The hippocampus as a predictive recurrent neural network

Stefano Recanatesi, Greg Wayne, Mattia Rigotti

Modern investigations on the role of the hippocampal system are typically being pursued along two largely independent venues: a *spatial navigation view*, positing that the hippocampus is part of a complex dedicated to localization in the environment, and a *declarative memory view*, focusing on rapid encoding and retrieval of episodic experience.

Here we investigate a theoretical proposal that tries to reconcile these apparently contrasting views: that the hippocampus supports *a semantic relational network* organizing semantically related episodes in the service of sequential planning (Eichenbaum and Cohen, 2014). In particular, we provide an algorithmic grounding to this idea in the form of a Recurrent Neural Network (RNN) whose task is to predict future observations in partially observable environments, based on current observations and the actions taken by an actor module.

We tested our model by simulating an agent navigating a linear track and a squared 2d arena, and receiving only partial observations about its location. We hypothesized that, in order to accurately predict future observations, our RNN model had to learn to generate recurrent states that usefully summarize the history of its inputs, which, in a Markovian navigation task, would correspond to representing the current location. Indeed, our predictive training procedure leads to recurrent activations that are reminiscent of the well-known hippocampal physiological observations of place cells, border cells, and head-direction cells. Moreover, we show that these learned representations provide good generalization performance in reinforcement learning navigation tasks in partially observable environments.

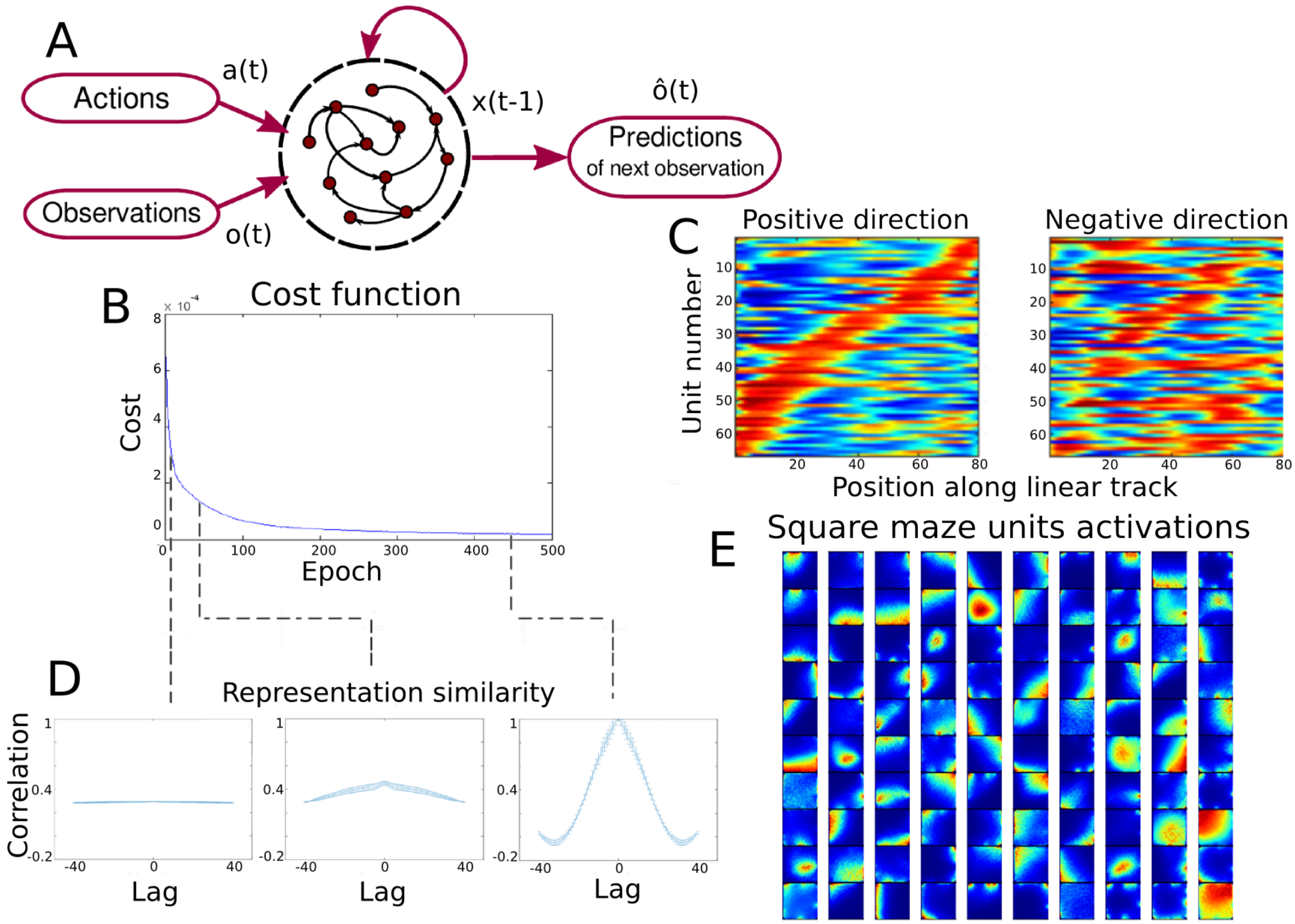
Our model supports the unifying view that both, place-related activity observed in the hippocampus during spatial navigation, as well as its involvement in episodic memory formation, could be a consequence of its role as a *semantic relational network* (Eichenbaum and Cohen, 2014), specifically via the computational mechanism of a recurrent system trained with predictive coding.

Methods.

Inspired by recent computational work that combines models of episodic memory formation (Howard and Kahana, 2012) and reinforcement learning (Dayan, 1993), we set up a RNN with the architecture depicted in Fig. A trained to predict future observations as described below. In addition to what's apparent from the figure, the RNN is structurally constrained, so as to simultaneously favor the learning of long term temporal dependencies (Mikolov et al. 2014) and to match the direct/indirect pathway anatomy of the hippocampus. At a given time t the RNN receives as input an observation vector $\boldsymbol{o}(t)$ and an action vector $\boldsymbol{a}(t)$. The output $\hat{\boldsymbol{o}}(t)$ of the network is trained to reproduce the observation vector at time $t+1$, $\boldsymbol{o}(t+1)$, by minimizing a quadratic cost function over the length T of an episode, $Cost = \frac{1}{T}\sum_{t=1}^{T}\left(\hat{\boldsymbol{o}}(t)-\boldsymbol{o}(t+1)\right)^2$, via the Back Propagation Through Time (BPTT) algorithm (Williams 1990). The cost function is such that, once training has converged (Fig. B), the network predicts future observations given the current observation and next action.

We simulated a linear track (a linear grid maze of length 80) through which the agent runs at constant speed, only switching direction at the edges. The agent also receives observations (random vectors spatially smoothed with a Gaussian kernel of size 2-steps) only for the first 5 and last 5 positions of the maze. We reasoned that the lack of external input in the central part of the track would be crucial to prompt the agent to form useful internal representations to correctly predict the observations at the edge of the run. In fact, the agent learns to predict future observations, and does so by building internal representations that are informative of its current location in the track. We demonstrate this by analyzing the activations of the recurrent units of the RNN for different locations, and showing in Fig. C that different units have a preference for different locations in space. Moreover, this spatial tuning (the specificity of the preference) increases during learning (Fig. D), and is directional, i.e. place selectivity depends on the direction of movement of the agent (Fig. C), consistently with what has been reported for place cells recorded in rodents in linear track experiments.

Next we summarize the results of the 2d environment simulation. The input to the network in this case is distance information from the wall in front of the agent, simulating somatosensory input (Sofroniew et al. 2015), as well as distal visual information in terms of the "color" of the wall in front of the agent. In Fig. E. we show the activation of all the 100 units in the RNN (one per panel) sorted and averaged according to the agent's location. Several units appear to be clearly place selective. These representations have interesting dependencies on the specific details of the action selection policy (which in this case is random memoryless exploration) and on the statistics of the observed inputs that we plan to further characterize in the future.



A. Network architecture. The network has a structurally-constrained recurrent layer that receives as input at each time *t* the action vector ***a****(t)* and the observations vector ***o****(t)*. The network is trained such that its output ***ô****(t)* predicts the observation ***o****(t+1)* in the next time step *t+1*. **B**. Learning curve: the cost function (the error in predicting the next observations) decreases as the learning progresses. **C**. Representation of space in the model trained on the linear maze. Left panel: The average activity of units in the recurrent layer sorted according to the position of their preferred location during runs in their positive direction (the direction for which they're the most active). Right panel: the activity of the same units sorted in the same way, during a run in their negative direction. **D**. Spatial tuning increases as a function of training. The three panels correspond to three different epochs during training. For each epoch we plot the average correlation between representations recorded at locations at different spatial distance. **E.** Average activations of all units in the network for the exploration of the agent in a 2d square maze. Each panel represent the average activity of one unit in all the positions of the maze.

References.

- Dayan P (1993). Neural Computation, 5(4), 613-624
- Eichenbaum, H. & Cohen, N. J. (2014). Neuron, 83(4):764–770.
- Gershman SJ, Moore CD, Todd MT, Norman KA, Sederberg PB (2012). Neural comp., 24(6), 1553-1568
- Howard MW, Kahana MJ (2002). Journal of Mathematical Psychology, 46, 269–299
- Mikolov, T., Joulin, A., Chopra, S., Mathieu, M., & Ranzato, M. (2014). arXiv:1412.7753
- Sofroniew NJ, Vlasov YA, Andrew Hires S, Freeman J, Svoboda K (2015). Elife. 2015 Dec 23;4. pii: e12559
- Williams RJ, Peng J. (1990). Neural Computation 2(4): 490-501