

Second Order Isomorphism: A Reinterpretation and Its Implications in Brain and Cognitive Sciences

Yoonsuck Choe (choe@tamu.edu)

Department of Computer Science

Texas A&M University

3112 TAMU

College Station, TX 77843-3112

Abstract

Shepard and Chipman's second order isomorphism describes how the brain may represent the relations in the world. However, a common interpretation of the theory can cause difficulties. The problem originates from the static nature of representations. In an alternative interpretation, I propose that we assign an *active* role to the internal representations and relations. It turns out that a collection of such active units can perform analogical tasks. The new interpretation is supported by the existence of neural circuits that may be implementing such a function. Within this framework, perception, cognition, and motor function can be understood under a unifying principle of analogy.

Introduction

One of the central tenets in neuroscience is that neurons receive incoming spikes, process that spatial or temporal information, and then pass on the transformed information for further analysis. Furthermore, neurons that fire together develop strong connections (Hebb 1949). Thus, the neurons represent input features, and connections establish a correlational context among neurons. This viewpoint is analogous to the second order isomorphism by Shepard and Chipman (1970; below, just S&C). However, a problem can arise depending on how we interpret S&C's theory. The difficulty comes from a more or less *static* role assigned to the representations.

In this paper, the representations and the relations are given an active role. When working as a collection, these active units can perform an analogical function. In fact, a similar active approach has been employed in previous work, resulting in the emergence of analogical (Hofstadter 1985; Mitchell 2001) or metaphorical (Narayanan 1999) functionality.¹ An important observation advanced in this paper is that the function of active representations and relations are very similar to that of neurons, and specific circuits in the cortex and the thalamus can actually *implement* analogical functions. Analogy is commonly attributed to higher cognitive faculties only, but it does not always have to be the case (Chalmers et al. 1992); it may be part of a larger set of human brain function including perception and motor function. I will discuss

¹Analogy and metaphor are closely related in that they refer to similarities in relations and attributes although the relative degree in each may differ (Gentner 1989). Thus, they may be sharing a common underlying mechanism.

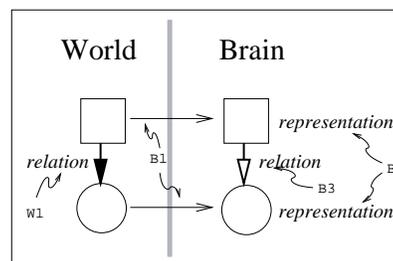


Figure 1: **S&C's Second Order Isomorphism.** There are two objects, one square and one round in the world (on the left). The internal representations in the brain of these two objects are shown on the right. The vertical arrows represent the relations between the objects. The two horizontal arrows represent mapping from the world to the brain which is initiated by sensory transduction. Note: The square and circle on the right (in the brain) are just there for the ease of reference. They can be removed without causing any change in content (this applies to the rest of the figures).

in the end how such an analogical framework can allow us to better understand the nature of cognition and brain function.

Common Interpretations

Under second order isomorphism the brain needs to find the relation between the (1) relations between external objects and (2) relations between internal representations. Figure 1 illustrates the theory. S&C's theory seems to be more appropriate in modeling how our brain represents the world than Locke's Isomorphism (Edelman 1998). In physical terms, we can interpret the figure as follows: (1) relation in the world (W1; coincidences in sensory events) (2) arrows from world to brain (B1; sensory transduction), (3) representations in the brain (B2; afferent connections), and (4) relation in the brain (B3; lateral connections). Of these, let us focus on what is available in our brain (B1–B3). If we take for granted the information our sensory transducers tell us, we can drop B1 from our discussion and focus on just B2 and B3.

An implicit message in figure 1 is that two objects are represented, and some brain process then judges the relationship between the two (the open arrow). Making this point more explicit, we can illustrate S&C's theory as in figure 2a (the diamond box). We can see that a difficulty

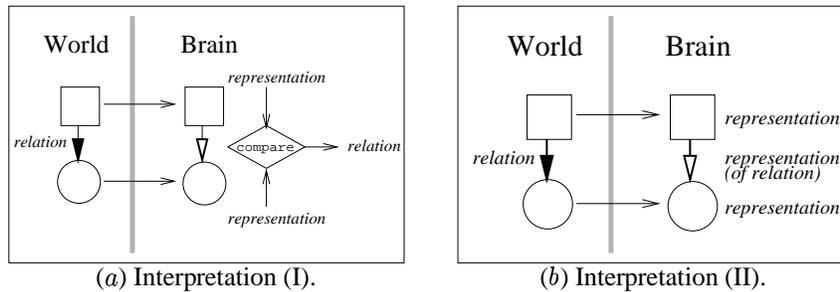


Figure 2: **Common Interpretations of Second Order Isomorphism.** (a) An explicit comparison mechanism is necessary to judge the relations between the two representations. (b) The comparison box is replaced by a representation of relation. However, this figure still requires a third-party to *evaluate* the representation of relation.

can arise in such an interpretation; something has to perform the comparison function, but this creates an ever increasing levels in a hierarchical way (i.e. higher areas judging the output relation in the lower areas). However, as Hilgetag et al. (1996) noted, it is hard to determine a strict (or even a loose) hierarchy among cortical maps (in this case, between visual areas). Thus, representing something and delaying the interpretation until later may not work very well.

One can argue that the lateral connections *represent* the similarity relation, not requiring a separate interpreter (figure 2b). However, we still need something to evaluate (or interpret) the representation of the relation. Thus, this reformulation just replaces the need for one kind of interpreter with another.

Assigning an Active Role to Relation

What can a relation be if it should not be a representation? The problem seem to come from representations and relations playing a *static* role. What if we assign an active role to the representations as shown in figure 3a? In the figure, I assigned an active role to the relation arrow itself, allowing one representation to invoke another. Thus the activation of the internal representation of the square *invokes* (or turns on) that of the circle, and vice versa.

Now consider how can we use this new active relation (note that it is *directional*) to describe the relations in the world. First, we have to know what kinds of relations exist in the world. There are two basic relations: spatial and temporal relations. Spatial relations are between objects, and they are causally bidirectional.² On the other hand, temporal relations are between events, and they are causally directional. When one event precedes the other, the reverse cannot happen simultaneously.

In the brain, action potentials only propagate in one direction along the axon, and the adaptation of synapses tend to learn causality (Song et al. 2000). Such connections are ideal for implementing temporal relations, but what about spatial relations? If we pair a unidirectional arrow from A to B with a reciprocal one from B to A,

²Note that *causal* simply means that one event precedes the other in time.

then we can indeed represent spatial relations with only directional arrows. If representations for object A and B simultaneously activate through mutual excitation, then they can represent the spatial relation between the two. So, let us update our figure again to include backward relational arrows (figure 3b). We can now think in terms of *temporal* relations only, because spatial relations turned out to be a special case of temporal relations (at least in the brain).

The Role of Active Relations and its Neural Basis

In the previous section, I replaced the representation for relation by an active relation. What about the representations for the objects (or events)? Representation is an inherently static term (like a symbol), thus, we should take a more active viewpoint, i.e. we can ask *what action occurs* when a neuron detects a pattern in its incoming input, rather than focusing only on what a neuron represents.

To discover the relationship between things in the world, we need the motor capabilities as much as we need sensors. Thus, between the world and the brain there must be a backward arrow from the brain to the world. The resulting diagram is shown in figure 3c. This addition is crucial in learning the relations in the world (O'Regan and Noe 2001). Surprisingly (or equally not surprisingly), the final diagram looks very similar to the basic circuitry in our brain. How can this final figure help us understand the mechanisms of the brain? The key is to understand what is the *action* taken by a neuron, no less than to know what it represents.

Active Relations: A Primitive for Analogical Processing

Now we have a single active functional unit: a neuron that fires a spike along the *active relational arrow* as soon as it detects a certain input feature. This 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. This may be an obvious line of thought because that is what our brains seem to do. However, it turns out that the collective effort of these

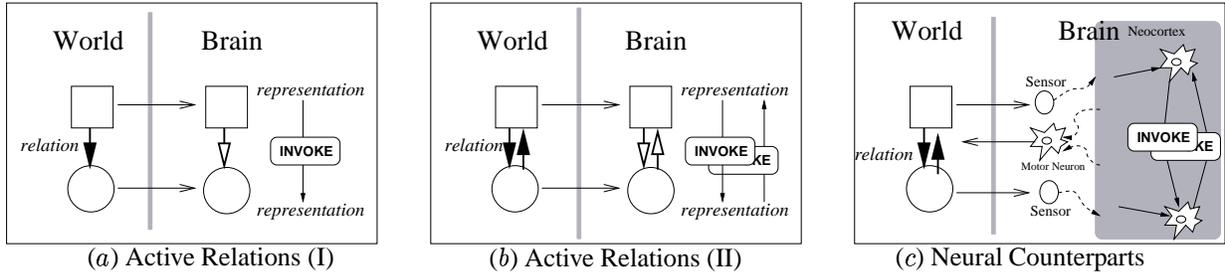


Figure 3: **A New Interpretation of Second Order Isomorphism.** (a) In this new interpretation of S&C's theory, an active role is assigned to the relation arrow. Notice that the INVOKE arrow is a single arrow, not one arrow going into and another leaving out of the box. Thus, the rounded box signifies that the arrow actually performs an *action*. (b) Backward relational arrows are added in the brain to account for the mutual, but directional nature of relations. With two relational arrows, both spatial and temporal relations can be implemented. (c) The neural counterparts of (b) are shown. The limiting term *representation* is removed, and the motor reaction (backward arrow from the brain to the world) is added. Sensory transducers are also explicitly shown.

simple units can embody a simple yet powerful function 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 4). 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 $\langle \text{apple} \rangle$, $\langle \text{orange} \rangle$, and $\langle \text{word-red} \rangle$ 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 we can view this as a filtering process). The result of the filtering is then $\langle \text{word-orange} \rangle$. 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: $\langle \text{apple} \rangle : \langle \text{orange} \rangle = \langle \text{word-red} \rangle : \langle ? \rangle$. The filtered cortical response $\langle \text{word-orange} \rangle$ can then be the *answer* to this query.³ 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 analogi-

³There is an issue of how the presence of $\langle \text{word-red} \rangle$ can affect the outcome at all. This problem will be discussed later in the discussion section.

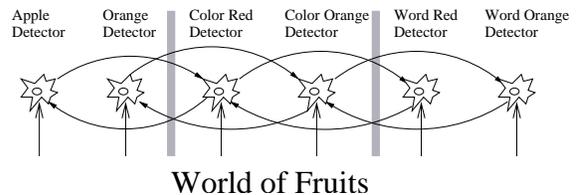


Figure 4: **World of Fruits.** A brain with object, color, and word detector neurons is shown. The six neurons each respond to these input features as labeled above. At the bottom is the fruit world, and the thick vertical arrows represent afferent input. The horizontal arcs are the relational arrows 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.

cal query like this: $\langle \text{big} \rangle \langle \text{apple} \rangle : \langle \text{small} \rangle \langle \text{apple} \rangle = \langle \text{big} \rangle \langle \text{orange} \rangle : \langle ? \rangle$. In this case, if we follow the same steps as above we come across a problem. Because the answer we expect (i.e. $\langle \text{small} \rangle \langle \text{orange} \rangle$) already appeared in the query, if we look for purely cortically-driven activations, the answer will be $\langle \text{word-red} \rangle \langle \text{word-orange} \rangle$. However, we can overcome this problem if we ask: *what are the most cortically-driven activities in each partition of the brain?* Because $\langle \text{big} \rangle$ and $\langle \text{apple} \rangle$ appeared in the input twice but $\langle \text{small} \rangle$ and $\langle \text{orange} \rangle$ 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 $\langle \text{color-orange} \rangle$ 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 multi-

it can even seem as simple pattern completion. For example, <orange> = ? will result in the same answer <word-orange> as in the *Active Temporal Relations:...* section. 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 4) that are activated by input gets higher preference. In this example, the fruit-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 among areas (or maps) needs to be investigated further.

Researchers regard the analogical capability as the crux of high-level cognition (see Gentner et al. 2001 for a collection of current work on analogy). However, analogy does not need to be limited to high-level cognition. Recent results suggest that analogy may be needed in perception as well (Davis and Goel 2001; Morrison 1998), and such an ability emerging in perceptual systems may even be a crucial *requirement* for cognitive development (Chalmers et al. 1992). Then it is not unthinkable that motor functions also employ analogy in a similar manner (cf. sensory motor contingency theory by O'Regan and Noe 2001), thus we can then start to understand perception, cognition, and motor functions under the unifying framework of analogy.

How can such a diverse functionality be integrated under a single principle of analogical processing? Massive connections exist within and across different functional areas in the brain, and the sensory/motor maps are topologically organized, i.e. nearby neurons are responsive to nearby features in the sensory space (Kohonen 1982; von der Malsburg 1973). Within each map, the feature detectors and the cortico-cortical connections learn to encode the relations (Choe 2001). It is possible that cognitive maps also have a topological organization where nearby areas learn to encode similar concepts, such as semantic maps or episodic memory maps (Miikkulainen 1993), or even temporal sequences (James and Miikkulainen 1995). When the sensory, cognitive, and motor maps are connected in an orderly way preserving their local topology, analogies can be drawn within and (more importantly) *across* different functional domains.

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 (Mumford 1995). Such a state can then pose as another analogical query, and the process can repeat. When the cascade reaches a motor area, behavior will be generated. Memory content may 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 enter 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.

Neuroscience research has revealed a lot about perception and motor abilities in the brain, but understanding the cognitive faculty still remains elusive. Investigation into cognitive functions can proceed under the analogical framework, where we can infer the functionality of the higher areas by backtracking the connections to the perceptual and motor areas and study their topology and analogical links. Specific predictions regarding the layout of the higher centers may be made based on the topology of the lower centers and the connection structure between the two, and experiments can then focus on verifying these predictions. For example, there are orientation maps with smoothly changing orientation preference in V1 (primary visual cortex; Blasdel 1992), and there are object maps in TE (temporal area E; Tanaka 1996) that also change smoothly (for example, rotation of a head). The analogical framework predicts that there will be an orderly mapping from V1 to TE that preserve such local topology across different representation spaces.⁶ Similar mappings may exist between sensory and cognitive areas, and if such a mapping is found, we can start to understand the abstract cognitive functions based on concrete perceptual architecture.

The strong connection made in this paper between analogical function and specific neural circuitry can help us better understand the two. We can study specific neural mechanisms of analogical function, and thus we can gain a concrete understanding of the abstract analogical process. More specifically, the functionality of the target area of a neuron can be studied to understand *what action occurs* when a neuron detects a certain feature in the input. Such a study can reveal the kinds of relations implemented in the brain, thus providing us with a hint of what kind of analogy is possible. The mechanisms of neural circuits can also be further revealed by carefully designed analogical tasks in perception, cognition, and motor function, and also in a combination of these different domains. Different types of unimodal and cross-modal analogical tasks can reveal how the different cortical areas are related and how they invoke each other.

Conclusion

In this paper, I analysed the difficulties that the common interpretations of S&C's second order isomorphism can cause in understanding the brain. I proposed an *active* role for representations and relations, and it turned out that collectively they can perform an *analogical function*. An important connection between analogical function and a specific brain circuit was then established, providing support for the new interpretation. This new framework allows us to understand perception, cognition, and motor function under a unifying principle of

⁶Although the pathway from V1 to TE is not direct, involving V2, V4, and TEO areas, but successive mappings within this path can reveal how V1 and TE are topologically mapped.

analogy, and it can help us take a more focused approach in brain and cognitive sciences.

Acknowledgments

I am greatly indebted to Sang Kyu Shin, James Bednar, Risto Miikkulainen, Un Yong Nahm, Marshall Mayberry, Bruce McCormick, and Jyh-Cham Liu for their feedback and encouragement. This research was supported in part by Texas A&M University, and the Texas Higher Education Coordinating Board ARP/ATP program under grant #000512-0217-2001.

References

- Blasdel, G. G. (1992). Orientation selectivity, preference, and continuity in monkey striate cortex. *Journal of Neuroscience*, 12:3139–3161.
- Chalmers, D. J., French, R. M., and Hofstadter, D. R. (1992). High-level perception, representation, and analogy. *Journal of Experimental and Theoretical Artificial Intelligence*, 4:185–211.
- Choe, Y. (2001). *Perceptual Grouping in a Self-Organizing Map of Spiking Neurons*. PhD thesis, Department of Computer Sciences, The University of Texas at Austin, Austin, TX. Technical Report AI01-292.
- Crick, F., and Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in The Neurosciences*, 2:263–275.
- Davis, J., and Goel, A. K. (2001). Visual analogy in problem solving. In *Proceedings of the International Joint Conference on Artificial Intelligence*, to appear.
- Edelman, S. (1998). Representation is representation of similarities. *Behavioral and Brain Sciences*, 21:449–498.
- Geisler, W. S., Perry, J. S., Super, B. J., and Gallogly, D. P. (2001). Edge Co-occurrence in natural images predicts contour grouping performance. *Vision Research*. 711–724.
- Gentner, D. (1989). The mechanisms of analogical learning. In Vosniadou, S., and Ortony, A., editors, *Similarity and Analogical Reasoning*, 199–241. New York, NY: Academic Press.
- Gentner, D., Holyoak, K. J., and Kokinov, B. N., editors (2001). *The Analogical Mind: Perspectives from Cognitive Science*. Cambridge, MA: The MIT Press.
- Harth, E., Unnikrishnan, K. P., and Pandaya, A. S. (1987). The inversion of sensory processing by feedback pathways: A model of visual cognitive functions. *Science*, 237:184–187.
- Hebb, D. O. (1949). *The Organization of Behavior: A Neuropsychological Theory*. New York: Wiley.
- Hilgetag, C.-C., O’Neill, M. A., and Young, M. P. (1996). Indeterminate organization of the visual system. *Science*, 271:776.
- Hofstadter, D. (1985). Waking up from the boolean dream, or, subcognition as computation. In *Metamagical Themas*, chapter 26. New York, NY: Basic Books.
- Hofstadter, D. R., and Mitchell, M. (1994). The copycat project: A model of mental fluidity and analogy-making. In Holyoak, K. J., and Barnden, J. A., editors, *Advances in Connectionist and Neural Computation Theory*. Norwood, NJ: Ablex Publishing Corporation.
- James, D. L., and Miikkulainen, R. (1995). SARDNET: A self-organizing feature map for sequences. In Tesauro, G., Touretzky, D. S., and Leen, T. K., editors, *Advances in Neural Information Processing Systems 7*, 577–584. Cambridge, MA: MIT Press.
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43:59–69.
- MacKay, D. (1956). The epistemological problem for automata. In Shannon, C. E., and McCarthy, J., editors, *Automata Studies*, 235–251. Princeton, NJ: Princeton University Press.
- Miikkulainen, R. (1993). *Subsymbolic Natural Language Processing: An Integrated Model of Scripts, Lexicon, and Memory*. Cambridge, MA: MIT Press.
- Mitchell, M. (2001). Analogy-making as a complex adaptive system. In Segal, L., and Cohen, A., editors, *Design Principles for the Immune System and Other Distributed Autonomous Systems*, to appear.
- Morrison, C. T. (1998). *Situated Representation: Solving the Handcoding Problem with Emergent Structured Representation*. PhD thesis, Bringhamton University; State University of New York.
- Mumford, D. (1992). On the computational architecture of the neocortex, pt. II, the role of the cortico-cortical loop. *Biological Cybernetics*, 65:241–251.
- Mumford, D. (1995). Thalamus. In Arbib, M. A., editor, *The Handbook of Brain Theory and Neural Networks*, 153–157. Cambridge, MA: MIT Press.
- Narayanan, S. (1999). Moving right along: A computational model of metaphoric reasoning about events. In *Proceedings of the National Conference on Artificial Intelligence (AAAI ’99, Orlando, FL)*, 121–128. AAAI Press.
- Newman, J., Baars, B. J., and Cho, S.-B. (1997). A neural global workspace model for conscious attention. *Neural Networks*, 10:1195–1206.
- O’Regan, J. K., and Noe, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24(5):to appear; ID: bs00000506.
- Shepard, R. N., and Chipman, S. (1970). Second-order isomorphism of internal representations: Shapes of states. *Cognitive Psychology*, 1:1–17.
- Song, S., Miller, K. D., and Abbott, L. F. (2000). Competitive hebbian learning through spike-timing-dependent synaptic plasticity. *Nature Neuroscience*, 3:919–926.
- Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience*, 19:109–139.
- von der Malsburg, C. (1973). Self-organization of orientation-sensitive cells in the striate cortex. *Kybernetik*, 15:85–100.