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

During navigation, animals dynamically create rich representations of the environment, forming personalized cognitive maps used for exploration and goal planning. The hippocampal area CA1 features spatial cells that adapt based on behavior and internal states. A possible modeling approach is a labeled graphs that, with the intent of avoiding a map with metric structure, relies on nodes enriched with spatial information only specified locally. Another popular direction if training of deep neural networks on spatial tasks, from which record network dynamics from emerging spatially tuned neurons.

In this study, we introduce a place-cells based architecture for developing cognitive maps in one-shot while exploring novel environments. We used a simulated agent for reward-driven navigation tasks, which operates online and forms spatial representations of its surroundings. Further, by means of neuromodulators it incorporates behaviorally relevant information, such as boundaries and reward location. Learning was involved the combination of a rapid Hebbian plasticity, lateral competition, and modulation of the place cells,

The agent proved successful in exploring, retrieving and reaching goal locations in a variety of environments, and displayed adaptability when the reward was moved. Further, the analysis of the neuromodulated place cells showed the importance of dynamically changing neuronal density and tuning field size following relevant events. These results align with experimental evidence of reward effects on hippocamapal spatial cells, and provides additional computational support to the labeled graph approach.

Animals naturally form personalized cognitive maps to support efficient navigation and goal-directed behavior. In the brain, the hippocampal CA1 region plays a key role in this process, hosting spatially tuned neurons that adapt based on behavioral context and internal states. Computational models of this ability include labeled graphs with locally specified spatial information, avoiding global metric structure, and deep neural networks trained on spatial tasks that exhibit emergent spatial tun-

ing. However, these approaches often struggle to model one-shot adaptive mapping.

We propose a biologically inspired place-cell architecture that enables rapid, on-the-fly construction of cognitive maps during exploration of novel environments. Our agent learns to navigate reward-driven tasks using an architecture that integrates neuromodulatory signals responsive to boundaries and rewards, with learning mechanisms that combine Hebbian plasticity, lateral inhibition, and modulatory gating of place-cell activity.

Compared to reinforcement learning (RL) models, our approach demonstrates superior efficiency, accomplishing in a single trial what RL agents typically require thousands of episodes to learn. It also generalizes robustly to novel environments, forming new maps without retraining. Simulations show that the agent adapts to changing reward locations and maintains successful navigation across diverse environments. Analysis of neuromodulated place cells reveals dynamic changes in tuning field size and spatial density after behaviorally salient events, mirroring experimental observations in hippocampal neurons. Our findings support the efficacy of biologically grounded models and locally structured graph representations in cognitive map formation.

1 Introduction

Survival in complex environments requires adequate agency abilities, namely the capacity to explore, learn, and navigate toward goals. This becomes especially challenging when destinations are out of sight, a scenario known as *wayfinding*, where effective behavior relies on an internal model of the world [1, 2].

To support such behavior, the mammalian brain has evolved powerful spatial representations of the environment, refined through repeated experiences. These internal representations—commonly known as *cognitive maps*—were first proposed by Tolman, who observed that rats could navigate mazes in a goal-driven fashion, even when typical cues were removed [3].

Several strategies have been proposed to explain how agents reach known targets using cognitive maps. A basic method is *route learning*, where paths are stored as action-position pairs. Though effective in small-scale settings, it struggles with scalability due to ambiguities at path intersections [4, 5, 6]. A more flexible alternative is *survey knowledge*, involving metric representations that support vector operations in Euclidean space [5, 7]. However, empirical evidence suggests real brain maps often violate strict geometric constraints [8, 9, 10, 11]. A compelling compromise is the *labeled graph* model: a topological network where spatial relations are encoded locally through experience-based labels [12, 5, 13, 14]. This structure tolerates global inconsistencies while supporting local vector-like operations, aligning with both behavioral studies [15, 9, 16] and *path integration* mechanisms observed across species [17, 18, 19].

The neural substrates behind cognitive maps have been predominatly traces to the hippocampus and its surroundings, where arious neuronal types have been linked to allocentric and egocentric spatial features. Border cells signal environmental boundaries; speed cells encode locomotion velocity; head-direction cells track angular orientation; place cells respond to specific locations; and grid cells exhibit hexagonal spatial tuning [20, 21, 22]. These cells are primarily found in the entorhinal cortex (especially medial EC) and hippocampal formation, including areas CA1, CA3, and the subiculum. Place cells in CA1, in particular, are often viewed as key components of cognitive maps [23]. The origin of their spatial tuning remains debated, with hypotheses focusing on input from entorhinal grid cells via the temporo-ammonic pathway or recurrent projections from CA3 through the Schaffer collaterals [24, 25, 26, 27].

Place cells in CA1 integrate spatial input from the medial EC and CA3, along with sensory and contextual signals from the lateral EC [28]. These inputs are modulated by neuromodulators which dynamically gate information and influence plasticity. In this regard, a primary role is covered by the dopaminergic projections from the VTA and LC target hippocampal subregions, including CA1, where they modulate excitatory input, regulate long-term potentiation

(LTP) via D1-like receptors, and influence memory consolidation [29, 30, 31, 32, 33]. Dopamine conveys reward-related signals that reshape place cell tuning [34, 35, 36, 37], supports novelty detection [38], and encodes reward prediction errors [39, 40], especially in relation to LEC inputs [41, 42].

Computational models of cognitive maps have evolved from early proposals—where the hippocampus encodes position and direction, and the parietal cortex supplies metric information [43]—to route-based topological graphs [6]. Recent models leverage predictive coding, such as the successor representation and Tolman-Eichenbaum Machine, which generalize across spatial and relational tasks while mimicking biological activity [44, 45, 46]. Path integration models trained on motion cues also generate grid- and place-like tunings [47, 48, 49], and some incorporate neuromodulation-driven Hebbian plasticity for reward-based learning [50].

Although many of these frameworks feature bio-realistic tuning, predictive learning, or neuromodulation, none integrates all elements simultaneously, and most depend on biologically implausible training like backpropagation.

In this study, we present a model that builds an online spatial representation of place cells, enriched with behaviorally relevant sensory information through neuromodulators.

The cognitive map is represented as a topological graph, where neighboring place cells are linked across experienced trajectories, going beyond traditional route learning [6]. Goal-directed navigation is then achieved through a path-finding algorithm that operates on this graph-level spatial data, aligning with the *labeled graph* framework [51].

Neuromodulators are modeled as analog sensor nodes that form synapses with place cells, defining scalar fields over the map [52]. Learning is conducted entirely online, mimicking animal behavior and avoiding the costly training processes typical of deep and reinforcement learning models, particularly in path integration tasks [47, 48, 49]. This is enabled by an architectural inductive bias: a predefined stack of grid cell layers with varying spatial frequencies, from which competing networks generate place cells [53, 54].

Our primary goal is to validate the emergence of a cognitive map that reflects the agent's experiences and to assess its effectiveness in supporting goal-directed navigation. A secondary objective is to explore the map's adaptability to environmental changes. To achieve this, we introduce a synaptic update mechanism based on prediction errors in future sensory inputs, inspired by predictive coding theories [40, 55, 56, 39]. This aligns with prior work on neuromodulation in spiking networks for exploration-exploitation dynamics [57, 58], and supports the view that modulators like dopamine are essential for maintaining and adapting neural representations tied to rewards [59, 60, 61, 62, 45].

Finally, we investigate how neuromodulation affects the density and spatial distribution of place fields, driven by experimental observations of remapping, resizing, and relocation in response to significant events [36, 63, 64].

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 constructs by freely moving in a closed environment and discovering target locations. The complete schema of its components is illustrated in the plot 1**a** below. The architecture relies on the core assumption that the external information is minimal, consisting solely of a reward and collision signals as two binary values.

The formation of the spatial representation is instead based on idiothetic information, which is the agent's perception of self-motion [65]. 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 believed to result from the integration of inertial and relative motion cues [66, 67].

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 shown in the plot 1**b**. A grid cell module l has been defined as a set of neurons N^{gc} with a Gaussian tuning curve evenly distributed over the surface of a two-dimensional torus $\mathbf{T}^2$. In plot 1**c** is shown the activity of a grid cell module on a trajectory, the periodicity being underlined by the cell in green.

The activity of all modules of grid cells, indicated as $\mathbf{u}^{\mathrm{GC}}$, is then projected down to two independent layers of initially silent cells, whose feedforward weights $\mathbf{W}^{\mathrm{GC,PC1}}$; $\mathbf{W}^{\mathrm{GC,PC2}}$ are initialized at zero. As the agent moves and the activity of the grid cells changes, if no neurons within a place cell layer are active, then one is randomly chosen and its weights are set to the population vector of current grid cells (at time t) $\mathbf{W}_{\mathrm{i}}^{\mathrm{GC,PC}} \leftarrow \mathbf{u}_{\mathrm{t}}^{\mathrm{GC}}$. For the plasticity process to be completed, the possible overlap with other cells in the same layer is also checked, 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 $\theta_{\mathrm{rep}}^{\mathrm{PC}}$, 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:

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

where ϕ is a generalized sigmoid function $\phi(\mathrm{z}) = [1 + \exp(-\beta(\mathrm{z} - \alpha))]^{-1}$ with gain β and threshold α. Recurrent connections between place cells are calculated

by the same cosine similarity, but compared against a different threshold $\theta_{\text{rec}}^{\text{PC}}$. In addition, for each layer, a neural trace $\mathbf{m}$ of activity is recorded, decaying with a fixed time constant.

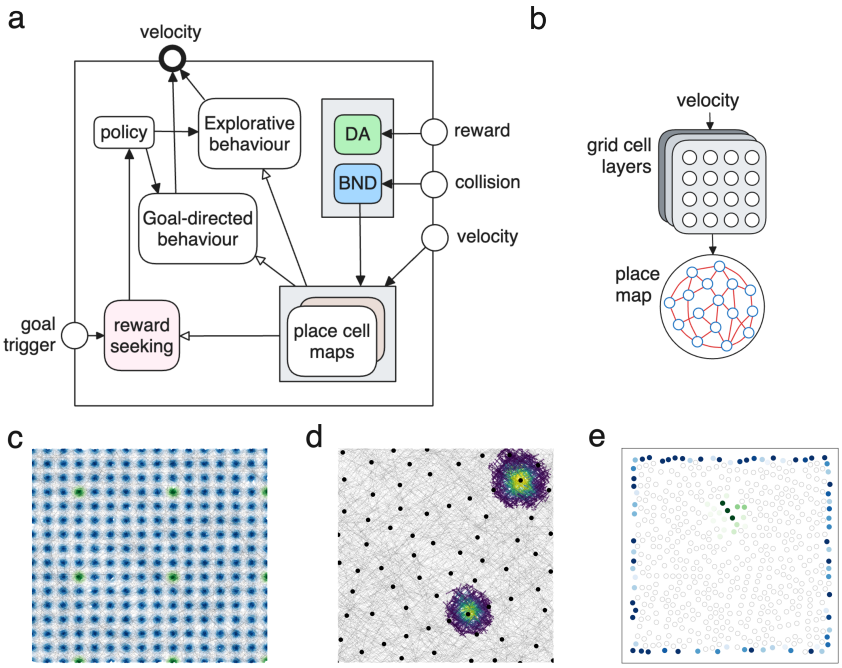

Figure 1: Model layout and spatial representations - **a**: *the full architecture of the model, consisting of three main sensory input, targeting the two modulators and the cognitive map module, and the executive components, represented by a policy module, two behavioral programs and a reward receiver.* **b**: *the cognitive map component, organized with a stack of grid cell modules receiving the velocity input and projecting to the layer of place cells.* **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.*

In figure 1**d** are shown the place cells centers and the activity of two cells are shown, with their place field highlighted as a heat map.

2.2 Neuromodulators

Neuromodulators are operationalized as analog sensors of meaningful environmental events, here to reward and collision, and map directly to the place cells through plastic connections $\mathrm{W}^{\mathrm{k,PC}}$. 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}:

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

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

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

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

On the one hand, the reward modulation, signed as $\mathbf{W}^{\mathrm{DA}}$, is sensitive to the instantaneous presence of the reward, defined as a Boolean value. Over time, its coupling with the population vector $\mathbf{u}_{\mathrm{t}}^{\mathrm{PC}}$ delineates a region of the environment where the reward has been experienced. On the other hand, the collision modulation, referred to as $\mathbf{W}^{\mathrm{BND}}$, 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 to efficiently planning routes to target locations.

At each moment during navigation, the weight matrices $\mathbf{W}^{\mathrm{DA,PC}}$, $\mathbf{W}^{\mathrm{BND,PC}}$ 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** the activity of the two neuromodulators is shown on place cells map, showcasing the limits of the environment and the reward location.

2.2.1 Online adaptation

For what concerns resilience with respect to environmental changes, such as reward location, the model was endowed with a prediction-base mechanism to correct for the accuracy of internal representations. During navigation, before execution of each movement towards a future location $\mathbf{x}_{\mathrm{t+1}}$, a prediction of the activity of the neuromodulator k is computed and cached: $\hat{\mathrm{v}}^{\mathrm{k}} = \sum_{\mathrm{i}}^{\mathrm{N}} \mathbf{W}_{\mathrm{i}}^{\mathrm{DA,PC}} \cdot \hat{\mathrm{u}}_{\mathrm{i}}^{\mathrm{PC}}$, where $\hat{\mathrm{u}}^{\mathrm{PC}}$ is the predicted population activity at $\mathbf{x}_{\mathrm{t+1}}$. Then, a prediction error with a specific learning rate is calculated and added to the weight update:

$$\Delta \mathbf{W}^{\mathrm{k}} = \Delta \mathbf{W}^{\mathrm{k}} + \hat{\eta}^{\mathrm{k}} (\hat{\mathrm{v}}^{\mathrm{k}} - \mathrm{v}^{\mathrm{k}} \mathbf{u}^{\mathrm{PC}}) \tag{4}$$

Further, since the rate $\hat{\eta}^{\mathrm{k}}$ is less than 1, it takes several erroneous prediction to effectively set connections to zero.

The rationale is to depress those synapses that no longer reliably predict the sensory experience referenced by the neuromodulator. This simple rule follows the framework of predictive coding and temporal-difference learning [68]. Previous computational models have used neuromodulatory signals to control neuronal dynamics, applied to the learning process [58, 69] or as a feedback error [70]. Furthermore, it aligns with the experimental evidence for the involvement of neuromodulation in dynamic update of internal beliefs [71, 45, 72].

2.2.2 Modulation of place fields

Lastly, in order to study the possibility of neuromodulatory alteration of neuronal properties, the cells were subjected to relocation and resizing of their place fields according to neuromodulatory input.

The movement of the place centers occurred after a reward or collision event and affected all nearby cells. More in details, each cell center was displaced, within the grid cells space, by a vector in the direction of the current position $\mathbf{u}^{\mathrm{GC}}$, and with magnitude proportional to the value v^{k} of the neuromodulator k and its distance:

$$\Delta \mathbf{W}_{\mathrm{i}}^{\mathrm{GC,PC}} = \mathrm{c}^{\mathrm{k}} \cdot \mathrm{v}^{\mathrm{k}} \cdot \varphi_{\sigma^{\mathrm{k}}} \left(\mathbf{u}^{\mathrm{GC}} - \mathbf{W}_{\mathrm{i}}^{\mathrm{GC,PC}} \right) \tag{5}$$

where c^{k} is a strength scale and φ is a Gaussian distance with width σ^{k}. In addition, lateral inhibition was still taken into account, to avoid overlaps. This approach has been inspired by the BTSP rule [63, 36, 73], which has been responsible for the dislocation of the CA1 place cells after reward events (or an external step current), changing their associated spatial position.

Concerning the field resizing, it again occurs when its reference event occurs. The neuromodulatory action was determined by scaling the baseline neural activation gain $\bar{\beta}$ of the most recently active neurons, identified by their neural trace $\mathrm{m_i}$, according to a constant γ^{k}:

$$\beta_{\mathrm{i}} = \gamma^{\mathrm{k}} \cdot \bar{\beta} \cdot \mathrm{m_i} + \bar{\beta} \cdot (1 - \mathrm{m_i}) \tag{6}$$

The rationale is to modulate the surface over which neurons are sensitive, with $\gamma < 1$ having a shrinking effect while $\gamma > 1$ an enlarging one.

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 exploration tasks and goal-directed navigation. To this end, we defined a simple hard coded policy that toggles between these two behaviors according to the presence of a goal signal, externally provided, and the presence of a real goal representation, cared for 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, namely exploitation, serves the dual purpose of going to a random position within the map to improve exploration and reach the actual reward location. In practice, goal navigation is achieved by calculating the shortest path between the current position, identified by the cells with the most active place, and a target position, picked randomly or considering the cells with the strongest dopaminergic weights, inspired by hippocampal replay and experimental observations [74, 75, 76]. The place cells are thus treated as nodes of a graph, and their connections constitute its edges. Then, we used a Dijkstra algorithm coupled with a value function over the nodes, such that boundary neurons are penalized and a planning near walls is avoided.

2.4 Comparison with previous architectures

Table 1: Comparison of neural network models for spatial navigation and representation

Model	**Architecture**	**Training Method**	**Ext. C.**
Banino et al. [47]	LSTM + linear layers + CNN	BPTT and deep RL, supervised	Yes
Cueva et al. [49]	RNN + linear layers	Hessian-free algorithm with regularization	Yes
Sorcher et al. [48]	RNN + linear layers	Backpropagation with regularization	Yes
Whittington et al. [46]	Attractor network and deep networks	Backpropagation and Hebbian learning	No
de-Cothi et al. [45]	Successor representation	TD-learning + eligibility traces	Yes
Brozsko et al. [58]	Spike Response Model	Online modulated Hebbian plasticity	Yes
Ours	Rate layers	Online neuromodulated plasticity	No

Model	**Task**	**Input**	**Output**
Banino et al.	Path integration, goal navigation	Velocity, visual input, reward	PC, HDC
Cueva et al.	Path integration	Velocity	Position
Sorcher et al.	Path integration	Velocity	PC
Whittington et al.	Relational graph knowledge	Observation and action	Observation
de-Cothi et al.	Planned navigation	Observation	–
Brozsko et al.	Goal navigation	Position, reward	Action
Ours	Goal navigation	Velocity, reward, collision	Action

Note: PC = Place Cells, HDC = Head Direction Cells, Ext. C. = External Spatial Coordinates

There are previous computational models bear similarity with the current work. One class of these made use of deep neural networks, mostly with recurrent units, and they were thus bound to utilize variants of gradient descent for learning of

the connection weights. This aspect constitues an important difference with our model, which relies on synaptic plasticity and only requires one episode for training. Nevertheless, a relevant similarity is the type of outputs and emergent internal representations, involving the place cells population vector.

Another class of models instead used spiking neurons [58] and esplicit neural representations [45], and using online learning rules closer to ours. Also the most explicit goal-directed navigation is more similar than solely path integration. However, both rely on external coordinate information for representing the current spatial position while our model exploits its tracking, obtained by integrating effective output velocities in an internal coordinate system.

2.5 Naturalistic task

The evaluation of the model to construct and use a cognitive map was defined as the total count of rewards collected during several trials. All environments used were closed boxes with different layouts, determined by the location and number of internal walls.

The testing protocol was inspired by the behavior of animals that venture into new territories in search of food. First, there was an exploration phase, in which the agent was placed in a random location and walked a pseudo-random route for 5000 time steps. Then an exploitation phase began, in which a circle of 5% of the total was designated to provide a binary reward $\mathrm{R} \sim \mathcal{B}(\mathrm{p_r})$ drawn from a Bernoulli with probability $\mathrm{p_r}$.

Optimization of the model parameters was carried out using the evolutionary Covariance-Matrix Adaptation strategy (CMA-ES) [77] with a population of 256 individuals for 100 generations.

3 Results

3.1 Performance in wayfinding

Our primary aim was to evaluate the formation of the cognitive map through neuromodulation in terms of the performance of the 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.

The left panel of plot 2**a** displays place cells associated with collisions and reward events, signaling boundaries (in blue) and reward (in green) locations. The overlap of these two representations and the place cells (in pink) is what we refer to as a cognitive map, since these are the main sources of spatial and contextual information used during planned navigation, whose path is depicted as a gray line. The right panel instead portrays the actual environment with walls (black), reward location (green), and multiple trajectories (red). During

exploration, the main areas were visited until the reward position was located and the goal-directed navigation dominated, as highlighted by the density of the path lines. Considering the position of the walls and corners, the layout of this environment does not always make the target locations visible, as it is a non-convex area and therefore can be classified as wayfinding [78]. The challenge of not being able to use straight lines is overcome by the graph approach using local data and the consideration of boundary place cells, allowing the agent to plan accordingly. In addition, by construction the path also minimized the length of the path, within the part of the space covered by the cognitive map.

In general, this result confirms the ability of the model to focus on navigation and obstacle avoidance. However, it is worth nothing that not all simulations resulted in a reward being found in the first place, due to the randomness of the exploratory process; this was more pronounced in complex environment.

3.1.1 Detour task

The planning ability and the plastic nature of the cognitive map should provide resilience against unexpected changes in the environment layout. In order to verify this we implemented a detour experiment. Initially, the agent was familiarized with a square environment with the reward in the middle and starting always from the same position. Then, a wall was placed in between the starting position and the reward, therefore forrcing new trajectory for reaching it. As expected, the agent was able to form a representation of the new obstacle and calculating new paths around it, succeeding the task. In plot 2**b** they are shown the trajectories before and after the wall placement, and it is manifesed the ability of detour in the new layout.

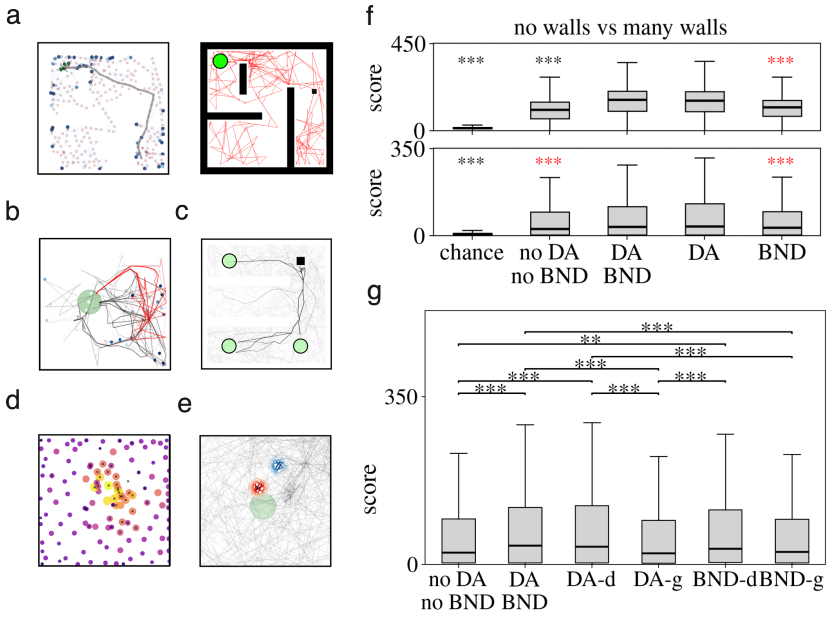

Figure 2: COGNITIVE MAPS AND PERFORMANCE RESULTS - **a**: *the plot on the left represent a cognitive map over a space, together with the plan (grey line) to reach a target location from a starting position. The plot on the right is a view of the same enviroment but with highlighted walls (black thick lines) the reward (green circle), trajectory (red line), agent position (black square).* - **b**: *plot of trajectories before (black) and after (red) the insertion of a wall between the starting and goal positions, the wall can be spotted from the boundary cells in blue* - **c**: *trajectories for multiple trials with the agent starting at the same position (black square) but with the reward location (green circles) periodically moving* - **d**: *place cells centers with size and color proportional to its node degree (number of neighbors); further, it notable is the correlation with the distance from the reward (in the center)* - **e**: *place fields of the same cell before and after several relocation of its center following reward events* - **f**: *performance comparison with same environmental conditions for five different models: one baseline as chance level (plasticity disabled, except for place cells generation), and four variants with different ablations of DA and BND modulations of place fields density and size. Results in terms of reward count and Bonferroni-corrected pairwise t-test, black stars stand for statistical difference with respect to all other groups, red starts only to the DA and DA+BND.* - **g**: *ablation performance comparison similar to* c, *and over the same environment for model variants with different active modulatory mechanisms: no DA and BND place field modulation, full DA and BND, DA-d for PC density, DA-g activation gain, BND-d for PC density,BND-g for activation gain.*

3.2 Adaptive goal representation through prediction error

Then, we tested the adaptability to environmental changes. In this scenario, the reward object was moved after being fetched a fixed number of times. Here, the difficulty was to unlearning previous locations and discovering new ones, in a protocol similar to [50]. In plot **2c** is reported the set of trajectories over many trials with the reward displaced in three possible locations. The agent was capable of planning behavior, as earlier, but also exploring and finding the new rewards, as shown by the density of lines. Whenever a goal path resulted in a failed prediction, the DA-based sensory error weakened the association between the place cells and the reward signal, leading to an extinction of its representation at that location.

This result validates the resilience of the model to changing sensory expectations, in this case the reward position.

3.3 Modulation of spatial resolution affects performance

Lastly, we investigated the effect of modulating the density of place cells and the size of the field. The goal position was fixed, but the agent was randomly relocated after fetching; performance was defined as the total number of reward counts within a time window.

Our working hypothesis is that these experience-driven neuronal changes would improve the quality of the cognitive map and be reflected in navigational 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 to form. All models were run in two different environments differing by number of internal walls, in total 2048 simulations were done for each case.

Plot **2d** showcases the distribution of place cells with the circle size and color represent the node degree, which aligns discretely with the reward position and the density. Further, in plot **2e** the place field of one cell is shown, before and after several reward occurences and consequent center relocation.

The statistical results are shown in plot **2f**. The top reports the scores in the setting without internal walls. All models performed above chance, but the main finding is the affirmation of the importance of neuronal modulation, as revealed by the statistical difference between those endowed with it and the one not. Furthermore, possession of DA modulation resulted in a significantly higher score than BND modulation alone (red stars). In fact, in a situation with a convex region such as this, once the reward has been located, the boundary information has limited utility.

A similar pattern emerged when more internal walls were introduced, as shown by the results at the bottom. In this environment, scores were overall lower as navigation became more difficult, and the difference among groups thin-

ner, although noticeable. Here, the sole presence of the boundary modulation of place cells did not result in better performance than the baseline. This outcome highlights how the main improvements are brought by reward-driven neuronal modulation (DA).

Then, in order to enquiry which particular modulatory action was mostly relevant, we conducted a similar test on the same environment but ablating not only specific neuromodulators, but also whether they affected the place cells density or the activation gain parameter. In plot 2**g** are reported the comparions among different variants. The general trend matched the previous results 2**c**, in that DA is the best neuromodulator in terms of correlation with performance. The additional finding was that density modulation is significantly the primary action behind the behavioural improvements, and alone does not perform worse that the model with all actions together.

Taken together, these findings support the hypothesis of practical utility of direct modulation of place-field structure for active navigation, even in these simple settings.

4 Discussion

Exploration and planning in known and past environment are essential behaviors of animals, directly affecting their success in world understanding 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. Numerous speculations have been made about the shape and neural foundations of such an object, varying in the types of modeling assumptions and experimental support.

The contribution of the present work was to propose a rate network model, inspired by the CA1 hippocampal region [23]. We used grid cells together with synaptic plasticity as a mechanism to develop information-rich representations based on place cells updated through experience, grouped with common perspectives on cognitive maps [79]. In the spirit of minimizing the geometric assumptions in the neural space, we treated the generated place network as a topological graph, with sensory information added locally through the action of neuromodulators. This idea aligned with the concept of a *labeled graph* [14, 9], however, it is also true that no metric violations were actually possible in these settings.

The tasks we applied the agent consisted of an exploratory and exploratory phase, in which it was prompted to plan and reach reward positions. For simplicity, the first stage relied on a random walk process, as it was outside the scope of this work. This choice had the side effect that the reward was not always discovered, leading to the formation of incomplete maps, and thus impairing performance. However, this issue was limited in frequency.

The simulation results validated the model, showing the expected emergence of cognitive maps and their encoding of information collected during the experi-

ence. The online nature of the formation of the locations on the map aligns with the idea of using only idiothetic velocity input, as in path integration [7, 18, 80]. Previous work followed a similar direction using recurrent networks, but required extensive gradient-based training [81, 49, 67]. Another important difference is that our resulting neural network was composed solely of place cells, although neuromodulated, and no other types of neuron were present. This distinction is justified by the partially different task structure, which did not involve supervised learning and did not receive visual information as in [47]. Furthermore, our model relied on predefined grid cells layers, which constituted a strong and sufficient inductive bias, and did not have to be learned from scratch.

An additional relevant aspect is also the consideration of the place cell layer as an explicit graph data structure, on which the path-planning and decision-making algorithm was applied. The adoption of this level of description lead to robustness and flexibility, enabling effective navigation in all tested environments, which varying in layout complexity. Nevertheless, this approach did act as another clear inductive bias, which lifted the need to learn an approximation of it through network dynamics and even more differently tuned neurons.

Adaptability was tested by occasionally moving the reward position, leading to the generation of an internal prediction error that was used to update its representation on the map. The agent was proved capable of unlearning previous associations, returning to exploration, and memorizing new reward locations. This behavioral protocol is similar to previous work [58], in which dopaminergic and cholinergic activity was utilized within a Hebbian plasticity rule to strengthen or weaken reward-associated spatial representations. However, alternatively to exploiting neuromodulators with opposite valence, we followed a preditive coding framework, a direction linked to hippocampal representations [45, 82] and explored various computational approaches [83, 84, 85]. This choice departed from our focus on using operations on the cognitive map itself by simulating future sensory experiences and learning from feedback. In fact, neuromodulation has been long associated with this functionality [52], especially dopamine [34, 59, 62, 39].

Lastly, the hypothesis of relevance of the active modulation of the neuronal properties of place cells was corroborated by simulating ablation experiments. These tests reported a significant impact of altering the place cells density on the total count of collected rewards. In general, these results are consistent with the experimental observations of alteration of place cells following reward events [36, 86], in particular in terms of increased clustering of cells [87, 88], reminiscent of changes in firing rate after contextual changes [89, 90].

Concerning the modulation of place fields, there is significant experimental evidence of their alternation during reward events [91, 92, 93], some reporting shrinkage near reward objects [94], and boundaries [95]. The coupling with higher local density could be explained by better optimization of the cell distribution for goal representation and planning [70]. However, in our settings, the fields become enlarged, especially in the direction of the target, although the performance improvements were not tested significant. A possible explanation can be the simplicity of our reward, which was solely defined as an area of space.

The lack of rich non-spatial features thus did not require the place cells to code for smaller spatial variation. Therefore, enlargement might have improved the stability of the representation, marking the nodes associated with rewards more solidly, given the stochasticity of its delivery. Further, the graph-path algorithm utilized the strength of the DA-modulated connections for determining the goal representation, stronger fields inherently developed stronger weights, making planning more reliable planning. Although these findings are limited within the limit of our simulation protocol, there have been experimental observations of elongation of place fields along trajectories over meaningful experiences [96, 97].

In conclusion, this work showed a possible architecture for coupling emergent spatial representations with neuromodulated plasticity to achieve an experience-driven cognitive map. The reliance of few spatial and algorithmic inductive biases, grid cells and planning algorithm, supports the idea of label graph for goal navigation. Future work can investigate the application to other spatial domains, such as motor control and three-dimensional navigation. In addition, a richer input feature can be added, such as visual information [98], as well as new neuromodulators that encode different sensory dimensions or internally generated signals.

Acknowledgements & Statements
The authors declare no competing interests.
The code is publicly available and can be found at `https://github.com/iKiru-hub/PCNN`.
This research was funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement Nº 945371 and the University of Oslo.
The research presented in this paper has benefited from the Experimental Infrastructure for Exploration of Exascale Computing (eX3), which is financially supported by the Research Council of Norway under contract 270053.

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] 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.

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

[6] 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.

[7] 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.

[8] 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.

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

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

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

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] 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.

[21] 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.

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

[23] 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.

[24] 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.

[25] 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.

[26] 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.

[27] 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.

[28] 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.

[29] 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.

[30] 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.

[31] 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.

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

[33] 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.

[34] 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.

[35] 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.

[36] 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.

[37] 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.

[38] 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.

[39] 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.

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

[41] 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.

[42] 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.

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

[44] 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.

[45] 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.

[46] 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.

[47] 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.

[48] 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.

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

[50] 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.

[51] 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.

[52] Marielena Sosa, Mark H. Plitt, and Lisa M. Giocomo. Hippocampal sequences span experience relative to rewards. *bioRxiv*, page 2023.12.27.573490, February 2024.

[53] 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.

[54] 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.

[55] 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.

[56] 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.

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

[58] 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.

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

[60] 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.

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

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

[63] 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.

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

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

[66] 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.

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

[68] Richard S Sutton and Andrew G Barto. The Reinforcement Learning Problem.

[69] 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.

[70] Pablo Scleidorovich, Jean-Marc Fellous, and Alfredo Weitzenfeld. Adapting hippocampus multi-scale place field distributions in cluttered environments optimizes spatial navigation and learning. *Frontiers in Computational Neuroscience*, 16:1039822, December 2022.

[71] P. R. Montague, P. Dayan, and T. J. Sejnowski. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience*, 16(5):1936–1947, March 1996.

[72] Seetha Krishnan, Chad Heer, Chery Cherian, and Mark E. J. Sheffield. Reward expectation extinction restructures and degrades CA1 spatial maps through loss of a dopaminergic reward proximity signal. *Nature Communications*, 13(1):6662, November 2022.

[73] Christine Grienberger and Jeffrey C. Magee. Entorhinal cortex directs learning-related changes in CA1 representations. *Nature*, 611(7936):554–562, November 2022.

[74] Colin G. McNamara, Álvaro Tejero-Cantero, Stéphanie Trouche, Natalia Campo-Urriza, and David Dupret. Dopaminergic neurons promote hippocampal reactivation and spatial memory persistence. *Nature Neuroscience*, 17(12):1658–1660, December 2014.

[75] Frédéric Michon, Esther Krul, Jyh-Jang Sun, and Fabian Kloosterman. Single-trial dynamics of hippocampal spatial representations are modulated by reward value. *Current biology: CB*, 31(20):4423–4435.e5, October 2021.

[76] Philip Shamash and Tiago Branco. Mice identify subgoal locations through an action-driven mapping process, December 2021.

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

[78] 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.

[79] Vincent Hok, Pierre-Pascal Lenck-Santini, Sébastien Roux, Etienne Save, Robert U. Muller, and Bruno Poucet. Goal-Related Activity in Hippocampal Place Cells. *Journal of Neuroscience*, 27(3):472–482, January 2007.

[80] Bruce L. McNaughton, Francesco P. Battaglia, Ole Jensen, Edvard I. Moser, and May-Britt Moser. Path integration and the neural basis of the 'cognitive map'. *Nature Reviews Neuroscience*, 7(8):663–678, August 2006.

[81] Ben Sorscher, Gabriel C. Mel, Samuel A. Ocko, Lisa M. Giocomo, and Surya Ganguli. A unified theory for the computational and mechanistic origins of grid cells. *Neuron*, 111(1):121–137.e13, January 2023.

[82] Fraser Aitken and Peter Kok. Hippocampal representations switch from errors to predictions during acquisition of predictive associations. *Nature Communications*, 13(1):3294, June 2022.

[83] Manu Srinath Halvagal and Friedemann Zenke. The combination of Hebbian and predictive plasticity learns invariant object representations in deep sensory networks. *Nature Neuroscience*, pages 1–10, October 2023.

[84] Alexander Ororbia. Spiking neural predictive coding for continually learning from data streams. *Neurocomputing*, 544:126292, August 2023.

[85] Kimberly L Stachenfeld, Matthew M Botvinick, and Samuel J Gershman. The hippocampus as a predictive map. *Nature Neuroscience*, 20(11):1643–1653, November 2017.

[86] Indrajith R. Nair, Guncha Bhasin, and Dipanjan Roy. Hippocampus Maintains a Coherent Map Under Reward Feature–Landmark Cue Conflict. *Frontiers in Neural Circuits*, 16, April 2022.

[87] Valerie L. Tryon, Marsha R. Penner, Shawn W. Heide, Hunter O. King, Joshua Larkin, and Sheri J. Y. Mizumori. Hippocampal neural activity reflects the economy of choices during goal-directed navigation. *Hippocampus*, 27(7):743–758, July 2017.

[88] Hannah S Wirtshafter and Matthew A Wilson. Differences in reward biased spatial representations in the lateral septum and hippocampus. *eLife*, 9:e55252, May 2020.

[89] Michael I. Anderson and Kathryn J. Jeffery. Heterogeneous Modulation of Place Cell Firing by Changes in Context. *Journal of Neuroscience*, 23(26):8827–8835, October 2003.

[90] Inah Lee, Amy L. Griffin, Eric A. Zilli, Howard Eichenbaum, and Michael E. Hasselmo. Gradual Translocation of Spatial Correlates of Neuronal Firing in the Hippocampus toward Prospective Reward Locations. *Neuron*, 51(5):639–650, September 2006.

[91] Marianne Fyhn, Sturla Molden, Stig Hollup, May-Britt Moser, and Edvard I. Moser. Hippocampal Neurons Responding to First-Time Dislocation of a Target Object. *Neuron*, 35(3):555–566, August 2002.

[92] P.-P. Lenck-Santini, B. Rivard, R.u. Muller, and B. Poucet. Study of CA1 place cell activity and exploratory behavior following spatial and nonspatial changes in the environment. *Hippocampus*, 15(3):356–369, 2005.

[93] David Dupret, Joseph O'Neill, Barty Pleydell-Bouverie, and Jozsef Csicsvari. The reorganization and reactivation of hippocampal maps predict spatial memory performance. *Nature Neuroscience*, 13(8):995–1002, August 2010.

[94] S.N. Burke, A.P. Maurer, S. Nematollahi, A.R. Uprety, J.L. Wallace, and C.A. Barnes. The Influence of Objects on Place Field Expression and Size in Distal Hippocampal CA1. *Hippocampus*, 21(7):783–801, July 2011.

[95] Sander Tanni, William De Cothi, and Caswell Barry. State transitions in the statistically stable place cell population correspond to rate of perceptual change. *Current Biology*, 32(16):3505–3514.e7, August 2022.

[96] Mayank R. Mehta, Carol A. Barnes, and Bruce L. McNaughton. Experience-dependent, asymmetric expansion of hippocampal place fields. *Proceedings of the National Academy of Sciences*, 94(16):8918–8921, August 1997.

[97] Jangho Lee, Jeonghee Jo, Byounghwa Lee, Jung-Hoon Lee, and Sungroh Yoon. Brain-inspired Predictive Coding Improves the Performance of Machine Challenging Tasks. *Frontiers in Computational Neuroscience*, 16:1062678, 2022.

[98] John H. Wen, Ben Sorscher, Emily A. Aery Jones, Surya Ganguli, and Lisa M. Giocomo. One-shot entorhinal maps enable flexible navigation in novel environments. *Nature*, 635(8040):943–950, November 2024.

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

[100] 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.

5 Appendix

5.1 Spatial representations

5.2 Grid cell module

Unlike other approaches to generate grid fields [99, 100], we defined a correspondence between the global environment in which the agent moves, a two-dimensional Euclidean space $\mathbf{R}^2$, and a bounded local space of a grid module, corresponding to the torus.

The global velocity $\mathbf{v} = \{\mathrm{x}, \mathrm{y}\}$ is then mapped to a local velocity, scaled by a speed scalar $\mathrm{s}_{\mathrm{l}}^{\mathrm{gc}}$ specific to the grid cell module l, which determines its periodicity in space. This approach has been used in previous works [26].

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 sizes 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.

A grid cell module l of size N^{gc} is identified by a set of positions defined over a square centered on the origin and size of 2, such that $\{(x_i, y_i) \mid i \in N^{gc} \wedge x_i, y_i \in (-1, 1)\}$. This local square space has boundary conditions for each dimension, such that, for instance, when $x_t + s_l^{gc} \cdot v_x > 2$ the position update is taken to the other side $x_{t+1} = x_t + s_l^{gc} \cdot v_x - 2$, where s_l^{gc} is the scale of the velocity in the local space of the module l with respect to the real global agent velocity $v = \{v_x, v_y\}$. When the module is initialized, the starting positions of its cells are uniformly distributed over the square forming a lattice. When the agent is reset in a new position at the beginning of new trial, a displacement vector is applied to the last cells positions such that the mapping between the module local space and the global environment is preserved.

The firing rate vector of each cell is determined with respect to a 2D Gaussian tuning curve centered at the origin at $(0, 0)$, and it is calculated as $r_i = \exp\left(-\frac{x_i^2+y_i^2}{\sigma_l^{gc}}\right)$, where σ_l^{gc} is the width of the tuning curve for module l.

The final population vector of the grid cell network GC is the concatenated and flattened firing rate vector of all modules $\mathbf{u}^{GC}$.

In our model, each grid cell had a tuning width of 0.04. They were defined as 8 modules of size 36, and the relative speed scales were $\{1., 0.8, 0.7, 0.5, 0.4, 0.3, 0.2, 0.1, 0.07\}$.

5.3 Place cells

Tuning formation The tuning of a new place cell is simply defined as the current GC population vector $\mathbf{u}_t^{GC}$, and its index is that of the first silent cell, which is added to the forward weight matrix $\mathbf{W}_i^{GCtoPC} \rightarrow \mathbf{u}^{GC}$.

In order to avoid overlapping of place fields, the tuning process is aborted in case the cosine similarity of the new pattern and the old ones is greater than a threshold θ_{rep}^{PC}.

Each cell represents a position in the GC activity space, which can be considered a node within a graph of place cells (PC). Although it is totally possible to only use the N^{GC}-dimensional tuning patterns and be agnostic about the dimensionality of the space in which the agent lives, to simplify the calculations, we mapped each pattern to 2D positions in a vector space. Then, the PC recurrent connectivity matrix is calculated with a nearest neighbors algorithm, which instead of a fixed number K of neighbors uses a distance threshold θ_{rec}^{PC}.

Activity The current firing rate of the PC population is determined by the cosine similarity between the GC input and the forward weight matrix, then passed through a generalized sigmoid $\phi(z) = [1 + \exp(-\beta(z - \alpha))]^{-1}$. The parameter α represents the activation threshold, or horizontal offset, while β the gain, or steepness. It is also defined as an activity trace, which has an upper value of 1 and decays exponentially:

$$\mathrm{m_i} = -\mathrm{m_i}/\tau^{\mathrm{PC}} + \mathrm{u_i} \tag{7}$$

It is used as a proxy for a memory trace.

In the model, a PC population is defined by its average place field size, determining the granularity of the representation of the place.

Below, a table of the parameters specific to the two PC populations is reported.

	β	α	$\theta^{\mathrm{PC}}_{\mathrm{rep}}$	$\theta^{\mathrm{PC}}_{\mathrm{rec}}$	τ^{PC}
pc	33.0	1.0	0.86	43	140

Table 2: PC LAYER PARAMETERS

5.4 Modulation

Neuromodulation is defined in terms of a leaky variable v whose state is perturbed by an external input x, whose qualitative meaning differs for each neuromodulator k.

$$\begin{aligned} \mathrm{v_k} &= -\mathrm{v_k}/\tau_{\mathrm{k}} + \mathrm{x_k} \\ \mathrm{v_k} &= \max(\mathrm{v_k}, 0) \end{aligned} \tag{8}$$

Learning rule The connection weights $\mathbf{W}^{\mathrm{k}}$ are updated according to a plasticity rule composed of an Hebbian term, involving the leaky variable and the place cells above a certain threshold θ^{k}, and a prediction error.

A prediction $\mathbf{p}$ is calculated before each time-step and signifies the expected neuromodulation activation for a given place cell $\mathrm{p_i} = \mathbf{W}^{\mathrm{k}}_{\mathrm{i}}\hat{\mathrm{u}}^{\mathrm{PC}}$, where $\hat{\mathrm{u}}^{\mathrm{PC}}$ is the PC population vector obtained by simulating the planned action. A prediction error is computed as the difference between the prediction and the current modulated place activation.

The full connection update then becomes:

$$\Delta\mathbf{W}^{\mathrm{k}}_{\mathrm{i}} = \eta^{\mathrm{k}}\mathrm{v}^{\mathrm{k}}\mathrm{u}^{\mathrm{PC}}_{\mathrm{i}} - \eta^{\mathrm{k}}_{\mathrm{pred}}(\mathrm{p_i} - \mathrm{v}^{\mathrm{k}}\mathrm{u}^{\mathrm{PC}}) \tag{9}$$

where η^{k}, $\eta^{\mathrm{k}}_{\mathrm{pred}}$ are the learning rates, interpretable as the weight contribution of the Hebbian coupling and prediction accuracy respectively. Additionally, connections values are kept non-negative and, in order to speed up the extinction of past place-reward associations, if a non-zero prediction error is below a 0.5 then it is set to 0.5.

Active neuronal modulation Neuromodulation acts on the neuronal profile of the place cells by affecting the value of the activation gain and relocate the center of their tuning.

Gain modulation is implemented using the activity traces and a constant reference gain value $\bar{\beta}$:

$$\beta_{\mathrm{i}} = \mathrm{c}_{\mathrm{a}}^{\mathrm{k}} \bar{\beta} \mathrm{m}_{\mathrm{i}} + (1 - \mathrm{m}_{\mathrm{i}}) \bar{\beta} \tag{10}$$

where $\mathrm{c}_{\mathrm{a}}^{\mathrm{k}}$ is a scaling gain parameter, and if it is 1 then no modulation takes place.

Concerning center relocation, it is applied to recently active neurons with non-zero trace m_{i}. For a place cell i with position $\mathbf{x}_{\mathrm{i}}$ (in the vector space), it is calculated a displacement vector q_{i} with respect to the current position $\mathbf{x}_{\mathrm{j}}$, identified as the most active place cell j.

$$\mathrm{q}_{\mathrm{i}} = \mathrm{c}_{\mathrm{b}}^{\mathrm{k}} \mathrm{v}^{\mathrm{k}} \exp\left(-\frac{\|\mathbf{x}_{\mathrm{i}} - \mathbf{x}_{\mathrm{j}}\|}{\sigma^{\mathrm{k}}} \right) \tag{11}$$

where $\mathrm{c}_{\mathrm{b}}^{\mathrm{k}}$ is a scaling relocation parameter, while σ^{k} the width of the Gaussian distance. This displacemented is used to move in GC activity space and get the new GC population vector to use as tuning pattern.

Also in this case, it is ensured that the new place field center is at a minimum distance $\theta_{\mathrm{min}}^{\mathrm{PC}}$ from the others; here Euclidean distance is used.

5.5 Decision making

5.5.1 Path-planning algorithm

The planning of a new route is implemented as a modified Dijkstra algorithm over the coarse-grained place cell graph, provided as connectivity matrix C. Its particularity is the use of a weighting $\widetilde{\mathbf{W}}$ of the nodes according to the neuromodulation map.

Algorithm 1 Modified Dijkstra algorithm

Require: Connectivity matrix $\mathbf{C} \in \{0,1\}^{\mathrm{N^{PC}} \times \mathrm{N^{PC}}}$, Node coordinates $\mathbf{x} \in \mathbb{R}^{\mathrm{N^{PC}} \times 2}$, Node weights $\widetilde{\mathbf{W}} \in \mathbb{R}^{\mathrm{N^{PC}}}$, Start node s, End node t
Ensure: Shortest path from t_0 to T

distances $\leftarrow [\infty, \infty, \ldots, \infty]$ ▷ Initialize distances
distances[t_0] $\leftarrow 0$
parent $\leftarrow [-1, -1, \ldots, -1]$ ▷ Parent pointers
finalized $\leftarrow$ [false, false, ..., false] ▷ Set of finalized nodes
PQ $\leftarrow \emptyset$ ▷ Priority queue
PQ.push($(0, \mathrm{t}_0)$) ▷ Insert start node with priority 0
while PQ $\neq \emptyset$ **do**
    (dist, j) $\leftarrow$ PQ.extractMin()
    if finalized[j] or dist $>$ distances[j] **then**
        continue
    end if
    finalized[j] $\leftarrow$ true
    if j = T **then**
        break ▷ Destination reached
    end if
    for each node i where $\mathbf{C}_{\mathrm{i,j}} = 1$ and not finalized[i] **do**
        if $\widetilde{\mathbf{W}}$[i] < -1000 **then**
            continue ▷ Skip nodes with high negative weights
        end if
        $\Delta \mathrm{x} \leftarrow \mathbf{x}_{\mathrm{j},0} - \mathbf{x}_{\mathrm{i},0}$
        $\Delta \mathrm{y} \leftarrow \mathbf{x}_{\mathrm{j},1} - \mathbf{x}_{\mathrm{i},1}$
        edge_dist $\leftarrow \sqrt{\Delta \mathrm{x}^2 + \Delta \mathrm{y}^2}$ ▷ Euclidean distance
        new_dist $\leftarrow$ d[j] + edge_dist
        if new_dist $<$ d[i] **then**
            d[i] $\leftarrow$ new_dist
            parent[i] $\leftarrow$ u
            PQ.push((new_dist, i))
        end if
    end for
end while
path $\leftarrow$ []
if d[T] $= \infty$ **then**
    return $\emptyset$ ▷ No path exists
end if
curr $\leftarrow$ t
while curr $\neq -1$ **do**
    path.append(curr)
    curr $\leftarrow$ parent[curr]
end while
path.reverse()
if path is empty or path[0] $\neq$ s **then**
    return $\emptyset$
end if
return path