Learning in a unitary coherent hippocampus

Charles Fox[1] and Tony Prescott[1]

Adaptive Behaviour Research Group, University of Sheffield S10 2TN, UK

Abstract. A previous paper [2] presented a model (*UCPF-HC*) of the hippocampus as a unitary coherent particle filter, which combines the classical hippocampal roles of associative memory and spatial navigation, using a Bayesian filter framework. The present paper extends this model to include online learning of connections to and from the CA3 region. Learning in the extended neural network is equivalent to learning in a temporal restricted Boltzmann machine under certain assumptions about neuromodulatory effects on connectivity and learning during theta cycles, which suggest detailed neural mappings for Bayesian inference and learning within sub-stages of a theta cycle. After-depolarisations (ADP) are hypothesised to play a novel role to enable reuse of recurrent prior information across sub-stages of theta.

1 Introduction

Anatomy. The principal input structures of the hippocampus are the superficial layers of Entorhinal Cortex (ECs). ECs projects to Dentate Gyrus (DG) which is believed to sparsify the encoding of ECs. Both ECs and DG project to CA3, which also receives strong recurrent connections that are disabled [3] by septal ACh. CA3 and ECs project to CA1, which in turn projects to the deep layers of Entorhinal cortex (ECd), closing a loop if ECd sends information back to ECs. ECs, CA1 and ECd outputs appear to share a coding scheme, as evidenced by one-to-one topographic projections. In contrast, DG and CA3 outputs are thought to work in a second basis or latent space. In a second loop, ECs and CA1 both project to Subiculum (Sub), which projects to the midbrain Septum (Sep) via fornix. Septal ACh and GABA fibres project back to all parts of hippocampus.

UCPF-HC model. A previous paper [2], mapped this hippocampal circuit onto a modified Temporal Restricted Boltzmann machine (TRBM, [8]). The TRBM assumes Boolean observation vectors (including a bias node), z'; Boolean hidden state vectors (including a bias node), x'; weight matrices $W_{x'z'}$ and $W_{x'x'}$, and specifies joints,

$$P(x_t, x_{t-1}, z_t) = \frac{1}{Z} \exp \sum_t (-x'_t W_{x'x'} x'_{t-1} - x'_t W_{x'z'} z'_t). \quad (1)$$

Unlike the standard TRBM, UCPF-HC uses the following deterministic update to obtain maximum *a posteriori* estimates:

$$\hat{x}_t \leftarrow \arg\max P(x_t | \hat{x}_{t-1}, z_t) = \{\hat{x}_t(i) = (P(x_t(i) | \hat{x}_{t-1}, z_t) > \frac{1}{2})\}_i \quad (2)$$

which is the zero-temperature limit of an annealed sequential Gibbs sampler.

The noisy inputs $z_t = y_t + \epsilon_t$ are mapped to the combined ECs and DG, where the DG activations are functions of the ECs activations, $z_t = (ECs_t, DG_t(ECs_t))$. CA3 is mapped to the hidden state, x_t. CA1 performs a partial decoding into the DG basis. Finally the estimated de-noised output is mapped to ECd, $\hat{y}_t = ECd_t$. Each neural population is a Boolean vector at each discrete time step t.

A major problem with UCPF-HC tracking is tracking loss, as it approximates the whole posterior with a single sample. To deal with this, performance of the filter is monitored to heuristically detect when tracking is lost – by thresholding a moving average of discrepancy between observed and denoised sensors – then the priors are disabled when lostness is detected. In UCPF-HC, the Subiculum-Septum circuit performs this monitoring. Sub then compares the partially decoded CA1 information against the original ECs input, receiving one-to-one connections from both regions. If they differ for an extended period of time, this indicates loss of tracking. *Tonic* cholinergic projections from Sep, activated via Sub, are well-placed to disable the CA3 priors when lostness occurs, as they are known [3] to disable the recurrent connections in CA3.

The present study presents a new version of UCPF-HC, using the plus maze environment detailed in [2], and extendend with ADP to perform learning in CA3. The plus-maze consists of 13 discrete locations as shown in fig. 4(a). The agent sees unique visual markers if facing two of the arms; it also has touch sensors to report walls to its immediate left, right and front. The original UCPF-HC model included mechanisms to perform path integration in the grid and heading cells using odometry and denoised ECd states – to simplify the present study we assume that grid and heading cells give uncorrelated noisy (Global Positioning System style) estimates of location and orientation, as would be obtained if the UCPF-HC's outputs were always perfect or known to be lost but the odometry was noisy.

ADP Physiology. CA3 pyramidal cells [1] exhibit a single cell short-term memory effect called after-depolarisation (ADP), illustrated in fig. 4(b). A spike (1) in membrane potential, V, is followed by a fast after-hyper-polarisation (AHP, 2), then an after-depolarisation (ADP, 3) and a second, slower AHP (4), before returning to its resting potential (5). (See [7] sections 5.2.5 and 5.3.5 for a detailed review.) ADP has previously been suggested [5] as a basis for multiplexed short-term memories in hippocampus, enabling around seven patterns to be stored simultaneously by re-activating themselves after other patterns, using the ADP gain plus an external excitatory oscillator. We will suggest a related but novel role for ADPs, allowing priors to be restored during separate wake and sleep cycles [4] in a temporal network. ADP is dependent on the presence of ACh or 5HT [6], and septal *phasic* ACh has been suggested to play a role in the hippocampal theta rhythm [3].

2 On-line learning for the UCPF-HC model

The previous version [2] of UCPF-HC did not perform any realistic learning. CA3 cell semantics were specified by hand – for example cells were specified to respond to conjunctions of places, headings and light states. Ideal CA3 responses were computed offline, from these hand-set specifications and ground truth data sets, then weights for

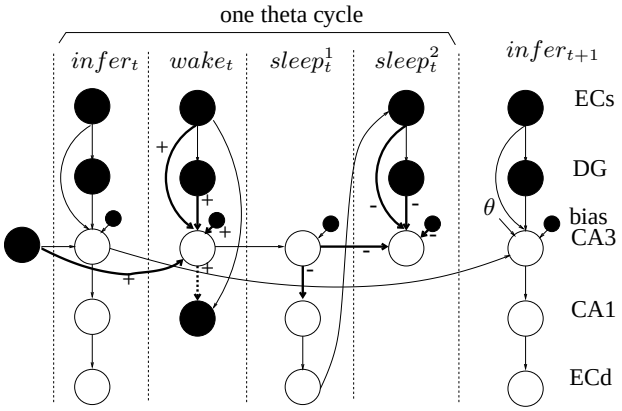


Fig. 1. The four substeps, $infer, wake, sleep^1, sleep^2$ within one theta cycle in the learning neural network model. Circles denote populations of neurons. Arrows indicate fully connected neural network projections (this is *not* a Bayesian network diagram). Thick arrows indicate connections whose weights are updated with Hebbian (+) and anti-Hebbian (-) rules. Dotted arrows indicate where learning occurs but no information is projected (i.e. when the child population is clamped from elsewhere). Filled-in nodes are fixed values at each substep, unfilled nodes are to be computed. The bias population contains a single neuron which is always on, and abstracts the threshold values in CA3. In the first substep of the next cycle, $infer_{t+1}$, CA3 receives a θ signal which disconnects the recurrent connections and switches to ADP recurrent activation.

each input population to CA3, *pop*, were set using independent wake-sleep [4] updates,

$$\Delta w_{ij} = \alpha(\langle CA3_i pop_j \rangle_{\hat{P}(pop,CA3|b)} - \langle CA3_i pop_j \rangle_{P(pop,CA3|b)}) \tag{3}$$

where $\hat{P}$ is the empirical data distribution including the hand-set ideal hidden values; P is the model's generative distribution; and b is the set of hidden nodes biases, preset empirically to model priors on the handset semantics. This was not indented as a realistic learning model, rather just a computational method to set the weights. In particular the computation was greatly simplified by having access to ground-truth hidden states, which made the weights mutually independent given the bias. In reality the agent does not have access to ground truth hidden states – only to sensors.

We do not give the new model access to ideal CA3 states or hand-set their semantics – this time the semantics must be learned. The semantics of DG and CA1, and hence the weights $W_{EC\to DG}$ and $W_{CA1\to EC}$ remain set by hand – we focus only on extending the model to learn all connections to and from CA3: namely $W_{ECs\to CA3}$, $W_{DG\to CA3}$, $W_{CA3\to CA3}$ and $W_{CA3\to CA1}$.

To simplify both the presentation and implementation of the learning model, we will first present the hippocampal learning algorithm for the UCPF-HC neural network as a *fait accompli*, then describe a graphical model simplification used in the implementation. The graphical model formulation also provides insight into the purposes of the neural network processes, which were in fact derived from the graphical model during development. The algorithm is based on the wake-sleep process but is now intended as a biological model.

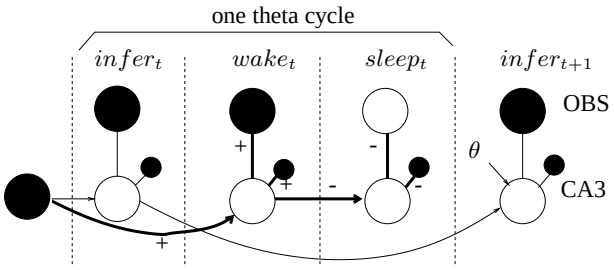


Fig. 2. The reduced TRBM model. This network is equivalent to the neural network, but includes undirected connections, and groups the information from EC,DG and CA1 into a single observable population, OBS.

2.1 Neural network model

The neural learning algorithm is based on the wake-sleep algorithm [4] and is illustrated in fig. 1. It assumes that for every discrete time step t there are four substeps, $infer_t, wake_t, sleep_t^1, sleep_t^2$ corresponding to different phases of one hippocampal theta cycle. These substeps have differing connectivity and learning dynamics, which might be controlled by neuromodulators during the theta cycle. The substeps occur sequentially. But importantly, CA3 activation during the $infer_t$ substep is required to directly influence CA3 at $wake_t$; and CA3 activation at $infer_t$ is required to directly influence CA3 at $infer_{t+1}$; as shown by the arrows in the figure. We tentatively suggest that ADP, discussed in section 1, might play a role in such temporally incontiguous transmission of information.

The $infer_t$ substep is identical to inference in the UCPF-HC model. ECs sensor data is observed; deterministic DG activations (via handset $W_{EC\to DG}$) are computed, and thus act as observations too. We assume that the state of CA3 at $infer_{t-1}$ was inferred exactly and correctly by the UCPF, and is available as an input via recurrent transmission weights $W_{CA3\to CA3}$. Using these inputs, CA3 is updated with a Gibbs sampling step at temperature $T = 0$. CA1 and ECd decode it to retrieve denoised sensor estimates.

In the $wake_t$ substep, the same input vector is maintained in EC and DG; and CA1 activation becomes clamped by a training signal from the ECs input. We assume that conjunctions of facts from ECs are represented perfectly in CA1 by this process, as in DG. We require a delayed copy of the recurrent CA3 input from $infer_{t-1}$ as was received in the $infer_t$ step – *not* a recurrent CA3 input from $infer_t$ – as the recurrent input to CA3 in $wake_t$. CA3 is resampled at $T = 1$ and Hebbian learning is performed at all synapses to and from CA3. In the $sleep_t^1$ substep, the recurrent CA3 connections are used directly so that CA3's state now is influenced by its previous state, $wake_t$. Its connections from ECs and DG are made ineffectual. CA3 is sampled again at temperature $T = 1$, then CA1 and ECd is decoded from it. In $sleep_t^2$ we assume that ECs becomes clamped to the ECd result – feeding back the denoised output into the input. CA3 is resampled again at $T = 1$ and antihebbian learning is performed in all synapses to and from CA3. The theta cycle is now complete, and the next one begins at $infer_{t+1}$.

2.2 Reduced undirected model

We next explain why the neural network is equivalent to the reduced graphical model shown in fig. 2. It is a new variant of the temporal restricted Boltzmann machine [8].

DG consists entirely of cells whose receptive fields are copies or conjunctions of ECs fields. In the reduced model, we form a single population, OBS, which contains both DG and ECs cells. CA1 in the neural model consists of cells with identical fields to DG cells, which are thus also implicitly contained in OBS. The weights $W_{OBS-CA3}$ are undirected as in the TRBM, though the steps of learning them correspond to the steps learning the weights in the neural model.

In phase $infer_t$, CA3 is driven by inputs from EC and DG in the directed neural model, which is equivalent to the undirected link to the observed OBS population in the reduced model. (The bias link is also changed from directed to undirected in the reduced model – again this is an equivalence as the bias is always observed.) In this phase, the temperature is zero so the inferences are always the MAPs. This gives the best denosied estimate of the state of the world.

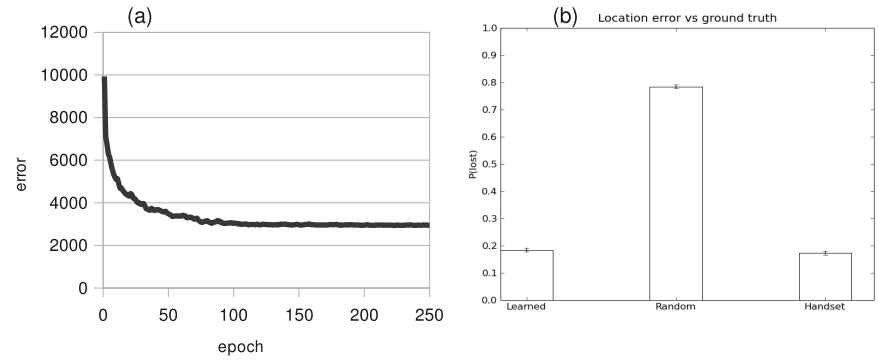


Fig. 3. (a) Training errors. Error is the sum of ECs-ECd discrepancies over all training data in each epoch. (b) Lostness probabilities in learned, random and handset-semantics weights. Error bars show one standard deviation of uncertainty about the population mean.

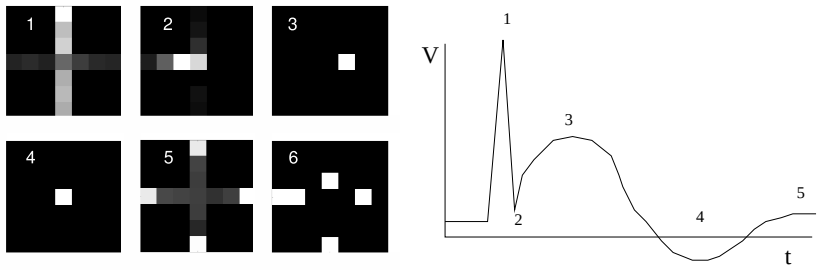
In phase $wake_t$, the drivers of CA3 are the same, but the temperature is $T = 1$. In the neural model, CA1 is clamped to EC, and Hebbian learning occurs in $W_{EC\to CA3}$, $W_{DG\to CA3}$ and $W_{CA3\to CA1}$. This is equivalent to clamping OBS again in the reduced model, and Hebbian learning on $W_{CA3-CA1}$. As in the neural model there is also Hebbian learning on the recurrent CA3 connections.

In phase $sleep_t$ of the undirected model, a CA3 sample is drawn conditioned on its recurrent state only. Then an OBS sample is drawn conditioned on CA3, and finally a new CA3 sample is drawn conditioned on its recurrents and on the OBS sample. Antihebbian learning is performed on all connections to CA3. This is equivalent to the process in the neural model's $sleep_t^1$ and $sleep_t^2$, and is a standard TRBM sleep step.

Phase $infer_{t+1}$ is the start of the next cycle, and like the neural model, requires historical CA3 input from $infer_t$, as might be obtained using ADP.

3 Results

We tested the learning algorithm in the plus-maze world (see [2]), using a path of 30,000 random walk steps. The path was replayed for several epochs until the weights converged. For computational simplicity, learning was performed used the equivalent reduced model, though inference was performed with the full neural model, sharing the learned weights. Python code for the simulation is available from the authors. There is some subtlety in handling learning for cases where the Sub-Sep lostness circuit is activated, which is detailed in the appendix. Fig. 5(a) shows the training errors during learning – using a learning rate of $\alpha = 0.001$ – most of the learning takes place in the first 10 epochs. As in the original [2] UCPF-HC model, the neural network is used to infer denoised ECd estimates of position and sensors. Fig. 5(b) shows the average rate of location errors using the learned weights, compared against the handset semantic of the original UCPF-HC model. A run with randomised, untrained weights is shown for comparison. Inspecting the receptive fields of CA3 cells learned by the training, we find cells in fig. 4(a) responding to individual places (3 and 4); regions around a place (2); the ends of the arms (5); and less well defined fields (1 and 6).



(a) Examples of learned CA3 receptive fields, over the plus maze.

(b) Typical membrane time course following a spike under ADP dynamics.

Fig. 4.

4 Discussion

We have presented a top-down mapping of a wake-sleep learning algorithm onto the biological hippocampal circuit and existing UCPF-HC model. The UCPF-HC model was extended by adding detailed substeps within theta cycles, which specify the connectivity and learning operations required by the algorithm, as biological hypotheses. Our revised model demonstrates a biological plausible online learning mechanism for CA3 pyramidal cells, and thus lends support to our general hypothesis that the hippocampal system may operate as a unitary coherent particle filter. This type of mapping necessarily makes strong predictions about what neurons would be required to do to implement the algorithm. In particular we have relied on specific timing features of ADP and on

ACh to switch between recurrent CA3 activation and ADP-based CA3 memories. It remains to be seen whether biological ADP and ACh are able to provide these functions.

The inference step was performed at zero temperature, separately from higher temperature wake and sleep steps. There are several possible variations on this theme. First, both wake and sleep could be performed at zero temperature, removing the need for a separate inference step, and resulting in a different type of optimisation during learning (minimising $KL[Q^T||P^T]$ rather than $KL[Q||P]$. In the limit $Q = P$ they would give the same result). Second, wake and sleep steps could be extended to run for several steps. This would result in longer sequences of uninterrupted tracking of observations, alternating with longer 'hallucinated' sequences of generated samples. The latter would resemble sequence replay and preplay known to take place in CA3.

Future work could implement the neural learning model directly, in place of the reduced simplification. It could also consider memories of the agent's own actions as way to increase the predictability of plus maze sequences. Finally the agent should iteratively estimate the amount of error in its own location estimates rather than rely on the artificial noisy GPS assumption used in this proof-of-concept implementation.

References

1. M. Ceaser, D.A. Brown, B.H. Gahwiler, and T. Knopfel. Characterization of a calcium-dependent current generating a slow afterdepolarization of CA3 pyramidal cells in rat hippocampal cultures. *Eur. J. Neuroscience*, 5:560–569, 1993.
2. C.Fox and T.Prescott. Hippocampus as unitary coherent particle filter. In *Proceedings of the International Joint Conference on Neural Networks (IJCNN)*, 2010.
3. M.E. Hasselmo, E. Schnell, and E. Barkai. Dynamics of learning and recall at excitatory recurrent synapses and cholinergic modulation in rat hippocampal region CA3. *Journal of Neuroscience*, 15:5249–62, 1995.
4. G.E. Hinton, P. Dayan, B.J. Frey, and R.M. Neal. The wake-sleep algorithm for unsupervised neural networks. *Science*, 268:1158–1161, 1995.
5. O. Jensen, M.A.P. Idiart, and J.E. Lisman. Physiologically realistic formation of autoassociative memory in networks with theta/gamma oscillations: role of fast NMDA channels. *Learning and Memory*, 3:243–256, 1996.
6. J. Lisman and M.A.P. Idiart. Storage of 7+/-2 short-term memories in oscillatory subcycles. *Science*, 267:1512–1515, 1995.
7. N. Spruston and C. McBain. Structure and functional properties of hippocampal neurons. In P. Andersen, R. Morris, D. Amaral, T. Bliss, and J. O'Keefe, editors, *The Hippocampus Book*. Oxford University Press, 2007.
8. G.W. Taylor, G.E. Hinton, and S.T. Roweis. Modeling human motion using binary latent variables. In B. Schölkopf, J. Platt, and T. Hoffman, editors, *NIPS 19*. 2007.