Computational Neuroscience: Bibliographic Annotations and Thoughts on Selected Texts

Charles Burton (cpburton18@earlham.edu)
Earlham College

May 2, 2022

References

[1] G. L. et al., "Neural coding of naturalistic motion stimuli," in *Network: Computation in Neural Systems*, January 2001.

In this paper, Lewen et al. seek to determine how noisy neural processing is. Specifically, they follow two experimental procedures to stimulate the H1 neuron in a blowfly and, through information theoretical analysis, determine that the imprecision of input data leads to imprecision in neural signals, contrary to the idea that entropic noise is generated within the nervous system itself.

The scientists in this experiment set up experimental procedures as follows: a fly is made stationary via embedding it in wax, and an electrode is placed proximal to the H1 neuron, a wide-field neuron that responds to horizontal motion. The fly-electrode apparatus is then mounted on a stepper motor, where horizontal motion can be sufficiently simulated. By rotating the fly at variable velocity while exposing it to both natural (lab) and artificial (outdoors) visual stimuli, different firing patterns of the H1 neuron were recorded and analyzed.

For data analysis, the researchers discretized the data by separating it into bins (binning strategy not mentioned), where the spikes in each bin formed "words". These words were used to form probability disributions that played a central role in analysis. From these word distributions came the Shannon entropy, the noise entropy, the information conveyed by words, and the coding efficiency of the spike train. What they found was that for low rotational velocity, the spike rates were similar between conditions (lab vs. nature). As the rotational velocities increased, however, the natural condition's motion response peaks at around a full order of magnitude greater (1000 degrees/second) than the lab condition's response (100 degrees/second). Additionally, they found from analyzing motion response at different times in the day and therefore different light intensities, the neuron's ”vocabulary”, or information content, increases in size as the input becomes better defined, which generally occurred with greater light intensity.

[2] A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," in *Journal of Physiology, Volume 117*, March 1952.

Possibly one of the most important papers in modern neuroscience (as seen by almost 30,000 citations on Google Scholar), Hodgkin and Huxley use the experimental data they collected from voltage clamp techniques on giant squid axons to attempt to build a quantitative framework for neuronal communication and action potentials. Because theoretical neuroscience was such an infant field in 1952, the beginning of this article postulated ideas for the biological mechanisms of what we now know as an electrochemical process in which a resting membrane potential is built up by transmembrane protein transporters who, with the help of ATP and ionic concentration gradients, actively build a potential of around -75 mV between the inside and outside of a membrane. Some of their ideas included ionic coupling of sodium with some negative sister molecule which served to alter the membrane potential when the time arose for signal propagation. After this, they dove into mathmatical models, seeking to represent data collected quantitatively with systems of equations that best governed communicative processes between neurons. First, they build up rate constants whose form follows that of a Boltzmann distribution function, or $Ce^{-E/kT}$. In place of ”E”, or energy, they input voltage. One of the conclusions of the models they build up for conductance and rate equations is that the rate that the membrane potential reaction progresses is dependent of the concentration and current variables of ions involved in reactions (primarily sodium and potassium ions) as well as the time variable T. They test this out experimentally, and find out that sure enough, temperature differences induce phase changes of potential curves. Overall, this paper is of paramount importance in theoretical neuroscience. Hodgkin and Huxley used systems of equations to build up models of action potentials and (mostly) sucessfully stacking these models up to experimental findings, where they held firm. In doing so, these neuroscientists paved the way for exploration of both microbiological theory of action potentials and a fundamental understanding of how nerve cells communicate.

[3] Hopfield and Tank, "Computing with neural circuits: A model," in *Science Magazine*, 1986.

This paper was written by Hopfield and Tank as an elaboration of the computational neural models suggested by Donald Hebb in his influential 1949 book titled "The Organization of Behavior". In it, they explore the existing Hebbian model of associative learning, with its binary states and an early model of synaptic plasticity based on the modulation of synaptic "weights" that takes place and is learned from feedback mechanisms. Hopfield and Tank use the idea of learned patterns and associative learning via Hebb's Law and apply continuous mathematics to it in order to create a more useful and robust model of a neural system that can be actualized into circuitry.

The paper begins with a discussion of the shortfalls of the model proposed by McCulloch and Pitts, who described neural bodies as two-state variable taking on value of "on" or "off" depending on preceding signals. To build more robust models, nonlinear dynamics and continuous variables had been integrated into the algorithms of neural networks in order to gain more biological semblance. Even so, issues arose in that despite the ability to predict results in complex systems, the underlying processes that happened in order to yield those results were yet unknown. Here, Hopfield and Tank attempt to use an understanding of optimization and analysis of dynamics to get a grasp of the process occurring in a nonlinear model of a neural network.

For neat analysis, simplifications must be made and they tried to make things more clean by omitting any complexity that might arise due to the shape of neurons in the network. Additionally, they dealt with only primary synaptic signaling, meaning that only initial signal propagation and electric conductance was built into the model. As a result of this, some change in potential would generate an instantaneous conductance change in the corresponding post-synaptic cell. Next, this is incorporated into an algebraic model that equates the input capacitance of some neuron to some potential as the sum of the postsynaptic currents, leakage current, and input currents from other neurons. The resulting firing rate of this model of a neuron can be graphed as a function of input potential, and the graph follows as a monotonic sigmoid curve of varying steepness. The sigmoid curve ends up allowing this model of a neuron to realize optimization utility and grants it a high degree of computational power. If I were more knowledgeable about control theory and optimization (mostly by being better at linear algebra, which is on my to-do list), I think that I would follow the processes that allow for this phenomenon to give rise, but I have to trust Hopfield and Tank for the time being.

The real kicker comes in when energy functions of a neural system made up by our model neurons are considered. The emergence of the energy function is independent of individual neurons and is seen at the level of the neural network. Hopfield and Tank draw lines of comparison to the entropy of a gas as it pertains to individual atoms versus to the gas as a whole. Next, they map out an associative memory network similar to Hebb's where natural networks adopt their own kind of associative symmetry. They posit that if desired "memories" can be grouped into stable states of a network, then computational process can be seen to move towards these desired states. In terms of the computational energy function, as a stable state is reached, the energy function is minimized. This is textbook optimization, and it follows that in an associative neural network, the goal is to minimize the energy function in the most efficient manner possible.

The paper goes into detail with examples of the practical usefulness of this computational model, such as with the traveling salesman problem (a classical combinatorial explosion problem), but the most important point of the paper is that by consideration of thee energy function and the nature of an associative memory model of neurons, the process by which said model goes through with computations is able to be realized. Today, we call this the Hopfield model of associative memory, and it is used in artificial intelligence models to parse through and find solutions to data of a large magnitude.

[4] Malsburg and Schneider, “A neural cocktail-party processor,” in *Biological Cybernetics*, June 1986.

Malsberg and Schneider investigate sensory segmentation, or the neurological clarification of discrete stimuli, with respect to synaptic modulation. Prior to this paper, factoring synaptic modulation into the sensory segmentation process had not been done, and their work is recognized as a novel application of correlation theory in neuroscience.

Prior theories cited peripheral and central evidence for segmentation as a kind of pattern-pass system. These ideas are said to be inapplicable in true neural context by the authors. Instead, they propose that through the use of selective attention mechanisms. Using the example of a cocktail party, where a listener is able to concentrate efforts on and successfully follow the words of a single speaker despite a high degree of distracting stimuli (and hence the name of the paper), the authors cite auditory segmentation via cochlear nerve processing. Essentially, the cochlear nerve parses and segments nonuniform signals and propagates that of greatest information. As a result, the listener at the cocktail party is able to follow a dialogue and not be overwhelmed by an auditory cacophony.

To test this ”Cocktail Party Processing” theory, Malsburg and Schneider got to work by designing a model network composed of multiple excitatory input cells, all of whom are connected to an inhibitory ”H” cell. Without getting into too much detail, the scientists used the model in simulations to determine if the model would be able to succeed in a kind

of sensory segmentation upon receiving afferent signals over time. It turns out that synaptic modulation plays a large role in sensory segmentation, and by recognizing input patterns from different sources over time, data can be process and relevant auditory stimuli is able to be propagated through to the auditory cortex while irrelevent data is disregarded as noise.

[5] W. McCulloch and W. Pitts, “A logical calculus of the ideas immanent in nervous activity.” in *The bulletin of mathematical biophysics 5.4*, 1943.

Regarded as a seminal paper in computational neuroscience as well as the basis of modern machine learning, the importance of the early formulations of McCulloch and Pitts are hard to overstate. By treating the biological existence of neurons as computational devices, these scientists applied logical theory to biological systems, most notably coming up with the idea that the communicative signaling of neurons was "all-or-nothing" and could, in theory, be quantified by algorithms that utilized AND, OR, and NOT gates. Excluding the mathematical portion of the model, the dendritic equivalent of this modeled neuron receives one or more input signals. Next, the collection of signals are processed according to weight (excitatory weight and/or inhibitory weight). After this, a binary output is determined, either 0 or 1 with zero representing no firing of the "action potential" and 1 representing firing of the "action potential". Though simplicstic in design, this model of a single neuron was profound in that it allowed for a quantitative basis of the neuron, demystifying the most integral component of the nervous system.

[6] Rao and Ballard, “Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects,” in *Nature Neuroscience*, 1999.

Rao and Ballard present evidence that in visual processing, there exists "non-classical" receptive field effects. The authors hypothesized that visual processing takes into account statistical normalities of daily life. This means that redundant or routine processes or input signals are the kind of baseline as recognized by visual processing. When deviations from this

baseline pop up in the form of abnormal stimuli, signals are sent to higher processing neural regions.

In order to get data, Rao and Ballard first build a model. Their network is a hierarchical predictive coding model, in which feedforward and feedback data are cycling through and predictive power is constantly being compared. Within this model are predictive estimators, which do this comparison. They receive afferent feedforward data error signals and efferent feedback predictions. This model is supposed to simulate a visual processing neural network, with predictive processing adjusting and changing based on error signals.

When the scientists ran their model through image recognition simulations, they found that there did exist non-classical receptive field effects. When peripheral image input matched receptive field image input, little error signaling was evident because the surrounding field was able to predict the center. However, when the image was presented in isolation without a kind of visual context, higher level processing was activated. This is the non-classical receptive field effect, and this finding has been backed by previous literature that has observed the same kinds of visual processes in monkeys. The next steps forward are to extrapolate the physiological structures that identify with each component of the model, or at least relate most closely.

[7] C. Shannon, "A mathematical theory of communication," in *Bell System Technical Journal*, 1948.

Claude Shannon wrote this groundbreaking article to explore the stochastic and statistical dependencies communication. He based the mathematical framework on Markov chains, which are extensions of Bernoulli's statistical formulations to dependent events through a kind of mathematical "memory". Markov postulated that dependent events have a memory of prior event(s) which affects the outcome of the current event. As a result, dependent events also converge to some distribution or function given sufficient data/input. This idea and Shannon's curiosity about human communication set the tone for the creation of his theory.

Shannon first builds up the language needed to tackle english communication and introduce some random sequence of letters (A,B,C) the we will seek to replicate using Markov chain principles. Next, the first model is build up. The first-order condition-dependent algorithm tries to emulate some sequence of letters by essentially weighting probabilities of each component letter so that the frequency of letters in the final sequence will be approximately the same between model and original sequence. To make things better, we can take into account which *pair* of letters are randomly selected and create a condition-dependent algorithm that selects a pair of letters, and the next letter picked will be dependent on the last letter of the previous pair. Now, we have gone from similar frequencies to similar frequencies *and* more similar looking sequences built up between model and original. For simple purposes, this is very good at creating similar-looking sequences and can be expanded upon in a third-order approximation which takes into account groups of three letters. This can continue on as many orders of approximation as computation allows, but at some point the complexity of the computation exceeds the scope of what we are trying to accomplish, so we are satisfied with some order less than too much.

This principle of creating condition-dependent algorithms for sequences of letters can be scaled up to sequences of words, which Shannon does and shows that comprehensible sentences emerge as our model increases in the order of complexity. This is all interesting, but what is most relevant to computational neuroscience is "information entropy" which was defined by Shannon in this same article. Entropy in this context is the expected value of the self-information of a variable. Just like in thermodynamics, the entropy tends to increase as the data grows and in information theory, this means that the more data we have, the more we can know about the dependencies of some events on others.

Learning this bit of information theory has given me access to the understanding of some mathematical models used to emulate sequences and patterns. As put by Timme and Lapish in their 2018 article on information theory in neuroscience: "...data from neuroscience experiments are multivariate, the interactions between the variables are nonlinear, and

the landscape of hypothesized or possible interactions between variables is extremely broad. Information theory is well suited to address these types of data, as it possesses multivariate analysis tools, it can be applied to many different types of data, it can capture nonlinear interactions, and it does not require assumptions about the structure of the underlying data..." Therefore, learning from information theory is important in figuring out the best way to process and think about neuronal data.

[8] Todorov and Jordan, "Optimal feedback control as a theory of motor coordination," in *Nature Neuroscience*, October 2002.

One of the most important practical functions of the human nervous system is coordination. Ensuring that the body is energetically and positionally optimized while completing a task allows for efficiency. On top of this, performing this kind of optimization reliably and in a reproducible manner is ideal with respect to behavior and automaticity of action. In their paper, Todorov and Jordan confront the question of how humans, as biomechanical systems, are able to achieve behavioral goals consistently, but with iterative attempts containing movements that are rarely reproducible due to a redundancy of degrees of freedom. To answer this question, they come up with a model of human biomechanics based on stochastic optimal feedback control.

Now, to get any further in this paper with any semblance of a comprehensive understanding, I had to do some background research on stochastic optimal control, a subdomain of control theory that utilizes differential equations and Wiener processes (Brownian motion). A shallow reading through of literature told me that in essence, stochastic optimal feedback uses differential equations to glean value and/or statistical relevance for stochastic generated data. In the context of Todorov and Jordan, this means that they apply these kinds of mathematics to come up with control laws that govern processing that governs human biomechanics. The most interesting part of this paper for me was the point at which they theorize that in the case of delayed and noisy sensors (which do exist in a human biomechanic system), the calculation of optimal con-

trol signals requires integrating feedback with a knowledge of plant dynamics (the mental processing of some task and the selective output techniques in order to address the task) and a record of past plant dynamics. I thought that this was quite interesting because in this analytical process, the analysis itself bears semblance to external processing in neuronal models where feedback data is constantly optimizing a process for a given task-dependent network.

After conducting an experiment where they collect data from human participants who performed multiple trials of goal-based tasks, Todorov and Jordan assembled a computational theory of coordination. What they found is that optimal performance in biomechanics utilizes the redundancy in the degrees of freedom or lack of movement constraints in humans. Basically, they say that motor variability not be disregarded as erroneous, but should be studied to gain more insight into how internal noise (disruption/error in neuronal communication and the process of perception to motor output) might be behind this variability. This paper is an example of the importance of interdisciplinarity as a neuroscientists. To understand control theory and emerging findings within fields such as this one is to see an opportunity of value, and to step into research looking at how internal noise does, in fact, fit into the framework developed by Todorov and Jordan.