Building a Demand based memory based on a Marr type Cortex, requires Columns, Mini-Columns, Attention and a Bottleneck Search

Graeme E. Smith, GreySmith Institute for Advanced Studies
http://en-wikiversity.org/Portal:GreySmith_Institute
http://en-wikiversity.org/User:GreySmith_Institute
grysmith@telus.net

When Marr first created a theory of the Cerebral NeoCortex, his model was quickly attacked because it failed to satisfy the current understandings on how the cerebral cortex worked, and also because it failed to explain much of the function of the cortex, because it limited itself to the first 4 layers of the cerebral cortex, and ignored the functions of the other layers completely. Marr abandoned his work, when he couldn't with his probabilistic mathematics explain the connections between Neurons. With Neural Darwinism, to fall back on, the architecture of the cerebral cortex is less formal than Marr's Codon would suggest, but may still for all of that function, as he suggested. To round out the Model however we need to add a few external support elements and explain columns and mini-columns.

While Marr [Marr 1970] captured the basic function of the first three layers of the Cerebral Neo-Cortex, we have reason to believe that his 4th layer model was flawed, if only because he assumed that it's purpose was to calculate and so he assumed that the basket neurons where divisional in nature. Today, we recognize that multiple inhibitive connections at the synapse is more characteristic of a shunt function, than a division function, and so the Mathematical model seems to fall apart, suggesting other roles for layer 4, but at Marr's time, Mathematical function was assumed, if only because we did not know that there was so much individualism in the connections at the neuron level that such calculations would be meaningless.

Marr himself assumed that his probabilistic mathematical treatment would explain away the individual variation, and was sorely disappointed when it failed to do so. However since then, we have Neural Darwinism, and the concepts of Non-Genetic Darwinism to explain why despite the fact that individual neuron connections are so individualized as to be impossible to map except for the individual themselves, at the gyrus and sulcys level, functional mapping seems accurate across species and with some caveats within a group of species.

It is therefore quite plausible that Marr's main thesis, that the Cerebral Cortex acted as a "self-classifying content addressable memory" might still be accurate even if the details of how it worked were not as accurate as some would like. In this article I suggest that the thesis is necessary but not sufficient to explain the complete cortex function, and thus with some additions, suggest a larger function I call the Dual Mode

Cortex Theory. I note that Marr's work was roundly criticized because it failed to explain certain signal data, which this model will more readily explain, but note that this theory is undoubtedly still missing important information that will be needed to completely explain the signal environment of the cerebral cortex, and thus while it might be necessary, will still not be sufficient to completely explain the function of the cerebral neocortex. However the other signals may prove to be the exceptions that prove this model to be accurate when other support mechanisms are included in the model. It is because I am trying to limit the model as much as possible to the cortex itself, that I must limit the involvement of the other areas of the brain.

To understand the why of the necessity for a larger model of the brain than Marr suggested, I need to capture the thoughts of Jerry A. Fodor [Fodor 2000] who suggested that there was a problem with both the phenomenal school that claimed that the brain was not computationally determinable, and with the Computational Theory of Mind, that claimed that computation was based on "Truth Preserving Functions", and that neither was sufficient to explain the full function of the mind. Other Philosophers have accepted a wider construction for computation, and have noted that simulations of the phenomenal school are still computational in nature, and that therefore there is no reason not to believe in the computational theory of mind, if you keep an open mind about the definition of computation [Chalmers undated]. Given this interpretation, however he had also noted that the nature of phenomenal methods based on the neural network models most often used by phenomenalists, left one at a disadvantage in that these methods did not lend themselves to demand-type addressing. As a result, Marr's content addressable memory presentation makes more sense, but the obvious ability of the brain to recall declarative memory, suggests a need for a more complex model to capture the function of the memory plausibly. Fodor suggested we look at the Columnar organization of the brain for ideas in how it might work as a demand type memory.

In fact, Columnar function, might possibly explain locality of memories within the brain, by promoting the concentration of memories into neural groups. But it does not fully explain the nature of demand type memory, within the brain, and as such, fails to complete the model. To understand how columnar organization within the brain explains locality in the cerebral cortex, we must first accept that the nature of neural networks is their phenomenal nature where location of a memory within the brain is difficult if only because there is no reason except for happenstance why a memory would be located in any particular area within the network, since the content of the network is distributed across many neurons.

Scientists have noted however that, especially within the cerebral cortex, neurons tend to form into groups, roughly equivalent to the columns of the brain, where no matter how the group is stimulated a center surround type function assures that one or two neurons in the group fire, and the rest remain insufficiently activated to fully fire. This it is thought is caused by some function of the sixth layer of the cortex, in which there is a pyramidal neuron that acts to suppress function in the peripheral of the group, while promoting it in the center. This suggests that the group as a whole is responsive to sub-activation level signals that vote on whether the group will fire. The output is controlled by the sixth level neuron which keeps the signals below firing threshold except for the centroid of the set where the neurons are allowed to fire.

If Firing, as seems to be the case is a requirement for some types of plasticity, then we will find that the centroid of the set is more likely to learn than the peripheral locations, thus promoting the location of certain types of data within a specific centroid area. According to Dr. Edelman [Edelman 1990] neural groups are interchangeable and thus any coding method used to address them would have to be degenerate or arbitrary and therefore not predictable or mappable outside the individual. Using this information, scientists have developed special pods based on Self-organizing maps[Kohonen 2001], that allow them to detect and thus decode individually the particular neural mapping of some functions, and thus have allowed them to teach a monkey to use a robotic arm, to give itself treats. Merely by controlling the motor control center in its own brain as if it were moving its own arm.

Obviously if we can map at the individual level, accurately motor center instructions to an arm, we can map other areas of the brain as well, and there is a chance we could learn the individuals address coding. Thus we can accept that the brain is capable of learning its own address coding at some level. All we need is some mechanism whereby it can address the columns, and we have a dual mode memory capable of acting as a demand memory. Unfortunately the sixth layer neurons seem to be internally connected, within the cerebral cortex, so their function seems to be limited to the promotion of the locality of knowledge.

However, David LaBerge, [LaBerge 2002] has noted a system that might have the required degeneracy to act as an addressing technique suitable for activation and thus addressing of neural groups, and this mechanism is centered not on the sixth layer pyramidal neurons but on the 5th layer pyramidal neurons and their Apical Dendrites.

If this is true, then except for the 4th layer neurons, which we dispute with Marr, we have covered enough function to explain the nature of the Cerebral Cortex, as a Dual ported memory with one port allowing access via content addressable format and necessarily implicit memory, since it can't be retrieved due to its phenomenal nature, and a second port allowing access via a degenerate addressing scheme based on the mini-column sub-architecture of the columnar architecture. This Dual Mode Memory, as I have named it, still does not adequately explain the signal nature of the cerebral cortex function. Thus I can say that it is necessary but not sufficient to explain that function.

However, the nature of the error does point to a plausible link between the RAS function and the cerebral function, that might explain at least partially, the missing signal functions. One of the main criticisms of the Marr Model, was that it didn't explain the frequencies that have been observed in the Cerebral Cortex. It has been noted also that the nature of the cortex is such that different sensory modalities are isolated from each other except in illness. There is a necessary binding function, missing from the architecture as described so far, that would explain why we can link cross modal information to achieve a coherent whole. Mere redescription from implicit to explicit memory, would not achieve these ends.

The connection, it seems, lies in the other ends of the neurons that link to the 5th layer Apical Dendrites. LaBerges work [LaBerge 1997] links these connections to the Bottom-up activation system which is part of what he describes in other works as a Triangular Attention Circuit. The neurons link back to the Thalamus, which has in other cases been implicated in redirection of sensory inputs.

Because the Thalamus is so complex, Dr. Edelman [Edelman 2005] has named it part of a "Dynamic Core" which resonates in synchrony with areas of the cerebral cortex, linking areas that are spread across the cortex, and thus, hopefully binding cross sensory information. Because we have detected the frequencies of Gamma Synchronous Oscillations in the Thalamus, and the thalamus is more closely linked to the Reticular Activation System, than the Cerebral Cortex, I believe we have a reason for suspecting that in fact the thalamus is involved in the imposition of the GSO signal onto signals being generated by the first three layers of the Cerebral cortex.

This is because the Reticular Activation system has long been implicated in lesion studies with the distribution of wave like signals from a source in the reticular formation to areas of the cerebral cortex. Dr. Edelman suggests that it is these GSO's that bind the cerebral outputs into Functional Clusters somehow.

If so, it is probable that it is the prefrontal cortex, (implicated in Laberges Triangular Circuit) that through its suppressive selection mechanism, isolates the signals from the areas that are not the focus of our attention. One way that this could be done, is if the selective function worked as a band-pass filter, allowing only one GSO at a time to pass. This would mean that we can separate the output of the cerebral cortex according to the functional clusters defined by the imposition of different GSO frequencies. However this theory of the function of the prefrontal cortex is not without its detractors, since if it operated in that manner we would expect certain frequency sensitivities that have not shown up in testing.

Part of the problem with the band pass filter analogy is that the cerebral signal is not modulated by the GSO it is simply added on, on top of the signal, tagging it. As a result, a true band-pass filter would lose the data content of the signal. However, even if the mechanism is different the model is suggestive enough that scientists are looking closely at the GSO, as a mechanism of selection for functional clusters. What those functional clusters define however, is still somewhat problematic, since evidence has shown up that no attempt is made to separate figure from background at the GSO level.

Laberges Triangular Circuit Theory has proven not to be sufficient to explain Attention, since there are distinct areas of the attention system that have been found to impact on it's operation, that are not even mentioned in his theory, but I believe it is necessary for us to understand how his triangular mechanism works, if we want to understand the nature of the cerebral cortex.

However our model is advanced enough to wonder what the filtering that we are describing is useful for. The model would not be complete without some method of redescribing implicit memory into explicit memory, so that it can be manipulated. The answer of how to do that, is not immediately obvious, although it is easy to cheat and assume that we understand without looking to see what is involved. The main problem is the nature of the output from phenomenal implicit (Content Addressable) memory. Until we characterize this memory, we might think that it, like computer memory is able to pick out specific elements, and limit its output to those elements. Francis Crick's [Crick 1990] idea of an attention system that acts like a spotlight, assumes this. However as Marr clearly noted but glossed over in his article, one of the main problems with Content Addressable Memory, is the fact that it is Voluntary, and necessarily redundant. In other words, the output is information rich, but selectively poor. With this understanding much of the problems associated with Qualia, and the "Hard Problem" of human consciousness become less difficult, but the redescription between implicit and explicit forms becomes more problematic, because we have to deal with a broad field of data that we can't isolate elements in, making it a distinct LUMP of data, called a Quale, rather than an organized listing. Not only that, but the large lump of data is presented in parallel, requiring large processing elements to deal with it. The relative size of the organs with respect to the cortex, suggests that each organ specializes in only a small functional role, and thus that the depth of the processing happens not in the other organs but in the Cerebral Cortex somehow.

With a broad field of parallel data even after filtering, the main requirement for the conversion between implicit memory and explicit memory is the detection of all the parallel mini-columns needed to address the functional cluster. This makes the most likely representation, something called a CLUMP, which is essentially just a list of mini-column addresses needed to simulate the implicit functional cluster.

This clump is kept in the working memory so that it can be further manipulated, because it doesn't make a whole lot of sense in its natural form. What is needed to make sense of it, is the ability to break it up into salient elements that can then be addressed separately. However we haven't yet dealt with how the CLUMP is created. Because we don't know which sub-elements of the functional clusters Quale will be, we need to scan all the mini-column addresses and determine if they are part of the quale. To do this we need only one quale at a time to be present which means we need to isolate the single functional cluster we are going to convert. We know how to do this using suppression of every other quale at the prefrontal cortex level, but what we don't know exactly how to do, is to detect whether or not a mini-column is part of the quale. I will postulate a Bottleneck device that is some sort of comparator, to take on this role.

Given the isolation and the comparator all that is needed is a serial scan through the mini-columns testing each with the comparator to find out if it is part of the quale, and noting its address if it is. Note that if we tried to do this in parallel we would create the situation where the comparator was measuring two mini-columns at once, and the comparator mechanism would be too simple to work. We can get partial parallelism by pipelining the mechanism, but the actual comparator step has to be done in sequence, in order to work. This explains the serial dependency of short term memory.

A CLUMP isn't much better than a functional clusters quale, because the output of the clump once it has been rehearsed is another quale, but because the clump is made up of independent addresses, it is possible to filter the addresses, in order to break the resulting quale down into more salient elements. All that is needed is to pass the modified address through both the rehearsal, and bottleneck stages, and out will pop a smaller quale that is measurably more or less salient. A feedback loop between the organization of neural groups and the salience of resulting quales, will allow the development of the perceptual system to specialize for the type of environment it is involved in. It is for this reason that most new robotic systems are being embodied and asked to respond to specific environments. It is probably also the reason why most advanced animals require a "childhood" in which to learn their environment.

The nature of this memory system, is that it develops over time, and specializes to the type of environment it lives in. I hope I have shown you how a more complex model involving attention and a bottleneck device is required to indicate the nature of the

Cerebral Cortex and that the nature of the Cortex is essentially a Dual Mode Memory system based on an implicit Marr like content addressable memory .

References:

[Chalmers Undated] Chalmers David J. *A Computational Foundation for the Study of Cognition* http://consc.net/papers/computation.html (Undated)

[Crick 1990] Crick, Francis and Koch C *Towards a neurobiologial theory of Consciousness* Seminars in Neuroscience 2 263-75

[Edelman 1990] Edelman, Gerald M. *The Remembered Present:A Biological Theory of Consciousness*, Basic Books, (1990) ISBN 046506910X, 9780465069101

[Edelman 2005] Edelman, Gerald M. *Wider than the Sky: The Phenomenal Gift of Consciousness*, Yale University Press (2005) ISBN 0300107617, 9780300107616

[Fodor 2000] Fodor, Jerry A. *The Mind doesn't Work That Way:The Scope and Limits of Computational Psychology*, MIT Press(2000) ISBN 0262561468, 9780262561464

[Kohonen 2001] Kohonent, Teuvo *Self-Organizing Maps, Edition 3* Springer(2001) ISBN 3540679219, 9783540679219

[LaBerge 2002] LaBerge, David *Attentional control: brief and prolonged,* Psychological Research (2002) DOI 10.1007/s00426-002-0097-2 [published online]

[LaBerge 1997] LaBerge, David *Attention, awareness and the Triangular Circuit.* Consc. And Cog. 6, 140-181(1997)

[Marr 1970] Marr, David *A Theory for Cerebral Neocortex, proc R Soc Lond B Biol Sci 176:161-234*