Modeling of Multi-Scale Spatial Navigation in Complex Environments

P. Scleidorovich[1], M., Llofriu[1] , B. Harland[2], M.Contreras[2], JM. Fellous[2], A. Weitzenfeld[1]

1. Department of Computer Science and Engineering, University of South Florida 2. Department of Psychology, University of Arizona

Abstract

Studies suggest that spatial navigation is supported by the multiscale neural system of the hippocampus and associated structures. The dorsal hippocampus (DH) plays a critical role in spatial learning and memory, but most of the supporting experiments have been conducted using simple spatial tasks in small arenas. Such conditions may not have required the integration of spatial representations at different scales along the dorso-ventral axis of the hippocampus, and hence may not have relied on ventral hippocampal (VH) computations.

Our overall hypothesis is that the interactions of spatial maps at multiple scales along the dorso-ventral axis of the hippocampus allows for spatial navigation in large and complex environments. We are developing a biologically-inspired computational model to simulate goal-directed navigation tasks in a large environment with different density of obstacles. The current work will present preliminary results related to this computational model and its evaluation with autonomous robotic systems.

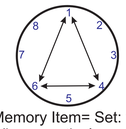
Methods

Animals

- Male brown norway rats, 7-8 months old.
- Bilateral cannulae in Dorsal or Ventral Hippocampus.

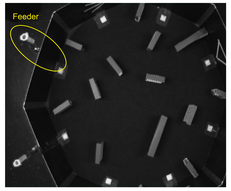
Behavioral Apparatus

- Open field arena with 8 equally spaced feeders containing sugar water.

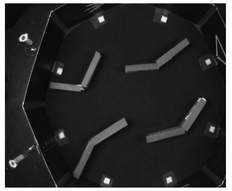


Memory Item= Set: 1,4,6
(in no particular order)
Objects: Lego blocks

Small objects

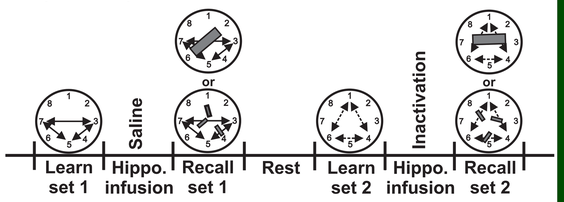


Large objects



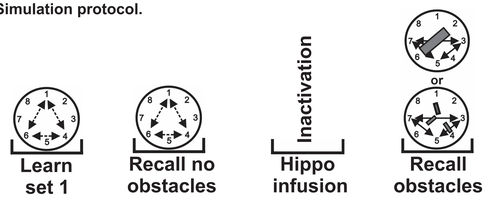
Hippocampus inactivation

- **Experimental protocol.**



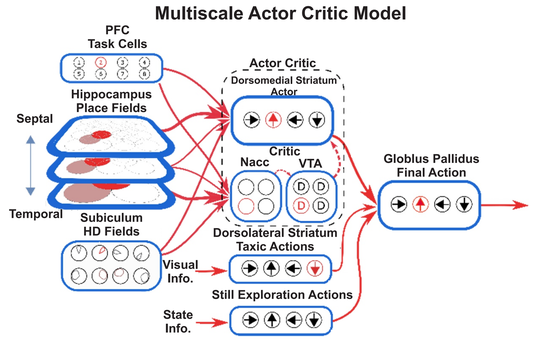
Infusion: Bupivacaine hydrochloride (bilateral, 2.5%, 1μl/side)

- **Simulation protocol.**



Model

Multiscale Actor Critic Model



- **Task Cells (TC)**

TCs model the target feeder. They allow the policy to depend on the intention of the rat.

- **Head direction cells (HD)**

Modeled as Gaussians centered on the angular distance between the rat's orientation and the cell's preferred direction.

$$HD_i(\theta) = N(distance(\theta, \theta_i), 0, \sigma_i)$$

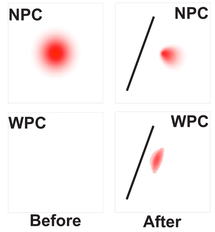
- **Place Cells (PC)**

Since the experiment requires the learning of the policy after introducing the walls, the state needs to encode wall information.

Based on the studies of Rivard et Al 2004, we model 2 types of cells "normal place cells" (NPC) and "wall place cells" (WPC).

NPCs are place cells whose field disappear when walls are introduced nearby.

WPCs are place cells whose field only appears after the introduction of a nearby wall.

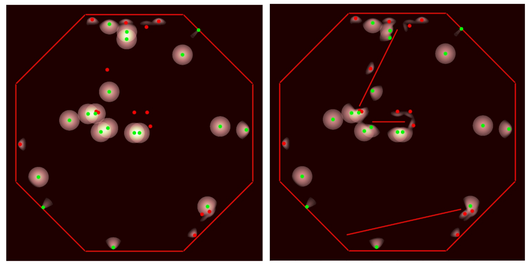


$$PC_j(d_j, d_w) = N(d_j, 0, \sigma_j) \cdot m(d_j, d_w)$$

Both types of PCs are modeled as the product of a Gaussian kernel of the distance to the PC's preferred location, multiplied by a modulator function that depends on both the distance to the cell's preferred location, and the distance to the closest wall. The modulator function differs for NPCs and WPCs.

- **Example of PC fields**

Red dots indicate WPC centers.
Green dots indicate NPC centers.



Before introducing walls After introducing walls

Model (continued)

- **Striatal Cells (SC)**

The actor critic algorithm uses striatal cells as soft states. Each striatal cell is the product of one PC, one HD cell and one task cell.

- **Reinforcement Learning (RL)**

State value function using soft states.

$$V^t = \frac{\sum_i SC_i^t * V_i^t}{\sum_i SC_i^t}$$

Action value function using soft states.

$$Q_j^t = \frac{\sum_i SC_i^t * Q_{ij}^t}{\sum_i SC_i^t}$$

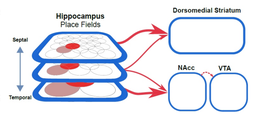
Error function and update equations using eligibility traces

$$e_t = r_{t+1} + \gamma \cdot V^t(s_{t+1}) - V^t(s_t)$$
$$V_i^{t+1} = V_i^t + \alpha \cdot \lambda_i^t \cdot e_t$$
$$Q_{ij}^{t+1} = Q_{ij}^t + \alpha \cdot \lambda_{ij}^t \cdot e_t$$

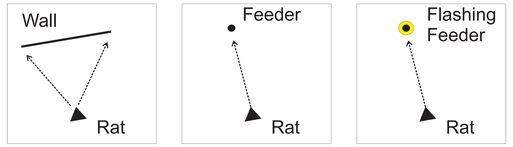
- **Network Connectivity**

In this model, striatal cells project to either the state value function or to the action value function, not to both.

- Septal layer project to action values
- Temporal layer project to state values
- Mid layer project half to actions values and half to state values



- **Taxic Modules**



Wall end taxic Feeder taxic Flashing Feeder taxic

Taxic modules guide navigation when the RL policy is being learned.
3 taxic modules provide votes for different actions.
Votes for an action are given by the difference between the positive reward expected at the taxic goal and the negative reward associated with taking each step times the number of steps to get to the goal after the action has been performed:

$$votes_j = pReward - nReward * steps$$

- **Exploration module**

Exploration is performed using an exponential decaying schema:
The module chooses a random action and assigns votes to that action weighted by a negative exponential function of the elapsed time.

- **Action Selection**

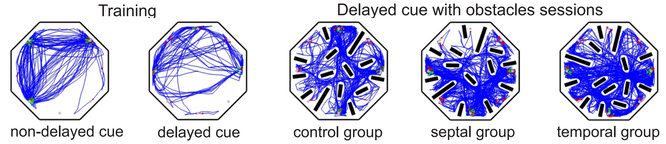
Action votes from RL, taxic and exploration modules are added together.
The action to be performed is selected using a WTA (winner take all) process.

- **Inactivation**

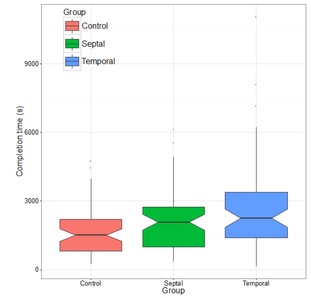
Simulation of injection to either the septal or temporal layer.
To simulate injection each cell in the respective layer is assigned a random distance to the site of injection.
Then the activation of each place cell is modulated by a function inversely proportional to the cube of the distance.

$$mod(PC_i) = \frac{k}{inj\_d_i^3}$$

Simulation Experiment Results



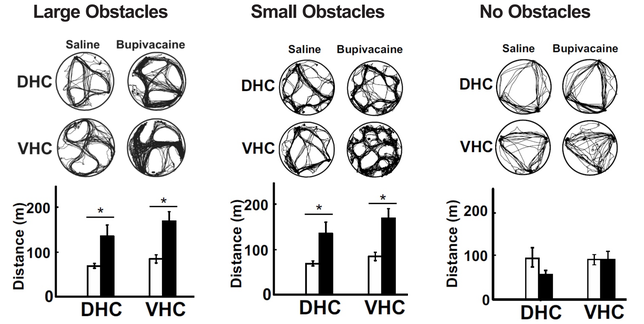
Typical paths followed by the simulated robot for different task phases and groups. Green dots indicate successful reward deliveries, whereas red dots show unsuccessful deliveries. Delayed cue paths were chosen from the individual with performance closest to its group median.



Completion times for the delayed cue phase with small obstacles for all three groups over 64 simulations. The septal portion of the HPC was inactivated in the Septal group and the temporal portion in the Temporal group. The plot shows box plots using the 1.5 IQ outlier criterium.

We note how deactivation of either the septal or temporal groups results in a lower performance than the control group (no inactivation). Additionally, inactivation of the temporal group results in a lower performance

Rat Experiment Results



Conclusions and future work

Our preliminary study shows that the model displays similar results to the experiments with rats. When using small obstacles, ventral and septal groups resulted in reduced performance when compared to controls, but no statistical difference was observed between both groups.

Future studies will further analyze the differences between ventral and septal roles, assessing the advantages of multiscale representation when varying PC size distributions.

Also, future studies will assess the role of using wall modulated place cells in the model.

References

- Rivard, B., Il, Y., Lenck-Santini, P.-P., Poucet, B., & Muller, R. U. (2004). Representation of Objects in Space by Two Classes of Hippocampal Pyramidal Cells. The Journal of General Physiology, 124(1), 9–25. https://doi.org/10.1085/jgp.200409015
- Contreras, M., Pelc, T., Llofriu, M., Weitzenfeld, A., & Fellous, J. M. (2018). The ventral hippocampus is involved in multi-goal obstacle-rich spatial navigation. Hippocampus, in press.

Acknowledgments

This work has been funded by NSF IIS RI #1703340: "Experimental and Robotics Investigations of Multi-Scale Spatial Memory Consolidation in Complex Environments".