Switches in the brain?

A potential mechanism for long-term memory storage

Dilawar Singh

We forget often. But we remember some memories as long as we live. This means that our brain is capable of protecting memories for years. This is a remarkable feat given that the *biochemical hardware* involved in creating new memories is a hostile place for its storage. What are the challenges involved? And what type of biochemical mechanisms may overcome them? This article explores a major hypothesis that molecular switches may be behind our remarkable ability to remember for a lifetime.

1. Introduction

Our brain is made up of roughly 100 billion neurons, joined together with over 100 trillion connections called **synpase**s. Each neuron on average makes 1000 connections. It is now widely accepted that memories are created by changing these connections.

Let's label these synapses as $s_1, s_2, \ldots s_n$. A subset of these synapses participate in a specific memory formation. For example, my memory of being chased by a ferocious street dog named *Lalu* (lets call it M_{Lalu}) is represented by the set of the synapses $M_{\text{Lalu}} = (s_{10}, s_{21}, s_{12}, \ldots, s_{331})$ i.e. these connections were changed during my troubling encounter with Lalu. I sometimes recall this memory whenever I see a similar looking dog.

Dilawar Singh is currently a graduate student at National Center for Biological Sciences (NCBS), Bengaluru. His hobby is to convince people to move to open-source softwares to live happily ever after.

I can recall an experience as long as the set of synapses in which the particular experience was stored remains *intact*. Therefore our ability to remember is contingent on our brain's ability to keep its connections intact. But on the other hand, our ability to learn is depends on our brain's ability to change its connections. And here is the first challenge!

1.1 *Learning quickly v/s forgetting slowly, a zero-sum game*

Ability to change (plasticity) is good for learning, but bad for remembering.
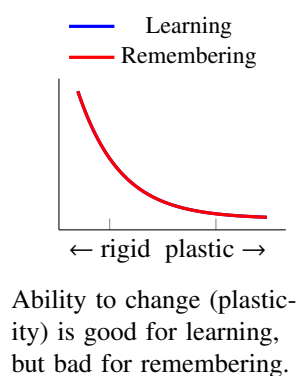
For M_{Lalu} (or any other memory) to remain intact, each of its component ($s_{10}, s_{21}, s_{12} \ldots$) should also remain intact. The longer a synapse can keep itself unchanged, the better it will be at keeping the memory. Lets assume that somehow I create a synapse which maintain its state for a very long time (i.e. a rigid synapse). This synapse will not *forget* easily, but it causes another problem. Rigid synapse will not participate in any memory formation anymore since learning requires change and it can't change. It behaves like a read-only compact disk. On the other hand, if I create a synapse which is easily changeable (i.e. a plastic synaspe), it will be good at learning new experiences but won't be able to retain it for long. Plastic synapse forgets easily. We know that not only we remember for long time, we are capable of learning quickly too. For example, honey bees can learn the location of food after one encounter with flowers. Indeed a good memory system is the one which learns quickly from new experiences and forgets old information as slowly as possible. Forgetting and remembering are the two sides of the same coin. They are conflicting demands – a zero-sum game. The challenge is to strike a balance.

2. Hopfield network – associative memory network

Memory storage and retrieval is trivially done by a computer. It will be helpful to compare memory storage in the computer and the brain. In the computer, we always know the address of every stored memory and we access it by providing this address. The file icon on your desktop is a graphical way of encoding this ad-

dressing scheme. This process is very similar to looking up the index page in a reference book to find a chapter. Our brain, on the other hand, is very unlikely to have such an indexing mechanism.

We recall when we are provided with *cues*. For example, when you see some part of of a familiar person in a wedding album – while the rest of the person may be hidden behind others – you could easily identify the person. And many other memories of that person will also be recalled. A famous class of recurrent neural network popularly known as Hopefield network can do just the same as shown in *Figure 2*.
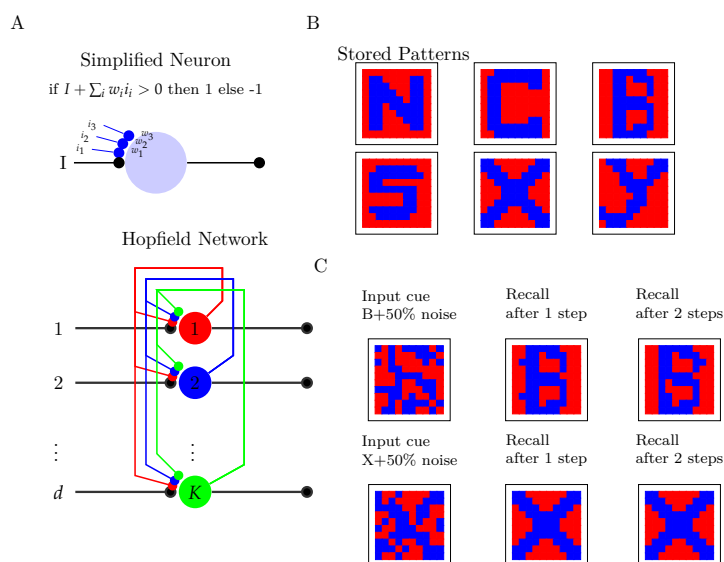

Figure 2. Hopfield network with 100 spiking neurons. These *recurrent* configurations give rise to interesting brain-like computation. **(B)** 6 patterns (memory) i.e. NCBSXY are stored in this network. **(C)** When a very distorted *cue* is applied to the network input, it *fetches* one of stored pattern which is the *closest* to the applied cue.

How does this recurrent network work is beyond the scope of article. Readers are encouraged to explore more by themselves. *"How well we can explain biological memory by these network"* is an active research area. Though these networks are extremely successful in accomplishing various *brain-like* computation (*a la* machine learning), I would like to advise the reader to be sceptical by noting the following:

- Neurons used in these networks are highly simplified. *Real* neurons are not this simple. Even though these simplified neurons

capture the essential *all-or-none* (electrical spike) way of communication and learning by changing synaptic connections, they do ignore rich local computations which can be accomplished by branches of these neurons called *dendrites*.

- There is no strong evidence that neurons make such dense recurrent connections. However some studies have shown that Hopfield network can work with very sparse recurrent connections as well.

- Activity in these networks does not match usually observed activity in the primate brain during memory-recall experiments.

Solutions contributed by other disciplines are helpful for comparison and contrast and often provide very useful insight. But in the end, these solutions must be tested under the constraints imposed by biology.

Nonetheless, these network provide us with a framework to concretely think about the problem of memory storage and its recall. We learn a great deal about these problem by pointing out the limitations and failure of these models. Hopfield network has properties which will sound very natural to us. Can you store as many memories as you like in these networks? No. There is an upper limit. Adding more patterns over maximum limit causes distortion in memories. When a cue is given, the network no longer fetch the right pattern. It often fetches a pattern which was not even stored; the retrieved pattern instead resemble some mixture of many stored patterns. When too many memories are stored, they corrupt each other by mixing up. One can ask more questions. What if connections are allowed to decay in these networks, and when memories starts disappearing, do the weakest memory disappearing first or the newest? When a new memory is added, what do happen to the old memories?

After this necessary detour, lets go back to the main theme: how do synapses maintain their state?

3. How does a synapse maintain its state?

Very complex biochemistry plays out during learning that changes the synapse. Surprisingly, the net effect of this complex biochemistry can be summarised by a simple mathematical expression.

Ah, *the unreasonable effectiveness of mathematics*[4]! Let's assume that synaptic strength *w* is tightly correlated with a chemical species *X* found at synapse i.e. *w* changes with *X*. The problem of maintenance of *w* can be rephrased as the problem of maintenance of the level – or the activity – of *X*. Therefore, the problem of "*synapse maintaining its state*" becomes the problem of "molecule *X maintaining its state*" – a more concretely defined problem.
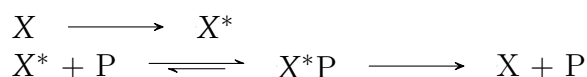
$$X \longrightarrow X^*$$
$$X^* + P \xrightleftharpoons{} X^*P \longrightarrow X + P$$

Figure 3. Phosphorylation and deposphorylation of X. P is phosphatase.

Lets assume that *X* is converted to its ACTIVE form X^* by adding a phosphoryl group (PO_4^{2-}). The phosphoryl group is removed by a phosphatase and X^* is turned back into INACTIVE *X*. The phosphorylation and its counterpart dephosphorylation are a very common motif for controlling various chemical reactions by *activating* and *inactivating* protein molecules. Once most if not all *X* has been turned into X^* during memory formation, how do we make sure that X^* does not turn back into *X* (lose memory)?

Let's mull over a solution to this problem of long term maintenance of X^*. Let's assume that somehow followings are true.

Can you think of other set of hypotheses? It must conform to laws of chemistry!

1. (**Amplification**) X^* **auto-phosphorylate** itself i.e. $X \xrightarrow{} X^*$. If we manage to get sufficient X^* somehow, it will act as a catalyst to its own production. X^* will always remain high.

2. Dephosphorylation of X^* is minimized by controlling the number of *P* or reducing the reaction rate.

Both (1) and (2) help in making X^* highly stable. Problem solved? No. Now we have constructed a rigid synapse. Recall the *rigid* v/s *pastic* synpase dilemma discussed previously (section **1.1**). This synapse will definitely remember for longer but it will not participate in any new learning anymore.

As long as we are in the realm of theory, let's propose a solution to this problem . We add another reaction say $P' + X^* \rightarrow P'X \rightarrow$

$P' + X$ which deactivates X^* when the *need* arise. Phosphatase P' is different than P. This adds another layer of control to an already complicated problem i.e. forgetting is now controlled by another process. This requires one more explanation: how does this new mechanism controlling *forgetting* works? And philosophically – if you care about it – it violates the principle of **parsimony** which recommends to pick the simplest explanation.

We still have two big problems hiding underneath. We have not considered the underlying biological hardware i.e. synapse in any detail where this biochemical network suppose to function. The first problem is chemical noise . For biochemical system operating in very small volumes, effect of chemical noise can be very strong. There are over 200 types of protein molecules in a typical synapse. Indeed, most of these protein molecules have few tens of molecules. The brain is always active and the chemical noise caused by the background activity in the brain will surely turn some molecules of X into X^*. Then due to auto-phosphorylation, sooner than later, all of the X will be turned into X^*. We have created a very stable memory of nothing but background noise. This is highly undesirable!
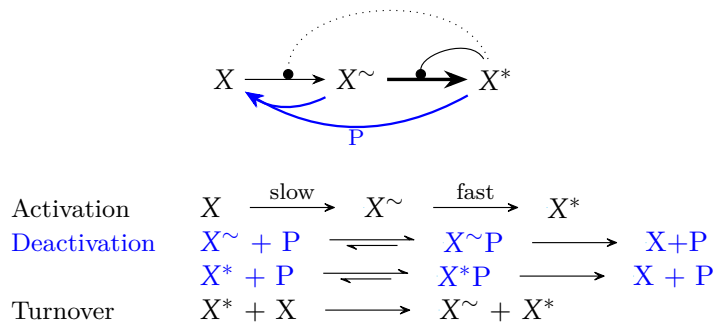
The second problem is *turnover* i.e. old molecules are constantly degraded and being replaced by newly minted molecules. Assume that at the time of memory formation, we had 100 molecules of X^* in synapse. And also assume that on average, every day one new molecule (i.e. X) replaces an old one (X^*). After 50 days, half of the synaptic strength is gone! We must have a *refresh* mechanism by which we make sure that the new molecule quickly changes its state according to the state of synapse i.e. newborn X becomes X^* most molecules at synapse are X^*.

Effectively, we want a stable `ACTIVE` state (where all X are X^*) and don't want chemical noise to turn X into X^*. We want a switch like behaviour – if it is *OFF* or ON it tends to stay *OFF* or `ON` respectively. A significant force is required to flip a switch; it does not get flipped by noise. If few X are turned into X^* by background noise, we expect them to be quickly turned back into X by phosphatase. And if during memory formation, a significant

> The volume of a typical synapse is $1 \times 10^{-20}\,\mathrm{m^3}$. At this volume, $1\,\mu\mathrm{M}$ concentration is roughly equal to 6 molecules.
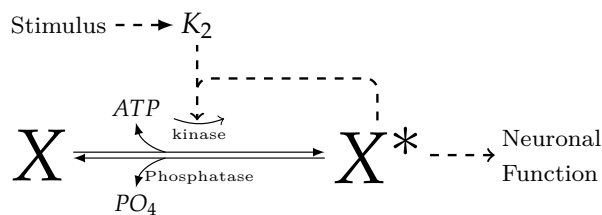
$$X \xrightarrow{\bullet} X^\sim \xrightarrow{\bullet} X^*$$

Activation	X $\xrightarrow{\text{slow}}$	$X^\sim$ $\xrightarrow{\text{fast}}$	X^*	
Deactivation	$X^\sim + P \rightleftharpoons$	$X^\sim P \longrightarrow$	$X + P$	
	$X^* + P \rightleftharpoons$	$X^* P \longrightarrow$	$X + P$	
Turnover	$X^* + X \longrightarrow$	$X^\sim + X^*$		

portion of X has been turned into X^* then we expect that any X^* turned back into X is quickly activated again into X^*. This system should operate like a switch which does not flip unless significant force is applied. These are called **bistable switch**es.

Is there any proof that bistable systems are even possible? Do they occur at all in living cells? Bistability (and its close relative oscillations) are very common in biology; from cellular level to population levels. So if it won't be surprising if we find bistable switches in syanpses as well. *Is there a set of chemical reactions which forms a bistable switch at synapse?* Various studies have shown that calcium/calmodulin dependant protein Kinase II (CaMKII) may form a bistable switch in synapse.

Figure 4. A hypothetical network which can solve the problem of chemical noise and turnover with suitable parameters (how?). The activation step is divided into slow and fast components such that fluctuations caused by background noise do not cause system to activate itself. X^* also partially activate X to $X^\sim$ to overcome *turnover*.

4. Molecular bistable switch at synapse

Figure 5. Reaction in a bistable switch proposed by John Lisman. Modified from [1]. Permission not required for reuse for educational purpose.
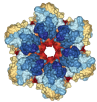
John Lisman hypothesised that a kinase and a phosphatase together (*Figure 5*) can form a bistable switch which is stable against *turnover*. CaMKII and its phosphatase protein phophatase 1 (PP1)

were identified as the hypothesised kinase and phosphatase. This chemical system had been extensively studied using computational models for over two decades [3]. There is evidence that CaMKII is bistable *in vitro* conditions.

CaMKII is known to play an important role in memory formation. In experiments involving mice, deactivating CaMKII in any way has always resulted in impairment of memory formation and learning. CaMKII molecule also has many interesting properties which makes it an attractive candidate for storing memory. 12 to 14 subunits of CaMKII make up one molecules, usually arranged in dodecameric (top view ✾) or tetradecameric form (top view ✿) . Activation of its first subunit is very slow. Once a subunit has been activated, it catalyses activation of its neighbours i.e. CaMKII auto-phosphorylates itself. Moreover fully active CaMKII holoenzyme can loose an *ACTIVE* subunit which can be picked up by another holoenzyme. This process is called subunit-exchange and this is a way by which CaMKII spreads its activation (*Figure 6*).
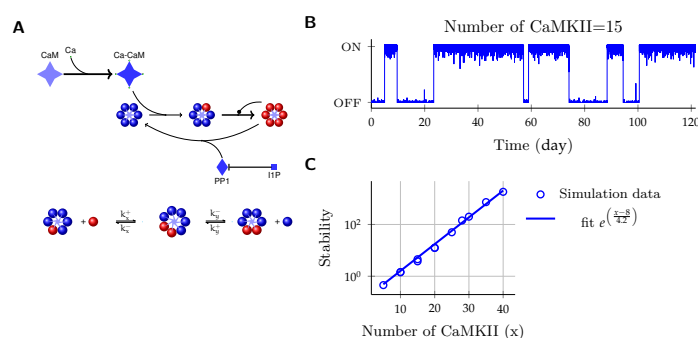
In our computational study of this pathway, we show that subunit-exchange improves information retention capacity of CaMKII. We also show that distributed clusters of CaMKII can form very stable bistable switches. And it also operate as integrator which is often observed in experiments. In short, we show the subunit-exchange makes CaMKII molecule better at retaining information and it is likely that CaMKII is bistable in special micro-environments. To prove it, one needs to observe single molecule activity in synapse near surface of synapse which is a very challenging experiment.

Top view of dodecameric CaMKII (12 subunits). From Protein Data Bank (https://www.rcbs.org).

5. Conclusion

In this article, we have discussed why bistable motif is an attractive candidate for storing biological memories. Most support for this idea has came from computational studies. To really prove it, we need experimental data supporting this hypothesis. There is growing experimental evidence that synapse change

Figure 6. **(A)** Graphical representation of CaMKII signalling pathway. **(B)** This pathway shows bistable behaviour when synapse is receiving background activity. One trajectory is shown for system with 15 molecules. **(C)** Stability of synapse increases exponentially with number of CaMKII molecules. With ~ 60 molecules, switch remain stable for roughly 150 years.

in *all-or-none* manner, a finding which is consistent with this idea. Some studies claim that changes are graded i.e. synapse changes in step-wise manner much like a **multi-stable** synapse. A multi-stable synapse is an ensemble of many bistable components. Whether CaMKII is bistable in synapse (or in some part of it) is still an open question. So far there is no concrete evidence that it is. There could be other still unknown mechanisms which can give rise to bistability. Given that bistable chemical motifs are widespread, it is reasonable to believe that there are indeed switches in our brain – much like flip-flops in your pen-drives and memory cards – which are evolved to keep our memories safe from the onslaught of time and noise.

Address for Correspondence
Bhalla Lab, National Center for Biological Sciences, Bengaluru GKVK Campus, Bellary Road Bengaluru - 560065.
Email: dilawars@ncbs.res.in

Suggested Reading

[1] **Lisman J. E.,** *A mechanism for memory storage insensitive to molecular turnover: a bistable autophosphorylating kinase*. **Proc. Natl. Acad. Sci. USA, May 1985**

[2] **Christof Koch** *Biophysics of computations*. **Oxford University Press, 1999.**

[3] **Malin Sandstrom,** *Models of CaMKII activation*, **Master Thesis, Royal Institute Of Technology Sweden**

[4] **Eugene Wigner,** *The Unreasonable Effectiveness of Mathematics in the Natural Sciences*, **Communications in Pure and Applied Mathematics, vol. 13, No. I (February 1960)**