Coordinated hippocampal-entorhinal replay as structural inference

Talfan Evans*  
Institute of Cognitive Neuroscience  
University College London  
talfan.evans.13@ucl.ac.uk

Neil Burgess  
Institute of Cognitive Neuroscience  
University College London  
n.burgess@ucl.ac.uk

Abstract

Constructing and maintaining useful representations of sensory experience is essential for reasoning about one's environment. High-level associative (topological) maps can be useful for efficient planning and are easily constructed from experience. Conversely, embedding new experiences within a metric structure allows them to be integrated with existing ones and novel associations to be implicitly inferred. Neurobiologically, the synaptic associations between hippocampal place cells and entorhinal grid cells are thought to represent associative and metric structures, respectively. Learning the place-grid cell associations can therefore be interpreted as learning a mapping between these two spaces. Here, we show how this map could be constructed by probabilistic message-passing through the hippocampal-entorhinal system, where messages are scheduled to reduce the propagation of redundant information. We propose that this offline inference corresponds to coordinated hippocampal-entorhinal replay during sharp wave ripples. Our results also suggest that the metric map will contain local distortions that reflect the inferred structure of the environment according to associative experience, explaining observed grid deformations.

1 Introduction

Localizing in an environment relies on two sources of information. Firstly, unique sensory inputs may indicate absolute location in space. Secondly, path integration (PI) can update previous location on a metric map by integrating self-motion. Sensory inputs are required to correct the accumulation of error by PI, but problems arise when their role in localization occurs simultaneously with learning of their correspondence to locations on the metric map (SLAM) [9]. In general, computing the joint map-location distribution requires probabilistic inference over previous sensory observations and movements given their respective uncertainties. Associative representations can be computationally cheaper when used to perform high-level planning [43]. However, organizing associative structure in a metric space allows for efficient integration of new experience and the inference of metric relationships between sensory states in the absence of physical experience. This 'short-cutting' ability is crucial for efficient exploration and navigation [45, 42].

1.1 Place and grid cells

Neurobiologically, grid cells (GC) in the medial entorhinal cortex (mEC), whose firing fields are arranged on a periodic hexagonal lattice in space, are thought to play a role in PI [14] and constitute a metric map of space [19]. Their firing patterns are stable over time suggesting stabilization by

*Use footnote for providing further information about author (webpage, alternative address)—not for acknowledging funding agencies.

Preprint. Under review.
environmental cues \[21 \text{11}\]. Conversely, place cells (PC) in the hippocampus (HPC) fire at distinct locations \[33\] and are thought to respond to specific sensory stimuli such as environmental geometry \[34 \text{26}\]. PCs represent states in sensory space such that their activity most often reflects the animal’s current location, their synaptic associations constitute an \textit{associative} map of an environment \[31\] and their connections to GCs stabilize the GC \textit{metric} map. Although PC and GC activity most often represents the current location, coordinated sequential ‘replay’ of remote cells (i.e. whose firing fields are non-local) also occurs \[13 \text{35}\].

1.2 Summary of contributions

We propose that the HPC-mEC system operates in two distinct regimes. When navigating using a known \textit{map} (i.e. locations of sensory states in \textit{metric} space), an \textit{online} system probabilistically integrates PI and sensory information for localization (Fig. 4A). A simple learning mechanism allows the \textit{online} system to learn initial priors over the \textit{map} structure (Fig. 4B). However, conflicts in the PI and sensory estimates of location necessitate more complex \textit{offline} inference to correct the \textit{map}, which requires inference over previous sensory observations and movements (Fig. 2) \[37\].

We show how this \textit{offline} system can use the \textit{associative} structure stored in the recurrent CA3 synapses between PCs to construct and correct a \textit{metric} map stored in the synaptic associations between GCs and PCs, corresponding to one-shot learning. The distribution over landmark locations is computed via message passing \[36\] between PCs. Scheduling messages to minimize the propagation of redundant information not only improves performance, but also produces structured reactivations of PCs resembling those observed during hippocampal ‘replay’ \[13 \text{8} \text{39}\]. Our model provides both a functional and mechanistic interpretation for observations of coordinated HPC-mEC replay \[35 \text{49}\] and makes novel experimental predictions. In contrast to reward-based interpretations \[30\], our model poses replay as structured coordinated information transfer though complementary \textit{metric} and \textit{associative} representations of the world. Sharp-wave ripples \[6 \text{32}\] which coincide with replay events may correspond to structural prediction errors.

Lastly, when the learned \textit{associative} structure is non-Euclidean, organization within a \textit{metric} space predicts recently observed local distortions in GC firing patterns, such that the underlying structure represented by GCs reflects the informational ‘similarity’ of distinct locations in stimulus space.

2 Model

2.1 Grid and place cells

A single GC will fire periodically at the vertices of a triangular lattice in 2D space (Fig. 1A), GCs exist in anatomical ‘modules’, groups of GCs whose firing patterns share the same spatial scale (distance between vertices) and orientation relative to the environment, but differ in their spatial offsets \[19\]. Moreover, the spatial scale increases in discrete ‘jumps’ along the dorso-ventral axis, suggesting that these modules encode a hierarchical representation of space \[2 \text{12} \text{29}\]. We assume that the firing rates of the \(N_G\) GCs in a single module \(G\) encode the probability distribution over the current location \(G(x)\) within a periodic, discretized region of state space \(x\). Although we only
consider a single grid scale, our results naturally extend to multi-scale architectures, theoretically allowing encoding of ranges up to \([13, 29, 47]\) or beyond \([12]\) the largest grid scale.

PC firing \(P\) represents the probability of the presence of specific sensory stimuli, which in a spatial context can be considered as ‘landmarks’ whose location in physical space is denoted by \(\mu_p\). In our simulations, the firing of each of the \(N_F\) PCs is described as a noisy Gaussian distribution \(p'_G = \mathcal{N}(x'_G, \mu_p, \sigma^2_G)\), where \(x'_G \sim \mathcal{N}(x_G, \sigma^2_G)\) is a noisy estimate of its position in physical space \(x_G\). The distribution over the location of landmark \(p\) in GC space (its belief) is described by its synaptic associations to the GCs, the \(p^{th}\) column of matrix \(B \in \mathbb{R}^{N_F \times N_G}\). We consider these synaptic associations to constitute the metric embedding of sensory experience (Fig. 1A).

\subsection{2.2 Online localization and learning}

Given a suitable representation of uncertainty and a known map, localization is achieved by a process of recursive Bayesian estimation (RBE), where a model-based prediction based on perceived movement is corrected by incoming sensory information (Fig. 1A).

**Movement update**  The location distribution (grid module activity) from the previous time-step \(G(x_{t-1})\) is updated according to perceived movement given a transition model \(T(x_t|x_{t-1}, \hat{u}_t)\):

\[
G'(x_t) = \int T(x_t|x_{t-1}, \hat{u}_t) \cdot G(x_{t-1})dx_{t-1}
\]

where \(G'(x_t)\) is the movement estimate and \(\hat{u}_t \sim \mathcal{N}(u_t, \sigma^2_{PI}u)\) is the noisy perceived movement at time \(t\), where \(\sigma_{PI}\) scales the noise with distance travelled. Since the firing of GCs are periodic across space, we use a wrapped Normal distribution defined over the triangular lattice to account for the probability of having transitioned from any of the infinite grid tilings:

\[
T(x_t|x_{t-1}, \hat{u}_t) = \sum_{m,n=-\infty}^{\infty} \mathcal{N}(x_t - x_{t-1} + c_{mn}, \sigma^2_G, \lambda)
\]

where and \(c_{mn} = 2\lambda(mn_v_1 + n_v_2)\) is a spatial offset of scale \(\lambda\) given the lattice basis vectors \(v_1 = [\cos(\phi), \sin(\phi)]\) and \(v_2 = [\cos(\phi + \pi/3), \sin(\phi + \pi/3)]\) and \(\phi\) is the global orientation of the grid pattern. Where grid space is represented discretely by the firing rates of a population of GCs, the periodic form of the transition function can be replaced by multiplication by a velocity dependent circulant matrix \(T(\hat{u}_t)\) (see Appendix C.1), linking to the eigendecomposition of diffusive transition matrices \([40]\) and generalizing a previous mechanism to the case of noisy PI \([4]\).

**Observation update**  The predicted estimate is refined by incoming sensory input to give the integrated estimate \(G(x_t)\):

\[
G(x_t) = \frac{1}{K_t} H(p_t|x_t) \cdot G'(x_t)
\]

where \(H(p_t|x_t) = P_tB\) is the observation model defining the likelihood of the current sensory inputs \(P_t\) given the predicted location. The normalization constant \(K_t\) is the sum over the current GC activity, implemented by a simple inhibitory feedback circuit: \(\tau \frac{dG}{dt} = -G + E\), where \(E = 1\) is a constant excitatory drive such that the sum of the firing rates \(\|G\|_1 = 1\) at steady-state. Note that the notation \(G\) refers to the case where the distribution over continuous metric space is represented discretely by a vector of \(N_G\) GC firing rates (see Appendix A).

**Online learning as prior formation**  A simple error-based learning rule with learning rate \(\alpha = 1e - 4\) modifies \(B\) to minimize the error between the observation and movement models (Fig. 1B):

\[
\frac{1}{\alpha} \frac{dB}{dt} = 2p_t^\top (G'_t - P_tB)
\]

\subsection{2.3 Offline message passing for probabilistic structural inference}

During exploration of a novel environment, the *online* model produces stable learning when PI noise is low and the transition structure is static (Fig. 1A). However, all learning is local: only the synaptic weights of the currently active cells are modified at each time-step. This is not a full solution to the SLAM problem, which requires finding the most likely configuration of sensory observations (landmarks) and current location (in grid space) given all historic observations and perceived movements, described by the joint map-location distribution \(p(x_t, B|x_{0:t-1}, \hat{u}_{0:t-1}, x_0)\) (see
Where all place fields have equivalent receptive field covariance, the inferred Euclidean distance of which converges in the steady-state to $A_{ij}$ where the $d_{ij}$ with mean $\mu_{ij}$ and variance $\sigma_{ij}^2$ are equal to and proportional to the perceived distance, respectively (the latter reflecting accumulation of PI noise in Eq. $2_t$ Fig. $2_f$; Appendix D.1). The PC-GC synaptic associations can then be viewed as priors over the locations of each landmark associated landmarks to also be updated without needing to be re-visited. Secondly, multiple weak (high variance) observations can together form strong hypotheses if those observations are consistent.

Appendix Fig. 1 for a summary of the anatomical mapping). Computing this requires integrating over all possible configurations of PC locations (encoded in B), which requires inference over previous and non-local observations. There are several advantages of a system capable of propagating information through non-local locations. Firstly, updates to the perceived location of a given landmark cause errors in relative pairwise distance measurements are correlated such that their variance decreases over time \cite{9}. Relaxing the ‘spring mesh’ is equivalent to finding the maximally likely configuration of landmarks, if pairwise distance observations (pairwise potentials $\psi_{ij}$) are described by Gaussians with mean $d_{ij}$ and variance $\sigma_{ij} = \sigma_{PC}^2 + d_{ij}\sigma_{PI}^2$ that are equal to and proportional to the perceived distance, respectively (the latter reflecting accumulation of PI noise in Eq. $2_t$ Fig. $2_f$; Appendix D.1). The PC-GC synaptic associations can then be viewed as priors over the locations of each landmark in metric space, ‘anchoring’ the inferred structure which would otherwise be translation / rotation invariant. Together, the associative structure and metric mapping, encoded in the PC-PC (A) and PC-GC (B) associations respectively, define the posterior distribution over the landmark locations $b_i$:

$$P(b_1, ..., b_{N_L}) = 2 \prod_{1 \leq i \leq N_L} \psi_{ij}(b_i, b_j) \prod_{1 \leq i \leq N} B_i(b_i)$$

where the $\psi_{ij}(b_i, b_j) = \psi_{ji}(b_j, b_i) = \sum_{m,n=-\infty}^{\infty} \exp \left( -\frac{1}{2} \sigma_{ij}^{-2} (d_{ij} - ||b_i - b_j + c_{mn}||)^2 \right)$ terms define the pairwise potentials between PCs and $\lambda$ is the grid scale. Note that $B_p(b_p)$ here defines the continuous distribution of the location of PC $p$ in metric (GC) space for consistency with the literature; in reality it is a discrete vector described by the $p^{th}$ row of $B$.

**The hippocampus as a cognitive graph** The structure of an environment can be inferred from pairwise distance observations between landmarks \cite{9,31}. Intuitively, consider a ‘spring network’ of connected landmarks, where the edges represent noisy pairwise observations with stiffness and length equal to the certainty and estimated pairwise distance, respectively (see Appendix D.1). Convergence is contingent on the fact that, despite large absolute errors in landmark location (due to noisy PI), errors in relative pairwise distance measurements are correlated such that their variance decreases over time \cite{9}. The resulting form for the recovered distance is also scaled by the receptive fields’ variance (the Bhattacharyya distance) \cite{3}, such that ‘closeness’ is related also to the ‘discriminability’ (Fig. $2_c$). We discuss this scaling constant later (see also Appendix D.3). Our approach differs subtly from typical graph-based SLAM systems \cite{27,44} which treat each

**Associative encoding in the hippocampus** We propose that these pairwise distance measurements are encoded in the recurrent synaptic associations between CA3 PCs, constituting an associative representation of the structure of space \cite{31}. Given Gaussian place fields, a simple modified Hebbian learning rule with constant decay learns the pairwise PC weights (associative map) A:

$$\frac{1}{\alpha} \frac{dA_{ij}}{dt} = p_i(t)p_j(t) - A_{ij}^2$$

which converges in the steady-state to $A_{ij} = \sqrt{p_i(t)p_j(t)}$, the square root of the correlation between the firing of two PCs (see Appendix D.1 for more details on the choice of learning rule). Where all place fields have equivalent receptive field covariance, the inferred Euclidean distance of PC $j$ from the perspective of $i$ is then proportional to the true distance given a simple transformation: $d_{ij}^2 = -\log(A_{ij}) = (\mu_i - \mu_j)^2 / 2\sigma_{PC}^2$. The resulting form for the recovered distance is also scaled by the receptive fields’ variance (the Bhattacharyya distance) \cite{3}, such that ‘closeness’ is related also to the ‘discriminability’ (Fig. $2_c$). We discuss this scaling constant later (see also Appendix D.3). Our approach differs subtly from typical graph-based SLAM systems \cite{27,44} which treat each
We first tested the ability of the offline system to infer the structure of three environments (Fig. 2). Given erroneous initial estimates corresponding to priors formed during noisy PI, the system is able to correctly infer the true structures as those that satisfied pairwise measurements between states (Fig. 2D). However, an immediate consequence of the system is that this inferred structure will be sensitive to topology. Although PI will impose metric priors, where these priors are unreliable (as in the case of navigating around an unfamiliar ring environment under noisy PI), the inferred structure is sensitive to the ‘closure’ of loops (Fig. 2D, "Broken Ring").
Figure 3: The loop-closure task. A. The agent navigates a novel circular track, accumulating PI error. Lap completion (iii) triggers an offline inference event (see main text and Supp. Video 1 for details). B. Structure inferred after loop-closure. C. PE is reduced on completion of subsequent laps. D. Offline inference allows one-shot learning when compared to the online system.

3.2 Loop closure experiment

In the loop closure task (Fig. 3; Supp. Video 1), place fields are distributed uniformly around a circular 1D track. Initial location confidence is high, such that place and GCs active at the start location (0 radians) form strong associations. As the agent navigates around the track, PI error accumulates and the confidence in location decreases, resulting in subsequent PC-GC associations becoming more diffuse and less likely to correspond to the true structure (Fig. 3A). Due to the accumulated error, when the agent completes a full lap it receives a sharp input from the PCs initially active at the starting location, producing a strong positive PE and triggering an offline inference event (Fig. 3Aii).

Offline inference allows one-shot learning As expected, structural error is reduced significantly following the triggered offline inference events. This reduction is markedly larger than in equivalent trials using only the online system, resembling a ‘one-shot’ learning process (Fig. 3D). Given the rapid map-learning, PEs on subsequent laps are also reduced (Fig. 3C).

Principled message scheduling produces structured reactivations BP seeks a solution whereby messages received from neighbouring nodes cause negligible change to the receiving nodes’ belief. The scheduling is therefore important from an energetic perspective: messages that do not produce changes in the beliefs of neighbours are redundant. Nodes which did not significantly update their self-beliefs following receipt of a message therefore do not need to re-broadcast a message at the next time-step (Fig. 4A).
In addition to the energetic advantages, the ‘Max-Entropy’ schedule also contributes to inference performance, converging faster than a simple ‘sequential’ scheme in which all nodes broadcast messages at each time-step, despite broadcasting fewer total messages (Fig. 4B).

The sequences of reactivations also contained significant structure, tending to propagate initially backwards along the track from the animal’s current position, resembling the PC reactivations during reverse hippocampal replay (Fig. 4C) [13]. Sequences did not always hop to adjacent fields, occasionally hopping to new locations where remote sequences were then initiated (Fig. 4D) [7, 24, 41]. Multiple sequences at different remote locations can be seen to occur simultaneously or in an alternating fashion. Both forward and reverse sequences were observed in equal proportion (Fig. 4D) [13, 8]. Lastly, the ‘hoppiness’ of the sequences was related to the confidence in the pairwise observations, information propagating more quickly and smoothly in a ‘stiffer’ graph (a graph with more confident pairwise observations; Fig. 4E, F) [38, 24, 41].

3.3 Local distortions to the cognitive map

Grid patterns undergo significant local distortions in open environments, decreasing in scale and becoming less uniform (more sheared) towards the corners [20]. We hypothesized that these distortions might reflect the underlying structure of the environment as captured in the associative structure in CA3 and manifested in its projections to metric GC space.

In the same study, scale was also positively correlated to behavioural occupancy (animals spent more time in the middle of the environment; Fig. 5B, E bottom row, Appendix Fig. 2B) [20]. This effect was mirrored in our model, since over-sampling of the tails of the place fields near the boundaries of the environment led the associated PCs overestimating their pairwise distances (Appendix Fig. 2B); their mean co-firing was lower than expected if the animal were to sample from the place field uniformly; Fig. 2A). Note that pairwise associative distance is inversely related to the scale of the grid pattern readout, since larger associative distance implies travelling further in metric space (see Appendix D.1 and Appendix Fig. 2C).

Given that PC firing is related to the confidence of the presence of specific sensory cues, we also explored the case where place fields were sharper near the edges of an environment, as would be the case if driven by strong geometric cues (Fig. 5A, top row) [5, 22, 21]. These non-uniformities also produced the same local warping of the grid pattern (Fig. 5C; D, E top row). This latter effect is attributed to the nature of Hebbian learning rules, whose learned synaptic strengths reflect the variance normalized distance between the fields, as opposed to the true Euclidean distance (Fig. 2A; Appendix D.1) [3]. Our model suggests that the cognitive ‘distance’ (or ‘discriminability’) between two sensory stimuli should be greater if the absolute confidence in the locations of each is greater.

3.4 A neural-level model of coordinated place-GC replay

How might belief propagation be implemented in the brain? More specifically, how might ‘message broadcasts’ correspond to spikes fired by PCs during replay, and how would GCs contribute to offline inference? Our proposed mechanism relies on coincidence detection by a ‘receiving’ PC $P_R$ of a direct spike from a ‘broadcasting’ PC $P_B$ and a travelling wave of activity across the GC population (Fig. 6A, C and Supp. Video 2).

A message broadcast is initiated by the firing of a spike from $P_B$ to synaptically connected PCs, with a transmission delay proportional to the inferred pairwise distance $d_{ij}$ (Fig. 6A, "Place cells").
Figure 6: A Neural model of coordinated HPC-mEC replay. A (1D) The broadcasting PC $P_B$ sends a spike to neighbour $P_R$ (CA3), at the same time initiating a travelling wave in the GCs (mEC) by virtue of its synaptic projections B. (Left) No learning occurs when the spike and travelling wave arrive at $P_R$ at the same time. (Right) If the CA3 spike arrives ahead of the travelling wave, the synaptic associations of $P_R$ are adjusted towards the currently active GCs. B Comparison of the 'algorithmic' \[28\] and neural BP implementations. C Travelling waves on the 2D GC sheet.

In parallel, the same spike from $P_B$ drives activity in the GC population via the PC-GC synaptic associations (Fig. 6A, "Grid cells"). This activity propagates radially outwards at a constant speed, accumulating noise in proportion to the distance travelled (i.e. identically to PI during online localization; Fig. 6C; see Appendix D.2). This can be viewed as activity propagating through two generative models of associative and metric space (see Appendix D.3).

When $P_R$ receives the spike from $P_B$, we assume that the depolarization causes learning between $P_R$ and the currently active GCs, even though $P_R$ does not necessarily fire a spike \[17\] \[15\] \[46\]. If the distance indicated by the relative propagation of activity between GCs corresponding to the synaptic projections of $P_B$ and $P_R$ is equal to the distance encoded by the recurrent association between $P_B$ and $P_R$, $P_R$ will receive the spike from $P_B$ at the same time that the travelling wave arrives at the GCs to which $P_R$ projects, so that no significant synaptic changes are produced (Fig. 6A, Left). If these two distances are in disagreement, $P_R$ will revise its belief, shifting its synaptic associations to ‘earlier’ or ‘later’ GCs, respectively (Fig. 6C, Right).

Lastly, firing of $P_R$ is triggered only if there is significant change in its synaptic weights to the GC population, i.e. only messages indicating belief changes are propagated (a similar condition to that used to initiate the offline system). We propose therefore that the ‘message tension’ term, which governs spiking, might correspond to the accumulation of a learning related neuromodulator.

4 Discussion

Ours is the first model to demonstrate how PI and sensory based estimates could interact probabilistically during online localization in the HPC-mEC system (Fig. 4). We also show how more complex probabilistic inference could be performed via the offline interaction of HPC and mEC (Fig. 2, 6) and propose a detailed mapping of the joint map-location distribution to physiological correlates (Appendix Fig. 1). We then show how prediction errors between predicted and observed sensory stimuli can be used to efficiently arbitrate between the two systems (Fig. 3).

Offline inference events based on principled message passing resemble PC reactivations during replay events, which occur during offline behaviours such as pausing or sleep \[13\]. Our model predicts therefore that replay events (and associated sharp wave ripples) should be more frequent during structural changes to the environment rather than being solely responsive to reward \[41\], although rewards themselves may constitute salient sensory stimuli. Our algorithmic and neural models \[28\] of this process are the first to predict the detailed interaction between PCs and GCs during coordinated replay events \[35\] \[49\]. Although investigated in a spatial context, structured information propagation may be a general mechanism for embedding associative experience in metric space \[25\] \[16\].

Our model is also the first to propose that observed local distortions to the grid pattern \[20\], reflect the underlying associative structure of the environment. Place fields are known to be smaller and more dense near to boundaries and salient locations \[23\]. Warping of the grid scale would thus be consistent with preserving a constant rate of change of sensory information \[49\]. The HPC-mEC
interaction can be interpreted therefore as the embedding of associative structure within a metric map, to allow the agent to determine shortcuts between previously unexperienced state transitions [40].

Acknowledgements

We acknowledge funding from European Union’s Horizon 2020 research and innovation programme Human Brain Project SGA2 (grant agreement no. 785907), Wellcome and ERC Advanced grant NEUROMEM.

References


