# Ornithopter or Concorde? A review of the understanding and imitation of biological neural networks in theoretical and computational approaches

Kenneth Schlax
PHYS 569

Submitted to:
Prof. Nigel Goldenfeld

December 19, 2012

**Abstract**

The emergence of various properties in neural networks (memory, robustness, computation, and processing) is approached through theory and computational methods. Neural networks are found in nature and created artificially in computer systems. Connections can be made between emergence in models of interacting and dynamical systems and interactions between neurons in the neural networks. Computer methods often attempt to replicate the emergent properties found in biological neural networks. This paper will examine and evaluate the study of biological and computational neural networks in the area of emergent behaviors.

# 1 Introduction

Before highlighting some important historical aspects of the development of the study of neural networks (indeed, this paper will be virtually entirely such a review), I would like to provide some motivation for the topic at hand. I wonder at the immense computing power and diversity of the human brain, which is a biologically occuring neural network. Even though, as is the case for the computer I am using right now, current technology has created computer chips capable of processing data with enormous speed and at ever-smaller scales, far faster than humans can calculate or comprehend, there are certain behaviors such as abstract thought, or pattern recognition, or playing baseball, which scientists struggle to replicate or improve upon. What amazes me most, however, is the natural origin of the brain itself - an information processing machine arisen from the same physical laws that govern chemical reactions and general relativity (at least, if there is some grand unified theory). The fact that an object made of carbon and oxygen, sitting in and directing a self-propogating, self-temperature-regulating, self-sustaining, and self-aware (to some extent)[1] environment (by which I mean the body), has persisted and evolved over millions of years is the real puzzle, but one that has so many aspects, not even one field of science can encompass them all, let alone a single paper.

# 2 Physical Neural Network Models, Act I

Instead, in this paper I propose to focus on simply the consequences of the architecture of the brain. It seems natural to assume that the form of the brain follows its function, or, if we assume that there are some physical laws governing the brain as well, then the brain is required to. So then we must ask what is the structure of the brain, or at the very least, how can we describe it? This question is approached by scientists in medicine, biology, physics, chemistry, psychology, and neuroscience, and they bring with them a further plethora of methods, both theoretical and experimental, to map, model, and understand the brain. This paper is restricted to physical models of what are called neural networks. (See section 4 for a definition of neural network, generally.)

I drew a parallel in the title between the development of the study of biological neural networks and the development of aerodynamics. The similarity is that in each field, there was interest in replication (birds in flight; brain functions), with differing approaches in how to achieve them. There are two ways in which this is studied[1]:

[1] But enough of this self-promoting.

The first are top-down approaches, which, like the qualitative ornithopter, focus on brain behaviors, and work out how neurons can exhibit these behaviors. These analyses I will call 'black-box' experiments, where input and output are known, and the goal is to match the input to the output with some model. The second are bottom-up approaches, like the quantitative Concorde, which focus on individual neurons, and how their interactions create these cognitive functions. This approach relies on emergent behavior, and it is striking to see how brain functions such as memory (explored first), and computation and processing (touched on later) emerge from even very simple models of neural networks.

# 3 The Brain, or a Biological Neural Network

The "Brain", as we refer to it here, is little more than a set of interacting immobile point particles. In reality, the brain is a natural product of evolution, by which I mean the brain itself is an emergent dynamical state of, at some level, the expression of DNA under some chemical impetus or environment. As we have noted, we will not delve too deeply into the realities of the brain, but it may help in our understanding of why neural networks are modeled in particular fashions, to look at the basics of brain function. The brain consists of neurons, which, for our purposes, are identical. Neurons are cells which can relay an electrical signal via an extension called an axon to a synapse, which is a connection to another neuron. The neuron is capable of transmitting a signal if an energy threshold, called the action potential, is reached. Each neuron will "test" the signal against this threshold, and transmit the signal if it is greater. As we will note in the next section, this structure is far more interesting when there can be nonlinear paths, i.e. loops. Their interesting interaction comes in the form of time-evolution, and path creation, not in the motion of the neurons themselves.

In discussing the brain and its functions, it may be helpful to analyze some explanations and experiments relating to brain activities from a behavioral viewpoint; that is, a phenomenological view.

In 1969 David Marr sought to create a series of testable hypotheses about the brain, and a theory which would describe the appropriate responses to inputs to the Cerebellum. [2]. Marr's theory relied on a changeable nature of the neurons between excitatory and inhibitory. This idea is called *synaptic plasticity*, and many subsequent papers have been written on the topic (for a current review of Marr's paper, see [3]). Synaptic plasticity indicates that experience changes the behavior of neurons over time. Importantly, this leads to patterns, or paths, preferentially taken over others. The robustness of these paths, enabled by synaptic plasticity, is

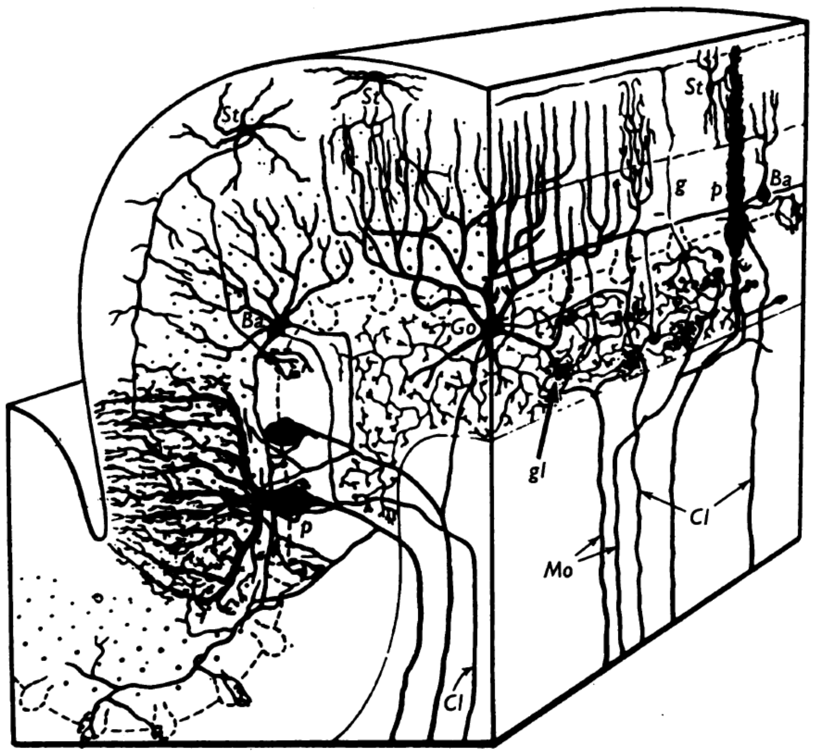


Figure 1: A diagram of the cerebellum, as studied by Marr. [2]. In Marr's simulations, the distribution of the mossy fibers (**Mo**) are randomly distributed among the granule cells **g**.

indicative of learning in the cerebellum, namely Hebbian learning. Although Marr's research was aimed at a neuroscience audience, his theory was based not on any particular brain structure but rather random distributions of neuron types and parts found in the cerebellum (see Figure 1). As such, it was not just an attempt to replicate the brain but an attempt to replicate the function of the brain, generally. The fact that Marr could hypothesize about the function of the cerebellum without giving it a particular structure[2] is particularly relevant to the successive bottom-up, emergent approaches we will examine in Section 5. As Strata notes in his review, this theory is still very relevant and discussed today in physiological circles, and is relevant too to our present goal.

In the 1970s and 80s, Bienenstock, Cooper (of BCS theory fame) and Munro

[2]However, Marr does note "It is regrettable that no data exist to give a better model."

spent some time theorizing about the development of neurological phenomena. As an example, we will look at their work on a particular aspect called stimulus selectivity. [4]. Stimulus selectivity is a preferential response to some aspect of an input, for example a particular color, shape, sound, etc. Cooper's paper is significant because it generalizes (like Marr) to any type of selectivity, and because the theory proposed relates the behavior of individual neurons to the environment of the neural network as a whole, indicating the presence of a long-range order. Another important result was the tendency of environments to tend towards stable state(s) from a wide range of initial conditions. This is important for cognitive features such as memory recall, if one is trying to recall a stable state, even a partially correct guess (i.e. a state near the stable state) will produce the correct memory (i.e. the stable state). This highlights a feature we will see again in the work of Hopfield and Amit (see Section 5). An updated version of this selectivity experiment was performed by Afek, et al., on the patterns of fly bristles, who compared schotastic models with actual fly bristle distributions. [5].

# 4 A Computer, or an Artificial Neural Network

Artificial neural networks are one type of computation. (see Figure 2). [6]. They can be theoretical, just as the toy models of the brain, or they can be realized in computer systems. They can also be simulated on computers, just as models of the brain can be simulated. Most computers today are serial in processing, or, in today's multicore processors, multiply parallel. The hallmark of artificial neural networks is massively parallel, asynchronous processing. [1]. The result are calculations that may take many nonlinear paths. We will continue from this point in Section 5, but first an aside as to the historical origins of artificial neural networks.

In 1943, McCulloch and Pitts wrote a seminal work [7] laying out ten theorems describing the necessary structure of the brain, in terms of neural nets and circles, which are cycles of those nets. The structure was based off of the idea, as laid out in Section 2, that the components must transmit or not transmit (a binary action), based on some comparison to an external evaluator (the action potential). Importantly, they assumed that the structure of the network did not change with time. Contrast this with the Hebbian learning and synaptic plasticity of the previous section. These cycles allow for temporal nonlinearity, i.e. access to information about some aspect of the state from the past. Already we see some connect to a function of the brain, that is, memory. McCulloch and Pitts were careful to note that their model was limited, in that it could not explain sleep, for example (though certainly since then, more research on sleep has been done, but that is for another paper).

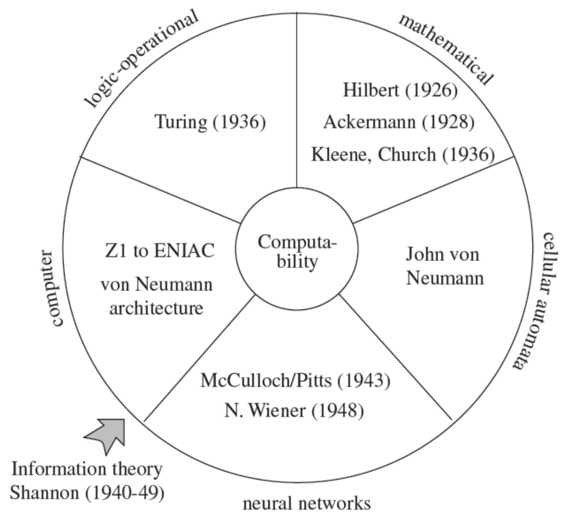


Figure 2: (Artificial) neural networks, as a type of computation [6].

The important result of McCulloch and Pitts was the fact that a system of simple components following prescribed rules can produce behaviors that were typically associated with complex systems.

# 5 Physical Neural Network Models, Act II

Continuing from the first paragraph of Section 4, here it is useful to mention the concept of chaotic attractors. If the equations that govern the neural network are nonlinear, then one must resort to a stochastic or averaged view of the process, to determine its behavior. If there exist some stable solutions to the system, then thermodynamics tells us that in the limit of infinite time the system will tend towards those stable states, i.e. those with minimum energy. These are the chaotic attractors. Taking an averaged view means that we must express the local interactions as a function of the overall system, and, as we noted, this indicates some long range order. We will see, in the following models, evidence for symmetry breaking and emergent phenomena as well.

We have looked at, first, a pair of black-box investigations into the working of the cerebellum [2] and the sensory cortex [4]. We then noted the previous work into a general mathematical theory of and idealized neural network by McCulloch and Pitts. [7] Hopfield, in 1982 [8], created something like an Ising model of a neural network. He proposed that each neuron was in a state of "firing" or "not firing", hence something like a spin-up or spin-down state. We can then connect the neuron network in some pair-wise fashion, with the relative strength of the pairs given by some factor, and a self interaction of zero (hence, there is no point storage, only storage in the connections). To introduce a time element, the neurons have a mean firing rate, so we have some sort of mean time field theory. If we want to store a set of states $V^s = V^{(1\cdots n)}$, the interaction can be expressed as $T_{ij} = \sum_s (2V_i^s - 1)(2V_j^s - 1)$ so that we have a Hamiltonian of the form

$$H_j^s = \sum_j T_{ij} V_j^s = \sum_{s'} (2V_i^{s'} - 1) \sum_j V_j^s (2V_j^{s'} - 1)$$

We want to find the stable solutions to this Hamiltonian. So, we will modify $T_{ij}$ by an average correlation $[V_i(t)V_j(t)]_{avg}$, to try to smooth the asynchronity and find a stable (at least, repetitive system). In the special case $T_{ij} = T_{ji}$, then the change in energy due to a particular change in state,$\Delta V_i$, is monotonically decreasing to some minimum energy by

$$\Delta E = -\Delta_i \sum_{j \neq i} T_{ij} V_j$$

In his work, Hopfield ran Monte Carlo simulations to test the memory recall of the system, or how likely it is to return a nearby stable state given a perturbed input. The variation he chose was the length of the state vector, $s$, so that a system with a higher $s$ would have more "memories", or stable states, to "remember". These results are presented in Figure 3. The Monte Carlo simulation shows that as the state vector length increases, the more errors are returned. This is indicative of a very commonly experienced occurrence, forgetfulness. Thus, the Hopfield model has gone beyond storing stable states, but has indicated the emergence of another cognitive (mis-)function, all without saying anything about brain structure, and merely making some assumptions about the interactions of the neurons.

In 1985, Amit, Gutfreund, and Sompolinsky expanded on the Hopfield Ising-type model, by first comparing it to a model by Little, and then by noting that the models were identical for temperatures below a critical temperature, and that at a temperature $T = T_c$, symmetry breaking would occur. [9]. This model is called the Ising-spin glass model, and the Hamiltonian is:

$$H = \frac{1}{2} \sum_{i,j,i \neq j} \left(\frac{1}{N} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu\right) S_i S_j$$

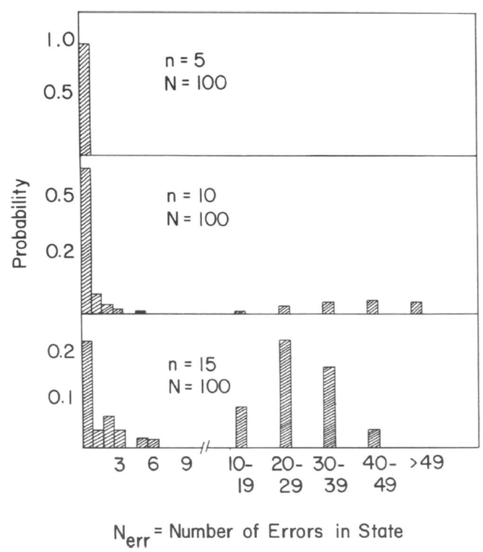


Figure 3: Monte Carlo simulations of the Hopfield model. N is the number of neurons, and n is the length of the state vector. It can be seen that as n increases, the average error increases.

where the $\xi$'s are independent random variables with zero mean, for example, randomly equal to +/-1. Amit et al. go on to derive a free energy, partition function, and order parameter, $m$, in terms of the quenched random variables $\xi$.

$$m = 1/N \sum_i \xi_i \tanh(\beta m \cdot \xi_i)$$

or, in the thermodynamic limit(i.e. mean field limit), $N \to \infty$,

$$m = \{\xi_i \tanh(\beta m \cdot \xi_i)\}$$

This is reminiscent of the form of the net magnetism (order parameter for a ferromagnet). Physically, $m$ is interpreted as the overlap of the local magnetization with the $\xi$ distribution. Expanding this and the free energy about $T_c = 1$, it can be seen that the system exhibits the paramagnetic state ($m = 0$) above $T_c$, and multiple nonzero $m$ states below $T_c$, and thus the spin-glass model exhibits symmetry breaking.

# 6 Imitation, Understanding, and Conclusion

In our previous examples of modeling neural networks as a thermodynamic system, we restricted ourselves to fairly simple models, and sought understanding more than perfect replication. It is through this aspect that the study of neural networks in physics has evolved away from the 'ornithopter' approach, and towards the 'Concorde' approach, leading from the purely top-down approaches of Marr and Bienenstock and towards the bottom-up approaches of Hopfield and Amit. Through such bottom-up approaches, neural networks have exhibited an order parameter, long range order, and symmetry breaking, and crucially to the study of biological neural networks, emergent phenomena consistent with that exhibited by brains.

I conclude by looking at a recent development in neural networks.

A large scale neuron network simulation by Eliasmith et al. [10] is, procedurally, very much a 'black-box' type experiment. In the example shown in the video, [11], handwritten inputs simulate the perturbations away from the typographical memory states that exist in the network. The system then has some time evolution, as it waits and/or processes other inputs. The error rate of the output of the system then, in a simple sense, reflects the same 'forgetfulness' we saw in the Hopfield model. However, this brain model is far more complicated (see Figure 4). Note in the brain model map the level of reciprocation and feedback. Hopfield noted in his paper that prior work had less success and that "all of his interesting results arise as consequences of the strong back-coupling". We can see a similar process here, despite its greater complexity. The complexity displayed here, though, seems to be more top-down oriented, so that the sum of the pieces is designed to look more like a brain (as Marr desired earlier, in the footnote). This model has far more available in terms of computing power and neuroscience when compared with the previous cases examined. The result of this complicated model is the ability to replicate many more of the brain's functions beyond mere memory, as shown in the videos of Figure 5.

# References


[1] W. Gerstner, H. Sprekeler, and G. Deco. Theory and Simulation in Neuroscience, Science vol. 338, pp. 60-65, (2012).

[2] David Marr, *A theory of cerebellar cortex*. J.Physiol. (1969), 202, pp.437-470.

[3] Piergiorgio Strata. *David Marr's theory of cerebellar learning: 40 years later*. December 1, 2009 The Journal of Physiology, 587, 5519-5520.

[4] E. Bienenstock, L. Cooper, and P. Munro. Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex, Journal of Neuroscience, Vol. 2, No. 1, pp. 32-48, January 1982.

[5] Y. Afek, et al. A Biological Solution to a Fundamental Distributed Computing Problem, Science, Vol. 331, no. 6014, pp. 183-185, January 2011.

[6] Raul Rojas. "Neural Networks: A Systematic Introduction". http://page.mi.fu-berlin.de/rojas/neural/neuron.pdf. Springer, Berlin, 1996.

[7] Warren S. McCulloch, and Walter Pitts. *A logical calculus of the ideas immanent in nervous activity*. The Bulletin of Mathematical Biophysics 5:115-133. 1943.

[8] J.J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities", Proc. Natl. Acad. Sci. USA, Vol. 79, April 1982.

[9] D. Amit, H Gutfreund, and H. Sompolinsky. Spin-glass models of neural networks. Phys. Rev. A, Vol. 32, No. 2, pp. 1007-1018, August 1985.

[10] Chris Eliasmith, et al. *A large-scale model of the functioning brain*. DOI: 10.1126/science.1225266, 1202 (2012); 338. Science.

[11] Chris Eliasmith, et al. "Videos for Spaun simulation". http://nengo.ca/build-a-brain/spaunvideos. Accessed Decemeber, 2012.

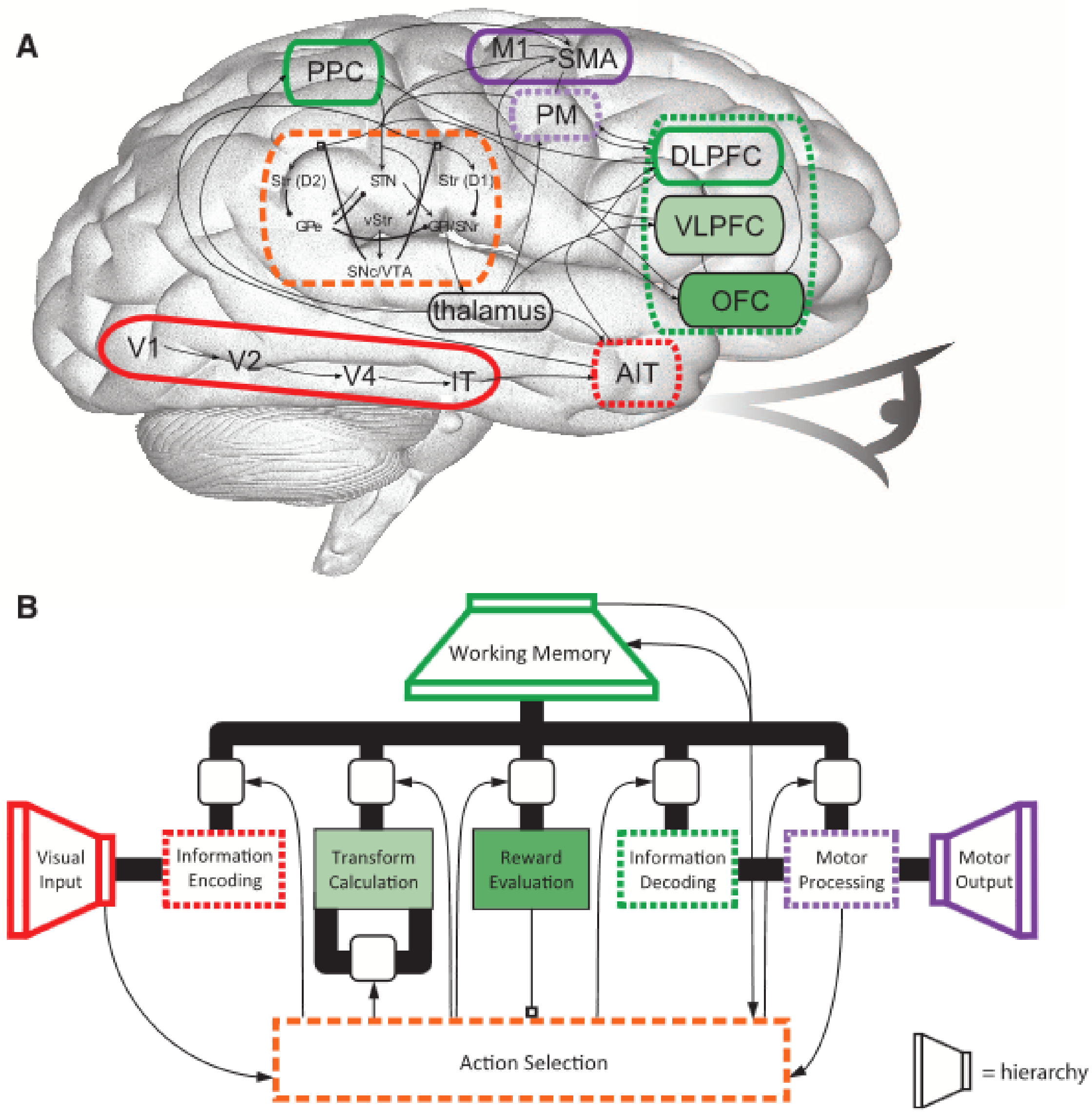


Figure 4: A diagram of the model brain of Eliasmith, et al. [10].

Click here.

And/or click here.

Figure 5: Movies providing a brief view of some capabilities of the Spaun artificial brain project. Click once to load and start the video, double click to stop. (If video does not load, go to: [11]). Note there is sound.