

Logical Extraction of Neo-cortex Structure

Ronald J. Swallow

Brain Research Corporation
7681 Cross Creek Circle
Breinigsville, PA 18031

Tel. 610 301 6629, 610 395 6716, brainresearch.bbnow.org
rswallow@ptd.net

Opinion

I do not understand why the neocortex is a mystery to everyone. Its neuron net circuits are repeated throughout the cortex. It consists of excitatory and inhibitory neurons whose functions have been known for decades. The neuron net circuit is repeated over layers whose axonal outputs feed on as inputs to other layers. The neurons of each layer, each receive axonal inputs from one or more sending layers and all that they need to do is to correlate the axonal input stimulus pattern with their axonal connection patterns from those inputs and produce output frequencies proportional to their resultant psp's (psp = Post Synaptic Potential). Axonal growth toward a neuron is definitely the mechanism for permanent memory formation and it is just what is needed to implement conditioned reflex learning. This axonal growth must be under the control of the glial cells which conditionally enable axon growth and must be a function of the signals surrounding the neurons.

The cortex is known to be able to do pattern recognition and the correlation between an axonal input stimulus and an axonal input connection pattern is just what is needed to do pattern recognition. However, pattern recognition needs normalized correlations and a means to compare these correlations so that the largest correlation in a group of neighboring neurons is recognized by one of the neurons of the group. Without normalization, the psp's relative values would not be constrained properly and could not be used to determine the best pattern match. In order to get psp's to be compared so that the maximum psp neuron would fire, an inhibitory neuron is needed. By having a group of excitatory neurons feed an inhibitory neuron that feeds back inhibitory axonal signals to those excitatory neurons, one is able to have the psp's of the excitatory neurons compared, with the neuron with the largest psp's firing before the others do as the inhibitory signal decays after each excitatory stimulus inhibiting the other excitatory neurons with the smaller psp's. This inhibitory neuron is needed in order to achieve psp comparisons, no question about it. For a meaningful comparison, the psp's must be normalized [1]. As unlikely as it may seem possible, it comes out that the inhibitory connections growing by the same rules as excitatory connections, grow to a value which accomplishes the normalization. That is, as the excitatory axon pattern grows via conditioned reflex rules, the inhibitory axon to each excitatory neuron grows to a value equal to the square root of the sum of the squares of the excitatory connections. This can be shown by a mathematical analysis of a group of mutually inhibiting neurons under conditioned reflex learning. This normalization does not require the neurons to behave differently from what has been known for decades about neurons, but rather requires that they interact with an inhibitory neuron as described.

Thus, by simply having the inhibitory neurons receive from neighboring excitatory neuron with large connection strengths where if the excitatory neuron fires, the inhibitory neuron fires, and by allowing the inhibitory axonal signals be included with the excitatory axonal input signals to the inputs to those excitatory neurons, the neocortex is able to do normalized conditioned reflex pattern recognition as its basic function.

If one thinks about it, layers of mutually inhibiting groups of neurons are all that are needed to explain the neocortex functions. The layers of neurons are able to exhibit conditioned reflex behavior between subpatterns, generating new learned behaviors between those subpattern as observed by the human. With layer to layer feedback, multi-stable behavior of layers of neurons results, forming short term memory patterns that become part of the stimulus to other neurons. With

normalized correlations, there are always axonal input stimulus patterns that will excite every excitatory neuron [1].

The only way to prove this cortex model is to build a simulator, modeling large nets of neurons and observing behaviors that appear human. Most certainly we will never be able to measure the neuron nets of the cortex due to their small sizes. This means, that projects must be formed that do these simulations and do not waste R&D efforts to try to measure a wide range of properties of the cortex which don't result in defining neuron nets that can be built and simulated. Certainly the area to area connection scheme is needed, but it likely can be varied from that of the human, still with intelligence being exhibited. Trials will be needed to determine the initial connection strengths when initiating the simulator. These connections will need to be simple, such as non-zero between corresponding neurons of the mutually inhibiting groups.

Axon growth toward pulsing neurons is the likely mechanism for memory alteration. Having neuron axons back away from neurons, rather than grow toward neurons, has no physical basis and is supported by the fact that the number of axons increases, not decreases, with age in the human. Certainly axon connection strengths never become proportional to axon pulsing frequencies, otherwise the nets of neurons will never exhibit permanent past memories, but rather would be a function of recent events only. Glial cells are likely participants to axonal growth control. It is likely that they will inhibit axonal growth physically, unless a chemical falls below a concentration. In particular, this would be when the excitatory stimulus (chemically emitted to a neuron by axons to that neuron) to a cell, falls below a critical level, where the correlation between stimulus and connection pattern falls below a limit. The result of such a rule is that learning would only occur if stimulus patterns are new and don't match well the connection patterns to neurons. The psychological effect would be a curiosity behavior, observed in humans. Also, it would result in old age reduction of ability to learn, also observed in humans.

Progress in understanding how the brain works has been basically non-existent over the last 40 years due to limits in measurement. Rather, progress requires simulation to work out the missing details. I predict that simulation will dominate the future efforts of researchers. Also, I predict that special purpose hardware will dominate the approach where using conventional computers to simulate nets of neurons in real-time will go out of style very soon because of their high costs.

Simulation permits an evolution process to arrive upon a successful brain understanding. If a logical conclusion of some property of a neuron net is wrong, simulation will eliminate it. If it is right, simulation will verify it.

I believe that I know how the neocortex works, permitting a detailed artificial neuron representation of the full cortex. In the above presentation, I present my logic for my neuron net model and there is direct evidence from neurological experiments to support the model (that the neocortex consists of mutually inhibiting neurons). I am looking for neuro-scientists who support my model and will help me acquire a home for a project to build a neocortex simulator. At least, I hope to find others to communicate with, who might share my theories.

For a derivation of the normalization of the axonal patterns to a group of mutually inhibiting neurons, see the bellow text, based on [2].

Abstract

The post-synaptic potential of a neuron for a long time has been known to be a cross-correlation between an input axonal frequency pattern and an excitatory synaptic strength connection strength pattern. In order for correlations to be useful, they must be able to be compared and must be normalized (covariance of the connection strengths must be constant over all the correlations). A biologically feasible net of neurons was studied (net of N excitatory neurons interacting with an inhibitory neuron) and a very simple Pavlovian rule used for connection strength variation (same rule for both excitatory and inhibitory neurons). The surprise was that the neurons of such a net are able to compare their correlations in a normalized way. Also, the net exhibited greater

learning for new input patterns than for old input patterns, thereby explaining the brain's curiosity drive and reduction of permanent memory plasticity as one ages.

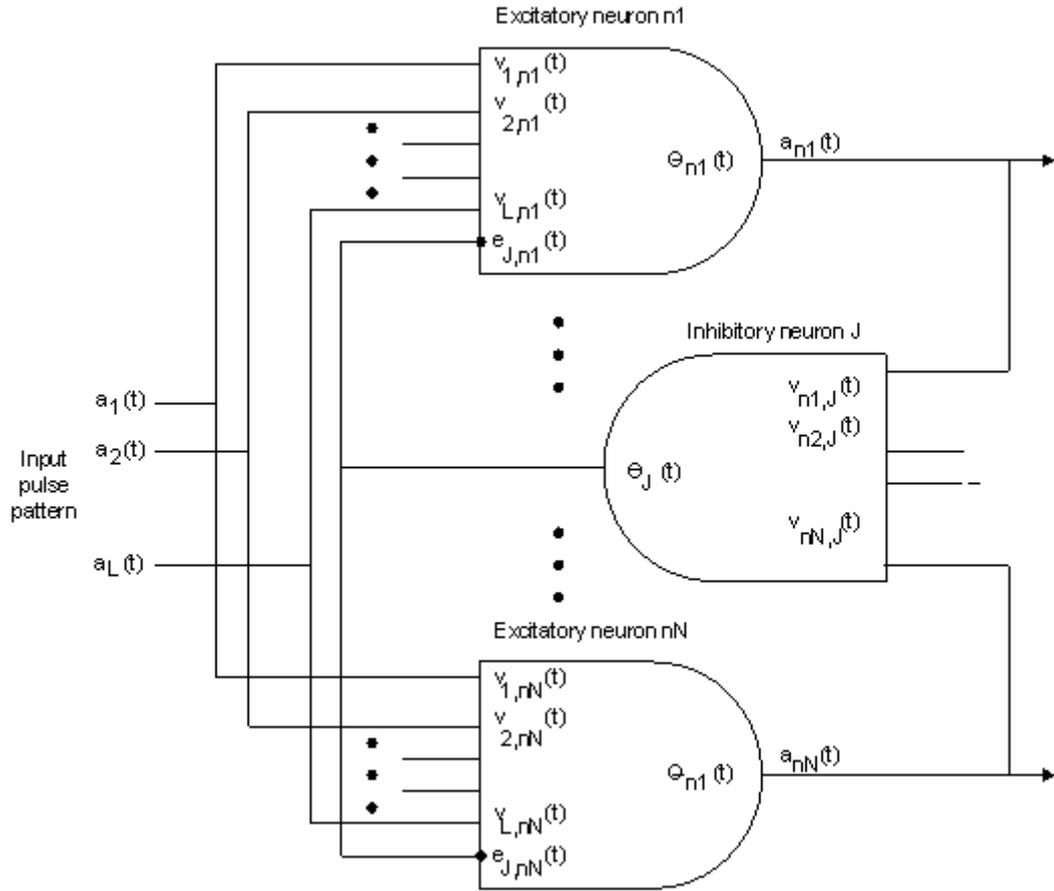


Figure 1. The structure

Referring to Figure 1, if each neuron n ($n = n_1, n_2, \dots, n_N$) receives the same input excitation pattern $a_1(t), a_2(t), \dots, a_I(t)$ where each of these I inputs to neuron n is weighted by a corresponding component of that neuron's connection strength pattern $v_{1,n}(t), v_{2,n}(t), \dots, v_{I,n}(t)$, then the neuron which will fire among the neurons is that with the maximum $P_n(t)$ where (1)

$$P_n(t) = k \frac{\sum_{i=1}^I a_i(t) v_{i,n}(t)}{\sqrt{\sum_{i=1}^I v_{i,n}^2(t)}} \quad (1)$$

and where the denominator actually is the inhibitory connection strength, $e_{J,n}(t)$, to neuron n from the inhibitory neuron J .

The model presented and studied here uses the same rule for varying inhibitory and excitatory connection strengths. A Pavlovian rule (for the growth of these connections) is used. Connection strengths increase only. The dynamic range of each connection variable must (for physical reasons) be finite so that the use of only increasing variables requires considerations of what prevents variables from approaching infinity. By requiring the excitatory psp (post synaptic component) proportional to the acetyl-choline around a neuron to fall below a threshold before any connections may increase, did the trick. Thus, axons of neurons only grow forward under control of glial cells which can detect the acetylcholine levels.

A group of excitatory neurons interacting with a single inhibitory neuron is observable in the human cortex. It can be shown that the model presented here still results in normalized correlations even if more than one inhibitory neuron is involved, as long as all of these inhibitory neurons are triggered by the pulsing of any one of the excitatory neurons.

1. The Model and its Analysis

Consider again the net of N excitatory neurons n_1, n_2, \dots, n_N interacting with the inhibitory neuron J in Figure 1. The $a_i(t)$ associated with the output of a neuron i does not denote the voltage waveform within neuron i 's output axon, but rather, it denotes the psp voltage waveform which would be produced within a receiving test neuron via a unit connection strength. As pointed out in the previous section, all N neurons receive the same input pulse pattern $a_1(t), a_2(t), \dots, a_N(t)$ via corresponding sets of connection strengths $v_{1,n}(t), v_{2,n}(t), \dots, v_{N,n}(t)$. Also, there are the connection strengths $e_{J,n1}(t), e_{J,n2}(t), \dots, e_{J,nN}(t)$ which denote the connection strengths from the inhibitory neuron J to neurons $n = n_1, n_2, \dots, n_N$ and there are the N excitatory connection strengths $v_{n1,J}(t), v_{n2,J}(t), \dots, v_{nN,J}(t)$ from each of the N neurons to neuron J . Because of the way the output state of a neuron is defined, the psp's within the N neurons can be written as follows (see Equation (2)),

$$psp_n(t) = \sum_{i=1}^I a_i(t) v_{i,n}(t) - a_J(t) e_{J,n}(t) \quad (2)$$

and that for neuron J , of course, can be written as in Equation (3).

$$psp_J(t) = \sum_{n=n1}^{nN} a_n(t) v_{n,J}(t) \quad (3)$$

The usual conditions for pulsing within a neuron n with threshold "theta" is (4)

$$psp_n(t) = \theta_n(t) \quad \text{when} \quad \frac{d}{dt}(psp_n(t) - \theta_n(t)) > 0 \quad (4)$$

where the threshold in any neuron n , θ_n , is infinite for a short period of time immediately after Equation 4 is satisfied, after which θ_n exponentially decays toward a resting level θ_{n0} with a time constant of around 1 millisecc.

If the inhibitory neuron J has fired at times t_m up to time t , where the psp's are exponential waveforms, we can write (5)

$$a_J(t) = \sum_m e^{-\frac{(t-t_m)}{T_i}} \quad T_i = 100m \text{ sec } s \quad (5)$$

where a relatively long time constant is assumed for the inhibitory psp waveforms (which has been measured for inhibitory psp waveforms and is rather important in this derivation).

A similar output equation applies for each neuron n (6):

$$a_n(t) = \sum_m e^{-\frac{(t-t_m)}{T_j}} \quad T_j = 3m \text{ sec } s \quad (6)$$

where the t_m are the times when neuron n has pulsed up to time t .

To make this a net of mutually inhibiting neurons, the connection strengths to inhibitory neuron area assumed to obey the following relation (7):

$$v_{n,J}(t) > \theta_0 \quad (7)$$

Thus, neuron J will pulse for each pulsing of one of the N neurons. If no two neurons among the N neurons pulse within a few milliseconds of each other, one pulse will be produced in the inhibitory neuron per pulse from one of the N neurons. This will be the case if the psp waveforms in the N neurons are not equal, since when neuron one fires, the inhibitory component to the psp waveforms of the N neurons suddenly goes more negative preventing the others from reaching threshold.

Being more specific, when one of the N neurons fires, shortly thereafter the following conditions is true in all the neurons (8):

$$psp_n(t) = -a_J(t)e_{J,n}(t) + \sum_{i=1}^I a_i(t)v_{i,n}(t) < \theta_0 \quad (8)$$

where $a_J(t)$ decays in an exponential way with a time constant of 100msecs until the first neuron meets the firing condition and adds another exponential component to $a_J(t)$, etc. Thus, the neuron that will pulse will always be the one which can pulse at the largest value of $a_J(t)$. Solving for that value of $a_J(t)$ for each neuron n (Equation 2 substituted into Equation 4) and calling it $P_n(t)$, the following results (9).

$$P_n(t) = \frac{\sum_{i=1}^I a_i(t)v_{i,n}(t) - \theta_0}{e_{J,n}(t)} \quad (9)$$

Assuming (10):

$$\sum_{i=1}^I a_i(t)v_{i,n}(t) \gg \theta_0 \quad (10)$$

then Equation (9) reduces to (11):

$$P_n(t) = \frac{\sum_{i=1}^I a_i(t)v_{i,n}(t)}{e_{J,n}(t)} \quad (11)$$

where $v_{i,n}(t) / e_{J,n}(t)$ can be interpreted as an effective excitatory connection strength, $w_{i,n}(t)$.

To complete the specification of the model, the synaptic weights are assumed to vary according to the conditioned reflex-like rule as follows: (12) and (13).

$$\frac{d}{dt} v_{i,n}(t) = c_1 a_i(t) a_n(t) L_n(t) \quad (12)$$

$$\frac{d}{dt} e_{J,n}(t) = c_2 a_J(t) a_n(t) L_n(t) \quad (13)$$

where $L_n(t)$ is positive (non-zero) if reinforcement is present, but otherwise this derivation of a normalized cross-correlation is not dependent upon it.

If it is assumed that the frequency of firing of the pulsing inhibitory neuron is greater than 20 pulses per second (determined by the 100 millisecc inhibit decay rate), then (14):

$$(\max P_n(t)) \approx a_J(t) \text{ when the frequency of } a_J \text{ is } \gg 1/T_i \text{ pulses per second (14)}$$

so that (13) can be written as (15):

$$de_{J,n}(t) = c_2 (\max P_n(t)) a_n(t) L_n(t) dt \quad (15)$$

which can be combined with (11) to eliminate $P_n(t)$, and then both sides of the resulting equation can be multiplied by $e_{J,n}(t) / c_2$ to yield the following equation (16):

$$\frac{1}{c_2} e_{J,n}(t) de_{J,n}(t) = \sum_{i=1}^I a_i(t) v_{i,n}(t) a_n(t) L_n(t) \quad (16)$$

Taking Equation (12) and multiplying both sides by $v_{i,n}(t) / c_1$ and summing over $i = 1$ to I , we get the following result (17):

$$\frac{1}{c_1} \sum_{i=1}^I v_{i,n}(t) dv_{i,n}(t) = \sum_{i=1}^I a_i(t) v_{i,n}(t) a_n(t) L_n(t) \quad (17)$$

From Equations (16) and (17) we get (18):

$$\frac{1}{2} \frac{1}{c_1} \sum_{i=1}^I d(v_{i,n}^2(t)) = \frac{1}{2} \frac{1}{c_2} d(e^2_{J,n}(t)) \quad (18)$$

which when integrated with initial conditions of zero, becomes the following equation (19):

$$e^2_{J,n}(t) = \frac{c_2}{c_1} \sum_{i=1}^I v_{i,n}^2(t) \quad (19)$$

Taking the square root of both sides we get the following result (20):

$$e_{J,n}(t) = \sqrt{\frac{c_2}{c_1}} \sqrt{\sum_{i=1}^I v_{i,n}^2(t)} \quad (20)$$

Thus, the effective connection strength becomes (21):

$$w_{i,n}(t) = \frac{v_{i,n}(t)}{e_{J,n}(t)} = \sqrt{\frac{c_1}{c_2}} \frac{v_{i,n}(t)}{\sqrt{\sum_{i=1}^I v_{i,n}^2(t)}} \quad (21)$$

which has constant covariance of c_2 / c_1 as derived by the following equation (22).

$$\text{covariance of } w_{i,n}(t) = \sum_{i=1}^I w_{i,n}^2(t) = \frac{c_2 \sum_{i=1}^I v_{i,n}^2(t)}{c_1 \sum_{i=1}^I v_{i,n}^2(t)} = \frac{c_2}{c_1} \quad (22)$$

Equation (11) becomes (23):

$$P_n(t) = \sqrt{\frac{c_1}{c_2}} \frac{\sum_{i=1}^I a_i(t) v_{i,n}(t)}{\sqrt{\sum_{i=1}^I v_{i,n}^2(t)}} \quad (23)$$

where the $\max P_n(t)$ ($n = n_1, n_2, \dots, n_N$) determines which neuron will fire and will do so with a normalized correlation.

2. Discussion

In Equations (12) and (13), the $a_n(t)$ term determines which neuron connections will alter. The $L_n(t)$ term determines whether they may alter. The following function for $L_n(t)$ is interesting (see (24)):

$$L_n(t) = u(\theta_1 - \sum_{i=1}^I a_i(t) v_{i,n}(t)) \quad (24)$$

because it permits learning if the excitatory correlation is less than a threshold. It is readily performed by the glial cells that support the neurons. This function represents the “newness” of the input pattern as seen by neuron n . This function will produce a tendency in a large net of neurons to seek newness, since its trials that seek less newness, are less remembered. This tendency toward remembering those trials that yield greater newness, explains man’s curiosity drive. Since connections only increase, it also explains that there are less input patterns which will produce sufficient newness for learning to occur. This explains man’s poorer permanent memory as he ages.

When this paper was first presented over twenty years ago in a remote publication and the audience was generally not interdisciplinary enough to understand it, there were no inquiries about its contents. Today, the typical biologists have calculus behind them, so that this paper should be within their comprehension.

It has been impossible to measure the behavior of even a few neurons because of the large number of inputs, their size, and probe damaging effects. This paper has developed this model by deductive reasoning where it was necessary to get neuronal correlations to be normalized. The model is extremely simple with a simple conditioned reflex rule for connection growth, yet it achieves this normalization which required a square-root of the sum of squares of a large number of connection strengths. The connection growth mechanism is very feasible since each pulse reaching a synapse causes the vesicles containing the transmitter chemical to break through the synaptic membrane junction to a neuron, and replacing a portion of that membrane with the membrane of the vesicles, thereby expanding the synaptic contact area. If the glial cells do not “relax”, no learning occurs. If they do relax, the connections increase.

Looking back at the equations, $L_n(t)$ could be negative where the glial cells instead of relaxing actually pressure the synapse to shrink as the vesicles penetrate the synaptic boundary, yet the normalization process would not be altered. Permanent memory properties would not necessary exhibit the “newness” seeking behavior, but at least the normalization process would be preserved.

3. Conclusion

Twenty years ago integrated circuits were not of sufficient size and low cost to be used to model a large net of neurons. Today this is not the case. An FPGA chip connected to 19 memory chips, can be produced to model five hundred thousand frequency neurons (or around 25 million pulsing neurons) in real-time at a cost of \$500. A 25 billion pulsing neuron net would cost \$500,000. Thus, researcher efforts are now feasible to study a full brain model if desired. However, neuronal nets should show very interesting behavior at the 500 million pulsing neuron size.

4. References

- [1] Swallow, R. (1980). A Model for Normalized Cross-Correlations at the Neuronal Level. In: Modeling and Simulation, Volume 11, Part 3, Proceedings of the Eleventh Annual Pittsburgh Conference, May, 1980.
- [2] Brain Research Corporation, Retrieved 1 February 2010, from <http://brainresearch.bbnow.org/>.