Topological map learning during pre-exposure and replay as an explanation of latent learning

P. Scleidorovich[1], M. Llofriu[1]
B. Harland[2], N. Cazin[3], P. F. Dominey[4], J.-M. Fellous[2], A. Weitzenfeld[1]
1 Dept. of Computer Sci. and Engin., Univ. of South Florida, Tampa, FL
2 Dept. of Psychology, Univ. of Arizona, Tucson, AZ;
3 Human and Robot Cognitive Systems,
4 INSERM U846, Bron, France;

USF UNIVERSITY OF SOUTH FLORIDA COLLEGE OF ENGINEERING

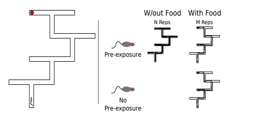
Abstract

In Blodgett's (1929) experiments, rats that were pre-exposed to a maze without rewards were able to learn a task on that maze faster. This phenomenon was termed latent learning. In this work we present a model of Blodgett's multiple-T maze experiments, and we hypothesize that latent learning could be explained by improved replay events at the start of the rewarding trials due to having a "better cognitive map".

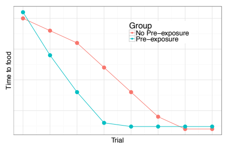
Our replay model, based on Johnson & Redish (2005), updates intra-hippocampal synaptic weights during navigation. Then, during resting periods these weights are used to generate replay events which, when rewarded trials begin, can then be used to drive off-line learning.

We show that pre-exposed artificial rats are able to learn the task significantly faster, validating the hypothesis of the model. We also show that the effect is increased with the number of pre-exposed trials. We also present an analysis of the shortcomings of this replay model and propose possible solutions. This work is part of our current project on assessing the role of the intra-hippocampal synapse modulation in tasks reminiscent of the "Traveling Salesperson Problem" (TSP), where rats have to optimize their navigation to multiple rewarded feeders.

Task

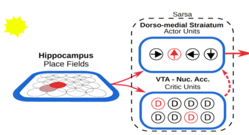


The task consists in training two groups of rats to find the food at the end of a 4x T maze. While one group receives food from the very first day, the other is allowed to explore the maze for the first few days before starting the rewarded trials. Each day only one trial is performed per rat.

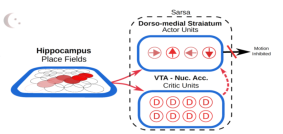


Blodget's original experiment showed that the pre-exposed rats learn the task faster.

Model

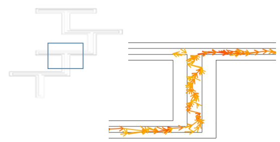


During navigation, the rat uses an actor-critic reinforcement learning algorithm to guide navigation. At the same time, intra-hippocampal synaptic weights (connection matrix weights) are updated thus generating a "summary of the traversed paths" as trials go by.

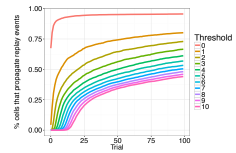


During resting periods, the connection matrix is used to generate replay events. To do so, first a random PC is activated, and its activation is propagated to its neighbor with maximum connectivity. This process is repeated until the replay reaches the rewarding location, the maximum connectivity of a cell does not reach a given threshold or a loop is formed. The generated sequence is given as a training sequence to the reinforcement learning algorithm.

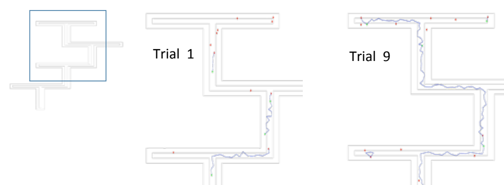
Replay Events Vs Trials



The connectivity matrix after one trial with food. The pattern already shows an ascending path to the food.

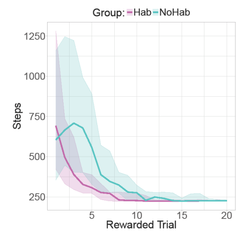


The percentage of cells that would propagate a replay event successfully as a function of trials and the propagating threshold.

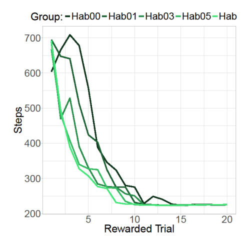


Replay paths generated using the connectivity matrix learned through experience. The paths become longer as trials go on. Green and red dots indicate start and end positions of replay events.

Latent Learning Observed

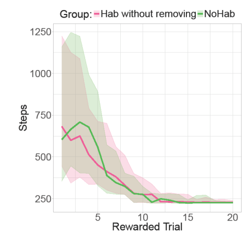


To see whether the model could account for latent learning, 100 rats were simulated with 10 habituation days, and another 100 hundred without. The group with habituation learns faster the path to the food. The lines shows the median and the shaded regions show upper and lower quartiles.



As seen in the plot, the learning rate increases with diminishing gains with the increase in the number of habituation days.

Limits of the explanation for the current model



Blodgett's original experiment removed the rats at the end of the maze both in rewarded and unrewarded trials alike. Thus, following the experiment we did the same.
Nonetheless, it is natural to expect rats to exhibit latent learning irrespectively if they are or if they aren't removed at the end of the maze during unrewarded trials. Thus, to test this condition, one extra group of 100 rats was trained with 10 days of habituation but letting the group explore freely without removing it at the end of the maze.
As seen on the graph, latent learning was not observed under this circumstances.

Discussion and Conclusions

1. The model shows how the latent learning phenomenon be explained in terms of replay events and a reinforcement learning algorithm.
2. The current replay model only serves to explain latent learning if the rat is removed at the end of the maze. This is due to the fact that the connectivity matrix generated by the model learns the trending paths performed by the rat rather than the actual map of the environment. When removing the rat at the end, the path from start to end gets consolidated as trial number increases. Thus when rewarded trials begin, the matrix is able to replay the path to the food sooner than non habituated rats. When removing the rat at the end, no one path is more frequently observed than the others, thus not biasing the replayed sequences and not showing the latent learning phenomenon.
3. For future work, matrix dynamics that learn "the map" rather than "the path" could be used, but this would require adapting the current replay model to use this information, and somehow also store the traversed path information on it.

References

Blodgett, Hugh Carlton. "The Effect of the Introduction of Reward upon the Maze Performance of Rats." University of California Publications in Psychology, 1929.
Johnson, Adam, and A. David Redish. "Hippocampal Replay Contributes to within Session Learning in a Temporal Difference Reinforcement Learning Model." Neural Networks 18, no. 9 (2005): 1163–1171.
Small, Willard S. "Experimental Study of the Mental Processes of the Rat. II." The American Journal of Psychology 12, no. 2 (1901): 206–39. doi:10.2307/1412534.

Supported by NSF Grants 1117303 and 1117302 to JMF, PD and AW

UNIVERSITY OF SOUTH FLORIDA

BioRobotics-Lab