

Flexible navigation with neuromodulated cognitive maps

Krubeal Danieli

1 Introduction

During navigation, animals dynamically create rich representations of the environment, forming personalized cognitive maps. The hippocampal area CA1 features spatial cells that adapt based on behavior and internal states. Computational models have usually obtained spatial tuning by training a deep recurrent network for solving navigation tasks such as path integration [1, 2, 3], lasting multiple numerous epochs and using backpropagation. However, these training methods do not closely align with real-time local learning paradigms used by animals.

In this study, it is introduced a rate model that generates place cells in one-shot as the agent navigates the environment by simply assigning the current spatial observation to a selected neuron while ensuring a sparse representation (*i.e.* spaced place fields).

An important ingredient for the learning dynamics of our model is neuromodulation. Neuromodulation is an important ingredient for biological neuronal dynamics, with different molecules covering a wide range of functions. Previous models [4, 5] inspired by experimental results crafted a simple spiking plasticity rule for reward-directed navigation where acetylcholine mediates explorative behaviour and dopamine reinforces memory of reward locations. Other approached using deep artificial networks have applied neuromodulation in conjunction with other training practices, such as dropout probability [6]. In this work, modulators are described as synaptic resources that are consumed by plasticity events, and their dynamics are modelled as leaky integrators. Further, acetylcholine is used to mediate the generation of new place fields, while dopamine mediates the slow remapping of the place centers in conjunction with a reward signal. The concentration of acetylcholine is affected by the presence of active neurons or by the occurrence of a weight update. Dopamine, on the other hand, is influenced by the presence of a reward.

This model successfully creates a representation of visited areas and recurrent connections are defined among similarly tuned cells. Importantly, plasticity hyper-parameters such as the equilibrium concentration and decay time-constant of modulators influence the density of place cells, impacting the encoding of behaviorally relevant information [7].

This network is then used to solve a goal-directed navigation task, where the agent is trained to reach a target location. The agent is equipped with a policy that modulates the exploration behaviour and the decision-making process.

2 Methods

The model is constructed around the concept of a cognitive map, which an agent builds by freely navigating a closed environment and reaching a discovered goal location. The architecture relies on the core assumption that the agent has receives minimal external information, consisting solely of a reward and collision input as two binary values. These two signals are used to enrich the cognitive map with experience-dependent data, which is then used to guide the agent's behavior.

The formation of the spatial representation is instead based on idiothetic information, which is the agent's perception of self-motion. In particular, here we assume this cue to be the factual velocity vector, namely the actual displacement of the agent in the environment. In the brain, this signal is thought to result from the integration of inertial and relative motion cues.

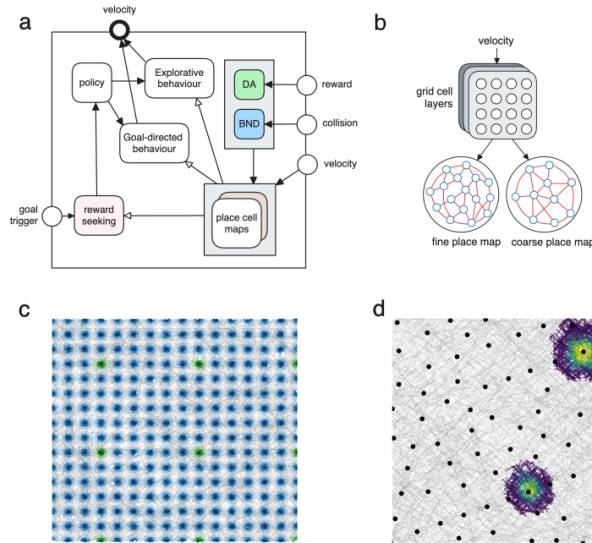


Figure 1: MODEL LAYOUT AND SPATIAL REPRESENTATIONS - **a**: the full architecture of the model, consisting of three main sensory input, targeting the two modulators and the cognitive map module, and the executive components, represented by a policy module, two behavioural programs and a reward receiver. **b**: the cognitive map component, organized with a stack of grid cell modules receiving the velocity input and projecting to two layer of place cells with different place field granularity. **c**: the neural activity of a grid cell module from a random trajectory; in blue the repeating activity of all cell, while in green the activity of only one, highlighting the periodicity in space. **d**: the distribution in space of the place cells centers, together with the activity of two cells showing the size of their place field

2.0.1 Place cell map

The formation of place cells is obtained from the activity of a set of grid cells organized into modules, or layers. This simple feed-forward architecture is depicted in plot 1-**b**. A grid cell module i has been defined as a set of N^{gc} neurons with gaussian tuning curve evenly distributed over the surface of a two dimensional torus \mathbf{T}^2 . When the agent moves in the environment, a two dimensional Euclidean space \mathbf{R}^2 , with a velocity $\mathbf{v} = \{x, y\}$, its position on the torus is updated by the same vector but scaled by a speed scalar s_i^{gc} local to the grid cell module i , which determines its periodicity in space. The initial position on the torus is randomly chosen at the beginning of each episode, since what matters is the sequence of displacements without any reference to a meaningful origin.

The choice of a toroidal space is motivated by consolidated experimental evidence of the neural space of grid cells, which are organized in modules of different size spanning the animal’s environment. However, the shape of their firing pattern is known to be hexagonal, which corresponds to the optimal tiling of a two dimensional plane, giving rise to a neural space lying on a twisted torus. In this work, for simplicity, we consider a square tiling and thus a square torus, without much loss of generality except for the slight increase of grid cells required for a sufficiently cover. In plot 1-**c** is shown the activity of a grid cell module over a trajectory, with the periodicity underlined by the cell in green.

The activity of all grid cell modules, indicated as \mathbf{u}^{GC} , is then projected down to two independent layers of initially un-tuned cells, whose feed-forward weights $\mathbf{W}^{\text{GC,PC1}}$, $\mathbf{W}^{\text{GC,PC2}}$ are initialized at zero. As the agent moves and the grid cells activity changes, if no neurons within a place cell layer are active, then one is randomly chosen and its weights are set to the current (at time t) grid cells’ population vector $\mathbf{W}_i^{\text{GC,PC}} \leftarrow \mathbf{u}_t^{\text{GC}}$. For the plasticity process to be completed, it is also checked the possible overlap with other cells in the same layer, effectively accounting for lateral inhibition. This mechanism is implemented by computing the cosine similarity with the weight vector of the other tuned cells and comparing it with a threshold $\theta_{\text{rep}}^{\text{PC}}$, with the possibility of aborting the plasticity process if the similarity is too high.

The activity of a tuned place cell i is given, again, by the cosine similarity between the current grid cells’ population vector and the weight vector of the cell:

$$\mathbf{u}_i^{\text{PC}} = \phi \left(\cos \left(\mathbf{u}^{\text{GC}}, \mathbf{W}_i^{\text{GC,PC}} \right) \right) \quad (1)$$

where ϕ is a generalized sigmoid function $\phi(z) = [1 + \exp(-\beta(z - \alpha))]^{-1}$ with gain β and threshold α . The two layers of place cells differ in the size of their place fields. This feature is affected by the sensitivity of a cell tuning with respect to the grid cell activation, determined by the parameters of the sigmoid, and the strength of the lateral inhibition, determined by the similarity threshold. One layer is set to be more fine-grained, with an overall higher density of place cells over the space, while the other is more coarse-grained, with overall large place field sizes.

In figure 1-**d** are shown the centers of one of the fine-grained layer and the

activity of two cells, with their place field highlighted as an heatmap.

2.0.2 Neuromodulators

References

- [1] Ben Sorscher, Gabriel Mel, Surya Ganguli, and Samuel Ocko. A unified theory for the origin of grid cells through the lens of pattern formation. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.
- [2] Christopher J. Cueva and Xue-Xin Wei. Emergence of grid-like representations by training recurrent neural networks to perform spatial localization, March 2018.
- [3] Andrea Banino, Caswell Barry, Benigno Uria, Charles Blundell, Timothy Lillicrap, Piotr Mirowski, Alexander Pritzel, Martin J. Chadwick, Thomas Degris, Joseph Modayil, Greg Wayne, Hubert Soyer, Fabio Viola, Brian Zhang, Ross Goroshin, Neil Rabinowitz, Razvan Pascanu, Charlie Beattie, Stig Petersen, Amir Sadik, Stephen Gaffney, Helen King, Koray Kavukcuoglu, Demis Hassabis, Raia Hadsell, and Dharshan Kumaran. Vector-based navigation using grid-like representations in artificial agents. *Nature*, 557(7705):429–433, May 2018.
- [4] Zuzanna Brzosko, Wolfram Schultz, and Ole Paulsen. Retroactive modulation of spike timing-dependent plasticity by dopamine. *eLife*, 4:e09685, October 2015.
- [5] Zuzanna Brzosko, Sara Zannone, Wolfram Schultz, Claudia Clopath, and Ole Paulsen. Sequential neuromodulation of Hebbian plasticity offers mechanism for effective reward-based navigation. *eLife*, 6:e27756, 2017.
- [6] Jie Mei, Rouzbeh Meshkinnejad, and Yalda Mohsenzadeh. Effects of neuromodulation-inspired mechanisms on the performance of deep neural networks in a spatial learning task. *iScience*, 26(2):106026, February 2023.
- [7] Katie C. Bittner, Aaron D. Milstein, Christine Grienberger, Sandro Romani, and Jeffrey C. Magee. Behavioral time scale synaptic plasticity underlies CA1 place fields. *Science*, 357(6355):1033–1036, September 2017.