Analogical Cascade: A Theory on the Role of the Thalamo-Cortical Loop in Brain Function [1]

Yoonsuck Choe [a]

[a]*Department of Computer Science*
Texas A&M University
College Station, TX 77843-3112

Abstract

The cortico-thalamic connections and the nucleus reticularis thalami (nRt) have gained much attention lately because of their integrative and modulatory functions. This particular architecture has been thought to perform analysis and synthesis of memories, work as active blackboards or a global workspace, or gate attention and give rise to consciousness. In this paper, I show that this circuitry can be implementing a general *analogical functionality*. The concrete connection between analogical function and its exact neural basis established in this paper can help us better understand the brain function.

Key words: Active Neurons, Analogy, Thalamus, Nucleus Reticularis Thalami

1 Introduction

Current computational neuroscience research is focused on discovering what kind of information is carried by neurons. In other words, we are interested in what the neurons represent. However, this viewpoint makes the neurons *passive*. In such a passive view, neurons require further interpretation (of what they represent) by others and that can cause the problem of infinite regress. To overcome this problem, I propose that we assign a more *active* role to the neurons. Active, in the sense that neurons receive input (temporal or spatial) and immediately *invoke* other neurons if the input was preferable. With this slight change of viewpoint, a surprising functionality can be derived; that of *analogy*. In this paper, I will describe how as a collection such active neurons can perform analogical function, and show that the inhibitory mechanism in the nucleus reticularis thalami (nRt) can carry out activity gating necessary for proper analogical function.

[1] This research was supported in part by the Texas Higher Education Coordinating Board ARP/ATP program under grant #000512-0217-2001. $Revision: 1.10$

2 Active Neurons: A Primitive for Analogical Computing

Instead of focusing on understanding what kind of information the neurons encode and process, we can ask what *action* is taken when they sense a certain feature in the incoming input, be that temporal or spatial. The action performed by neurons is basically *invoking* activity in other neurons. In this way, neurons represent a certain input feature, and take immediate action by invoking other neurons once the feature is detected. An important observation here is that such invocation establishes a relational context among neurons, because connections between neurons tend to strengthen when they are causality related [9].

The question then is what kind of general principle can such active neurons implement? Such a unit alone cannot achieve much, neither can a serial chain of such units. The true power of this simple unit is revealed when it is used in a massively parallel way. It turns out that the collective effort of these simple units can embody a simple yet powerful functional principle of analogy. We have to simplify matters to see how such neurons can process analogy.

Let us assume there are six neurons in an imaginary creature's brain inhabiting the world of fruits (figure 1). After the fruit brain experiences the world of fruits, it will learn the co-occurrences between features and establish relational arrows as shown in the figure (arcs with arrows). Also suppose that the brain is partitioned into several specialized map areas (or partitions), as in cortical maps. Now, suppose <apple>, <orange>, and <word-red> were presented to the creature simultaneously. If we track the activation, we can see that these detectors will turn on: apple detector, orange detector, color-red detector, color-orange detector, and finally, word-red detector. These activations are *input-driven*. Because the neurons are active, as soon as they detect what they are familiar with, they send out signals through the relational arrows horizontally across the cortex. As a result of this second order activation, the word-orange detector turns on, even without input. Now, here is the crucial moment. We can ask this question: *which neuron's firing was purely cortically-driven?*. Note that this question can be viewed as a filtering (or a gating) process. The result of the filtering is then <word-orange>. The significance of this observation is that this process is very similar to solving analogical problems. The input presented to the creature is basically an analogical query: <apple>:<orange> = <word-red>:<?>. The filtered cortical response <word-orange> can then be the *answer* to this query. [2] Thus, active neurons can perform a rudimentary analogical function when the responses are filtered properly.

However, things can get complicated when combinations of objects are used as a query. Let us extend the creature's feature detectors to include concepts of small and big (not shown in the figure). Then we can allow the creature to learn the relations again. We can then present an analogical query like this: <big><apple> :

[2] There is an issue of how the presence of <word-red> can affect the outcome at all. This problem will be discussed later in the discussion section.

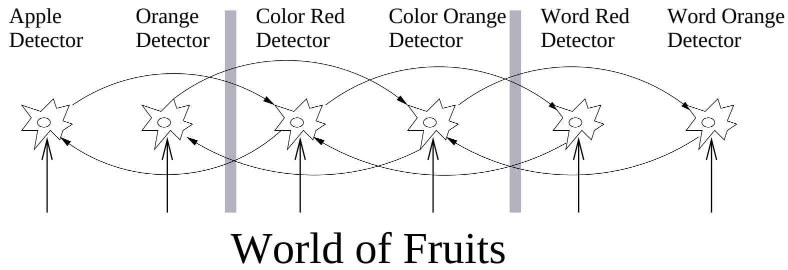

Fig. 1. **World of Fruits.** A brain with object, color, and word detector neurons is shown. The six neurons each respond to these input features as labeled above. At the bottom is the fruit world, and the thick vertical arrows represent afferent input. The horizontal arcs are the relational arrows that point to their most frequently co-occurring counterparts that have been learned through experience. The gray vertical bars represent the partitioning of the brain into separate map areas (from the left to right, object map, color map, and word map). Note that for simplicity, the word-orange detector connects only to the color-orange detector, but not the orange detector, i.e. it is a word-color-orange detector, not a word-object-orange detector.

<small><apple> = <big><orange> : <?>. In this case, if we follow the same steps as above, we come across a problem. Because the answer we expect (i.e. <small> <orange>) already appeared in the query, if we look for purely cortically-driven activations, the answer will be <word-red> <word-orange>. However, we can overcome this problem if we ask: *what are the most cortically-driven activities in each partition of the brain*? Because <big> and <apple> appeared in the input twice but <small> and <orange> appeared only once, the latter two can be selected, as well as the purely cortically driven activities listed above. Thus, even for derived activities that are input-driven, those that are less input-driven can survive and the correct analogical response can still be found among such activities that are more cortically-driven within each partition (or area). Note that <color-orange> also survives the filtering, but what is more important here is that a simple filtering process as described above can generate a *small subset of potential answers* to analogical queries. Although the simple analogical query presented above has a straight forward answer, in more complex analogical problems, there can be multiple answers depending on the interpretation [6].

In this section, I have shown that active neurons that encode input features and relational contexts can collectively perform rudimentary analogical functions. [3] But does the brain function in such a way? In fact, an exact circuit that may be implementing such a function exists in the brain.

3 Neural Basis of Analogical Completion and Filtering

Two basic mechanisms are needed to account for the proposed analogical function: completion and filtering. Below, I will discuss how the cortico-cortical connections

[3] Analogical tasks can become much more complex than the ones shown here. The example in this paper is decidedly simple to clearly illustrate the basic mechanism.

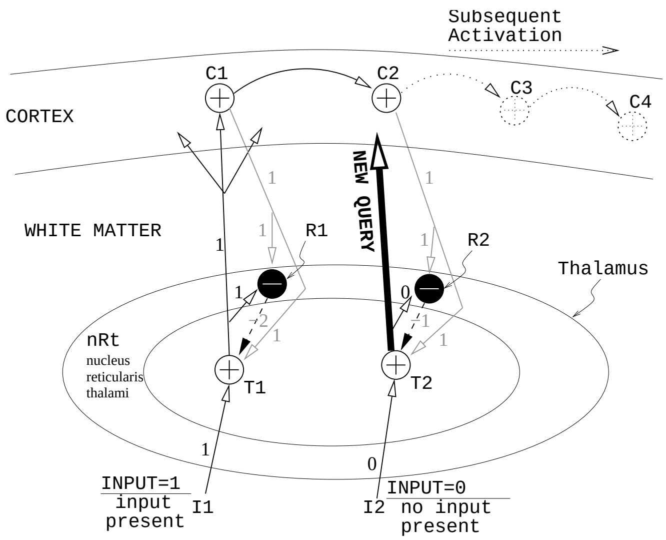

Fig. 2. **Analogical Filtering in the Thalamus.** The diagram shows a simplified thalamo-cortical loop that can perform analogical completion and selection, and propagate the selection back to the cortex. All connections shown are based on known anatomy of the thalamus and the cortex [7]. I1 and I2 are input fibers, T1 and T2 are thalamic relay cells, R1 and R2 are inhibitory nRt cells. C1, C2, C3, and C4 are cortical neurons (each is a set of neurons ranging multiple layers in a single cortical column). The neurons are either excitatory (+) or inhibitory (–), and the arrows are axons (pointing in the direction of action potential propagation). The numbered labels on each arc show the activity being carried. Black solid arrows are ascending fibers to the cortex and the cortico-cortical connections (relational arrows), and gray solid arrows are cortico-thalamic feedbacks. Black dashed arrows are inhibitory connections. The diagram shows a scenario when input was presented to C1, which excites C2, and in turn generates the feedback from C2 to T2, which is then retransmitted to the cortex as a new query (ascending thick black arrow). The selection decision for further propagation to the cortex depends on the relative excitation and inhibition T1(T2) receive from C1(C2) and R1(R2). On the right of C2 (dotted) in the cortex is the subsequent cascade of analogical completions. Note that to avoid clutter, reciprocal connections in the cortex, as well as disinhibiting connections within the nRt layer are not shown.

and the thalamo-cortical loop can implement these two mechanisms.

Completions may be accomplished by the long-range cortico-cortical connections. Synapses are strengthened when the presynaptic activity precede postsynaptic activity [9], thus the connections can implement causal relations. Also, specific patterns of connections observed in animals (e.g. in the primary visual cortex of monkeys; [1]) show how such patterns can implement specific completion functions. Computational models also showed how such connectivity patterns can encode feature co-occurrence and how they can dictate the performance of the model [3, 4].

For filtering, a separate mechanism is necessary. It turns out that the feedforward and feedback connections from the thalamus to the cortex together with the nRt inhibitions can filter feedbacks from the cortex to promote the most cortically-driven feedback, i.e. the analogical answers. Let us first see how the purely cortically-driven activities are selected (figure 2). In the thalamus, ascending fibers (T1 to C1) branch out and excite the inhibitory nRt neuron R1 (T1 to R1). When the feedback from C1 to T1 comes back, it branches and stimulates R1. As a result, if the descending feedback had a matching ascending signal, the inhibition T1 receives is twice as high as other neurons in the thalamus that are activated by purely cortically-driven feedback (e.g. T2). If the synaptic weights are appropriate (i.e. $w_{\mathrm{TC}} = 2$ and $w_{\mathrm{TR}} = 1$)[4], at T1 the feedback will cancel out, but at T2 the feedback will survive the inhibition and will be retransmitted to the cortex (the *new query* arrow). Such a surviving cortical feedback, together with the input stimulus at the next moment form a new analogical query to the cortex, and the same process is repeated. That is, C2 elicits activities in C3, and in turn C4 through the thalamo-cortical loop (note that they can be quite far away). For the selection of the *most cortically driven* feedback, the mutual inhibitions in the nRt layer (e.g. between R1 and R2) may disinhibit (inhibiting an inhibitory neuron results in less net inhibition) each other and allow the more cortically driven feedback to go back to the cortex, even when all current cortical activities are input driven.

4 Discussion

The neural mechanisms described in this paper can only account for simple kinds of analogies, and in some case it can even seem as simple pattern completion. For example, <orange> = ? will result in the same answer <word-orange> as in section 2. How can the term <word-red> in the original query affect the outcome at all? For this, I believe that among many possible completions, the general map area (i.e. the partitions in figure 1) that are activated by input gets higher preference. In this example, the object-map, word-map and color-map will turn on, thus purely cortical activations in other general maps (say odor-map, etc.) will not be as salient as that of <word-orange>. Thus, in this way, the presence of <word-red> can indeed affect the outcome of the analogical query. A more precise neural mechanism for this kind of selection of areas (or maps) needs to be investigated further.

Researchers regard the analogical capability as the crux of high-level cognition (see [5] for a collection of current work on analogy). However, analogy does not need to be limited to high-level cognition [2]. With this new view, we can start to understand perception, cognition, and motor functions under the unifying framework of analogy instead of trying to understand those as embodying separate functional principles. How can such a diverse functionality be integrated under the general principle of analogical processing? Massive connections exist within and across different functional areas in the brain, and the sensory/motor maps are topologically organized [8]. When the sensory, cognitive, and motor maps are connected in an orderly way preserving their local topology, analogies within and (more importantly) *across* different domains can be drawn [5].

[4] Here, w_{YX} is the synaptic connection strength from neuron X to neuron Y.

Within this huge number of maps specializing in different tasks, a cascade of multiple analogical completions can be going on in parallel, synchronized at each moment by the 40Hz rhythm to hold an instantaneously coherent state [7]. Such state can then pose as another analogical query, and that process can repeat. When that cascade reaches a motor area, behavior will be generated. Memory content can also enter the analogical cascade, and this quasi-static contribution can prevent the continuously changing input stream from causing random cascades, thereby maintaining a more goal-directed and stable behavior. Specific mechanisms of how the memory content enters the thalamo-cortical loop, and how completed analogies are archived in memory through the interactions with subcortical centers such as the hippocampus should be studied further.

5 Conclusion

In this paper, I assigned an active invocational role to neurons, and it turned out that collectively they can perform an *analogical function*. The thalamic and thalamo-cortical circuits were found to be ideal for implementing such a function, and especially, the gating inhibition in the nRt was found to be crucial for analogical processing. With analogy, different domains can be mapped and related, thus we can start to understand perception, cognition, and motor function under a unifying framework. This new framework will enable us to take a more focused approach in studying the brain function.

References

[1] Gary G. Blasdel. Orientation selectivity, preference, and continuity in monkey striate cortex. *Journal of Neuroscience*, 12:3139–3161, August 1992.

[2] D. J. Chalmers, R. M. French, and D. R. Hofstadter. High-level perception, representation, and analogy. *Journal of Experimental and Theoretical Artificial Intelligence*, 4:185–211, 1992.

[3] Yoonsuck Choe. *Perceptual Grouping in a Self-Organizing Map of Spiking Neurons*. PhD thesis, Department of Computer Sciences, The University of Texas at Austin, Austin, TX, 2001. Technical Report AI01-292.

[4] W. S. Geisler, J. S. Perry, B. J. Super, and D. P. Gallogly. Edge Co-occurrence in natural images predicts contour grouping performance. *Vision Research*, 2001. 711–724.

[5] D. Gentner, K. J. Holyoak, and B. N. Kokinov, editors. *The Analogial Mind: Perspectives from Cognitive Science*. The MIT Press, Cambridge, MA, 2001.

[6] Douglas G. Hofstadter. *Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought*. Basic Books (a division of Harper Colins), 1995.

[7] David Mumford. Thalamus. In Michael A. Arbib, editor, *The Handbook of Brain Theory and Neural Networks*, pages 153–157. MIT Press, Cambridge, MA, 1995.

[8] Joseph Sirosh, Risto Miikkulainen, and Yoonsuck Choe, editors. *Lateral Interactions in the Cortex: Structure and Function*. The UTCS Neural Networks Research Group, Austin, TX, 1996. Electronic book, ISBN 0-9647060-0-8, http://www.cs.utexas.edu/users/nn/web-pubs/htmlbook96.

[9] Sen Song, Kenneth D. Miller, and L. F. Abbott. Competitive hebbian learning through spike-timing-dependent synaptic plasticity. *Nature Neuroscience*, 3:919–926, 2000.