

Modeling of Hippocampal Behavior

Diana Ponce-Morado, Venmathi Gunasekaran and Varsha Vijayan

Abstract

The hippocampus is identified as an important structure in the cerebral cortex of mammals for forming new memories, storing these memories independently and retrieving them. The ability to associate to and distinguish between similar experiences forms an important component of learning. This project will attempt to model the learning dynamics of the Dentate Gyrus (DG) and CA3 regions in the hippocampus which revolve around two major methods of Pattern Separation and Pattern Completion. Based on previous research, our computation model incorporates Hebbian learning rule. We attempt to implement the key concept of Pattern Separation and highlight the constraints involved in our model.

Keywords: Hippocampus, Memory, Pattern Separation, Pattern Completion, Hebbian learning

1 Introduction

Electrophysiological experiments, lesion studies and immediate early-gene tests performed by Gold and Kesner [1] on mice have proven that the hippocampus is critical in pattern separation and pattern completion. Specifically, the DG and CA3 regions of the hippocampus are responsible in helping mammals identify different objects or patterns and recognize patterns when a partial or distorted input is given. This phenomenon is called associative memory. David Marr [2] suggested the presence of recurrent loops within the hippocampus that enable auto-association and hence, pattern completion. Since then, many research groups have attempted to model the hippocampus.

1.1 Pattern separation and Pattern Completion

Pattern separation [3] is the process of distinguishing between two similar objects and storing them in an orthogonalized fashion. Pattern completion deals with reconstruction of an experience based on previously stored memory given partial or distorted information. These are seen in Figure 1. In an experiment conducted by Kesner et al. [2], rats trained to spot food in one of a given set of four possible locations were let inside a black box. Even after the removal of all available cues amongst a control group, the rats exhibited excellent pattern completion. In a second test group, neurotoxic injections that affected the CA3 subregion of the hippocampus was administered. These CA3 lesioned rats showed progressive increase of errors as the number of given cues were reduced. This experiment proved to be a hallmark in hippocampal studies for a few significant reasons – CA3 was identified to be the specific region of Hippocampus responsible for spatial pattern completion; the importance of hippocampus in memory recall and its application under novel situations was reinforced. The setup is shown in Figure 2.

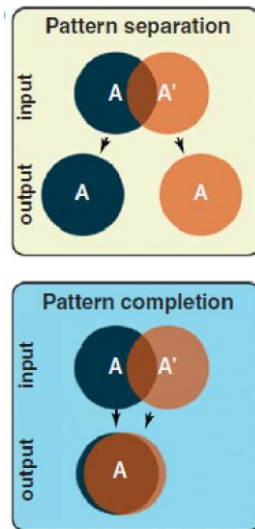


Figure 1.1.1 Conceptual representation of Pattern Separation and Completion. [3]

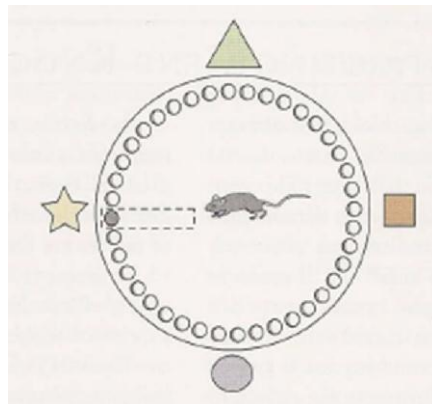


Figure 1.1.2: Experimental setup used by Gold A. et. al. [1]

1.2 Basic Hippocampal Circuitry

Hippocampus implements information transmission unidirectionally in three distinct stages [5] as seen in Figure 3. Projections of the axon from layer 2 of the Entorhinal Cortex (EC) reach the Dentate Gyrus (DG) and also proceed to make synapses with the next stage, CA3. These happen via Perforant Path (PP), acting as the crucial interface between the cortex and the hippocampus. The unmyelinated axons of the DG project onto CA3 with the help of mossy fibres. An interesting aspect of region CA3 is that it receives three major excitatory inputs [1] – from (i) EC, (ii) DG and (iii) its own recurrent collateral (RC) input. The third stage, CA1 has inputs from the third layer of EC and CA3 by means of perforant path and Schaffer Collaterals (SC) respectively. A schematic of the network in the hippocampus is shown in Figure 1.

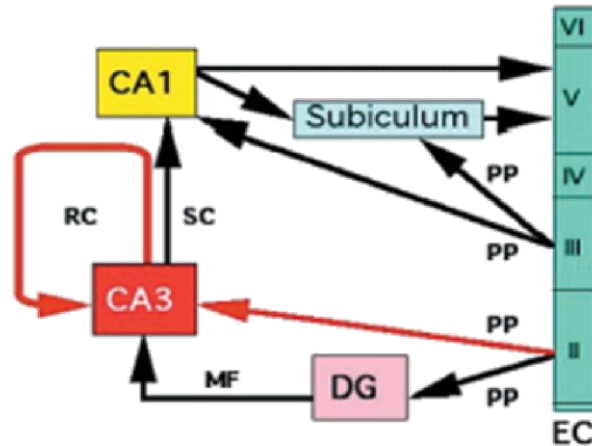


Figure 1.2.1 Schematic representation of Hippocampal circuitry

It is known from previous research [5] that there are about 10^7 DG granule cells in man which sparsely, yet strongly synapse with the CA3 pyramidal cells ($\sim 2.3 \times 10^6$ in man). It has been concluded that each CA3 cell receives not more than 50 mossy synapses and the maximum inputs for CA3 comes from RC.

Earlier computational models have suggested that DG performs Pattern Separation; CA3 is involved in both spatial pattern separation and completion and CA1 takes part in pattern completion and temporal pattern separation.

1.3 Goals

In this project, we have attempted to model the hippocampal regions involved in spatial pattern separation. We analyzed two different computational approaches to demonstrate how pattern separation and completion might occur. Both the methods are a form of Hebbian learning. The first method uses logic for learning simple binary inputs by updating the CA3 memory bank. The second method applies principles of neural networks to implement modified Hebb learning rule. We are exploring the possibilities of extending the same to spiking neurons.

2 Methods

2.1 Feedforward associative network model: Binary-valued synapses

The feedforward network model is a simple approach to trace how information is relayed from the entorhinal cortex to the hippocampal circuit. Figure 2.1.1.a shows three entorhinal neurons (red, blue, and green) synapsing onto eight dentate gyrus neurons (labeled purple). Before each synapse between the entorhinal and dentate gyrus neurons show learning by increasing their synaptic conductances, a weight matrix is constructed containing weights (Fig.

2.1.1.b), or the strength of synapses, denoted by a randomly generated number between 0 and 1. For synapses containing a value of 0.8 or higher, it will denote a strong synaptic connection otherwise synapses are labeled as weak synaptic connections. Each weight is stored and accessed in a two dimensional matrix, representing the number of entorhinal and dentate gyrus neurons in a network model. Synaptic conductance values between the dentate gyrus and CA3, and between CA3 and CA1 neurons, do not contain starting weights that are randomized. These weights are initialized by setting their conductance values to zero.

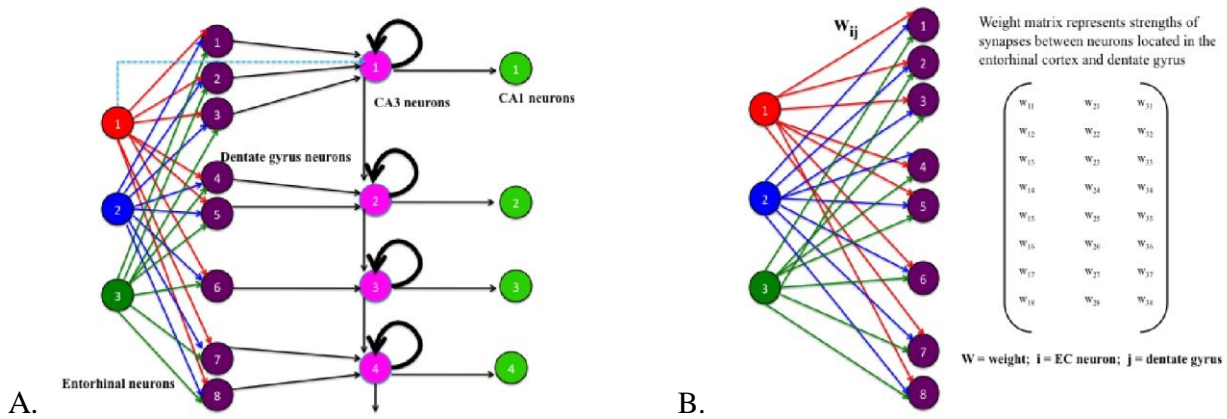


Figure 2.1.1 (a) Generalized circuit of binary network, and (b) illustrating a weight matrix to denote how the strengths of each synapse is stored and archived.

The inputs of the model are given as binary inputs to entorhinal neurons. An input of one indicates that a particular neuron has been stimulated and an input of zero indicates the neuron did not receive a stimulus. Thereafter, for each trial, the network stores a history of its strengthened conductances to determine how much the network has learned over time.

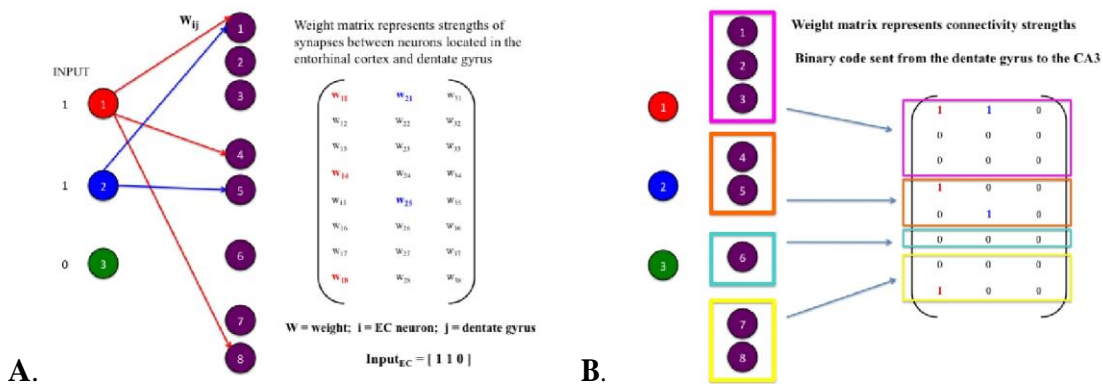


Figure 2.1.2 (a) Example of a trial with inputs 1, 1, 0, and (b) Binary information consolidation at the site of the CA3 neuron subregion.

Figure 2.1.2.a presents how the synaptic connections are mapped and represented by the weight matrix, specifically for input sequence 1, 1, and 0. Before the code 1, 1, 0 arrives to CA3 neurons via the mossy fiber pathway, the code is sent directly to CA3 neurons via the perforant path from the entorhinal neurons to determine if a similar input already exists in the CA3 memory bank. The network will perform a search and comparison computation on the CA3 memory bank. This computation will search for a match between each stored array and the input sequence code to determine if the pattern has been previously learned. In the case of an identified learned pattern, the network will relay the matched stored information from the memory bank to the CA3 neuron via the recurrent collaterals. This information will then be relayed forward to the CA1 neurons. However, before a pattern can be learned by the network and its information stored in the memory bank, a series of repeated trials must occur in order to reach a threshold where the correct code is stored in memory. Figure 2.1.2.b shows how the CA3 neuronal subnetwork consults binary code from each subpopulation in the CA3 area.

2.2 Artificial Neural Networks model

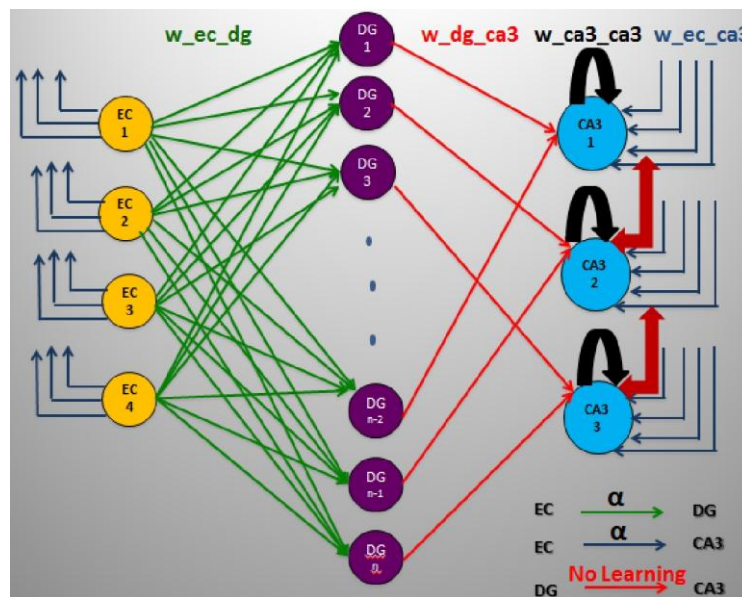


Figure 2.2.1 A snapshot of the network used

We also realized this network using concepts from neural networks and implemented learning using a modified Hebb rule. All neurons in the EC were connected to every neuron in the DG as well as CA3. However, as studies by Kesner et al [5] suggested, the neurons in DG are grouped in such a way that no two neurons in CA3 receive an input from the same subset of neurons in DG. Also, the number of neurons in DG is about a hundred times greater than that of the number of neurons in CA3. This indicates that pattern separation in the DG is achieved by the physiological manner in which the neurons are connected. Learning, in the hippocampus, occurs in the perforant pathways between the EC and DG and the EC and CA3. There's no associative learning between DG and CA3. Learning also occurs in the

recurrent network between the neurons in the CA3. These help complete patterns by exhibiting associativity. Hebb rule was used to implement long term potentiation and depression by training the neurons with a set of patterns.

$$\Delta W_{ij} = \alpha X_i(T_j - Y_j)$$

When the input and output neurons of a synapse are both on, the synapse strengthens. Else, it weakens. This way, when a synapse is repeatedly strengthened, long term potentiation occurs. The neural network that was created was tested using similar and slightly different partial inputs. In addition, to test the variability in the input that the system can tolerate, gaussian noise was introduced. This was done to check if the network gave a reasonably meaningful match with the trained input. The results are shown in the section to follow.

3 Results and Discussion

3.1 Feedforward associative network model: Binary-valued synapses

The network was trained to learn a pattern through multiple trials. Figure 3.1.1 shows an example of the network attempting to learn the pattern 1, 1, 0. This example underwent six trials in order to learn the specific input sequence. As explained previously, each consecutive trial archived information about its previous history until the network fully learned the sequence, in this case 1, 1, 0. This particular behavior in the model illustrates how hippocampal neurons in-vivo might process information.

<p>Learning has not occurred yet: CA3_memory = 0 1 0</p> <p>Novel stimulus.</p> <p>CA3_memory = 0 1 0 0 1 0</p>	<p>Final result in CA3 memory bank: CA3_memory =</p> <p>0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 1 1 0</p>
--	--

Figure 3.1.1 Example of a training program to learn inputs 1, 1, 0. There are six trials that were performed in order for the program to learn the pattern.

3.2 Application of artificial neural networks model

The system was designed with 4 neurons at the Entorhinal cortex and 3 neurons at the CA3 region. The number of neurons in the dentate gyrus was very high (1000). When we gave a lesser number of neurons in the DG, we noticed that either proper learning never occurs. The system either entered into an infinite loop, changing the weights according to the training inputs presented or it claimed to have trained the system but was unable to complete patterns or even recognize the correct pattern.

The network was trained for a set of 3 inputs as shown in Table 1. When the inputs given were very similar with drastically different outputs, we observed that the system took long to train, sometimes even more than a thousand epochs, where an epoch is defined as the number of times the entire training set is presented to the system repeatedly so as to ensure that it has learned them. Hence, we needed to find inputs that were quite different before we could train the system to generate the desired outputs. With the training set given in Table 1, the system required no more than 4 epochs to train.

Table 1: Training Input and Output patterns used

Input Pattern	Output Pattern
[1, 0, 0, 1]	[-1, 1, 1]
[0, 1, 1, 0]	[1, -1, -1]
[0, 1, 0, 1]	[1, -1, 1]

When a partial input was given, the system was able to identify it as a part of one or more of the training inputs. Screenshots of the system's performance are given below, in Figure 3.2.1.

```
Enter Input as [a, b, c, d]
[0, 0, 0, 1]
out: [1, -1, 1]
Closest pattern(s):
[1, 0, 0, 1]
Closest pattern(s):
[0, 1, 0, 1]
Enter Input as [a, b, c, d]
[0, 0, 1, 0]
out: [1, -1, 1]
Closest pattern(s):
[0, 1, 1, 0]
Enter Input as [a, b, c, d]
[0, 0, 0, 0]
out: [1, -1, 1]
Unfamiliar pattern
Enter Input as [a, b, c, d]
```

Figure 3.2.1 Performance of the system when partial inputs were given. Pattern Completion is exhibited by the system.

Noise analysis

To test the performance of the system when noise was introduced into the input, we added Gaussian noise (using the `random.gauss(mean, standard_deviation)` function in Python) to the input at the EC. We noticed that the system took longer to train. With zero noise the system learned the patterns in no more than 5 epochs but when noise was introduced, a minimum of 5 epochs was required for it to learn them. However, as the amount of noise was increased, the number of epochs needed did not seem to increase significantly.

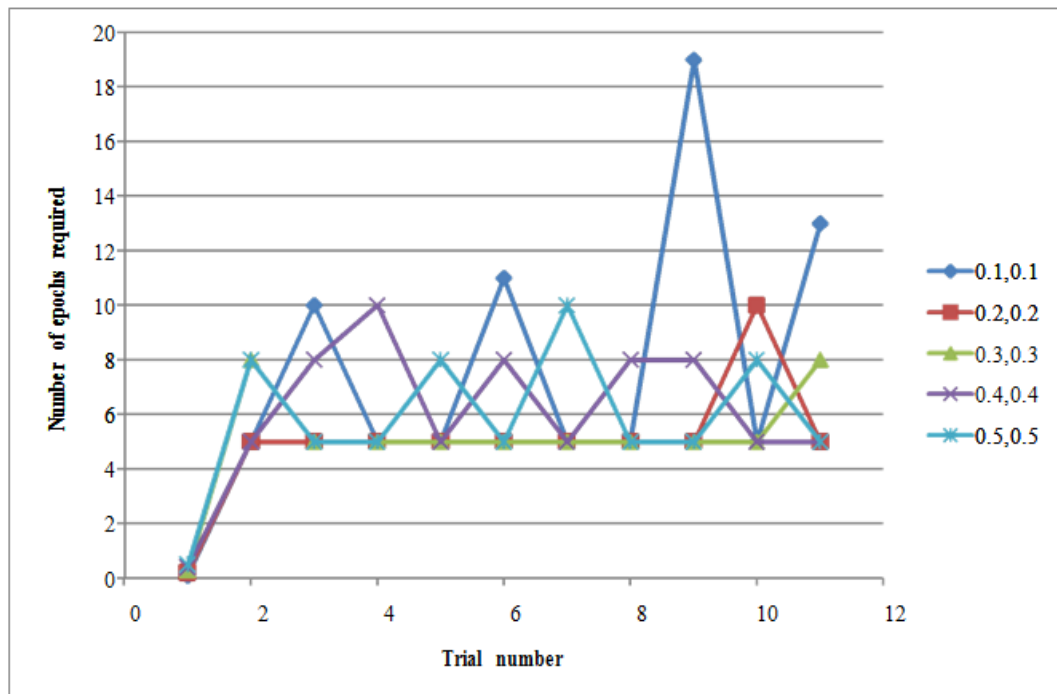


Figure 4.2.1: Number of epochs to train the system at each trail for different amplitudes of Gaussian noise introduced

Also, when noise was introduced, the system wasn't able to detect all the closest patterns.

```
Enter Input as [a, b, c, d]
[0, 0, 0, 1]
out: [ 1. -1.  1.]
[2]
Closest pattern(s):
[0, 1, 0, 1]
Enter Input as [a, b, c, d]
```

Figure 4.2.2 Output when noise with mean 0.5 and standard deviation of 0.5 was given to the system. Pattern completion occurs, but not all closest patterns are recognized.

4 Conclusion

In this project, we were able to model a simplified version of the network in the hippocampus that is responsible for associative learning and accomplishes pattern completion. As mentioned previously, pattern separation occurs physiologically as the neurons in the dentate gyrus are clustered into groups, each of which activate or synapse onto a different set of CA3 neurons. Also, the CA3 recurrent loop plays a key role in pattern completion. During training, if two or more output neurons are simultaneously activated by a certain input pattern, the recurrent collaterals between them strengthen over time. When presented with a partial input that activates only part of the output, the recurrent collaterals recognize the pattern from memory and help activating the rest of the output neurons.

5 Future Directions

The next step is to implement backpropagation rule from CA3 to EC and DG to EC layers will enhance the robustness of the network to partial inputs. Inclusion of inhibitory interneurons to the model will bring it closer to a biophysical system that doesn't just operate on excitatory signals. To simulate brain activity more realistically, the artificial neurons can be replaced with spiking neurons. If the challenging task of training spiking neurons is overcome, a reasonably accurate hippocampal model can be obtained. In addition to the above mentioned possible improvements, modeling CA1 region will round up and complete the hippocampal model with pattern completion capabilities.

6 References

- [1] Gold, A. E., & Kesner, R. P. (2005). The role of the CA3 subregion of the dorsal hippocampus in spatial pattern completion in the rat. *Hippocampus*, 15(6), 808-814.
- [2] Marr, D. (1971). Simple memory: a theory for archicortex. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 23-81.
- [3] Yassa, M. A., & Stark, C. E. (2011). Pattern separation in the hippocampus. *Trends in neurosciences*, 34(10), 515-525.
- [4] Treves, A., & Rolls, E. T. (1994). Computational analysis of the role of the hippocampus in memory. *Hippocampus*, 4(3), 374-391.
- [5] Kesner, R. P., Lee, I., & Gilbert, P. (2011). A behavioral assessment of hippocampal function based on a subregional analysis. *Reviews in the Neurosciences*, 15(5), 333-352.