau, Navarro Schröder, Bellmund, & Doeller, 2018; Bao et al., 2019; Julian, Keinath, Frazzetta, & Epstein, 2018; Viganò, Rubino, Di Soccio, Buiatti, & Piazza, 2021; Killian, Jutras, & Buffalo, 2012; Constantinescu et al., 2016). How can these same cells, seemingly built for organizing physical space, also be used to organize abstract cognitive environments? Theoretical models have demonstrated that grid cells are capable of representing high-dimensional Euclidean spaces (Klukas, Lewis, & Fiete, 2020), and can thus be used to represent non-spatial domains. Practically however, if grid cells are to map out these abstract domains, they must receive a low-dimensional velocity that characterizes changes in the input space. Here, we build a computational model that extracts such velocity signals from temporal variations in abstract input spaces, thereby enabling grid cells to build maps of these domains. Collectively, this work explores how the brain integrates complex information types of distinct modalities into cognitive maps and presents
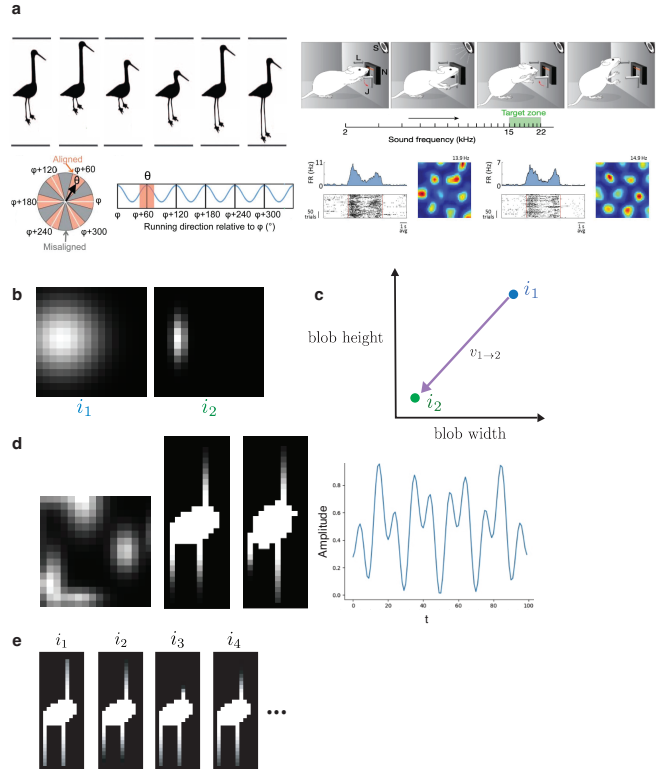

Figure 1: a. "Stretchy bird" and sound manipulation tasks (Constantinescu et al., 2016; Aronov et al., 2017). Experimental reviews observe grid tuning while operating through these non-spatial domains. b. An example of a similar non-spatial domain – a "blob" placed randomly in the visual field that can deform along the x-axis and/or y-axis. c. The difference between states in b. is a difference vectory describing velocity in the abstract blob width-height space. d. Procedurally generated cognitive environments (from left to right): Gaussian-blobs that all shift in space, 2D stretchy-bird whose neck and legs stretch/shrink, 3D stretchy-bird whose legs also deform independently, and 1D sound modulation task where a sum of sine waves all change in frequency uniformly. e. A segment of a random trajectory in 2D stretchy-bird space.

a model explaining how grid cells represent abstract cognitive domains, extending their functions beyond spatial navigation.

Methods

Problem Setup

To answer this question, we generate analogous environments in both visual and auditory modalities as seen in Fig. 1. The Gaussian "blobs" environment shows a group of randomly generated 2D Gaussians that each uniformly translate in 2D space corresponding to a given velocity. A 2D stretchy-bird environment shows a bird whose neck and legs stretch according to a two-dimensional velocity describing the extent of the stretch. A 3D stretchy-bird space follows similarly except each leg also stretches independently. The 1D sound modu-

lation task shows a sum of sine waves that uniformly change in frequency. Random trajectories in any of these "abstract spaces" are also procedurally generated.
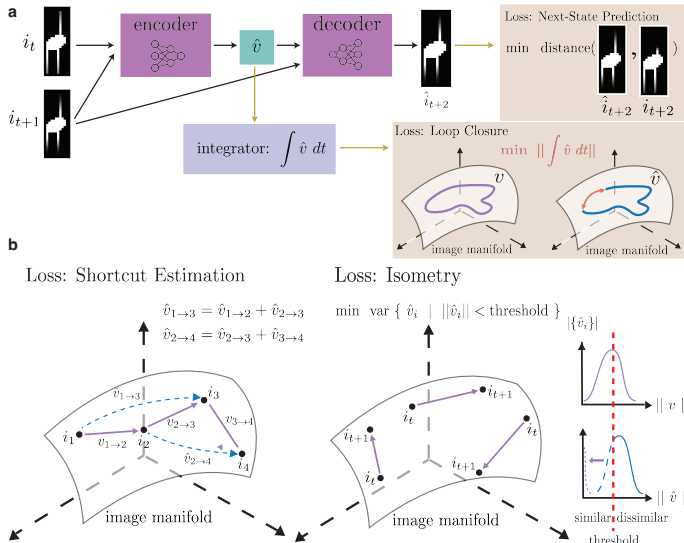

Figure 2: a. Our encoder and decoder model neural network with two critical loss terms: next-step prediction and "loop closure". b. Losses like shortcut estimation and isometry further refine the predicted latent space.

Model & Learning Dynamics

Our model consists of an encoder neural network which processes pairs of states of any generated trajectory and projects these states onto a low-dimensional latent space (i.e., 1-3 dimensions) describing the underlying velocity between the states. A decoder neural network uses this latent representation along with the input to predict the next state in the trajectory, as seen in Fig. 2a. The model is trained end-to-end via backpropagation with a series of self-supervised losses. We note that the model never has access to the true velocity distribution for any task.

The losses that we consider (visualized in Fig. 2a) consist of two crucial loss terms (image reconstruction and "loop closure") that enable the model to successfully learn, and additional auxiliary loss terms (visualized in Fig. 2b) that refine the learned solution ("shortcut estimation" and isometry). The image reconstruction loss minimizes the discrepancy between the predicted and true next states along a trajectory. The loop closure loss maintains geometric consistency by ensuring that the sum of predicted velocities for a trajectory starting and ending at the same point equals zero, facilitated by using grid cells' path integration to compute this sum, with training data arranged into closed loops for ease. The shortcut estimation loss ensures that the direct predicted velocity between two inputs matches the cumulative velocities through an intermediate input. Finally, the isometry loss reduces the variance in magnitudes of predicted velocities below a threshold, ensuring similarly small velocities have comparable sizes.
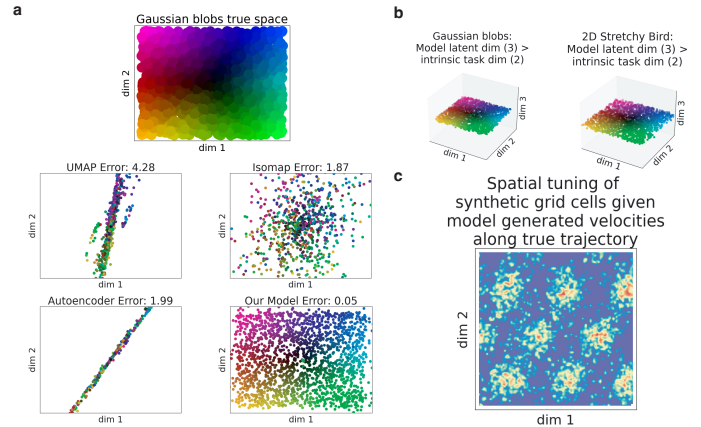
Results & Discussion

Figure 3: a. True velocity space of the shifting Gaussian-blobs task colored by x and y components, followed by estimations of this space with UMAP, Isomap, deep autoencoder, and our model. b. Our model's latent space converges to a minimally low-dimensional representation when the model latent dimensionality is higher than the intrinsic task dimensionality. c. Using our model's predicted velocities in a synthetic grid cell network shows gridlike firing patterns.

Our model finds a low-dimensional representation of input data by representing velocities between successive inputs, similar to dimensionality reduction methods. We compare our model to UMAP, Isomap, and a deep autoencoder, using a normalized error score to evaluate performance (see Fig. 3a). Our model significantly outperforms baseline models in accurately estimating velocities, particularly in the Gaussian-blobs task, by using grid cells themselves for error and loss computation. Additionally, our model effectively identifies a minimally low-dimensional representation of the velocity space even when the model is trained with a higher dimensional latent space (see Fig. 3b). Fig. 3c confirms the accuracy of our model when its outputs serve as inputs to a grid cell network, producing velocities that enable grid-like spatial tuning for a cell in a module. These findings strongly support the utility of grid cells in mapping abstract spaces given accurate velocity representations in these domains.

Conclusion

The model provides a compelling hypothesis for how animals extract velocities in diverse non-spatial contexts and how grid cells encode such abstract spaces. We show that the use of these velocities by grid cells depends on their own computational output, allowing them to map onto any space for path integration. Our work also suggests that a fixed integrator circuit can navigate between spaces of various modalities using common velocity representations, as long as these spaces share a common manifold on which velocities are defined.

Acknowledgments

References

Aronov, D., Nevers, R., & Tank, D. W. (2017). Mapping of a non-spatial dimension by the hippocampal-entorhinal circuit. *Nature*, *543*(7647), 719–722.

Bao, X., Gjorgieva, E., Shanahan, L. K., Howard, J. D., Kahnt, T., & Gottfried, J. A. (2019). Grid-like neural representations support olfactory navigation of a two-dimensional odor space. *Neuron*, *102*(5), 1066–1075.

Constantinescu, A. O., O'Reilly, J. X., & Behrens, T. E. J. (2016, June). Organizing conceptual knowledge in humans with a gridlike code. *Science*, *352*(6292), 1464–1468.

Julian, J. B., Keinath, A. T., Frazzetta, G., & Epstein, R. A. (2018). Human entorhinal cortex represents visual space using a boundary-anchored grid. *Nature neuroscience*, *21*(2), 191–194.

Killian, N. J., Jutras, M. J., & Buffalo, E. A. (2012, November). A map of visual space in the primate entorhinal cortex. *Nature*, *491*(7426), 761–764.

Klukas, M., Lewis, M., & Fiete, I. (2020). Efficient and flexible representation of higher-dimensional cognitive variables with grid cells. *PLoS computational biology*, *16*(4), e1007796.

Nau, M., Navarro Schröder, T., Bellmund, J. L., & Doeller, C. F. (2018). Hexadirectional coding of visual space in human entorhinal cortex. *Nature neuroscience*, *21*(2), 188–190.

Park, S. A., Miller, D. S., & Boorman, E. D. (2021). Inferences on a multidimensional social hierarchy use a grid-like code. *Nature neuroscience*, *24*(9), 1292–1301.

Park, S. A., Miller, D. S., Nili, H., Ranganath, C., & Boorman, E. D. (2020, July). Map making: Constructing, combining, and inferring on abstract cognitive maps. *Neuron*.

Viganò, S., Rubino, V., Di Soccio, A., Buiatti, M., & Piazza, M. (2021). Grid-like and distance codes for representing word meaning in the human brain. *NeuroImage*, *232*, 117876.