The Hippocampus

These notes are about the hippocampus; for convenience some of the figures here repeat the figures already used in the notes on the McCulloch-Pitts neurons.

Anatomy of the hippocampus

The **hippocampus** is situated at the medial ("toward the middle") edge of the temporal lobe, along the floor of the temporal horn of the lateral ventricle. It is divided into two main areas. The name *hippocampus* means seahorse in Latin, and comes from the appearance of the hippocampal formation and it's output tract. The hippocampus complex, and contains sub-regions also named for their shape:

- **Cornu Ammonis (CA)** - meaning the *horn of Ammon*, an Egypian god of ferility 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.

The role of the hippocampus

The role of any brain region is complex and the hippocampus is no exception. It may play a role in olfaction for example, however, it is widely believed that its principle role is in memory and in constructing spatial maps. Studying patients with hippocampal damage shows that it is involved in declarative memory, that is the sort of memory that can be described in words. It does not play a role in procedural memory, the memory process which allows us to learn new motor skills. It appears, again from patients with hippocampal damage, that some long term memories are stored outside the hippocampus.

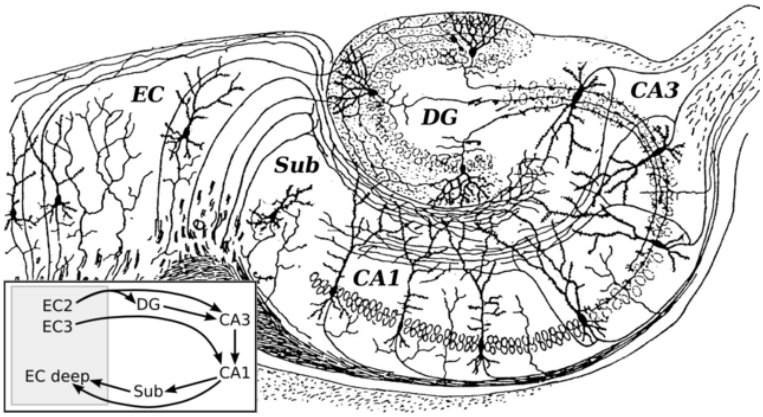


Figure 1: The hippocampus. This is a modified image originally due to Cajal, probably from his 1911 book, the inset shows the approximate connectivity. [From `http://en.wikipedia.org/wiki/File:CajalHippocampus_(modified).png`]

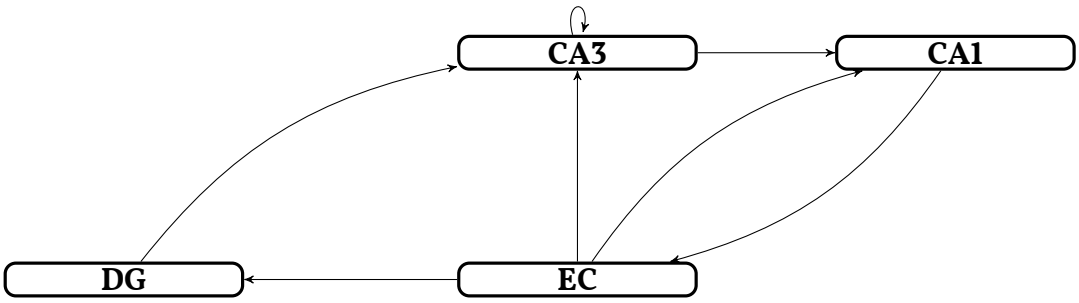


Figure 2: Connectivity of the hippocampus. A rough diagram showing the major connections between the areas of the connectivity. The set of axons running from EC to DG, CA3 and CA1 is called the perforant pathway, the mossy fibres run from DG to CA3 and the Schaffer collateral fibers go from CA3 to CA1.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

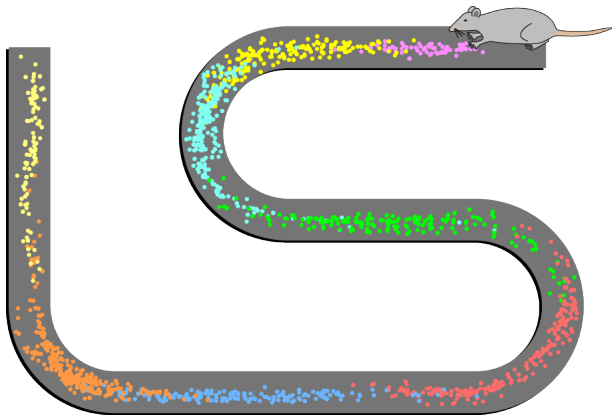


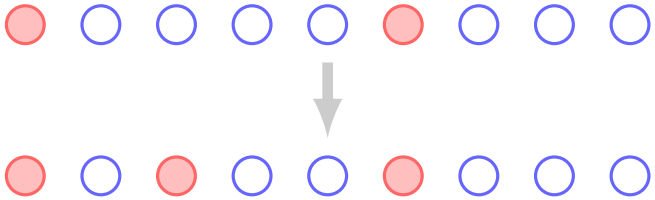
Figure 3: Place cells. This shows the firing activity of eight different cells in CA1 of a rat moving along a path, the dots correspond to spikes. [From `http://en.wikipedia.org/wiki/Place_cell`]

The best studied function of the hippocampus is spatial memory. There are cells in rat CA1 known as place cells which fire in response to specific location, an example is shown in Fig. 3. These were first discovered in 1971 [1] and show both that the hippocampus has a role in spatial memory and that the memory encoding is very sparse. Another, striking, example was discovered in 2005 in humans [2]. Recordings were taken from electrodes implanted for medical reasons in patients with focal epilepsy, these patients are required as part of an investigation into their epilepsy to spend time with the electrodes implanted; during this time many generously agree to take part in scientific investigations. They were shown many images while the activity of individual cells were monitored. It was found that some cells respond to very specific stimuli and that this stimulus can be quite abstract. In the most quoted example a cell in hippocampus in one patient was found which responded to pictures of Jennifer Aniston. It did this irrespective of how she appeared, but did not respond to other famous and non-famous faces, or curiously to pictures of Jennifer Aniston with Brad Pitt, her spouse at that time. Again, this demonstrates a role in memory, but one that involves very sparse responses.

Auto-associative memory

The standard paradigm for memory in the hippocampus is *auto-associative* memory. As we discussed in the last chapter auto-associative memories are patterns representing memories along with some dynamics that complete partial patters. 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 that performs auto-associative memory. This view is in some ways associated with David Marr, although his work preceded our modern knowledge of hippocampal anatomy. The most plausible site for this auto-associative network is CA3: this is the only area with recurrent connections between the principal cells. Often when discussing brain regions, the connectivity is classified as feed-forward or recurrent according to the connectivity of the principal cells (the most common cell type whose axons extend to other brain regions). There are almost always recurrent connections with other cells, such as inhibitory interneurons. This is true in the CA3 and CA2: In both cases the principal cells are pyramidal cells whose axons extend to other brain regions. There are also inhibitory neurons called basket cells, which are connected to the pyramidal cells. However, only in CA3 are the pyramidal cells typically connected to other pyramidal cells.

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. As before 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

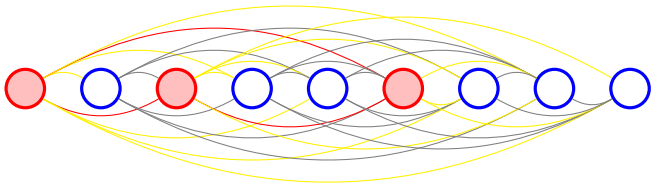


Figure 4: Learning in the associate network. The pattern has been imposed and connection strengths are changed. The red links increase by $\eta(1-a)^2$ and the gray by $\eta(-a)^2$, the yellow links decrease by $\eta a(1-a)$.

active at any one time is α, this is believed to be very small in actual neurons. For this simple model $w_{ij} = w_{ji}$.

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

$$\Delta w_{ij} = \frac{\eta}{4}(x_i + 1 - 2\alpha)(x_j + 1 - 2\alpha) \tag{1}$$

where $\eta/4$ is the learning rate, often a small number, the four is just for notational convenience, but, in hippocampus were memories need to be learned quickly, possibly during a single presentation, η is large. Since α is very small too for real networks there will be a large increase, $\tilde{\eta}$ for the connection between two neurons that are active at the same time, a tiny increase $\tilde{\eta}\alpha^2$ for pairs neurons that are inactive at the same time and a medium size decrease $-\eta\alpha$ for pairs of neurons where one is active and one inactive. See Fig. 4.

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

$$h_i = \sum w_{ij} x_j \tag{2}$$

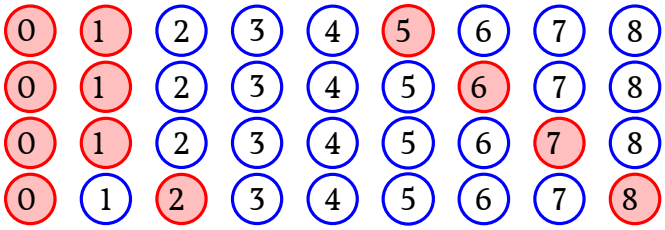
and is set in the active state if $h_i > \theta$ where θ is a threshold which is set to different values for different networks. The idea is that after learning the pattern $\{0, 2, 5\}$

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

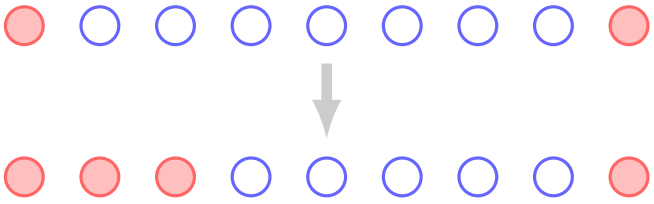
the value $h_2 = w_{12} + w_{52}$ will be larger than threshold and the subsequent dynamics will switch neuron 2 on. However, in this network, if a different initial set of neurons are activated, the activity will die away because the h_i will all be sub-threshold.

Correlated patterns

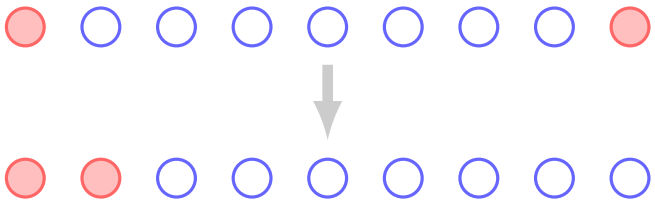
The estimates of capacity assume that the patterns are all independent. If they aren't the capacity is reduced. If patterns share some fragments or subpatterns then the connections in these subpatterns become very strong, perhaps dominating other elements in the patterns, the elements that make them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

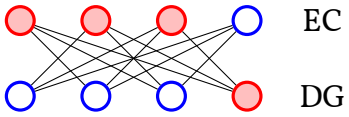


or even

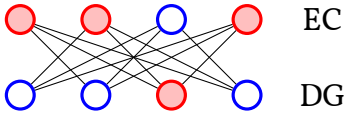


This means that auto-associative networks are not able to effectively store anything except random patterns! This is why they have never proved useful for machine learning; in cortical memory, since there are multiple presentations supported by a hippocampal representation, the memory system has the 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 randomise 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. This reduces overlap through a k-winner takes all mechanism. Roughly, k-winner-take-all assumes that local inhibition ensures that the k most active neurons in the DG layer "win" and suppress the activity of the other neurons [6]. The way this might reduce overlap is shown in this cartoon where $k = 1$ and two similar patters result in a different neuron being active:



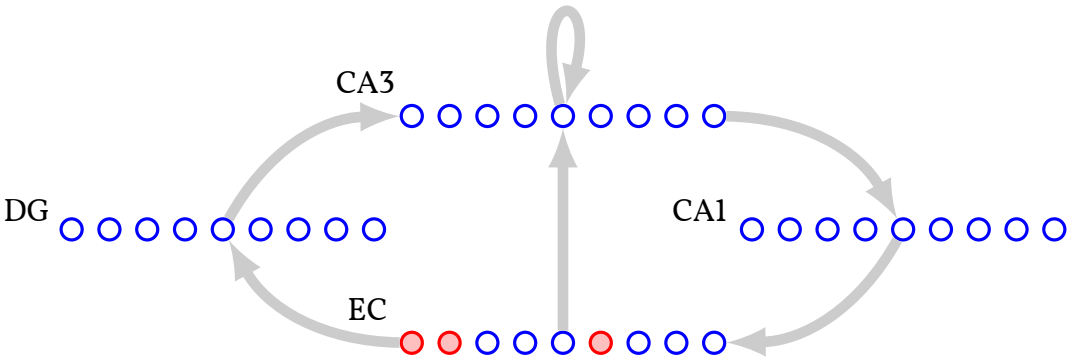
versus



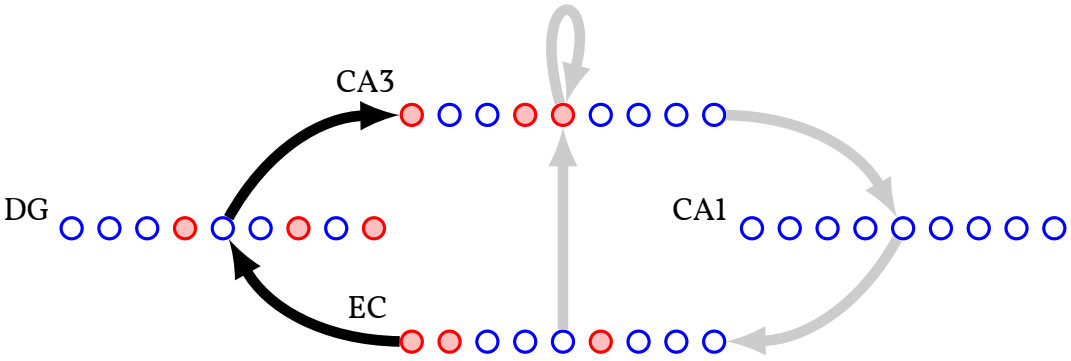
with the idea being that this randomization might be repeated in the subsequent connection between DG and CA3. This mechanism may explain why neurons in DG are being born all the time, perhaps their role is to create these random connections.

Models of the whole hippocampus

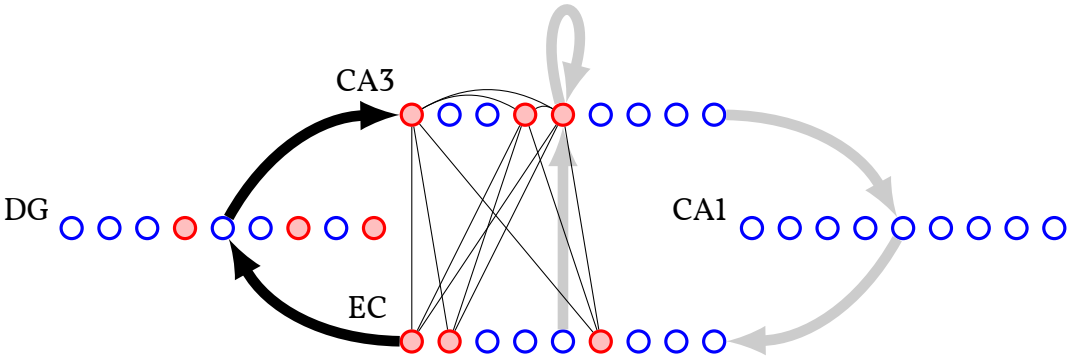
We are now almost in a position to consider models of the hippocampus as a whole; so far CA1 hasn't been mentioned. It has been suggested that the role of CA1 is to relay patterns back to EC. In learning, according to the standard model, a pattern is presented by EC:



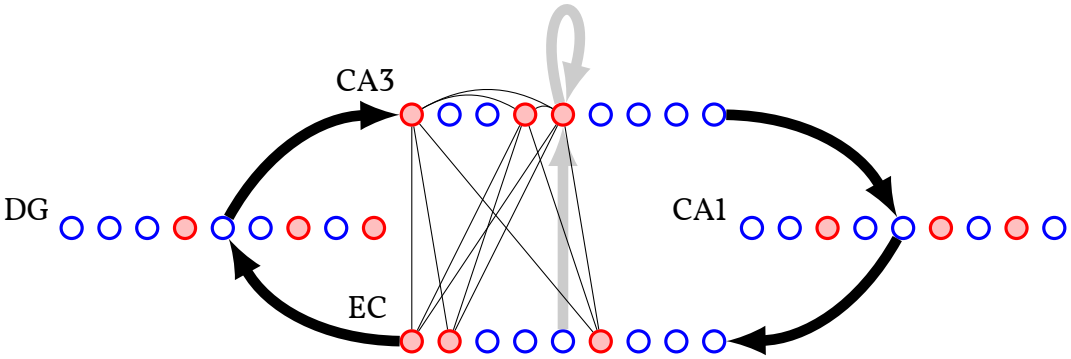
Connections, possibily random, between EC and DG and between DG and CA3, along with the 'k-winner takes all' mechanism, causes activity in CA3.



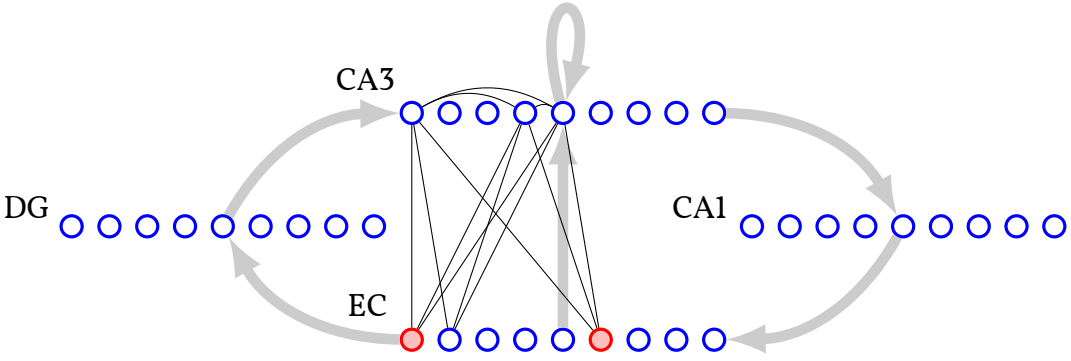
This representation is then learned by Hebbian plasticity between EC and CA3.



Hebbian learning also stregthens links to map the pattern to CA1 and link that to EC.

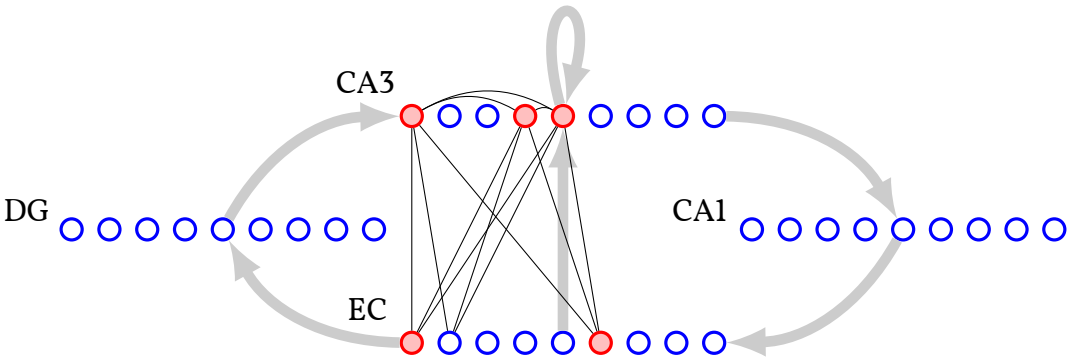


Now, during retrieval, part of the pattern is presented to EC.

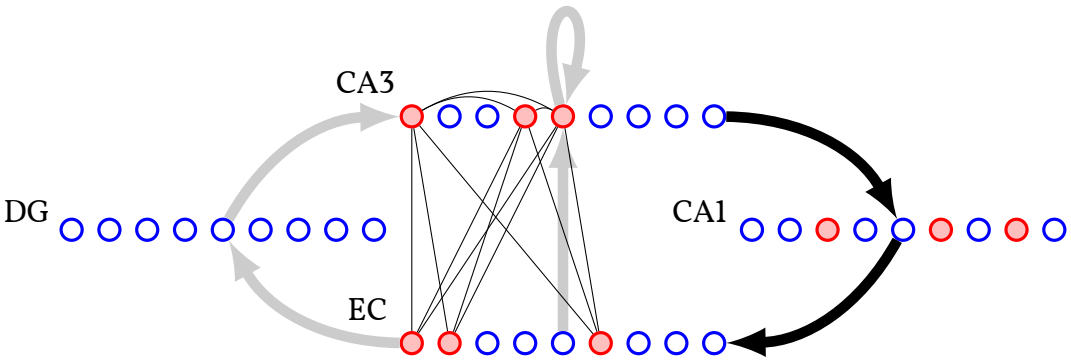


Because of the connections from EC to CA3 and the recurrent connections

in CA3, this excites the pattern in CA3:



This memory is sent back to EC via CA1, and recall has occurred!



This doesn't explain how the hippocampus switches between the learning and retrieval phase. One suggestion that the level of stimulus is different—that larger activity during learning excites the pathway that goes via DG [6], another is that the neuromodulator acetylcholine [7] or dopamine [8] is important.

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] Quian 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.