

Figure 1: The hippocampus. This is a modified image originally due to Cajal, probably from his 1911 book, the inset shows the approximate connectivity. [From [http://en.wikipedia.org/wiki/File:CajalHippocampus_\(modified\).png](http://en.wikipedia.org/wiki/File:CajalHippocampus_(modified).png)]

Introduction

These notes are about the hippocampus and introduce a more top down style of modelling.

Anatomy of the hippocampus

The hippocampus is situated at the edge of the cortex and is divided into two main areas. The hippocampus complex is distinctive in shape and these are named for the shape

- Cornu Ammonis (CA) - meaning the *horn of Ammon*, an Egyptian god of fertility with curved horns. The CA is usually divided into four regions, labelled CA1 through to CA4.
- Dentate Gyrus (DG)- gyrus is the name given to the ridges in the cortex, dentate means *with teeth*. The dentate gyrus is one of the few areas of the adult brain that exhibits neurogenesis.

In addition, the main input to the hippocampus comes from the

- Entorhinal Cortex (EC) - entorhinal means *near the smell processing area*.

and in this discussion this will be treated along with the hippocampus since it participates in hippocampal processing.

One key aspect of the anatomy we are going to consider here is that the excitatory neurons in DG are feedforward, that means they lack lateral connections, that is connections between themselves. CA3 in contrast is a recurrent network, the excitatory neurons are interconnected. DG is often called feedforward and CA3 recurrent even though in a way this ignores the recurrent connections in DG between excitatory and inhibitory cells.

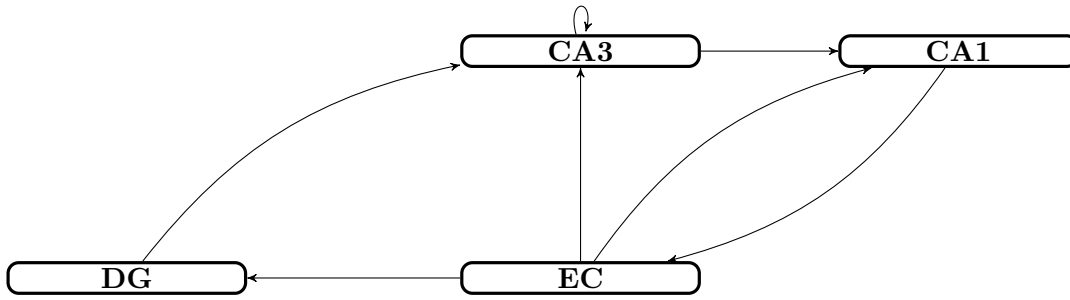


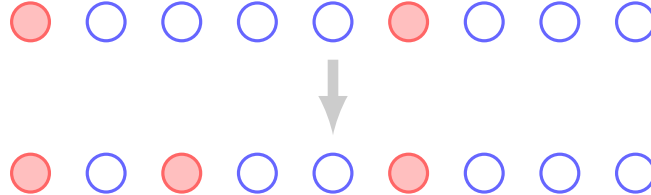
Figure 2: Connectivity of the hippocampus. A rough diagram showing the major connections between the areas of the connectivity. The set of axons running from EC to DG, CA3 and CA1 is called the perforant pathway, the mossy fibres run from DG to CA3 and the Schaffer collateral fibers go from CA3 to CA1. The loop on CA3 is supposed to represent the high level of recurrent connections in that regio.

Auto-associative memory

A standard paradigm for memory in the hippocampus is *auto-associative* memory. Auto-associative memories are patterns representing memories along with some dynamics that complete partial patterns. Imagine a sequence of on-off neurons



where the filled circles correspond to on. Recall occurs when the network is presented with a partial pattern and evolves into the complete patterns.



The idea is that the hippocampus implements a network which performs auto-associative memory.

A model of CA3

Here a highly simplified model of CA3 is presented [4]. In this model CA3 is all-to-all connected and made up of McCulloch-Pitts neurons [3]. We will consider the dynamics of these neurons shortly, for now it is enough to note that they are binary neurons with on and off states, this would correspond, biologically, to firing and not firing and obviously abstracts away all the details of firing along with the possibility of there being different firing rates. We will call the on state '1' and the off state '0'. Now let N be the number of neurons, x_i the activity of neuron i and w_{ij} the strength of the connection for i to j . The sparseness, the average proportion of neurons active at any one time is a , this is believed to be very small in actual neurons. For this simple model $w_{ij} = w_{ji}$.

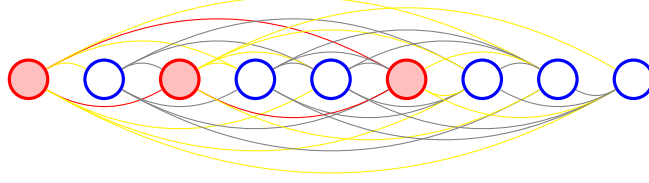


Figure 3: Learning in the associate network. The pattern has been imposed and connection strengths are changed. The red links increase by $\eta(1 - a)^2$ and the gray by $\eta(-a)^2$, the yellow links decrease by $\eta a(1 - a)$.

During learning the patterns are activated and plastic changes are made to the synapse strength according to a simple correlation based Hebbian plasticity rule.

$$\Delta w_{ij} = \eta(x_i - a)(x_j - a) \quad (1)$$

where η is the learning rate, often a small number, but, in hippocampus where memories need to be learned quickly, possibly during a single presentation, η is large. Since a is very small too for real networks there will be a large increase for the connection between two neurons that are active at the same time, a tiny increase for pairs neurons that are inactive at the same time and a medium size decrease for pairs of neurons where one is active and one inactive. See Fig. 3.

During recall some of the neurons are held in the active state and the rest of the network evolves according to a threshold input rule. That means each neuron has an input given by

$$h_i = \sum w_{ij}x_j \quad (2)$$

and is set in the active state if $h_i > \theta$ where θ is a threshold which is set to different values for different networks. The idea is that after learning the pattern $\{0, 2, 5\}$



the connections between these nodes will be strong, so if the network has nodes $\{0, 5\}$ activated



the value $h_2 = w_{12} + w_{52}$ will be larger than threshold and the subsequent dynamics will switch neuron 2 on. However, in this network, if a different initial set of neurons are activated, the activity will die away because the h_i will all be sub-threshold.

Capacity

When many patterns are stored it is likely that there will be interference between them. This is illustrated in Fig. 4. Although the figure shows how a single neuron fails to participate in two patterns, for larger networks some overlap is possible, but too much overlap prevents retrieval. In fact the capacity is proportional to the number of neurons, N . A hand-waving argument goes like this: the number of connections is roughly N^2 and the amount of information in a pattern is N so the number of patterns that can be stored is $N^2/N = N$ [4].

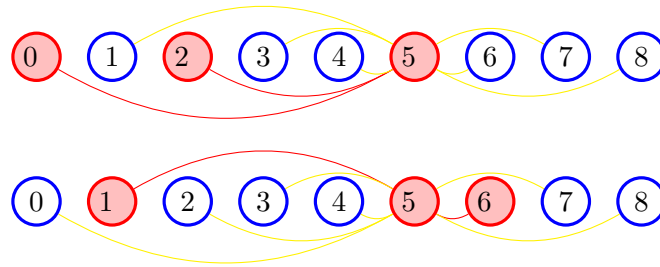
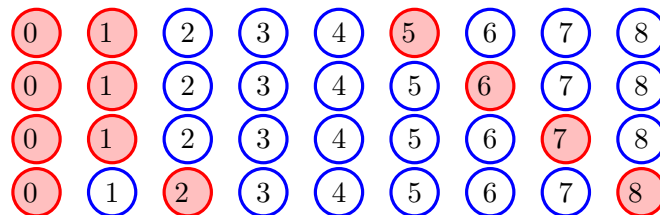


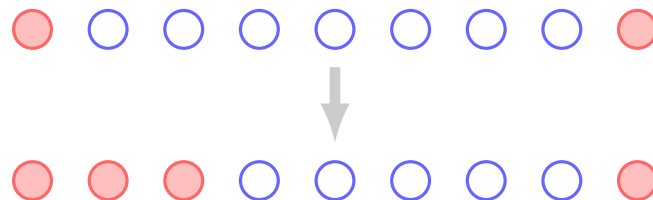
Figure 4: Interference in an associate network. Neuron 5 is involved in two patterns and, as a consequence, some of its connections are strengthened for one pattern and weakened for the other, if these strengthening and weakening effects are similar in size it makes it unlikely that either pattern will be accurately retrieved.

Correlated patterns

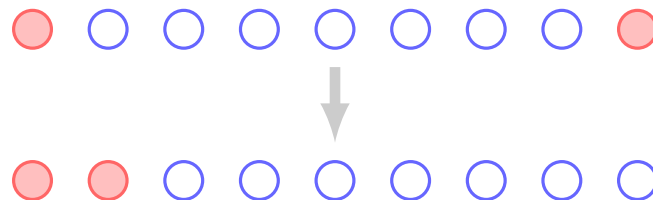
The estimates above assume that the patterns are all independent. If they aren't the capacity is reduced. If patterns share some fragments or subpatterns then the connections in these sub-patterns become very strong, perhaps dominating other elements in the patterns, the elements that make them different. Consider the four patterns



The connection between neurons 0 and 1 will become very strong because this connection is present in three patterns out of four. It is likely that this will result in this erroneous completion



or even



This means that auto-associative networks are not able to effectively store anything except random patterns! This is why they have never proved useful for machine learning; we will

see later, when looking at cortical memory, that if there are multiple presentations, the memory system would have the opportunity to learn different, similar, memories. The goal here, however, is to learn the memory quickly after a small number of presentations.

In the case of hippocampus it has been proposed, [6], that this problem is solved through the EC-DG-CA3 pathway and that the role of the dentate gyrus is to randomize the connectivity between EC and CA3. In short, during learning neurons in EC and CA3 are matched via DG and that the connections from EC to DG and from DG to EC are essentially random. In other words in this simple model CA3 is an auto-associative memory store and DG, as a feedforward network just passes the pattern along, but randomizes it a bit on the way to help keep similar patterns separate. This is referred to as pattern separation.

References

- [1] O’Keefe J, Dostrovsky J. (1971) The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Research* 34: 171–5.
- [2] Quiroga R, Reddy L, Kreiman G, Koch C, Fried I. (2005) Invariant visual representation by single neurons in the human brain. *Nature* 435: 1102–7.
- [3] McCulloch W, Pitts W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* 7: 115–33.
- [4] Amit D. (1992) *Modeling Brain Function: The World of Attractor Neural Networks*. Cambridge University Press, Cambridge England.
- [5] O’Reilly RC, Munakata Y (2000) *Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain*. MIT press, Cambridge MA.
- [6] O’Reilly RC, McClelland JL (1994) Hippocampal conjunctive encoding, storage, and recall: Avoiding a tradeoff. *Hippocampus* 4: 661–82.
- [7] Hasselmo ME. (2006) The role of acetylcholine in learning and memory. *Current Opinion in Neurobiology* 16: 710–71
- [8] Lisman JE, Grace AA. (2005) The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron*, 46: 703–713.