

---

# Flexible navigation with neuromodulated cognitive maps

---

**Anonymous Author(s)**

Affiliation

Address

email

1

## Abstract

2 Animals naturally form personalized cognitive maps to support efficient  
3 navigation and goal-directed behavior. In the brain, the CA1 subregion  
4 of the hippocampus plays a key role in this process, hosting spatially  
5 tuned neurons that adapt based on the behavioral context and internal  
6 states. Computational models of this ability include labeled graphs  
7 with locally specified spatial information, which avoid global metric  
8 structure, and deep neural networks trained on spatial tasks that exhibit  
9 emergent spatial tuning. However, these approaches often struggle to  
10 model one-shot adaptive mapping and typically rely on plasticity rules  
11 that lack biological plausibility.

12 We propose a neural architecture inspired by place-cell dynamics that  
13 enables rapid on-the-fly construction of cognitive maps during explo-  
14 ration of novel environments. The model relies on velocity inputs and  
15 grid cell modules to generate spatial representations and integrates neu-  
16 romodulatory signals responsive to boundaries and rewards. Learning  
17 combines synaptic plasticity, lateral inhibition, and modulatory gating  
18 of place-cell activity. For reward-driven navigation, the agent uses a  
19 graph-based algorithm to plan paths on the emergent cognitive map,  
20 treating place cells as nodes in a locally structured graph.

21 We tested the model on different environments, achieving high sample  
22 efficiency and solving tasks in a single episode, for which usual RL  
23 agents require thousands of training steps. This performance advantage  
24 arises from biologically inspired inductive biases embedded in the model  
25 architecture. In simulation, the agent adapts to dynamic reward locations  
26 and changes in the environment layout. Ablation experiments and  
27 analysis of neuromodulated place cells reveals task-dependent changes  
28 in tuning field size and spatial density, aligning with experimental  
29 findings from hippocampal recordings. These results highlight the  
30 promise of biologically grounded computation and locally structured  
31 graph representations for flexible and data-efficient cognitive mapping.

32 

## 1 Introduction

33 Survival in complex environments requires efficient navigational strategies. From desert ants to  
34 humans, successful wayfinding—navigating toward goals that are not directly visible depends on  
35 emergent internal spatial representations, known as cognitive maps [1, 2]. Understanding how  
36 these maps are constructed from ongoing experiences, and how they can be exploited for flexible  
37 goal-directed navigation remains an active area of research in both neuroscience and reinforcement  
38 learning (RL).

39 The hippocampus (HP) and entorhinal cortex (EC) serve as the primary neural substrates for spatial  
40 representation in the brain. They containing specialized neurons that encode spatial and contextual  
41 information—including grid, border, speed, and place cells [3, 4, 5]. Place cells in the CA1 region  
42 have attracted particular interest due to their convergence of inputs from periodically tuned grid  
43 cells, the CA3 region, and the lateral EC [6, 7, 8, 9]. This strategic integration of diverse spatial and  
44 contextual signals suggest CA1 place cells may play a critical role in the formation and maintenance  
45 of cognitive maps [10].

46 Another important component are neuromodulators. Their actions include modulation of neuronal  
47 dynamics, for instance adjusting the synaptic strength, tuning place fields [11, 12, 13]. Further,  
48 neuromodulators such as dopamine transmit reward signals reshaping place cell tuning [14, 15, 16, 17],  
49 support novelty detection [12], and encode prediction errors [18, 13], particularly via LEC inputs  
50 [19, 20]—mechanisms closely related to reinforcement learning principles.

51 Traditional cognitive map theories propose multiple strategies for spatial navigation based on map-like  
52 representations. Route learning encodes paths as sequences of action–position pairs, but it limited  
53 in scalability and generalization, especially at route intersections [21, 22, 23]. In contrast, survey  
54 maps, which rely on Euclidean geometry, offer greater flexibility [22, 24]; however, their strong  
55 geometric assumptions often conflict with neural and behavioral evidence pointing to geometric  
56 distortions and topological biases in spatial neural representations [25, 26, 27, 28]. As a middle  
57 ground, labeled graphs encode landmarks and transitions within a topological network, enabling  
58 vector-like operations, planning, and prediction [29, 30, 31].

59 Computational models have captured some of these aspects individually, showing new ways the  
60 brain might use for addressing spatial navigation tasks. Early work proposed that the hippocampus  
61 encodes spatial position and direction [32], while topological graph models based present scalability  
62 challenges [23]. More recent approaches draw inspiration from predictive coding and reinforcement  
63 learning, including successor representations and the Tolman-Eichenbaum Machine, which generalize  
64 across spatial and relational tasks while mimicking biological neural activity patterns [33, 34, 35].  
65 Path integration models trained on velocity inputs give rise to spatial-like receptive fields [36, 37, 38].  
66 Others incorporate reward-driven Hebbian plasticity modulated by neuromodulators [39]. However,  
67 these architectures mostly fail to unify these ingredients into a biologically grounded system that  
68 at the same time learns a map of the environment online without relying on an external coordinate  
69 system, and flexibility perform goal-directed navigation.

70 In this work, we present a biologically inspired model of cognitive map formation that integrates place  
71 cell representations, neuromodulatory signals, and graph-based spatial computations. Our aim is to  
72 demonstrate an architecture capable of building a content-rich topological map of the environment on  
73 the fly, and leveraging it for efficient, goal-directed navigation—without requiring offline training.

74 Critically, neuromodulators play a central role, as they form scalar fields over the map [40], drive  
75 local Hebbian plasticity in response to sensory updates [41, 18, 42], and support the formation and  
76 adaptation of reward-modulated neural representations used for planning [43, 44, 45, 46, 34]. We  
77 show how this system dynamically adapts to environmental changes and how neuromodulation shapes  
78 place field allocation and remapping [47, 48], linking cognitive flexibility to underlying physiological  
79 mechanisms.

80 The remainder of the paper is organized as follows: Section 2 details the model and experimental  
81 setup; Section 3 presents results; Section 4 discusses broader implications and future directions.

## 82 **2 Methods**

83 We propose a model of cognitive map formation driven by an agent’s experience within a closed  
84 environment.

85 The architecture operates with minimal external inputs—limited to binary reward and collision  
86 signals—as illustrated in Figure 1a. Instead of relying on exteroceptive cues, spatial representations  
87 emerge from idiothetic information, i.e., the agent’s internal perception of self-motion [49], consistent  
88 with prior path integration frameworks. Concretely, we use the agent’s ground-truth velocity vector,  
89 i.e. its actual displacement within the environment, as the primary navigational signal, reflecting  
90 the integration of inertial and proprioceptive cues observed in biological systems [50, 51]. Since no  
91 visual information is used, the agent is effectively navigating in the dark.

92 **Place Cell Formation** The primary spatial representation is formed by a set of grid cell modules,  
 93 each encoding a periodic tiling of 2D space, which directly maps to a toroidal manifold  $\mathbf{T}^2$  (Fig.  
 94 1b). Departing from traditional grid cell modeling approaches [52, 53], we generate population  
 95 activity directly via Gaussian tuning over the torus, continuously updated using the agent's velocity  
 96 vector—an approach used in prior work [8].

97 The grid cell population vector  $\mathbf{u}^{GC}$  is forwarded to a place cell network with initially zeroed synaptic  
 98 weights. When no place cell is sufficiently active for a given input, a silent unit is randomly selected  
 99 and imprinted with the current grid activity pattern. To enforce representational sparsity and tuning  
 100 specificity, lateral inhibition is implemented by comparing the cosine similarity between the new  
 101 weight vector and any existing one and a threshold  $\theta_{inh}^{PC}$ .

102 Each place cell's activation is computed via a bounded cosine similarity function, determining its  
 103 corresponding place field (Fig. 1d). Further implementation details, including lateral inhibition and  
 104 recurrent connectivity, are provided in the Appendix.

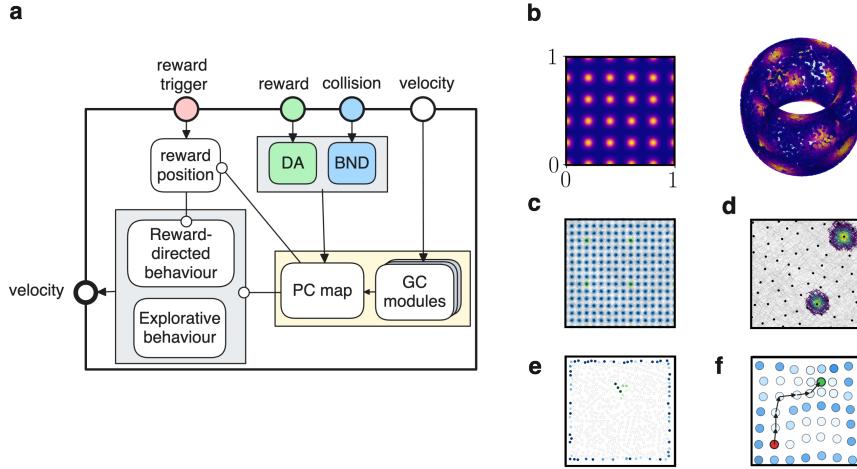


Figure 1: **Model layout and spatial representations** - **a:** the full architecture of the model, consisting of three main sensory input, targeting the two modulators and the cognitive map module, and the executive components, represented by a policy module, two behavioral programs and a reward receiver. **b:** a module of grid cells defined in a bounded square space of length 1, and an activity representation of their receptive field over a torus. **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. **f:** the place cells layer can be regarded as a graph with values assigned to each node according to the modulation strength; a path-finding algorithm can then be used to connect any two nodes taking into account the node values.

105 **Neuromodulation** Neuromodulators deliver event signals: rewards, denoted DA (for dopamine),  
 106 and boundary collisions, denoted BND. They are driven by binary inputs and are defined through a  
 107 leaky variable with exponential decay.

108 To remain resilient to environmental changes (e.g., moving rewards), the model uses a predictive  
 109 mechanism to correct keep internal representations updated. Each modulator  $k$  updates synaptic  
 110 weights to place cells through Hebbian plasticity:

$$\Delta \mathbf{W}^k = \eta^k \mathbf{u}^{PC} \left( \mathbf{v}^k - \mathbf{W}^k \right) \quad (1)$$

111 The term in brackets can be regarded as an error, implementing a simple form of predictive coding  
112 and is inspired by temporal-difference learning [54], aligning with evidence that neuromodulatory  
113 systems signal prediction errors and update beliefs [55, 34, 56].

114 Weight vectors are constrained to remain non-negative. Reward modulation tags cells near rewarded  
115 locations, while boundary modulation builds a representation of environmental edges. These scalar  
116 fields form the core of the cognitive map (Fig. 1e). See Appendix for full learning rules and parameter  
117 settings.

118 **Modulation of Place Fields** We further tested whether neuromodulators could directly alter spatial  
119 tuning. Place fields were dynamically shifted and resized based on recent salience signals.

120 Following a salient event (reward or collision), place field centers were displaced in grid cell space,  
121 with magnitude scaled by the neuromodulator  $v^k$  and proximity to the event:

$$\Delta \mathbf{W}_i^{GC,PC} = c^k v^k \varphi_{\sigma^k}(\mathbf{u}^{GC} - \mathbf{W}_i^{GC,PC}) \quad (2)$$

122 Here,  $\varphi_\sigma$  is a Gaussian function, and  $c^k$  a scaling factor. This rule is inspired by BTSP plasticity  
123 [16, 47], which shifts CA1 place fields following salient experiences. Lateral inhibition prevents field  
124 overlap during remapping.

125 In addition to dislocation, field size was modulated by scaling the gain of recently active neurons.  
126 Such mechanism allows neuromodulators to transiently enhance or suppress spatial sensitivity for  
127 specific cells. This modulation rule involves the gain  $\beta_i$  of each cell being adjusted proportionally to  
128 its activity trace  $m_i$ , a reference gain constant  $\bar{\beta}$ , and a modulatory scaling variable:

$$\beta_i = c_a^k m_i \bar{\beta} + (1 - m_i) \bar{\beta} \quad (3)$$

129 where  $c_a^k$  is a scaling gain parameter, for which a value of 1 signifies that no modulation takes place.

130 **Policy and Behavior** To evaluate the model’s utility in navigation, we implemented a simple policy  
131 toggling between exploration and reward-seeking behavior, depending on an external reward trigger  
132 and the internal map.

133 Exploration consisted of two possible strategies: a random walk, for purely stochastic movements,  
134 and periodic goal-directed navigation towards a random but visited location, aimed at preventing  
135 stagnation. In contrast, exploitation—defined as reward-directed navigation—involved identifying  
136 the reward location within the cognitive map. This location corresponded to the average position  
137 of DA-modulated place cells, reflecting mechanisms such as hippocampal replay and value-based  
138 navigation [57, 58, 59]. A graph-based pathfinding algorithm was then used to compute the route  
139 from the current location to the reward. In this graph, place cells served as nodes, while synaptic  
140 connections acted as edges. Additionally, a cost function was introduced over the graph nodes,  
141 assigning lower values to BND-modulated cells. This was designed to discourage proximity to the  
142 environment’s walls, as shown in plot 1f.

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

Model	Architecture	Training method	Ext. C.
Banino et al. [36]	LSTM + linear layers + CNN	BPTT and deep RL, supervised	Yes
Cueva et al. [38]	RNN + linear layers	Hessian-free algorithm with regularization	Yes
Sorcher et al. [60]	RNN + linear layers	Backpropagation with regularization	Yes
Whittington et al. [35]	Attractor network and deep networks	Backpropagation and Hebbian learning	No
de-Cothi et al. [34]	Successor representation	TD-learning + eligibility traces	Yes
Brozsko et al. [61]	Spike Response Model	Online modulated Hebbian plasticity	Yes
<b>Ours</b>	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
<b>Ours</b>	Goal navigation	Velocity, reward, collision	Action

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

143 **Comparison with previous architectures** Several previous computational models show structural  
 144 and conceptual similarities with the present work. A prominent category among them employs deep  
 145 neural networks—often with recurrent components—and relies on gradient-based learning strategies  
 146 such as backpropagation through time. These models typically require multiple training episodes or  
 147 large datasets for convergence. In contrast, our model adopts biologically inspired, synaptically local  
 148 plasticity rules, and requires only a single training episode for adaptation.

149 Other models utilize spiking neurons [61] or explicit neural representations [34], and incorporate  
 150 online learning rules more closely aligned with ours. These models also focus more directly on  
 151 goal-directed navigation, in contrast to purely path integration tasks. However, both of these rely on  
 152 external spatial coordinates to represent current position. Our model instead constructs an internal  
 153 coordinate system by integrating its own velocity output, enabling endogenous spatial tracking.

154 **Naturalistic task** The model was evaluated on a biologically inspired navigation benchmark  
 155 involving exploration and goal-seeking behavior in closed environments. Performance was measured  
 156 as the total number of rewards collected over multiple trials.

157 Optimization of the model hyper-parameters was carried out using the evolutionary Covariance-Matrix  
 158 Adaptation strategy (CMA-ES) [62] with a population size of 90. The choice of the hyper-parameters  
 159 to evolved took into account the type of dynamcis for which it was difficult to select pre-defined  
 160 values; the number of evolved variables was 10. The use of such method derived from the impracticility  
 161 of other optimization algorithms, also given the non-differentiability of several dynamic. Viable  
 162 alternatives were based on reinforcement learning, Bayesian and grid search, but evolution was more  
 163 appealing for exploration and visualization of the parameter space exploration. For evaluating the

164 individuals it was used a multi-objective fitness function: maximization of the reward count, and  
 165 minimization of the collision count from the time the reward position was discovered.

### 166 3 Results

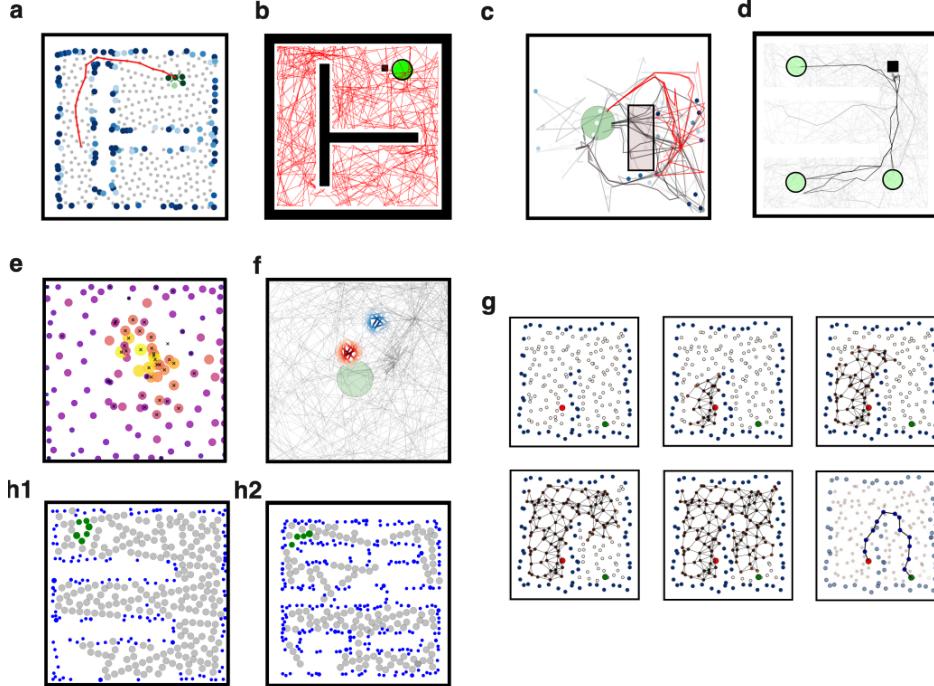


Figure 2: **Cognitive maps and performance results** - **a**: a cognitive map over a space, together with the plan (red line) to reach a target location from a starting position.. **b**: the same environment but with the reward (green circle), trajectory (red line), agent position (black square). - **c**: plot of trajectories before (black) and after (red) the insertion of a wall (rectangle) between the starting and goal positions, the wall can also be spotted from the boundary cells in blue - **d**: trajectories for multiple trials with the agent starting at the same position (black square) but with the reward location (green circles) periodically moving - **e**: place cells centers with size proportional to their node degree - **f**: place fields of the same cell before and after several relocation of its center following reward events - **g**: visualization of part of the path-finding algorithm, propagation of an activity wave through the place cells network from top-left to bottom-center, and the calculated path visualization in the bottom-right. - **h1-h2**: place cells centers with size inversely proportional to their gain value; in blue boundary cells with the highest average gain, in green reward cells with second smallest gain, the others in grey.

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

172 The left panel of plot 2a-b displays place cells associated with collisions and reward events, signaling  
 173 boundaries (in blue) and reward (in green) locations. The overlap of these two representations and the  
 174 non-modulated place cells (in grey) is what we refer to as a cognitive map, since these are the main  
 175 sources of spatial and contextual information used during planned navigation, whose path is depicted  
 176 as a gray line. The right panel instead portrays the actual environment with walls (black), reward  
 177 location (green), and multiple trajectories (red). During exploration, the main areas were visited  
 178 until the reward position was located and the goal-directed navigation dominated, as highlighted

179 by the density of the path lines. Considering the position of the walls and corners, the layout of  
180 this environment does not always make the target locations visible, as it is a non-convex area and  
181 therefore can be classified as wayfinding [63]. The challenge of not being able to use straight lines  
182 is overcome by the graph approach using local data and the consideration of boundary place cells,  
183 allowing the agent to plan accordingly. In addition, considering the Gaussian receptive fields and  
184 the approximately homogenous distribution of place field centers supported by lateral inhibition, the  
185 calculated path accounting for node-length also implicitly minimizes effective path-length, although  
186 not necessarily exactly. Figure 2g visualizes part of the path-finding process.

187 In general, this result confirms the ability of the model to focus on navigation and obstacle avoidance.  
188 However, it is worth noting that not all simulations resulted in a reward being found in the first place,  
189 due to the randomness of the exploratory process; this effect was more pronounced in environment  
190 with more walls and narrow passages.

191 **Detour task** The planning ability and the plastic nature of the cognitive map should provide re-  
192 silience against unexpected changes in the environment layout. In order to verify this we implemented  
193 a detour experiment. Initially, the agent was familiarized with a square environment with the reward  
194 in the middle and starting always from the same position. Then, a wall was placed in between the  
195 starting position and the reward, therefore forcing new trajectory for reaching it. As expected, the  
196 agent was able to form a representation of the new obstacle and calculating new paths around it,  
197 succeeding the task. In plot 2c they are shown the trajectories before and after the wall placement,  
198 and it is manifested the ability of detour in the new layout.

199 **Adaptive goal representation through sensory error** Then, we tested the adaptability to environ-  
200 mental changes. In this scenario, the reward object was moved after being fetched a fixed number  
201 of times. Here, the difficulty was to unlearning previous locations and discovering new ones, in a  
202 protocol similar to [39]. In plot 2d is reported the set of trajectories over many trials with the reward  
203 displaced in three possible locations. The agent was capable of planning behavior, as earlier, but  
204 also exploring and finding the new rewards, as shown by the density of lines. Whenever a goal path  
205 resulted in a failed prediction, the DA-based sensory error weakened the association between the  
206 place cells and the reward signal, leading to an extinction of its representation at that location.

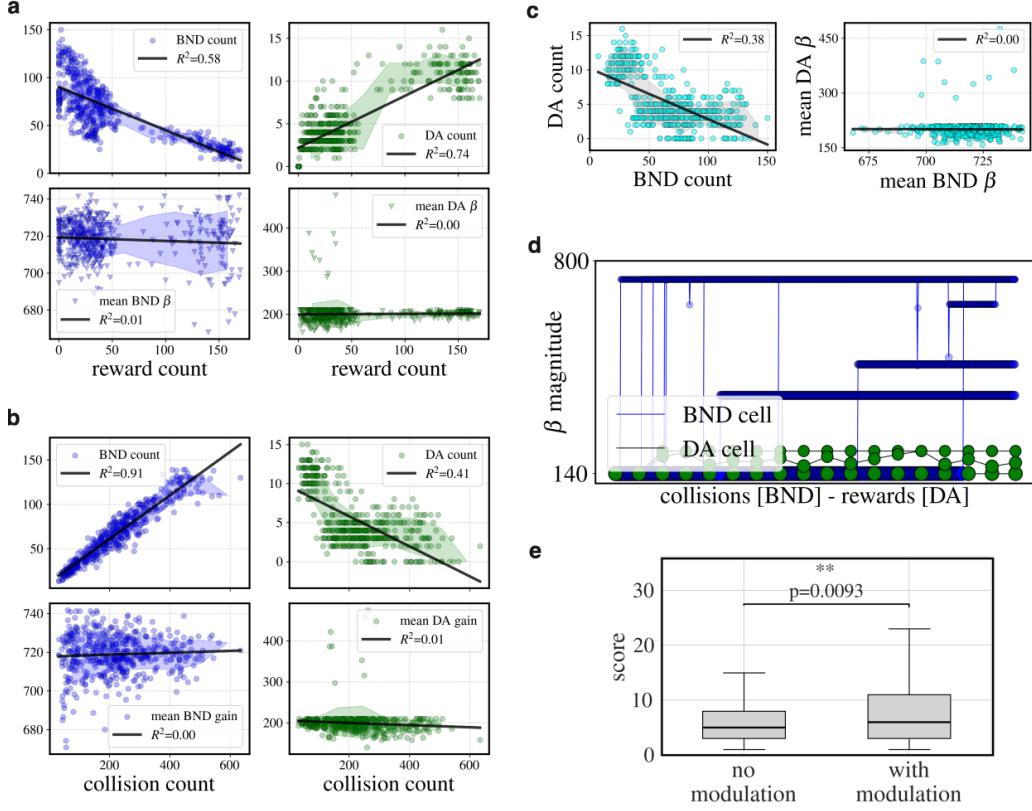
207 This result validates the resilience of the model to changing sensory expectations, in this case the  
208 reward position.

209 **Modulation of place field size** The construction of model is such that the experience of environ-  
210 mental events can impact the neuronal properties of the generated place cells. In particular, collision  
211 and reward events have the effect of affecting the neural activation gain  $\beta$  of BND and DA-modulated  
212 cells through an hyperparameter  $\beta$ . The hyperparameter values  $c_a^{\text{BND}}$ ,  $c_a^{\text{DA}}$  that yielded the best results  
213 were both larger than 1., meaning a shrinking of field size. In plots 2h1-h2 are showed the cognitive  
214 maps with relative place field sizes for two environments, showcasing the differences between bound-  
215 ary, reward, and non-modulated cells. in plots 3d is showed the evolution of the gain magnitude for  
216 a sample of boundary cells and reward cells over their corresponding modulatory events. Notable  
217 is the possible decrease in value, case that occurs when a cell is active but the modulation is absent  
218 and the gain thus is pushed down towards the baseline beta. This feature can be considered another  
219 adaptation property.

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

224 Our working hypothesis is that these experience-driven neuronal changes would improve the quality  
225 of the cognitive map and be reflected in navigational abilities. The assessment of this claim was  
226 conducted by comparing two variants of the model: with or without the modulation mechanisms,  
227 namely the density and gain modulation from reward and collision events.. Both models were ran in  
228 the same environment, with numerous wall, for a total of 128 simulations for each case.

229 The statistical results shown in the box plot of Figure 2g demonstrate a significant difference between  
230 the two groups, supporting the important of place field modulation.



**Figure 3: Cognitive maps and performance results - a:** effect of reward count on reward and boundary modulated cells (green and blue respectively), both in total count (top row) and average gain modulation magnitude (bottom row); simulation of 512 independent runs. - **b:** similar plot but with respect to collision count. - **c:** relation between count of reward and boundary modulated cells, and between gain modulation magnitude. - **d:** gain magnitude of boundary and reward cells over sequences of collision and reward events respectively - **e:** performance comparison for the same models on the same environment with and without modulation enabled. Pair-wise t-test over 128 iterations and Bonferroni correction

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

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

237 **Convergence of evolved hyperparameters** The use of an evolutionary algorithm for selecting  
238 the model dynamics had the convenience of providing a distribution of hyper-parameters over a  
239 population. In Figure 4, each dot represent the the value for one individual plotted with respect to  
240 its average reward count. A marked convergence is visible for most variables, with some displaying  
241 a quasi-bimodal distribution. The hyper-parameters related to the gain  $\beta$  of the neural activation  
242 function  $\phi$  have an high clustering index. Importantly, there is a strong tendency of having modulation  
243 increasing the magnitude of the gain for both reward and collision events, with the effect of reducing  
244 the size the active place fields. This result is in the direction of experimental observation showing  
245 shrinking of place fields near objects and walls [64, 65].

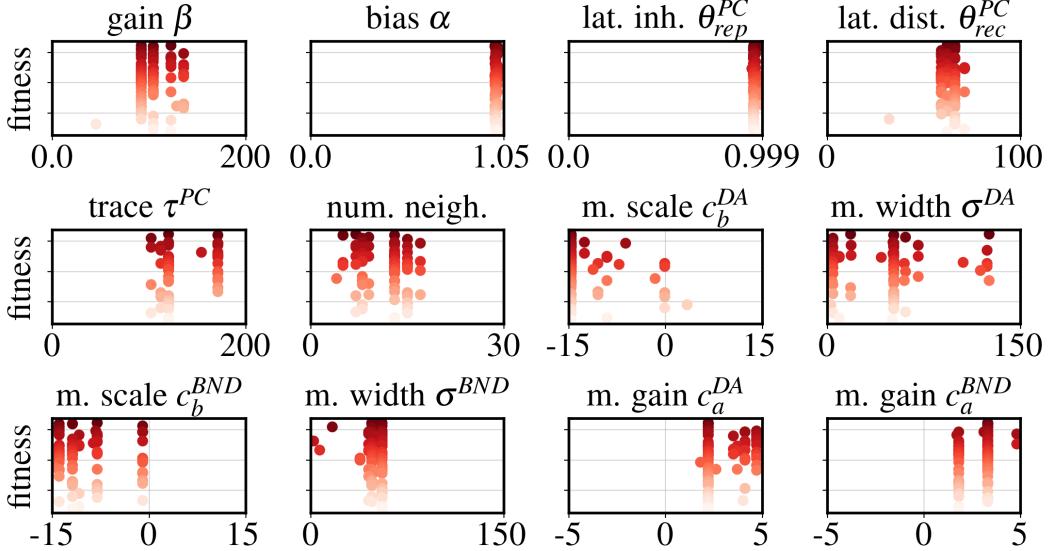


Figure 4: **Distribution of evolved hyper-parameters** - Results relative to the last generation, from a run with population size of 90 individuals. The hyper-parameters are: neural gain  $\beta$ , lateral inhibition threshold  $\theta_{inh}^{PC}$ , lateral distance threshold  $\theta_{rec}^{PC}$ , activity trace time constant  $\tau^{PC}$ , reward modulation scale  $c_b^{DA}$ , reward modulation spread  $c_b^{DA}$ , boundary modulation scale  $c_b^{BND}$ , boundary modulation spread  $c_b^{BND}$ , reward gain modulation  $c_a^{DA}$ , and boundary gain modulation  $c_a^{BND}$ .

## 246 4 Discussion

247 Exploration and planning in the known and past environment are essential behaviors of animals,  
248 directly affecting their success in world understanding and goal reaching.

249 An important element behind these abilities is the formation of a map of their surroundings as they  
250 make new experiences, known as a cognitive map. Numerous speculations have been made about the  
251 shape and neural foundations of such an object, varying in the types of modeling assumptions and  
252 experimental support.

253 The contribution of the present work was to propose a rate network model, inspired by the CA1  
254 hippocampal region [10]. We used grid cells together with synaptic plasticity as a mechanism to  
255 develop information-rich representations based on place cells updated through experience, grouped  
256 with common perspectives on cognitive maps [66]. In the spirit of minimizing the geometric  
257 assumptions in the neural space, we treated the generated place network as a topological graph, with  
258 sensory information added locally through the action of neuromodulators. This idea aligned with the  
259 concept of a *labeled graph* [67, 26], however, it is also true that no metric violations were possible in  
260 these settings.

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

266 The simulation results validated the model, showing the expected emergence of cognitive maps and  
267 their encoding of information collected during the experience. The online nature of the formation  
268 of the locations on the map aligns with the idea of using only idiothetic velocity input, as in path  
269 integration [24, 68, 69]. Previous work followed a similar direction using recurrent networks, but  
270 required extensive gradient-based training [60, 38, 51]. Another important difference is that our  
271 resulting neural network was composed solely of place cells, although neuromodulated, and no other  
272 types of neuron were present. This distinction is justified by the partially different task structure, which  
273 did not involve supervised learning and did not receive visual information as in [36]. Furthermore,

274 our model relied on predefined grid cells layers, which constituted a strong and sufficient inductive  
275 bias, and did not have to be learned from scratch.

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

282 Adaptability was tested by occasionally moving the reward position, leading to the generation of  
283 an internal prediction error that was used to update its representation on the map. The agent was  
284 proved capable of unlearning previous associations, returning to exploration, and memorizing new  
285 reward locations. This behavioral protocol is similar to previous work [61], in which dopaminergic  
286 and cholinergic activity was utilized within a Hebbian plasticity rule to strengthen or weaken reward-  
287 associated spatial representations. However, alternatively to exploiting neuromodulators with opposite  
288 valence, we followed a predictive coding framework, a direction linked to hippocampal representations  
289 [34, 70] and explored various computational approaches [71, 72, 73]. This choice departed from our  
290 focus on using operations on the cognitive map itself by simulating future sensory experiences and  
291 learning from feedback. In fact, neuromodulation has been long associated with this functionality  
292 [40], especially dopamine [14, 43, 46, 18].

293 Lastly, the hypothesis of the relevance of the active modulation of the neuronal properties of place  
294 cells was corroborated by simulating ablation experiments. These tests reported a significant impact  
295 of altering the place cells density on the total count of collected rewards. In general, these results are  
296 consistent with the experimental observations of alteration of place cells following reward events  
297 [16, 74], in particular in terms of increased clustering of cells [75, 76], reminiscent of changes in  
298 firing rate after contextual changes [77, 78].

299 Concerning the modulation of place fields, there is significant experimental evidence of their alter-  
300 nation during reward events [79, 80, 81], some reporting shrinkage near reward objects [64], and  
301 boundaries [65]. The coupling with higher local density could be explained by better optimization of  
302 the cell distribution for goal representation and planning [82]. However, in our settings, the fields  
303 become enlarged, especially in the direction of the target, although the performance improvements  
304 were not tested significantly. A possible explanation can be the simplicity of our reward, which was  
305 solely defined as an area of space. The lack of rich non-spatial features thus did not require the place  
306 cells to code for smaller spatial variation. Therefore, enlargement might have improved the stability  
307 of the representation, marking the nodes associated with rewards more solidly, given the stochas-  
308 ticity of its delivery. Further, the graph-path algorithm utilized the strength of the DA-modulated  
309 connections for determining the goal representation; stronger fields inherently developed stronger  
310 weights, making planning more reliable. Although these findings are limited within the limits of our  
311 simulation protocol, there have been experimental observations of elongation of place fields along  
312 trajectories over meaningful experiences [83, 84].

313 In conclusion, this work showed a possible architecture for coupling emergent spatial representations  
314 with neuromodulated plasticity to achieve an experience-driven cognitive map. The reliance on a few  
315 spatial and algorithmic inductive biases, grid cells, and a planning algorithm supports the idea of a  
316 label graph for goal navigation. Future work can investigate the application to other spatial domains,  
317 such as motor control and three-dimensional navigation. In addition, a richer input feature can be  
318 added, such as visual information [85], as well as new neuromodulators that encode different sensory  
319 dimensions or internally generated signals.

320

321

### 322 Acknowledgements & Statements

323 The authors declare no competing interests.

324 The code is publicly available and can be found at <https://github.com/iKiru-hub/PCNN>.

325 This research was funded by the European Union’s Horizon 2020 research and innovation programme  
326 under the Marie Skłodowska-Curie grant agreement N° 945371 and the University of Oslo.

327 The research presented in this paper has benefited from the Experimental Infrastructure for Exploration

328 of Exascale Computing (eX3), which is financially supported by the Research Council of Norway  
329 under contract 270053.

### 330 References

- 331 [1] Reginald Golledge, Dan Jacobson, Rob Kitchin, and Mark Blades. Cognitive Maps, Spatial  
332 Abilities, and Human Wayfinding. *GEOGRAPHICAL REVIEW OF JAPAN SERIES B*, 73:93–  
333 104, December 2000.
- 334 [2] Russell A. Epstein and Lindsay K. Vass. Neural systems for landmark-based wayfind-  
335 ing in humans. *Philosophical Transactions of the Royal Society B: Biological Sciences*,  
336 369(1635):20120533, February 2014.
- 337 [3] Francesca Sargolini, Marianne Fyhn, Torkel Hafting, Bruce L. McNaughton, Menno P. Witter,  
338 May-Britt Moser, and Edvard I. Moser. Conjunctive Representation of Position, Direction, and  
339 Velocity in Entorhinal Cortex. *Science*, 312(5774):758–762, May 2006.
- 340 [4] Emilio Kropff, James E. Carmichael, May-Britt Moser, and Edvard I. Moser. Speed cells in the  
341 medial entorhinal cortex. *Nature*, 523(7561):419–424, July 2015.
- 342 [5] Trygve Solstad, Edvard I. Moser, and Gaute T. Einevoll. From grid cells to place cells: A  
343 mathematical model. *Hippocampus*, 16(12):1026–1031, 2006.
- 344 [6] Daniel Bush, Caswell Barry, and Neil Burgess. What do grid cells contribute to place cell firing?  
345 *Trends in Neurosciences*, 37(3):136–145, March 2014.
- 346 [7] Torsten Neher, Amir Hossein Azizi, and Sen Cheng. From grid cells to place cells with realistic  
347 field sizes. *PLOS ONE*, 12(7):e0181618, July 2017.
- 348 [8] Tianyi Li, Angelo Arleo, and Denis Sheynikhovich. *Modeling Place Cells and Grid Cells in  
349 Multi-Compartment Environments: Hippocampal-Entorhinal Loop as a Multisensory Integra-  
350 tion Circuit*. April 2019.
- 351 [9] Olesia M. Bilash, Spyridon Chavlis, Cara D. Johnson, Panayiota Poirazi, and Jayeeta Basu.  
352 Lateral entorhinal cortex inputs modulate hippocampal dendritic excitability by recruiting a  
353 local disinhibitory microcircuit. *Cell Reports*, 42(1):111962, January 2023.
- 354 [10] Flavio Donato, Anja Xu Schwartzlose, and Renan Augusto Viana Mendes. How Do You Build  
355 a Cognitive Map? The Development of Circuits and Computations for the Representation of  
356 Space in the Brain. *Annual Review of Neuroscience*, 46(Volume 46, 2023):281–299, July 2023.
- 357 [11] John E. Lisman and Anthony A. Grace. The Hippocampal-VTA Loop: Controlling the Entry of  
358 Information into Long-Term Memory. *Neuron*, 46(5):703–713, June 2005.
- 359 [12] Adrian J. Duszkiewicz, Colin G. McNamara, Tomonori Takeuchi, and Lisa Genzel. Novelty  
360 and Dopaminergic Modulation of Memory Persistence: A Tale of Two Systems. *Trends in  
361 Neurosciences*, 42(2):102–114, February 2019.
- 362 [13] Wolfram Schultz, Peter Dayan, and P. Read Montague. A Neural Substrate of Prediction and  
363 Reward. *Science*, 275(5306):1593–1599, March 1997.
- 364 [14] Kimberly A. Kempadoo, Eugene V. Mosharov, Se Joon Choi, David Sulzer, and Eric R.  
365 Kandel. Dopamine release from the locus coeruleus to the dorsal hippocampus promotes spatial  
366 learning and memory. *Proceedings of the National Academy of Sciences*, 113(51):14835–14840,  
367 December 2016.
- 368 [15] Aude Retailleau and Thomas Boraud. The Michelin red guide of the brain: Role of dopamine  
369 in goal-oriented navigation. *Frontiers in Systems Neuroscience*, 8, March 2014.
- 370 [16] Katie C. Bittner, Aaron D. Milstein, Christine Grienberger, Sandro Romani, and Jeffrey C.  
371 Magee. Behavioral time scale synaptic plasticity underlies CA1 place fields. *Science*,  
372 357(6355):1033–1036, September 2017.

- 373 [17] Alexandra Mansell Kaufman, Tristan Geiller, and Attila Losonczy. A Role for the Locus  
374 Coeruleus in Hippocampal CA1 Place Cell Reorganization during Spatial Reward Learning.  
375 *Neuron*, 105(6):1018–1026.e4, March 2020.
- 376 [18] Denis Sheynikhovich, Satoru Otani, Jing Bai, and Angelo Arleo. Long-term memory, synaptic  
377 plasticity and dopamine in rodent medial prefrontal cortex: Role in executive functions. *Frontiers*  
378 in *Behavioral Neuroscience*, 16, January 2023.
- 379 [19] Kei M. Igarashi, Hiroshi T. Ito, Edvard I. Moser, and May-Britt Moser. Functional diversity  
380 along the transverse axis of hippocampal area CA1. *FEBS Letters*, 588(15):2470–2476, August  
381 2014.
- 382 [20] Hiroshi T. Ito and Erin M. Schuman. Functional division of hippocampal area CA1 via  
383 modulatory gating of entorhinal cortical inputs. *Hippocampus*, 22(2):372–387, 2012.
- 384 [21] Michael Peer, Iva K. Brunec, Nora S. Newcombe, and Russell A. Epstein. Structuring Knowl-  
385 edge with Cognitive Maps and Cognitive Graphs. *Trends in cognitive sciences*, 25(1):37–54,  
386 January 2021.
- 387 [22] Elizabeth R. Chrastil and William H. Warren. From Cognitive Maps to Cognitive Graphs. *PLoS*  
388 *ONE*, 9(11):e112544, November 2014.
- 389 [23] Steffen Werner, Bernd Krieg-Brückner, and Theo Herrmann. Modelling Navigational Knowl-  
390 edge by Route Graphs. In Christian Freksa, Christopher Habel, Wilfried Brauer, and Karl F.  
391 Wender, editors, *Spatial Cognition II: Integrating Abstract Theories, Empirical Studies, Formal*  
392 *Methods, and Practical Applications*, pages 295–316. Springer, Berlin, Heidelberg, 2000.
- 393 [24] C. R. Gallistel and Audrey E. Cramer. Computations on Metric Maps in Mammals: Getting  
394 Oriented and Choosing a Multi-Destination Route. *Journal of Experimental Biology*, 199(1):211–  
395 217, January 1996.
- 396 [25] Michael Peer, Catherine Nadar, and Russell A. Epstein. The format of the cognitive map  
397 depends on the structure of the environment. *Journal of Experimental Psychology: General*,  
398 153(1):224–240, January 2024.
- 399 [26] William H. Warren. Non-Euclidean navigation. *Journal of Experimental Biology*,  
400 222(Suppl\_1):jeb187971, February 2019.
- 401 [27] Mark Wagner. Comparing the psychophysical and geometric characteristics of spatial perception  
402 and cognitive maps. *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society*,  
403 15(1):6–21, 2008.
- 404 [28] Rainer Rothkegel, Karl F. Wender, and Sabine Schumacher. Judging Spatial Relations from  
405 Memory. In Christian Freksa, Christopher Habel, and Karl F. Wender, editors, *Spatial Cognition:*  
406 *An Interdisciplinary Approach to Representing and Processing Spatial Knowledge*, pages 79–  
407 105. Springer, Berlin, Heidelberg, 1998.
- 408 [29] Tobias Meilinger. The Network of Reference Frames Theory: A Synthesis of Graphs and  
409 Cognitive Maps. In Christian Freksa, Nora S. Newcombe, Peter Gärdenfors, and Stefan Wölfl,  
410 editors, *Spatial Cognition VI. Learning, Reasoning, and Talking about Space*, pages 344–360,  
411 Berlin, Heidelberg, 2008. Springer.
- 412 [30] Jane X. Wang, Zeb Kurth-Nelson, Dhruva Tirumala, Hubert Soyer, Joel Z. Leibo, Remi Munos,  
413 Charles Blundell, Dharshan Kumaran, and Matt Botvinick. Learning to reinforcement learn,  
414 January 2017.
- 415 [31] Victor R. Schinazi, Daniele Nardi, Nora S. Newcombe, Thomas F. Shipley, and Russell A.  
416 Epstein. Hippocampal size predicts rapid learning of a cognitive map in humans. *Hippocampus*,  
417 23(6):515–528, 2013.
- 418 [32] Bruno Poucet. Spatial cognitive maps in animals: New hypotheses on their structure and neural  
419 mechanisms. *Psychological Review*, 100(2):163–182, 1993.

- 420 [33] Paul Stoewer, Achim Schilling, Andreas Maier, and Patrick Krauss. Neural network based  
421 formation of cognitive maps of semantic spaces and the putative emergence of abstract concepts.  
422 *Scientific Reports*, 13(1):3644, March 2023.
- 423 [34] William de Cothi, Nils Nyberg, Eva-Maria Griesbauer, Carole Ghanamé, Fiona Zisch, Julie M.  
424 Lefort, Lydia Fletcher, Coco Newton, Sophie Renaudineau, Daniel Bendor, Roddy Grieves,  
425 Éléonore Duvelle, Caswell Barry, and Hugo J. Spiers. Predictive maps in rats and humans for  
426 spatial navigation. *Current Biology*, 32(17):3676–3689.e5, September 2022.
- 427 [35] James C. R. Whittington, Timothy H. Muller, Shirley Mark, Guifen Chen, Caswell Barry, Neil  
428 Burgess, and Timothy E. J. Behrens. The Tolman-Eichenbaum Machine: Unifying Space and  
429 Relational Memory through Generalization in the Hippocampal Formation. *Cell*, 183(5):1249–  
430 1263.e23, November 2020.
- 431 [36] Andrea Banino, Caswell Barry, Benigno Uria, Charles Blundell, Timothy Lillicrap, Piotr  
432 Mirowski, Alexander Pritzel, Martin J. Chadwick, Thomas Degris, Joseph Modayil, Greg  
433 Wayne, Hubert Soyer, Fabio Viola, Brian Zhang, Ross Goroshin, Neil Rabinowitz, Razvan  
434 Pascanu, Charlie Beattie, Stig Petersen, Amir Sadik, Stephen Gaffney, Helen King, Koray  
435 Kavukcuoglu, Demis Hassabis, Raia Hadsell, and Dharshan Kumaran. Vector-based navigation  
436 using grid-like representations in artificial agents. *Nature*, 557(7705):429–433, May 2018.
- 437 [37] Ben Sorscher, Gabriel C. Mel, Samuel A. Ocko, Lisa M. Giocomo, and Surya Ganguli. A unified  
438 theory for the computational and mechanistic origins of grid cells. *Neuron*, 111(1):121–137.e13,  
439 January 2023.
- 440 [38] Christopher J. Cueva and Xue-Xin Wei. Emergence of grid-like representations by training  
441 recurrent neural networks to perform spatial localization, March 2018.
- 442 [39] Zuzanna Brzozko, Susanna B. Mierau, and Ole Paulsen. Neuromodulation of Spike-Timing-  
443 Dependent Plasticity: Past, Present, and Future. *Neuron*, 103(4):563–581, August 2019.
- 444 [40] Marielena Sosa, Mark H. Plitt, and Lisa M. Giocomo. Hippocampal sequences span experience  
445 relative to rewards. *bioRxiv*, page 2023.12.27.573490, February 2024.
- 446 [41] Abdullahi Ali, Nasir Ahmad, Elgar de Groot, Marcel A. J. van Gerven, and Tim C. Kietzmann.  
447 Predictive coding is a consequence of energy efficiency in recurrent neural networks, November  
448 2021.
- 449 [42] Jacopo Bono, Sara Zannone, Victor Pedrosa, and Claudia Clopath. Learning predictive cognitive  
450 maps with spiking neurons during behavior and replays. *eLife*, 12:e80671, March 2023.
- 451 [43] Wolfram Schultz. Dopamine reward prediction error coding. *Dialogues in Clinical Neuroscience*,  
452 18(1):23–32, March 2016.
- 453 [44] Jeffrey B. Inglis, Vivian V. Valentin, and F. Gregory Ashby. Modulation of Dopamine for  
454 Adaptive Learning: A Neurocomputational Model. *Computational brain & behavior*, 4(1):34–  
455 52, March 2021.
- 456 [45] Philippe N. Tobler, Christopher D. Fiorillo, and Wolfram Schultz. Adaptive Coding of Reward  
457 Value by Dopamine Neurons. *Science*, 307(5715):1642–1645, March 2005.
- 458 [46] Roshan Cools. Chemistry of the Adaptive Mind: Lessons from Dopamine. *Neuron*, 104(1):113–  
459 131, October 2019.
- 460 [47] Aaron D Milstein, Yiding Li, Katie C Bittner, Christine Grienberger, Ivan Soltesz, Jeffrey C  
461 Magee, and Sandro Romani. Bidirectional synaptic plasticity rapidly modifies hippocampal  
462 representations. *eLife*, 10:e73046, December 2021.
- 463 [48] André A. Fenton. Remapping revisited: How the hippocampus represents different spaces.  
464 *Nature Reviews Neuroscience*, 25(6):428–448, June 2024.
- 465 [49] Luxin Zhou and Yong Gu. Cortical Mechanisms of Multisensory Linear Self-motion Perception.  
466 *Neuroscience Bulletin*, 39(1):125–137, July 2022.

- 467 [50] Steven J. Jerjian, Devin R. Harsch, and Christopher R. Fetsch. Self-motion perception and  
468 sequential decision-making: Where are we heading? *Philosophical Transactions of the Royal*  
469 *Society B: Biological Sciences*, 378(1886):20220333, August 2023.
- 470 [51] Ian Q. Whishaw and Brian L. Brooks. Calibrating space: Exploration is important for allothetic  
471 and idiothetic navigation. *Hippocampus*, 9(6):659–667, 1999.
- 472 [52] Yuri Dabaghian. Grid Cells, Border Cells and Discrete Complex Analysis.
- 473 [53] Vemund Sigmundson Schøyen, Kosio Beshkov, Markus Borud Pettersen, Erik Hermansen,  
474 Konstantin Holzhausen, Anders Malthe-Sørensen, Marianne Fyhn, and Mikkel Elle Lep-  
475 perød. Hexagons all the way down: Grid cells as a conformal isometric map of space. *PLOS*  
476 *Computational Biology*, 21(2):e1012804, February 2025.
- 477 [54] Richard S Sutton and Andrew G Barto. The Reinforcement Learning Problem.
- 478 [55] P. R. Montague, P. Dayan, and T. J. Sejnowski. A framework for mesencephalic dopamine  
479 systems based on predictive Hebbian learning. *Journal of Neuroscience*, 16(5):1936–1947,  
480 March 1996.
- 481 [56] Seetha Krishnan, Chad Heer, Chery Cherian, and Mark E. J. Sheffield. Reward expectation  
482 extinction restructures and degrades CA1 spatial maps through loss of a dopaminergic reward  
483 proximity signal. *Nature Communications*, 13(1):6662, November 2022.
- 484 [57] Colin G. McNamara, Álvaro Tejero-Cantero, Stéphanie Trouche, Natalia Campo-Urriza, and  
485 David Dupret. Dopaminergic neurons promote hippocampal reactivation and spatial memory  
486 persistence. *Nature Neuroscience*, 17(12):1658–1660, December 2014.
- 487 [58] Frédéric Michon, Esther Krul, Jyh-Jang Sun, and Fabian Kloosterman. Single-trial dynamics  
488 of hippocampal spatial representations are modulated by reward value. *Current biology: CB*,  
489 31(20):4423–4435.e5, October 2021.
- 490 [59] Philip Shamash and Tiago Branco. Mice identify subgoal locations through an action-driven  
491 mapping process, December 2021.
- 492 [60] Ben Sorscher, Gabriel Mel, Surya Ganguli, and Samuel Ocko. A unified theory for the origin of  
493 grid cells through the lens of pattern formation. In *Advances in Neural Information Processing  
494 Systems*, volume 32. Curran Associates, Inc., 2019.
- 495 [61] Zuzanna Brzosko, Sara Zannone, Wolfram Schultz, Claudia Clopath, and Ole Paulsen. Se-  
496 quential neuromodulation of Hebbian plasticity offers mechanism for effective reward-based  
497 navigation. *eLife*, 6:e27756, 2017.
- 498 [62] Christian Igel, Nikolaus Hansen, and Stefan Roth. Covariance Matrix Adaptation for Multi-  
499 objective Optimization. *Evolutionary Computation*, 15(1):1–28, March 2007.
- 500 [63] Tobias Meilinger, Marianne Strickrodt, and Heinrich H. Bühlhoff. Qualitative differences in  
501 memory for vista and environmental spaces are caused by opaque borders, not movement or  
502 successive presentation. *Cognition*, 155:77–95, October 2016.
- 503 [64] S.N. Burke, A.P. Maurer, S. Nematollahi, A.R. Upadhyay, J.L. Wallace, and C.A. Barnes. The Influ-  
504 ence of Objects on Place Field Expression and Size in Distal Hippocampal CA1. *Hippocampus*,  
505 21(7):783–801, July 2011.
- 506 [65] Sander Tanni, William De Cothi, and Caswell Barry. State transitions in the statistically stable  
507 place cell population correspond to rate of perceptual change. *Current Biology*, 32(16):3505–  
508 3514.e7, August 2022.
- 509 [66] Vincent Hok, Pierre-Pascal Lenck-Santini, Sébastien Roux, Etienne Save, Robert U. Muller,  
510 and Bruno Poucet. Goal-Related Activity in Hippocampal Place Cells. *Journal of Neuroscience*,  
511 27(3):472–482, January 2007.
- 512 [67] Toru Ishikawa and Daniel R. Montello. Spatial knowledge acquisition from direct experience  
513 in the environment: Individual differences in the development of metric knowledge and the  
514 integration of separately learned places. *Cognitive Psychology*, 52(2):93–129, March 2006.

- 515 [68] Sabine Gillner and Hanspeter A. Mallot. Navigation and Acquisition of Spatial Knowledge in a  
 516 Virtual Maze. *Journal of Cognitive Neuroscience*, 10(4):445–463, July 1998.
- 517 [69] Bruce L. McNaughton, Francesco P. Battaglia, Ole Jensen, Edvard I. Moser, and May-Britt  
 518 Moser. Path integration and the neural basis of the ‘cognitive map’. *Nature Reviews Neuroscience*,  
 519 7(8):663–678, August 2006.
- 520 [70] Fraser Aitken and Peter Kok. Hippocampal representations switch from errors to predictions  
 521 during acquisition of predictive associations. *Nature Communications*, 13(1):3294, June 2022.
- 522 [71] Manu Srinath Halvagal and Friedemann Zenke. The combination of Hebbian and predictive  
 523 plasticity learns invariant object representations in deep sensory networks. *Nature Neuroscience*,  
 524 pages 1–10, October 2023.
- 525 [72] Alexander Ororbia. Spiking neural predictive coding for continually learning from data streams.  
 526 *Neurocomputing*, 544:126292, August 2023.
- 527 [73] Kimberly L Stachenfeld, Matthew M Botvinick, and Samuel J Gershman. The hippocampus as  
 528 a predictive map. *Nature Neuroscience*, 20(11):1643–1653, November 2017.
- 529 [74] Indrajith R. Nair, Guncha Bhasin, and Dipanjan Roy. Hippocampus Maintains a Coherent Map  
 530 Under Reward Feature–Landmark Cue Conflict. *Frontiers in Neural Circuits*, 16, April 2022.
- 531 [75] Valerie L. Tryon, Marsha R. Penner, Shawn W. Heide, Hunter O. King, Joshua Larkin, and  
 532 Sheri J. Y. Mizumori. Hippocampal neural activity reflects the economy of choices during  
 533 goal-directed navigation. *Hippocampus*, 27(7):743–758, July 2017.
- 534 [76] Hannah S Wirtshafter and Matthew A Wilson. Differences in reward biased spatial representa-  
 535 tions in the lateral septum and hippocampus. *eLife*, 9:e55252, May 2020.
- 536 [77] Michael I. Anderson and Kathryn J. Jeffery. Heterogeneous Modulation of Place Cell Firing by  
 537 Changes in Context. *Journal of Neuroscience*, 23(26):8827–8835, October 2003.
- 538 [78] Inah Lee, Amy L. Griffin, Eric A. Zilli, Howard Eichenbaum, and Michael E. Hasselmo. Gradual  
 539 Translocation of Spatial Correlates of Neuronal Firing in the Hippocampus toward Prospective  
 540 Reward Locations. *Neuron*, 51(5):639–650, September 2006.
- 541 [79] Marianne Fyhn, Sturla Molden, Stig Hollup, May-Britt Moser, and Edvard I. Moser. Hippocam-  
 542 pal Neurons Responding to First-Time Dislocation of a Target Object. *Neuron*, 35(3):555–566,  
 543 August 2002.
- 544 [80] P.-P. Lenck-Santini, B. Rivard, R.u. Muller, and B. Poucet. Study of CA1 place cell activity and  
 545 exploratory behavior following spatial and nonspatial changes in the environment. *Hippocampus*,  
 546 15(3):356–369, 2005.
- 547 [81] David Dupret, Joseph O’Neill, Barty Pleydell-Bouverie, and Jozsef Csicsvari. The reorgani-  
 548 zation and reactivation of hippocampal maps predict spatial memory performance. *Nature  
 549 Neuroscience*, 13(8):995–1002, August 2010.
- 550 [82] Pablo Scleidovich, Jean-Marc Fellous, and Alfredo Weitzenfeld. Adapting hippocampus  
 551 multi-scale place field distributions in cluttered environments optimizes spatial navigation and  
 552 learning. *Frontiers in Computational Neuroscience*, 16:1039822, December 2022.
- 553 [83] Mayank R. Mehta, Carol A. Barnes, and Bruce L. McNaughton. Experience-dependent,  
 554 asymmetric expansion of hippocampal place fields. *Proceedings of the National Academy of  
 555 Sciences*, 94(16):8918–8921, August 1997.
- 556 [84] Jangho Lee, Jeonghee Jo, Byounghwa Lee, Jung-Hoon Lee, and Sungroh Yoon. Brain-inspired  
 557 Predictive Coding Improves the Performance of Machine Challenging Tasks. *Frontiers in  
 558 Computational Neuroscience*, 16:1062678, 2022.
- 559 [85] John H. Wen, Ben Sorscher, Emily A. Aery Jones, Surya Ganguli, and Lisa M. Giocomo. One-  
 560 shot entorhinal maps enable flexible navigation in novel environments. *Nature*, 635(8040):943–  
 561 950, November 2024.

562 **5 Appendix**

563 **5.1 Grid cell module**

564 It is defined a correspondence between the global environment in which the agent moves, a two-  
 565 dimensional Euclidean space  $\mathbf{R}^2$ , and a bounded local space of a grid module, corresponding to a  
 566 torus.

567 The global velocity  $\mathbf{v} = \{x, y\}$  is then mapped to a local velocity, scaled by a speed scalar  $s_l^{gc}$  specific  
 568 to the grid cell module  $l$ , which determines its periodicity in space.

569 The choice of a toroidal space is motivated by consolidated experimental evidence of the neural space  
 570 of grid cells, which are organized in modules of different sizes spanning the animal's environment.  
 571 However, the shape of their firing pattern is known to be hexagonal, which corresponds to the optimal  
 572 tiling of a two-dimensional plane, giving rise to a neural space lying on a twisted torus. In this work,  
 573 for simplicity, we consider a square tiling and thus a square torus, without much loss of generality  
 574 except for the slight increase of grid cells required for a sufficiently cover.

575 A grid cell module  $l$  of size  $N^{gc}$  is identified by a set of positions defined over a square centered  
 576 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  
 577 space has boundary conditions for each dimension, such that, for instance, when  $x_t + s_l^{gc} \cdot v_x > 2$   
 578 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  
 579 velocity in the local space of the module  $l$  with respect to the real global agent velocity  $\mathbf{v} = \{v_x, v_y\}$ .  
 580 When the module is initialized, the starting positions of its cells are uniformly distributed over the  
 581 square forming a lattice. When the agent is reset in a new position at the beginning of new trial, a  
 582 displacement vector is applied to the last cells positions such that the mapping between the module  
 583 local space and the global environment is preserved.

584 The firing rate vector of each cell is determined with respect to a 2D Gaussian tuning curve centered  
 585 at the origin at  $(0, 0)$ , and it is calculated as

586  $r_i = \exp\left(-\frac{x_i^2+y_i^2}{\sigma_i^{gc}}\right)$ , where  $\sigma_i^{gc}$  is the width of the tuning curve for module  $l$ . An illustration of the  
 587 receptive field over a 2D environment and a toroidal space is reported in Figure 5a-b.

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

590  
 591 In our model, each grid cell had a tuning width of 0.04. They were defined as 8 modules of size 36,  
 592 and the relative speed scales were  $\{1., 0.8, 0.7, 0.5, 0.4, 0.3, 0.2, 0.1, 0.07\}$ .

593 **5.2 Place cells**

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

597 In order to avoid overlapping of place fields, lateral inhibition is implemented. More specifically, the  
 598 tuning process is aborted in case the cosine similarity of the new pattern and the old ones is greater  
 599 than a threshold  $\theta_{inh}^{PC}$ .

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

606 **Activity** The current firing rate of the PC population is determined by the cosine similarity between  
 607 the GC input and the forward weight matrix, then passed through a generalized sigmoid  $\phi(z) =$

608  $[1 + \exp(-\beta(z - \alpha))]^{-1}$ . The parameter  $\alpha$  represents the activation threshold, or horizontal offset,  
 609 while  $\beta$  the gain, or steepness.

$$u_i^{PC} = \phi \left( \cos \left( u^{GC}, W_i^{GC,PC} \right) \right) \quad (4)$$

610 It is also defined as an activity trace, which has an upper value of 1 and decays exponentially:

$$m_i = -m_i/\tau^{PC} + u_i \quad (5)$$

611 It is used as a proxy for a memory trace.

612  
 613 In the model, a PC population is defined by its average place field size, determining the granularity  
 614 of the representation of the place. In plot 5b it is illustrated an example of place cells layer tuning  
 615 obtained from a continuous trajectory over a square environment.

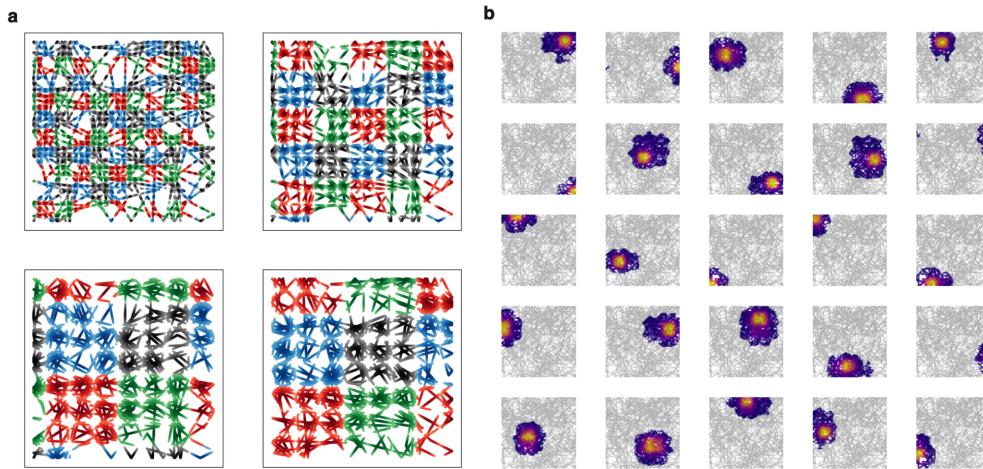


Figure 5: **Place fields obtained from grid cells activity** - **a**: grid cell modules with different granularity represented over a continuous trajectory in an open space. For visualization purposes, each module is represented as composed of four sub-modules of 9 grid cells each, whose periodic tuning generates activity that repeats in space. **b**: place cells whose spatial tuning has been obtained from the concatenation of the grid cells population vector.

### 616 5.3 Modulation

617 Neuromodulation is implemented as a reward-sensitive signal, represented as DA (mimicking the  
 618 function of dopamine), and a collision-sensitive signal, represented as BND (for boundary). Its  
 619 dynamics are defined in terms of a leaky variable  $v$  whose state is perturbed by an external input  $x$ ,  
 620 whose qualitative meaning differs for each neuromodulator  $k$ .

$$\begin{aligned} v^k &= -v^k/\tau^k + x^k \\ v^k &= \max(v^k, 0) \end{aligned} \quad (6)$$

621  $\tau_{DA} = 2, \tau^{BND} = 1$

622 **Learning rule** The connection weights  $W^k$  are updated according to a plasticity rule composed of  
 623 an Hebbian term, involving the leaky variable, the place cells that are above a certain threshold  $\theta^k$ ,  
 624 and the current connection weights value:

$$\Delta \mathbf{W}^k = \eta^k \mathbf{u}^{PC} \left( \mathbf{v}^k - \mathbf{W}^k \right) \quad (7)$$

625 where the weight contribution of the Hebbian update, and  $\eta^k$  is the learning rate:  $\eta^{DA} = 0.9$ ,  $\eta^{BND} =$   
626 0.9. Additionally, connections values are kept non-negative.

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

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

$$\beta_i = c_a^k m_i \bar{\beta} + (1 - m_i) \bar{\beta} \quad (8)$$

630 where  $c_a^k$  is a scaling gain parameter, and if it is 1 then no modulation takes place.

631 Concerning center relocation, it is applied to recently active neurons with non-zero trace  $m_i$ . For a  
632 place cell  $i$  with position  $\mathbf{W}_i^{GC,PC}$  (in the grid cell space), it is calculated a displacement vector  $q_i$   
633 with respect to the current position  $\mathbf{u}^{GC}$ :

$$q_i = c_b^k v^k \exp \left( - \frac{\|\mathbf{W}_i^{GC,PC} - \mathbf{u}^{GC}\|}{\sigma^k} \right) \quad (9)$$

634 where  $c_b^k$  is a scaling relocation parameter, while  $\sigma^k$  the width of the Gaussian distance. This  
635 displaced is used to move in GC activity space and get the new GC population vector to use as  
636 tuning pattern.

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

#### 639 5.4 Decision making

##### 640 Behaviour selection logic

641 The possible behaviours are *exploration* and *exploitation*, and an action is defined as a 2D velocity  
642 vector. For exploration, an action can be generated either as random navigation, using a polar vector  
643 of fixed magnitude (the speed) and angle sampled from a uniform distribution, or as a step within  
644 a goal-directed navigation plan to reach a random destination, which corresponds to a randomly  
645 sampled existing place cells. In the goal-directed navigation the magnitude of the velocity vector  
646 is less or equal than a fixed speed value, depending on the distance from the next target position in  
647 the plan. Instead, for the exploitation behaviour, the action is a step within a goal-directed navigation  
648 towards the reward location. The behaviour selection process depends on the experience of collision,  
649 the presence of a plan, and the success in the navigation planning. A diagram of this logic is reported  
650 in Figure 6.

651 The positions of the agent and of the target location for planning are identified by the place cells  
652 population vector. In particular, the reward position  $(x_r, y_r)$  is determined by the weighted average of  
653 the centers  $x_i, y_i$  of the place cells with respect to their DA-modulated connections weights. Further,  
654 only the top 5 place cells are considered.

$$x_r = \sum_i^5 \mathbf{W}_i^{DA} x_i \quad (10)$$

$$y_r = \sum_i^5 \mathbf{W}_i^{DA} y_i$$

##### 655 Path-planning algorithm

656

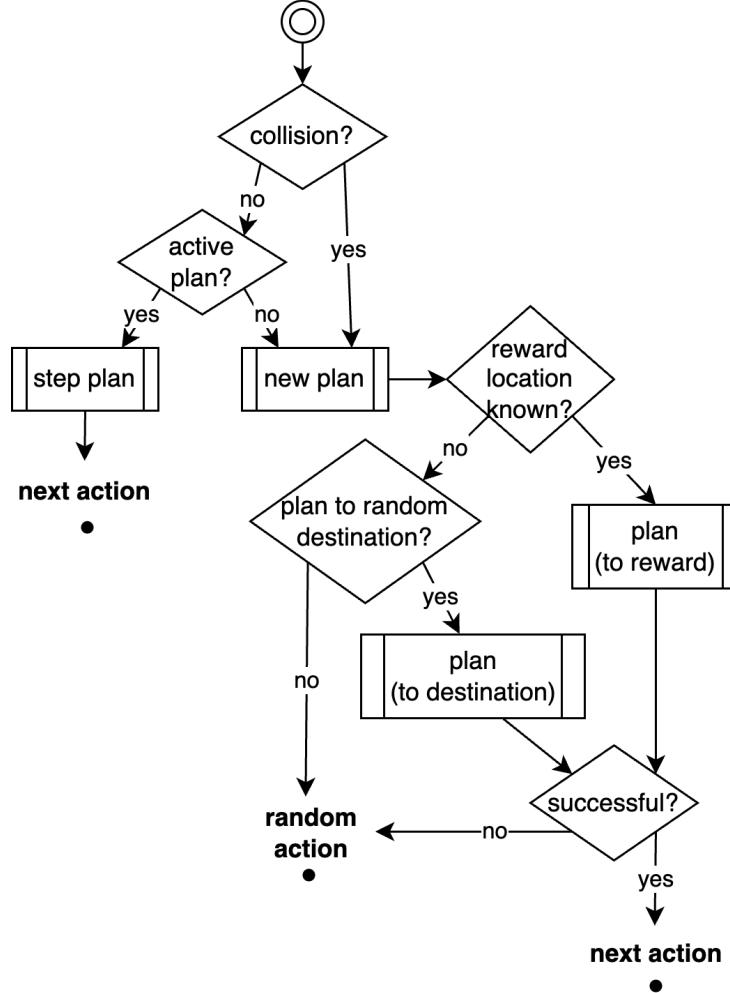


Figure 6: Diagram of the behaviour selection process

---

#### Algorithm 1 ACTIVITY-BASED PATHFINDING

---

**Require:** Connectivity matrix  $C \in \mathbb{R}^{n \times n}$ , node weights  $w \in \mathbb{R}^n$ , start node  $s$ , end node  $e$

**Ensure:** A list of nodes forming a (short) path from  $s$  to  $e$ , or empty if none found

- 1: Initialize activity vector  $a \leftarrow \mathbf{0} \in \mathbb{R}^n$ ; set  $a_s \leftarrow 1$
- 2: Initialize history list  $H \leftarrow []$   $\triangleright$  — Forward propagation phase —
- 3: **for**  $t = 1$  to  $\text{MAX\_PATH\_DEPTH}$  **do**
- 4:      $a \leftarrow C \cdot a$
- 5:      $a \leftarrow a \circ w$   $\triangleright$  Element-wise multiplication with node weights
- 6:      $a \leftarrow \sigma(4(a - 0.6))$   $\triangleright$  Apply sigmoidal activation:  $\sigma(x) = \frac{1}{1+e^{-x}}$
- 7:      $a_i \leftarrow 0$  if  $a_i < 0.1$  (thresholding)
- 8:     Append  $a$  to  $H$
- 9:     **if**  $a_e > 0$  **then**
- 10:         **break**
- 11:     **end if**
- 12: **end for**
- 13: **if** maximum depth reached **then**
- 14:     **return** []  $\triangleright$  No path found
- 15: **end if**
- 16: **if**  $|H| < 3$  **then**
- 17:     **return**  $[s, e]$   $\triangleright$  Path is trivially short
- 18: **end if**
- 19: Initialize path index stack  $G \leftarrow [[e]]$
- 20: **for**  $t = 1$  to  $\text{MAX\_PATH\_DEPTH}$  **do** 19
- 21:     Let  $m \leftarrow C_{G[t-1]}$   $\triangleright$  Get neighbors of current group
- 22:     **if**  $m_s > 0$  **then**
- 23:         **break**
- 24:     **end if**

657 The planning of a new route is implemented as a path-finding algorithm based on the place cell graph,  
 658 provided as connectivity matrix  $C$ . Its particularity is the use of a weighting  $\tilde{W}$  of the nodes according  
 659 to the neuromodulation map. A description is reported in algorithm 5.4.

## 660 5.5 Environments

661 The game in which test the model has been developed with the python library Pygame, used under  
 662 license GNU LGPL version 2.1 and available at <https://github.com/pygame/pygame>. The  
 663 environment layout consisting in a customizable arrangement of vertical and horizontal hard walls  
 664 with variable length and fixed width. Below in Figure 7 some samples are shown.

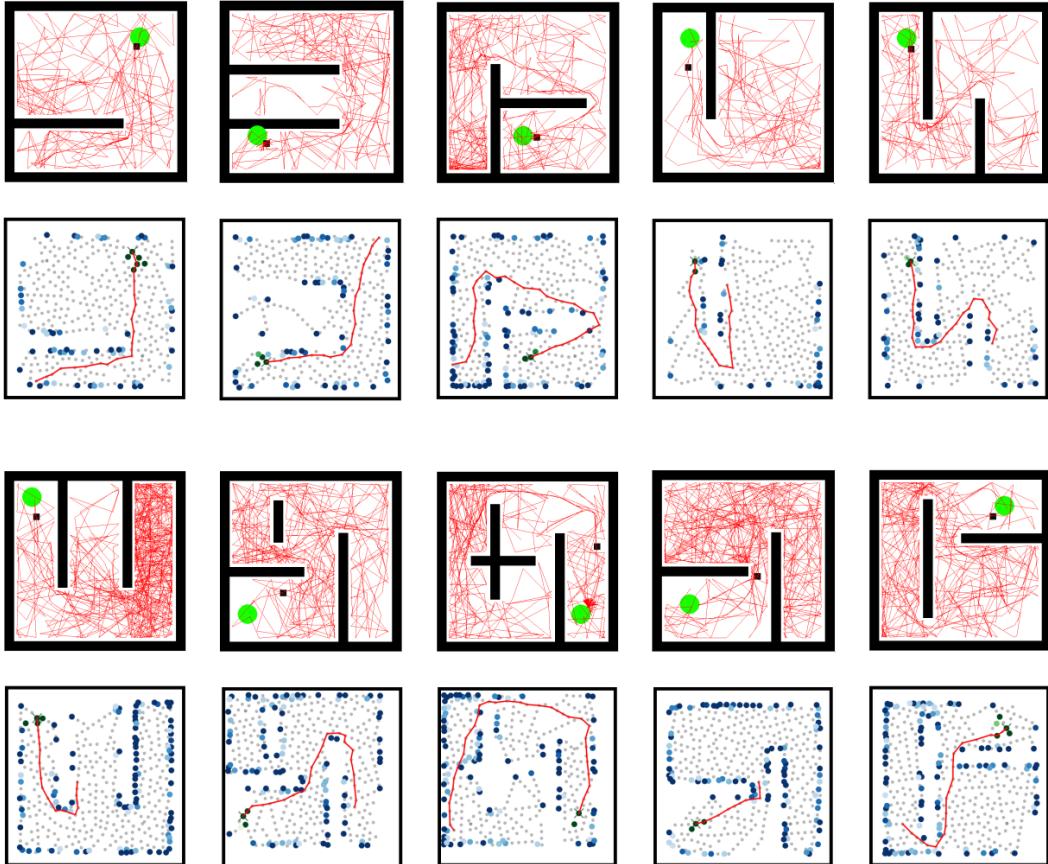


Figure 7: Sample of generated environments

665 The reward object is defined as a circle with size the 5% of the total environment area. When the  
 666 agent position is within its boundary, it is provided a binary signal  $R \sim \mathcal{B}(p_r)$  drawn from a Bernoulli  
 667 with probability  $p_r = 0.6$ . The duration of the reward fetching is set to 2 time steps.

668 The agent object is defined as a square with size the 3.3% of the total environment area.

669 The testing protocol was inspired by the behavior of animals that venture into new territories in search  
 670 of food. It was divided into two parts:

- 671 • **exploration phase:** the agent was placed in a random location within the environment for  
 672 10'000 time steps. In this phase the reward is not present. Further, in order to force greater  
 673 exploration of the environment, every 3'000 steps it was teleported to another random  
 674 location. This external intervention was meant to mitigate the randomness in the exploratory  
 675 behavioural strategy of the agent.
- 676 • **reward phase:** a reward is insert in a random location, and it is available to be discovered.  
 677 When it is encountered, the agent is teleported to a random location within the environment,

678 and after a fixed about 100 time step it is enabled its reward-seeking behaviour, in the form  
679 of goal-directed navigation. The total duration of this phase is set to 20'000 time steps.

680 An episode is defined as a continuous trajectory during the reward phase, namely a set of time steps  
681 starting from when the agent is place in a position until either it finds the reward or the simulation  
682 ends.

683 **Detour experiment** The protocol is modified such that after a fixed number of episodes the layout  
684 of the environment is changed, *e.g.* a wall is inserted. This experiment is meant to test the ability  
685 to reach the reward location by using the same cognitive map, and possibly update it with the new  
686 sensory information, such as the detection of the new boundaries.

687 **Changing reward experiment** During the reward phase, the reward location is changed after a  
688 fixed number of fetches.

689 **Optimization** The model hyper-parameters such as the constants for the neural dynamics and  
690 the behaviour selection have been optimized through a evolutionary algorithm. Initially, an initial  
691 population of individuals with different random genomes (string of hyper-parameters values) is  
692 sampled and evaluated. Then, the population of a new generation is constructed from the first by  
693 combining and mutating the genomes of the top ranked individuals from the previous generation.  
694 More in details, for the sampling of the new generation we used the Covariance Matrix Adaptation  
695 algorithm, in which the shape of distribution of genome values is iteratively adapted according to  
696 the recent performances. The fitness used for evaluating the individuals was of a tuple consisting  
697 of the number of collected rewards and the number of collisions. In particular, the latter was  
698 calculated starting from the time when the reward position was first discovered, in this way at least  
699 the stochasticity of the exploratory behaviour was excluded.

700 The evolved hyper-parameters are: neural gain  $\beta$ , lateral inhibition threshold  $\theta_{inh}^{PC}$ , lateral distance  
701 threshold  $\theta_{rec}^{PC}$ , activity trace time constant  $\tau^{PC}$ , reward modulation scale  $c_b^{DA}$ , reward modulation  
702 spread  $x_b^{DA}$ , boundary modulation scale  $c_b^{BND}$ , boundary modulation spread  $c_b^{BND}$ , reward gain  
703 modulation  $c_a^{DA}$ , and boundary gain modulation  $c_a^{BND}$ .