

7 Conclusion

*You perceive I generalize with intrepidity from single instances.
It is the tourist's custom.*

-Mark Twain

To my mind, there are two main reasons for creating a computational model of a cognitive behavior: 1) to help cognitive scientists understand the inner workings of the phenomena by providing necessary concepts and high-level explanations, and 2) to actually create a program capable of performing the phenomena. We might hope that these two goals could be satisfied with a single model, but there is no reason that the goals need be related; a model that is easy for us to understand may not perform well or completely, and one that does perform well may not be easy to understand. I believe that

the two camps of analogy-making models described in Chapter 6 are divided by these two conflicting goals. In this way, those that subscribe to the traditional view model analogy-making in terms of easy-to-understand concepts such as “mappings”, “objects”, “relations”, and “structured representations”. On the other hand, those who wish to create more realistic, truly functional, stand-alone models end up with explanations that are nearly as complex as the actual cognitive system being modeled. These models tend to explain analogy-making not in terms of easy-to-understand representations, but rather in terms of complicated, emergent processes.

Tabletop and the rest of the FARG models strike an interesting balance between these two viewpoints, straddling the boundary between an explanatory and functional model. In designing Analogator, I completely ignored the goal of explaining analogy at the level of easy-to-understand representations. For this reason, one might consider Analogator to be one of a growing number of “radical connectionist” models (Cummins and Schwarz, 1987; Dorfner, 1990; Peschl, 1991).

That is not to say that Analogator does not help to explain analogy-making, but it does so at a level other than of the traditional analogical concepts (i.e., mappings, objects, relations, etc.) By breaking with tradition, it de-emphasizes concepts and representations – those ideas that Hofstadter claims that we need to be focusing on in the field of computational analogy-making (Hofstadter *et al.*, 1995). I do not disagree. However, I believe that we can not *directly* program concepts and representations; the representations that we create by hand are far too rigid. It would be like trying to construct clouds with hammers and nails. We need to develop the tools that allow us to *indirectly* build representations. I believe that we can accomplish this through learning systems; this is the level at which Analogator helps explain analogy-making.

One might think that I believe that the best method for modeling human cognition is one big associative network, starting *tabula rasa*, which learns everything. Charniak and McDermott, in their Introduction to Artificial Intelligence, put it this way:

One idea that has fascinated the Western mind is that there is a general-purpose learning mechanism that accounts for almost all of the state of an adult human being. According to this idea, people are born knowing very little, and absorb almost everything by way of this general learner. (Even a concept like “physical object,” it has been proposed, is acquired by noticing that certain visual and tactile sensations come in stable bundles.) This idea is still powerfully attractive. It underlies much of behavioristic psychology. AI students often rediscover it, and propose to dispense with the study of reasoning and problem solving, and instead build a baby and let it just learn these things. We believe that this idea is dead, killed off by research in AI (and linguistics, and other branches of “cognitive science”). What this research has revealed is that for an organism to learn anything, it must already know a lot.

(Charniak and McDermott, 1985 pp. 609-610)

This belief may have been more dead in 1985 when they wrote the above passage (which, in their defense, was right before the resurgence in connectionism); however, one may feel that that there is a bit of this “let it learn everything” flavor to Analogator. Certainly I do not think that one large associating network is the whole story. But, I must admit, I do agree with some of the philosophy of the behavioral psychologists. After the fall of behaviorism, cognitive scientists could not only assume the existence of internal representations, but could also insert them into their models wherever they felt necessary with little or no justification. I believe that researchers should look at a high-level cognitive ability, such as analogy-making, from a purely behavioristic perspective before making any assumptions regarding proposed internal representations.

Creating systems capable of learning abstract cognitive abilities, such as analogy-making, is currently a very difficult goal to reach for many reasons. For instance, learning such a task imposes additional constraints that force researchers to frame a problem in a

learnable way. In addition, cognitive modelers are typically only peripherally interested in explaining how a model could come to be; they are more interested in the abilities of a model as an end product. However, learning is an interesting topic in and of itself, and will hopefully give us the appropriate tools with which to build flexible representations.

7.1 Analogator Limitations

Currently, Analogator suffers from a few major limitations. Probably the most serious is its lack of any mechanism of *attention*. Because of this, Analogator has problems with certain types of generalization, especially those involving invariance based on position. This problem was evident in both of the experiments involving spatial relations. Having some form of attention would, I believe, allow Analogator to scale-up to larger problems, such as being able to view scenes larger than 7 x 7 pixels.

Of serious concern is that fact that Analogator only operates in “pop-out mode”, never pondering a problem for more than a single propagation. In other words, analogies immediately come to Analogator’s “mind”, or not at all; it never *reasons* about a problem.

One possible extension to the basic Analogator framework could include a method of controlling its flow of recurrent activations by an attentional mechanism. I believe that Analogator could be adapted so that it could self-control recurrent activation, so that it could “explore” problems more thoroughly, and could solve more logic-based analogical reasoning problems (see Rumelhart, 1989, for other future visions of such a system). Eventually, problems that initially required sequential reasoning could be “compiled” into operating in the pop-out mode. This type of recurrent self-control could also be used to handle more complex, hierarchical problems that are currently out of reach of Analogator’s simple associative network. Recursively attending to objects (or

groups of objects) might possibly give Analogator the power and flexibility to address other cognitive abilities, as suggested by Hinton's discussion of representing part-whole hierarchies in connectionist networks (1988).

I believe that a more complex architecture is needed to hardwire in certain other invariances, such as object rotation. Currently, backprop is used to learn everything. Although this makes the model quite general, it has been shown to limit Analogator's abilities, which required "hints" to overcome. I believe that the basic architecture could be adapted in a manner similar to Fukushima *et al.*'s Neocognitron (1983).

Some might suggest that the large number of trials needed to allow Analogator to make analogies is a serious drawback of the model. However, this does not concern me. After all, humans are years old before they are capable of making less superficial analogies. On the other hand, Analogator does currently need a teacher to point out the right answer at every step. It would be desirable to have an unsupervised training phase, to partially eliminate this requirement. This might be accomplished by hardwiring into the network certain "interesting" features in the world that seem to naturally to be the focus of attention (i.e., the figure).

Analogator's performance has not yet been correlated in any way with human analogy-making performance. A preliminary set of experiments was made with the help of Rob Goldstone and his lab. A set of analogy problems similar to those from Analogator's geometric-spatial domain was presented to college-age students, and choices and response-times were recorded. The subjects did show a preference for making analogies, even in situations that did not require, or even suggest, it. These initial experiments showed that humans naturally made analogies, but the exact analogies made with geometric shapes were affected by low-level criteria such as objects encountered first in sequentially scanning a scene, and objects that were brightly colored. Such criteria were, of course, not built into Analogator. Currently, the only way to correlate Analogator

with this type of performance would be to train it to respond in such a manner; however, that would be a superficial method of gaining correlation, even if it could learn all of the subtleties. For proper correlation the model should incorporate sequential scanning, as well as many other low-level abilities.

Recall that Analogator is completely deterministic; it does not involve any randomness, or exploration of a problem. Because of that, it produces only a single answer for each problem, unlike Tabletop. Tabletop produces many answers for a single problem, and the distribution can be compared to human performance (Hofstadter *et al.*, 1995). This is impossible for Analogator in its current form. In addition, the one answer that Analogator does provide is based on the details of how it was trained. In short, Analogator is not quite ready for a fine-grained comparison with humans.

7.2 Contributions of this research

The contributions of this dissertation were made in two categories: the creation of a representation, and the creation of a training process. In the first category, iconic representations were developed which allow relations between objects to be easily compared and contrasted to single objects. For instance, a circle of triangles can be represented in a way that allows easy comparison to the representation of a single circle object. In this manner, the iconic representation creates a single, smooth landscape with which to represent objects, relations between objects, and object attributes.

A new training process was designed called the recurrent figure-ground associating procedure. This process associates the figure-ground components of source-target pairs in a connectionist network. As an associating process, it has much in common with other low-level perception models, and thus provides a unification with recognition and categorization. The training method also provides a unique means for “symbol

grounding” by forcing hidden patterns to share the same activation space with retinotopic representations. Of central importance, the associating method was shown to be able to learn to make analogies in widely differing domains. In addition, the methodology was shown to train much more quickly than comparative feed-forward networks. Most importantly, the networks trained in this style exhibited good generalization ability by being able to make intra- and cross-domain analogies.