Flexible navigation with neuromodulated cognitive maps

Krubeal Danieli[1] and Mikkel Elle Lepperød[2]

[1]Center for Integrative Neuroplasticity, FYSCELL, University of Oslo, Norway
[2]Simula Research Laboratory, Oslo, Norway

Abstract

something cool

1 Introduction

Agency in a large environment is a challenging task for any organism. Productive exploration and effective goal-directed navigation are essential in this regard, and become even more crucial when target locations are out of sight. In this scenario, often labelled as *wayfinding*, an adequate understanding of one's surroundings is necessary [1, 2].

For this scope, one of the tools evolved by the mammalian brain are neural spatial representations of the environment, enriched by repeated experiences. In the literature, they are often referred to as cognitive maps, which where first introduced by Tolman who noticed the reward-driven navigation ability of rats in a maze [3].

There are multiple approaches that can be used for solving the task of reaching a known target in its environment with a cognitive map, and the choice might depend on the task as well as personal traits, at least in humans [4, 5]. However, which are the ones the brain generally prefers the most is still debated. One of straighforward possibilities is *route learning*, where the gist is to memorized the previous paths as simple action-position pairs. This approach is effective in small environments, but it is impractical for constructing a proper map to exploit more generally, given the ambiguity at combining intersecting paths [6, 7, 8]. In contrast, a more general strategy is to rely on *survey knowledge*, where the subject builds an explicitly general representation of the environment by endowing it with an Euclidean metric, namely a global coordinate system based on distances and angles in which perform vector operations [7, 9]. The resulting map is highly flexible, and allows for the computation of short-cuts and arbitrary routes. However, it lacks convicing experimental evidence, with several studies in humans and mice highlighting how real brain maps often violate the strict algebraic constrains [10, 11, 12, 13]. Another possibility is a so-called *labeled graph* [7, 14, 15], which consists of a topological graph of the environment over locations in space. Importantly, it does not possess a universal metric but instead relies on local labels, built through experiences [16], for supporting information about distances and angles. The lack of a rigid operational structure leads to more tolerance for global spatial incosistencies, while still allowing for affine vector operations, at least locally. Multiple studies have shown support for this type of representation, in which the graph-based map and its labels are generated and successfully exploited for goal-reaching tasks, despite clearly infringing Euclidean geometry [17, 11, 5]. This online formation of position nodes is also aligned with *path integration*, the process of integrating velocity vectors for tracking the trajectory leading home, supported by observations in multiple species [18, 19] and theoretical analysis [20].

Regarding the main neural substrates of the cognitive maps, a variety of neuronal types has been associated to allocentric and egocentric spatial features.

Among others, border cells have being associated to boundary detection, speed cells with the magnitude of the perceived velocity, head-direction cells with the anglular position of the head as per vestibular perception, place cells with uni-modal tuning for spatial positions, and grid cells with an hexagonal periodic tuning [21, 22, 23]. Parts of this neuronal ecosystem has been found in various part of the cortex, but more predominatly in the entorhinal cortex (EC), in particular the medial region (MEC), and sub-regions of the hippocampal formation, particularly the cornus ammonis area CA3 and CA1, and the subiculum (SB). Specifically, place cells in CA1 have been often considered as the component of the cognitive maps [24]. The origin of the spatial tuning of these cells is still debated. Some theories trace it back to competitive dynamics over projections from entorhinal grid cells through the temporo-ammonic pathway, while others point at the Schraffer's collaterals from CA3 [25, 26, 27, 28].

1.1 Neuromodulation of cognitive maps

The upstream afferences to these place cells populations carry a variety of inputs, ranging from spatial information from the medial EC and CA3, to sensory and contextual data from the lateral EC [29]. In addition, they are also targeted by a multiple neuromodulators, molecules gating information and affecting neuronal dynamics in a multitude of ways, the most prominent of which in this context are dopamine and acetilcholine.

There are several set of dopaminergic innervations to the hippocampus, targeting all its main sub-components. The ones more directly involved with CA1 place cells are those from the ventral tegmental area (VTA) and locus coeruleus (LC) [30, 31, 32]. Their actions have been linked to synaptic excitation, modulation of theglutamatergic projections from EC, and regulation of long term potentiation (LTP) through D1-like receptors [33, 34]. One important function of dopamine concerning spatial navigation is the delivery of reward-related information, supporting memory consolidation of salient locations [35, 36] and active reshaping of the place cells tuning [37, 38]. Another extensively documented role is novelty detection [39], especially for contextual inputs from LEC [34, 40, 41], as well as encoding prediction errors of reward events [42, 43]

The cholinergic innervation to the hippocampus is mainly provided by the medial septum (MS). Through the action of different muscarinic recptors, it modulates the afferences to CA1 from MEC and CA3, and promotes both long-term synaptic potentiation and depression [44, 45, 46]. Its role in spatial navigation includes the support of explorative behaviour, contributions to novelty detection, and attentional regulation by increasing the input signal-to-noise ratio [47, 48, 44].

1.2 Computational models

There have been several efforts for devising computational models capturing the construction and functionality of cognitive maps. In relation to purely spatial navigation, a notable modelling framework is the one proposed by [49], in which

the hippocampus was envisoned to encode positions with directional information, while metric properties were brought by the pariental cortex. Later work has extended it into a route-based formalism [8], centered on the idea of a route graph formed by concatenation of episodic routes. Further, these graphs could have different granularity, organized hierarchially in layers. Another direction has seen the use of the successor representation (SR) algorithm, based on predictive modeling, for learning navigation tasks in spatial and semantic spaces, with good performance [50], also in comparison to humans and rats [51]. A similar approach for generating task-dependent is the Tolman-Eichenbaum Machine (TEM), based on artificial neural networks [52], which has been applied to navigation and relational reasoning. Interestingly, some of these artificial neurons resembled the firing pattern of actual hippocampal and entorhinal cells. Such bio-realistic neuronal tunings have also been consistently obtained in recurrent neural networks trained to solve *path integration* problems, using only idiothetic information such as the agent's velocity vector [53, 54, 55]. This line of normative research has been particularly successful in generating a rich ecosystem of spatial cells, such as grid cells and place cells, and reasoning through their functional importance for navigation.

However, what these models have in common is the reliance on backpropagation for training, which is not biologically plausible, and the lack of a proper real-time learning dynamics. Although they rely solely on locally available information during inference, the type of cognitive map they can support is limited to purely spatial data. For more diffult behaviours such as goal-directed navigation, a richer representation of the environment should be more advantageous. Furthermore, the role of neuromodulation in the formation of cognitive maps has been largely overlooked, despite its relevance in actively regulating neuronal dynamics and gating important information.

In this study, we introduced a model architecture capable of generating online a spatial representation made of two place cells layers, which can be endowed with additional behaviourally-relevan sensory inputs by means of neuromodulators. Our primary goal was to successfully construct a bio-realistic cognitive map and the effectivess of neuromodulation in enriching it with sensorial information, and evaluating their impact on the agent's performance in goal-directed navigation. One important modeling assumption was to solely rely on the spatial information available within the local graph, namely without invoking a global coordinate system, justified by the controversial evidence for a proper metric space in brain cognitive maps. This is thus more aligned with the *labelel graph* approach, although metric data is not explicitly attached to the nodes but rather inferred on-the-fly with an implicit and weak embedding into an affine space, similarly to other proposals [56]. Another assumption was that learning occurs entirely online, as animals do, avoiding a costly network training, which is instead popular in standard deep and reinforcement learning as well as in most models applied to *path integration* [53, 54, 55]. This feature is achieved by introducing a structural inductive bias in the form of a pre-defined stack of grid

cells layers with different spatial frequencies, from which multiple layers of place cells are generated through network competition [57, 58]. A map is then obtained by connecting neighboring place cells obtained over several trajectories, thus going beyond route learning [8] and constructing a topological graph. For what concerns neuromodulators, in this work they are best described by nodes working as analog sensors of external events. Their role is to form synaptic connections with the place cells, effectively determing a scalar field over their representation.

A secondary goal was to investigate the possibility of updating the cognitive map to environmental changes, accounting for the adaptation and flexibility in animals behaviour. For this scope, the model was endowed with a way to correct the synaptic weights between neuromodulators and place cells, based on prediction error of future sensory observations. This mechanism is aligned with principles of neural predictive coding [43, 59, 60, 42], according to which prediction-based learning is a relevant part of brain activity. We took inspiration from previous work which has applied neuromodulation in spiking neurons for controlling explorative-exploitative behaviour [61, 62]. Further, it is well established the role of modulators such as dopamine in actively maintaining and updating the state of neural representations, particularly those associated with reward features [63, 64, 65, 66, 51].

A last interest of ours regarded the effect of neuromodulation on the density of place cells, and in particular the reshaping of their place field, with respect to task performance. The motivation stems in the experimental observations of modulation of place cells representation following meaningful events, in terms of remapping, fields resizing and dislocation on the map [37, 67, 68].

The rest of the paper is organized as follows. In Section 2, we describe the model architecture and the task. In Section 3, we present the results of the simulations. In Section 4, we discuss the implications of the results, suggest future directions for research, and make conclusive observations.

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 full schema of its components its illustrated in plot 1-**a** below. 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.
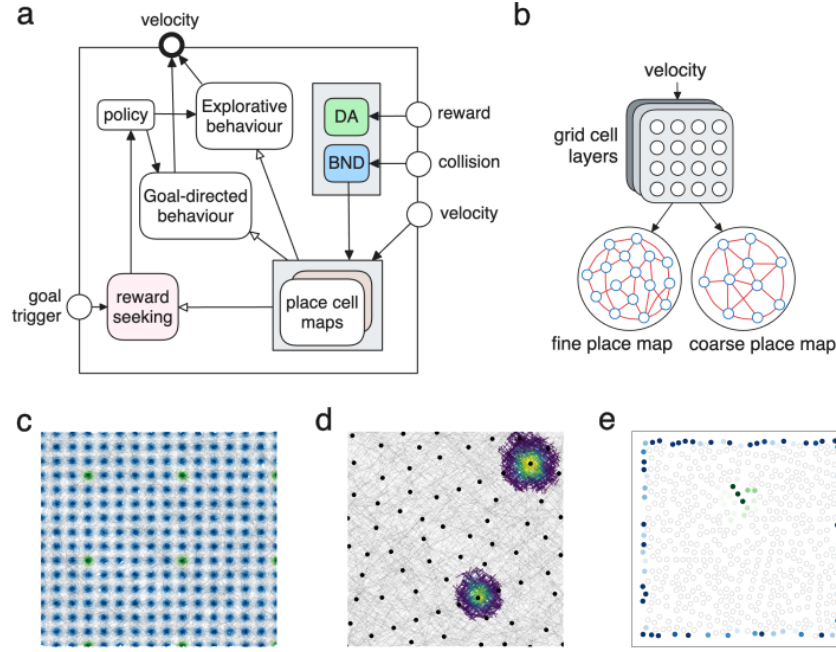
5

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.* **e**: *neuromodulation activity over the place cells map, with in blue the cells tagged by the collision modulation, and in green the ones targeted by reward modulation.*

The formation of the spatial representation is instead based on idiothetic information, which is the agent's perception of self-motion [69]. 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 [70, 71].

2.1 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 $\mathbf{T}^2$. Unlike other approaches for generating grid fields [72, 73], we defined a correspondance between the global environment in which the agent moves, a two dimensional Euclidean space $\mathbf{R}^2$, and a grid module bounded local space, corresponding to the torus. Then, the global velocity $\mathbf{v} = \{x, y\}$ is then mapped to a local velocity, scaled by a speed scalar s_i^{gc} specific to the grid cell module i, which determines its periodicity in space. This approach has been used in previous works [27]. 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 $\mathbf{W}^{GC,PC1}$; $\mathbf{W}^{GC,PC2}$ 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 $\mathbf{W}_i^{GC,PC} \leftarrow \mathbf{u}_t^{GC}$. 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 θ_{rep}^{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:

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

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. Within a layer, the connections between cells are calculated by the same cosine

similariy, but compared against a different threshold $\theta_{\text{rec}}^{\text{PC}}$. 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.

2.2 Neuromodulators

The fine-grained layer of place cells constitutes the main cognitive map of the agent, since it captures the environment with greater detail. The neuromodulators are operationalized as analog sensors of meaningful environmental events, here reward and collision, and map directly to the place cells through plastic connections $\text{W}^{\text{k,PC}}$. For each neuromodulator k, it is defined a leaky variable v^{k} that accumulates the corresponding signal I over time, and decays exponentially to zero in the absence of inputs with time constant τ^{k}:

$$\dot{\text{v}}^{\text{k}} = -\text{v}^{\text{k}}/\tau^{\text{k}} + \text{I} \tag{2}$$

This variable is then paired with the activity of each place cell i for updating the synaptic weights in a Hebbian fashion:

$$\Delta \mathbf{W}_{\text{i}}^{\text{k}} = \eta^{\text{k}} \text{v}^{\text{k}} \mathbf{u}_{\text{i}}^{\text{PC}} \tag{3}$$

where η^{k} is the neuromodulator-specific learning rate, and weights are kept ≥ 0.

On the one hand, the reward modulation, signed as $\mathbf{W}^{\text{DA}}$, is sensitive to the instantaneous presence of reward, defined as a boolean value. Over time, its coupling with the population vector $\mathbf{u}_{\text{t}}^{\text{PC}}$ delineates a region of the environment where the reward has been experienced. On the other hand, the collision modulation, referred to as $\mathbf{W}^{\text{BND}}$, signals the occurrence of a collision with a boundary, which is again given as a boolean. After enough events, the profile of the resulting weight matrix with the place cells provides an approximation of the shape of the environment given by its boundaries. From the perspective of the agent, this intuition of the topology of its surroundings is crucial for effectively planning routes to target locations.

At each moment during navigation, the weight matrices $\mathbf{W}^{\text{DA,PC}}$, $\mathbf{W}^{\text{BND,PC}}$ act as scalar fields over the neural space of the place cells, and their simultaneous contributions delineate what in this work is referred to as a cognitive map. In plot 1-**e** is shown the activity of the two neuromodulators over the fine-grained place cells map, showcasing the bounds of the environment and the reward location.

2.2.1 Online adaptation

Lastly, we introduced a simple prediction-base mechanism for supporting the ability to correct the accuracy internal representation. During navigation, before the execution of each movement towards a future location $\mathbf{x}_{\text{t}+1}$, it is computed and cached a prediction of the activity of neuromodulator k: $\hat{\text{v}}^{\text{k}} = \sum_{\text{i}}^{\text{N}} \mathbf{W}_{\text{i}}^{\text{DA,PC}}$.

$\hat{u}_i^{PC}$, where $\hat{u}^{PC}$ is the predicted population activity at $\mathbf{x}_{t+1}$. Then, a prediction error with a specific learning rate is calculated and added to the weight update $\Delta \mathbf{W}^k = \Delta \mathbf{W}^k + \hat{\eta}^k (\hat{v}^k - v^k \mathbf{u}^{PC})$. Further, since the rate $\hat{\eta}^k$ is less than 1, it takes several erroneous prediction to effectively set connections to zero.

The rationale is to depress those synapses no longer reliably predicting the sensory experience referenced by the neuromodulator. This simple rule aligns with predictive coding [51], and it is inspired by previous computational work on the role of dopaminergic and cholinergic activity in regulating exploration-exploitation behaviours through synaptic plasticity [74, 62].

2.3   Policy and behavior

In this work, the first interest was to test the usefulness of our simple cognitive map built from minimal assumptions for the tasks of exploration and goal-directed navigation. To this end, we defined a simple hard-coded policy that toggles between these two behaviours according to the presence of a goal signal, externally provided, and the presence of an actual goal representation, taken care of by a special component called *reward seeking*, depicted in the pink box of plot 1-**a**. Exploration is accomplished by a random walk, with a variable number of steps in the same direction to avoid stagnation, and occasional plans to visit random positions within the known map, again for limiting stagnation. Goal-directed navigation, either for reaching a random position for exploration or the actual reward location, is achieved by calculating the shortest path between the closest place cells centers of the current and target positions. The place cells are hence treated as nodes of a graph, and their connections constitute its edges. We use a Dijkstra algorithm applied to the coarse-grained layer, which contains less and more spread out cells and it is thus cheaper to compute, to derive a coarse-grained plan. However, in the case the agent gets stuck or the distance to the target is shorter than the cells' distance, the planning switches to the other layer for devising a fine-grained plan, which it is followed until either the target or the next node in the coarse-grained are reached.

The advantage of this dual-layer planning lies in its flexibility, as it lightens the computational load of planning by exploring the sparser and lighter map and only invokes the detailed one when necessary. In the process of behaviour, learning does not occur explicitly, but it is instead accounted for in the online formation of the cognitive map.

3   Results

3.1   Naturalistic task

The evaluation of the model ability to construct and utilize a cognitive map was defined as the total rewards collected during several rewards trials in different environments. The protocol was inspired by the behaviour of animals who are end up in a new territory in search for food in various locations.

The testing environments varied in layout, determined by the position and number of walls.

First there was the exploration phase, in which the agent was placed in a random location and let roam through a pseudo-random walk for 5000 steps. Depending on environment layout, the formed map was more or less matching the the actual topology. Then the exploitation phase began, in which a circle the 5% of the total are was designated to provide a binary reward $R \sim \mathcal{B}(p_r)$ drawn from a Bernoully with probability p_r. When the reward was successfully fetched, the agent was teleported to a random location in the environment. Further, in order to mimic the depletion of a food source the reward area was modeled with as leaky variable $\dot{v}_r = (E_r - v_r)/\tau_r + R$, and its location moved whenever its level went below a threshold $v_r < \theta_r$.

The optimization of the parameters of the model was carried out using the Covariance-Matrix Adaptation evolutionary strategy (CMA-ES) [75] with a population of 256 individuals for 100 generations.

3.2 Performance in wayfinding

Our primary aim was to evaluate the formation of the cognitive map through neuromodulation in terms of performance of goal-navigation in different environments. The best model resulting from evolution reached solid navigation and adaptation skills. The agent was able to visit a significant portion of the environment during exploration, and use neuromodulation to produce useful spatial representations.

In figure 2, three emerging cognitive maps are shown different environments. The left panel of plot 2-**a** displays fine-grained place cells associated with collisions and reward events, signaling boundaries (in blue) and reward (in green) locations. The overlapping of these two representations and the coarse-grained place cells (in pink) is what we refer to as a cognitive map, since these are the main source of spatial and contextual information used during planned navigation, whose path is depicted as a grey line. The right panel instead portraits the actual enviroment with walls (black), reward location (green), and multiple trajectories (red). During exploration, the main areas were visited, until the reward position was located, and goal-directed navigation dominated, as highlighted by the density of path lines. Considering the position of the walls and corners, the layout of this environment does not make target locations always visible, as it is a non-convex area, and can be thus be classified as wayfinding [76]. The challenge of not being able to use straight lines is overcome by the graph approach using local data and the switching between the differently grained place cells layers, allowing the agent to successfully going around obstructions and avoiding being stuck. Further, the plans were also minimizing path length by construction, within the part of the space covered by the cognitive map.

However, it is worth nothing that not all simulations resulted in reward being found in the first place, due to the randomness of the exploratory process; this was more pronounced in complex enviroment.
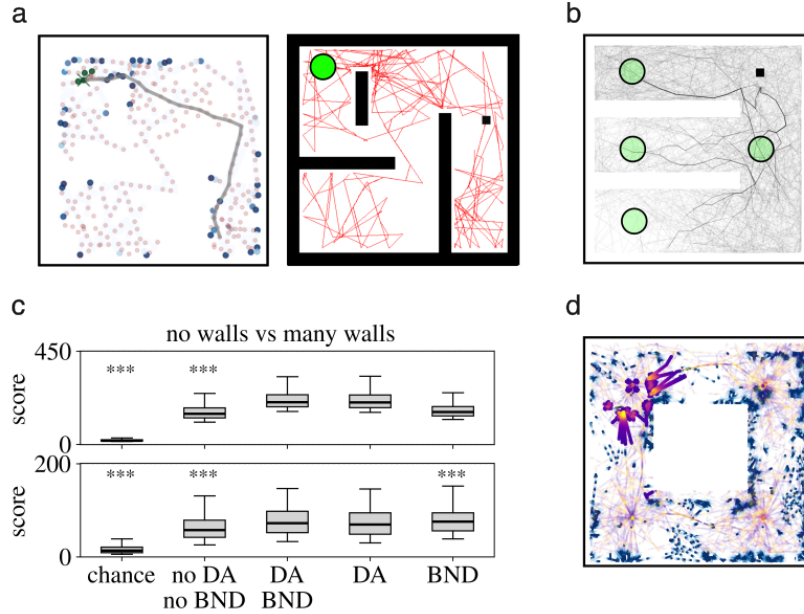
Figure 2: COGNITIVE MAPS AND PERFORMANCE RESULTS - **a**: *the plot on the left represent a cognitive map over a space, together with the plan (grey line) to reach a target location from a starting position. The plot on the right is a view of the same enviroment but with highlighted walls (black thick lines) the reward (green circle), trajectory (red line), agent position (black square). -* **b**: *trajectories for multiple trials with the agent starting at the same position (black square) but with the reward location (green circles) periodically moving -* **c**: *performance results over a common environment settigs for five different models: chance (plasticity disabled) being the lowest all all, no modulation of place fields for DA and BND being meaningfully lower than any variant with modulation of either kind enabled. -* **d**: *cognitive map but in terms of the activity of the place cells over multiple trajectories with a fixed reward location. Further, it is marked with high opacity the activity of the DA-modulated coarse-grained place cells, with their place fields enlarged near the reward.*

3.3 Adaptive goal representation through prediction error

Then, we tested the adaptability to environmental changes. In this scenario, the reward object was moved after it was fetched a fixed amount of times. Here, the difficulty lied in the unlearning of past locations and the discovery of the new ones, in a protocol similar to [74]. In plot 2-**b** it is reported the set of trajectories over many trials with the reward displaced in three possible locations. The agent was capable of planning behaviour, as earlier, but also to explore and find the new rewards, as shown by the density of lines. Whenever a goal path resulted in failed prediction, the DA-based sensory error weakened the association between

the place cells and reward signal, leading to an extinction of its representation at that location.

3.4 Modulation of spatial resolution affects performance

Lastly, we investigated the effect of modulating the place cells density and field size. The goal position was fixed, but the agent was randomly relocated after fetching; performance was defined as total number of reward count within a time window. Our working hypothesis is these experience-driven neuronal changes would improve the quality of the cognitive map, and be reflected in the navigation abilities. The assessment of this claim was conducted by comparing variants of the model, obtained by progressively ablating the reward (DA) and collision (BND) modulation of density and field size, but a cognitive map was still possible as their representations of goals and boundaries were preserved. We also defined a chance level by instead blocking all modulation-based plasticity and allowing only place cells formation. All models were ran in two different environments differing by number of internal walls, in total 2048 simulations were done for each case. However, due to the issue of the reward not being found, only the top 1024 were considered; the number of faulty runs affected all models in equal measure, since it depended solely on exploration noise.

The results are shown in plot 2-**c**. The top reports the scores in the setting without internal walls. All models performed above chance, but the main finding is confirmation of the importance of DA modulation, supported by the statistical difference between those endowed with it and not. Indeed, in a convex region such as this one, once the reward has been located boundary information has a limited utility.

A different situation was when more walls are present, as revealed by the results at the bottom. In this scenario, the boundary modulation turned into the statistically more dominant, although than reward one still remains better than no modulation at all. In this environment, scores are lower as navigation became more difficult, and adequately tuning place cells fields and density improves wall representation.

4 Discussion

Exploration and planning in known and past environment are essential behaviours of animals, directly affecting their success in world udnerstanding and goal reaching.

An important element behind these abilities is the formation of a map of their surroundings as they make new experiences, known as a cognitive map. The hippocampus, and especially CA1, is one of the principal brain areas involved in this processes [24]. Numerous speculations have been made about the shape and neural foundations of such object, varying in the type of modeling assumptions and experimental support. In terms of the map structure, an important aspect is the rigidity of the metric constraints, which concerns the type of operations

that can be applied to the underlying spatial representations [9, 7, 11]. Regarding its formation, computational approaches relying on training artificial neural networks have been used for reproducing experimental observations, such as path integration [53, 55, 54, 52] and navigation [49, 51]. Another important ingredient for neuronal dynamics is neuromodulation. Its implications for spatial cognition have been in part associated to the role of dopamine, especially in the context of plasticity, predictive learning, and reward representation [35, 39, 63].

The objective of the present work was to devise a network model, inspired by the CA1 region, embedded into an agent exploring and reaching target position online in different enviornments. We used grid cells, together with synaptic plasticity for developing context-rich maps based on place cells, and updating them through experience. In the spirit of minimizing the geometric assumptions on the neural space, we treated it as a topological graph with information added locally through the action neuromodulators. This aligned with the concept of a *labeled graph* [16, 11]. For the practical utilization of the map, graph based algorithms were used that rely on local spatial measures.

References

[1] Reginald Golledge, Dan Jacobson, Rob Kitchin, and Mark Blades. Cognitive Maps, Spatial Abilities, and Human Wayfinding. *GEOGRAPHICAL REVIEW OF JAPAN SERIES B*, 73:93–104, December 2000.

[2] Russell A. Epstein and Lindsay K. Vass. Neural systems for landmark-based wayfinding in humans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1635):20120533, February 2014.

[3] Edward C Tolman. COGNITIVE MAPS IN RATS AND MEN. *Psychological Review*, 55(4), July 1948.

[4] Steven M. Weisberg and Nora S. Newcombe. Cognitive Maps: Some People Make Them, Some People Struggle. *Current Directions in Psychological Science*, 27(4):220–226, August 2018.

[5] Victor R. Schinazi, Daniele Nardi, Nora S. Newcombe, Thomas F. Shipley, and Russell A. Epstein. Hippocampal size predicts rapid learning of a cognitive map in humans. *Hippocampus*, 23(6):515–528, 2013.

[6] Michael Peer, Iva K. Brunec, Nora S. Newcombe, and Russell A. Epstein. Structuring Knowledge with Cognitive Maps and Cognitive Graphs. *Trends in cognitive sciences*, 25(1):37–54, January 2021.

[7] Elizabeth R. Chrastil and William H. Warren. From Cognitive Maps to Cognitive Graphs. *PLoS ONE*, 9(11):e112544, November 2014.

[8] Steffen Werner, Bernd Krieg-Brückner, and Theo Herrmann. Modelling Navigational Knowledge by Route Graphs. In Christian Freksa, Christopher Habel, Wilfried Brauer, and Karl F. Wender, editors, *Spatial Cognition II: Integrating Abstract Theories, Empirical Studies, Formal Methods, and Practical Applications*, pages 295–316. Springer, Berlin, Heidelberg, 2000.

[9] C. R. Gallistel and Audrey E. Cramer. Computations on Metric Maps in Mammals: Getting Oriented and Choosing a Multi-Destination Route. *Journal of Experimental Biology*, 199(1):211–217, January 1996.

[10] Michael Peer, Catherine Nadar, and Russell A. Epstein. The format of the cognitive map depends on the structure of the environment. *Journal of Experimental Psychology: General*, 153(1):224–240, January 2024.

[11] William H. Warren. Non-Euclidean navigation. *Journal of Experimental Biology*, 222(Suppl_1):jeb187971, February 2019.

[12] Mark Wagner. Comparing the psychophysical and geometric characteristics of spatial perception and cognitive maps. *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society*, 15(1):6–21, 2008.

[13] Rainer Rothkegel, Karl F. Wender, and Sabine Schumacher. Judging Spatial Relations from Memory. In Christian Freksa, Christopher Habel, and Karl F. Wender, editors, *Spatial Cognition: An Interdisciplinary Approach to Representing and Processing Spatial Knowledge*, pages 79–105. Springer, Berlin, Heidelberg, 1998.

[14] Tobias Meilinger. The Network of Reference Frames Theory: A Synthesis of Graphs and Cognitive Maps. In Christian Freksa, Nora S. Newcombe, Peter Gärdenfors, and Stefan Wölfl, editors, *Spatial Cognition VI. Learning, Reasoning, and Talking about Space*, pages 344–360, Berlin, Heidelberg, 2008. Springer.

[15] Jane X. Wang, Zeb Kurth-Nelson, Dhruva Tirumala, Hubert Soyer, Joel Z. Leibo, Remi Munos, Charles Blundell, Dharshan Kumaran, and Matt Botvinick. Learning to reinforcement learn, January 2017.

[16] Toru Ishikawa and Daniel R. Montello. Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places. *Cognitive Psychology*, 52(2):93–129, March 2006.

[17] R. W. Byrne. Memory for Urban Geography. *Quarterly Journal of Experimental Psychology*, 31(1):147–154, February 1979.

[18] Rüdiger Wehner, Barbara Michel, and Per Antonsen. Visual Navigation in Insects: Coupling of Egocentric and Geocentric Information. *Journal of Experimental Biology*, 199(1):129–140, January 1996.

[19] Sabine Gillner and Hanspeter A. Mallot. Navigation and Acquisition of Spatial Knowledge in a Virtual Maze. *Journal of Cognitive Neuroscience*, 10(4):445–463, July 1998.

[20] Júlia V. Gallinaro, Benjamin Scholl, and Claudia Clopath. Synaptic weights that correlate with presynaptic selectivity increase decoding performance. *PLOS Computational Biology*, 19(8):e1011362, August 2023.

[21] Francesca Sargolini, Marianne Fyhn, Torkel Hafting, Bruce L. McNaughton, Menno P. Witter, May-Britt Moser, and Edvard I. Moser. Conjunctive Representation of Position, Direction, and Velocity in Entorhinal Cortex. *Science*, 312(5774):758–762, May 2006.

[22] Emilio Kropff, James E. Carmichael, May-Britt Moser, and Edvard I. Moser. Speed cells in the medial entorhinal cortex. *Nature*, 523(7561):419–424, July 2015.

[23] Trygve Solstad, Edvard I. Moser, and Gaute T. Einevoll. From grid cells to place cells: A mathematical model. *Hippocampus*, 16(12):1026–1031, 2006.

[24] Flavio Donato, Anja Xu Schwartzlose, and Renan Augusto Viana Mendes. How Do You Build a Cognitive Map? The Development of Circuits and Computations for the Representation of Space in the Brain. *Annual Review of Neuroscience*, 46(Volume 46, 2023):281–299, July 2023.

[25] Daniel Bush, Caswell Barry, and Neil Burgess. What do grid cells contribute to place cell firing? *Trends in Neurosciences*, 37(3):136–145, March 2014.

[26] Torsten Neher, Amir Hossein Azizi, and Sen Cheng. From grid cells to place cells with realistic field sizes. *PLOS ONE*, 12(7):e0181618, July 2017.

[27] Tianyi Li, Angelo Arleo, and Denis Sheynikhovich. *Modeling Place Cells and Grid Cells in Multi-Compartment Environments: Hippocampal-Entorhinal Loop as a Multisensory Integration Circuit*. April 2019.

[28] John L. Kubie and Steven E. Fox. Do the spatial frequencies of grid cells mold the firing fields of place cells? *Proceedings of the National Academy of Sciences*, 112(13):3860–3861, March 2015.

[29] Olesia M. Bilash, Spyridon Chavlis, Cara D. Johnson, Panayiota Poirazi, and Jayeeta Basu. Lateral entorhinal cortex inputs modulate hippocampal dendritic excitability by recruiting a local disinhibitory microcircuit. *Cell Reports*, 42(1):111962, January 2023.

[30] John E. Lisman and Anthony A. Grace. The Hippocampal-VTA Loop: Controlling the Entry of Information into Long-Term Memory. *Neuron*, 46(5):703–713, June 2005.

[31] Theodoros Tsetsenis, Julia K. Badyna, Rebecca Li, and John A. Dani. Activation of a Locus Coeruleus to Dorsal Hippocampus Noradrenergic Circuit Facilitates Associative Learning. *Frontiers in Cellular Neuroscience*, 16:887679, April 2022.

[32] Zhewei Zhang, Yuji K. Takahashi, Marlian Montesinos-Cartegena, Thorsten Kahnt, Angela J. Langdon, and Geoffrey Schoenbaum. Expectancy-related changes in firing of dopamine neurons depend on hippocampus. *Nature Communications*, 15(1):8911, October 2024.

[33] Elke Edelmann and Volkmar Lessmann. Dopaminergic innervation and modulation of hippocampal networks. *Cell and Tissue Research*, 373(3):711–727, September 2018.

[34] Akiko Wagatsuma, Teruhiro Okuyama, Chen Sun, Lillian M. Smith, Kuniya Abe, and Susumu Tonegawa. Locus coeruleus input to hippocampal CA3 drives single-trial learning of a novel context. *Proceedings of the National Academy of Sciences*, 115(2):E310–E316, January 2018.

[35] Kimberly A. Kempadoo, Eugene V. Mosharov, Se Joon Choi, David Sulzer, and Eric R. Kandel. Dopamine release from the locus coeruleus to the dorsal hippocampus promotes spatial learning and memory. *Proceedings of the National Academy of Sciences*, 113(51):14835–14840, December 2016.

[36] Aude Retailleau and Thomas Boraud. The Michelin red guide of the brain: Role of dopamine in goal-oriented navigation. *Frontiers in Systems Neuroscience*, 8, March 2014.

[37] 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.

[38] Alexandra Mansell Kaufman, Tristan Geiller, and Attila Losonczy. A Role for the Locus Coeruleus in Hippocampal CA1 Place Cell Reorganization during Spatial Reward Learning. *Neuron*, 105(6):1018–1026.e4, March 2020.

[39] Adrian J. Duszkiewicz, Colin G. McNamara, Tomonori Takeuchi, and Lisa Genzel. Novelty and Dopaminergic Modulation of Memory Persistence: A Tale of Two Systems. *Trends in Neurosciences*, 42(2):102–114, February 2019.

[40] Kei M. Igarashi, Hiroshi T. Ito, Edvard I. Moser, and May-Britt Moser. Functional diversity along the transverse axis of hippocampal area CA1. *FEBS Letters*, 588(15):2470–2476, August 2014.

[41] Hiroshi T. Ito and Erin M. Schuman. Functional division of hippocampal area CA1 via modulatory gating of entorhinal cortical inputs. *Hippocampus*, 22(2):372–387, 2012.

[42] Denis Sheynikhovich, Satoru Otani, Jing Bai, and Angelo Arleo. Long-term memory, synaptic plasticity and dopamine in rodent medial prefrontal cortex: Role in executive functions. *Frontiers in Behavioral Neuroscience*, 16, January 2023.

[43] Wolfram Schultz, Peter Dayan, and P. Read Montague. A Neural Substrate of Prediction and Reward. *Science*, 275(5306):1593–1599, March 1997.

[44] Tanja Fuchsberger and Ole Paulsen. Modulation of hippocampal plasticity in learning and memory. *Current Opinion in Neurobiology*, 75:102558, August 2022.

[45] E. Sugisaki, Y. Fukushima, M. Tsukada, and T. Aihara. Cholinergic modulation on spike timing-dependent plasticity in hippocampal CA1 network. *Neuroscience*, 192:91–101, September 2011.

[46] Siobhan H. Dennis, Francesca Pasqui, Ellen M. Colvin, Helen Sanger, Adrian J. Mogg, Christian C. Felder, Lisa M. Broad, Steve M. Fitzjohn,

John T.R. Isaac, and Jack R. Mellor. Activation of Muscarinic M1 Acetylcholine Receptors Induces Long-Term Potentiation in the Hippocampus. *Cerebral Cortex*, 26(1):414–426, January 2016.

[47] Michael E Hasselmo. The role of acetylcholine in learning and memory. *Current Opinion in Neurobiology*, 16(6):710–715, December 2006.

[48] Jon Palacios-Filardo, Matt Udakis, Giles A. Brown, Benjamin G. Tehan, Miles S. Congreve, Pradeep J. Nathan, Alastair J. H. Brown, and Jack R. Mellor. Acetylcholine prioritises direct synaptic inputs from entorhinal cortex to CA1 by differential modulation of feedforward inhibitory circuits. *Nature Communications*, 12(1):5475, September 2021.

[49] Bruno Poucet. Spatial cognitive maps in animals: New hypotheses on their structure and neural mechanisms. *Psychological Review*, 100(2):163–182, 1993.

[50] Paul Stoewer, Achim Schilling, Andreas Maier, and Patrick Krauss. Neural network based formation of cognitive maps of semantic spaces and the putative emergence of abstract concepts. *Scientific Reports*, 13(1):3644, March 2023.

[51] William de Cothi, Nils Nyberg, Eva-Maria Griesbauer, Carole Ghanamé, Fiona Zisch, Julie M. Lefort, Lydia Fletcher, Coco Newton, Sophie Renaudineau, Daniel Bendor, Roddy Grieves, Éléonore Duvelle, Caswell Barry, and Hugo J. Spiers. Predictive maps in rats and humans for spatial navigation. *Current Biology*, 32(17):3676–3689.e5, September 2022.

[52] James C. R. Whittington, Timothy H. Muller, Shirley Mark, Guifen Chen, Caswell Barry, Neil Burgess, and Timothy E. J. Behrens. The Tolman-Eichenbaum Machine: Unifying Space and Relational Memory through Generalization in the Hippocampal Formation. *Cell*, 183(5):1249–1263.e23, November 2020.

[53] 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.

[54] 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.

[55] Christopher J. Cueva and Xue-Xin Wei. Emergence of grid-like representations by training recurrent neural networks to perform spatial localization, March 2018.

[56] Tristan Baumann and Hanspeter A. Mallot. Metric information in cognitive maps: Euclidean embedding of non-Euclidean environments. *PLOS Computational Biology*, 19(12):e1011748, December 2023.

[57] Charlotte B. Alme, Chenglin Miao, Karel Jezek, Alessandro Treves, Edvard I. Moser, and May-Britt Moser. Place cells in the hippocampus: Eleven maps for eleven rooms. *Proceedings of the National Academy of Sciences*, 111(52):18428–18435, December 2014.

[58] Jake Ormond and Bruce L. McNaughton. Place field expansion after focal MEC inactivations is consistent with loss of Fourier components and path integrator gain reduction. *Proceedings of the National Academy of Sciences of the United States of America*, 112(13):4116–4121, March 2015.

[59] Abdullahi Ali, Nasir Ahmad, Elgar de Groot, Marcel A. J. van Gerven, and Tim C. Kietzmann. Predictive coding is a consequence of energy efficiency in recurrent neural networks, November 2021.

[60] Jacopo Bono, Sara Zannone, Victor Pedrosa, and Claudia Clopath. Learning predictive cognitive maps with spiking neurons during behavior and replays. *eLife*, 12:e80671, March 2023.

[61] Zuzanna Brzosko, Wolfram Schultz, and Ole Paulsen. Retroactive modulation of spike timing-dependent plasticity by dopamine. *eLife*, 4:e09685, October 2015.

[62] 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.

[63] Wolfram Schultz. Dopamine reward prediction error coding. *Dialogues in Clinical Neuroscience*, 18(1):23–32, March 2016.

[64] Jeffrey B. Inglis, Vivian V. Valentin, and F. Gregory Ashby. Modulation of Dopamine for Adaptive Learning: A Neurocomputational Model. *Computational brain & behavior*, 4(1):34–52, March 2021.

[65] Philippe N. Tobler, Christopher D. Fiorillo, and Wolfram Schultz. Adaptive Coding of Reward Value by Dopamine Neurons. *Science*, 307(5715):1642–1645, March 2005.

[66] Roshan Cools. Chemistry of the Adaptive Mind: Lessons from Dopamine. *Neuron*, 104(1):113–131, October 2019.

[67] Aaron D Milstein, Yiding Li, Katie C Bittner, Christine Grienberger, Ivan Soltesz, Jeffrey C Magee, and Sandro Romani. Bidirectional synaptic plasticity rapidly modifies hippocampal representations. *eLife*, 10:e73046, December 2021.

[68] André A. Fenton. Remapping revisited: How the hippocampus represents different spaces. *Nature Reviews Neuroscience*, 25(6):428–448, June 2024.

[69] Luxin Zhou and Yong Gu. Cortical Mechanisms of Multisensory Linear Self-motion Perception. *Neuroscience Bulletin*, 39(1):125–137, July 2022.

[70] Steven J. Jerjian, Devin R. Harsch, and Christopher R. Fetsch. Self-motion perception and sequential decision-making: Where are we heading? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 378(1886):20220333, August 2023.

[71] Ian Q. Whishaw and Brian L. Brooks. Calibrating space: Exploration is important for allothetic and idiothetic navigation. *Hippocampus*, 9(6):659–667, 1999.

[72] Yuri Dabaghian. Grid Cells, Border Cells and Discrete Complex Analysis.

[73] Vemund Sigmundson Schøyen, Kosio Beshkov, Markus Borud Pettersen, Erik Hermansen, Konstantin Holzhausen, Anders Malthe-Sørenssen, Marianne Fyhn, and Mikkel Elle Lepperød. Hexagons all the way down: Grid cells as a conformal isometric map of space. *PLOS Computational Biology*, 21(2):e1012804, February 2025.

[74] Zuzanna Brzosko, Susanna B. Mierau, and Ole Paulsen. Neuromodulation of Spike-Timing-Dependent Plasticity: Past, Present, and Future. *Neuron*, 103(4):563–581, August 2019.

[75] Christian Igel, Nikolaus Hansen, and Stefan Roth. Covariance Matrix Adaptation for Multi-objective Optimization. *Evolutionary Computation*, 15(1):1–28, March 2007.

[76] Tobias Meilinger, Marianne Strickrodt, and Heinrich H. Bülthoff. Qualitative differences in memory for vista and environmental spaces are caused by opaque borders, not movement or successive presentation. *Cognition*, 155:77–95, October 2016.