1 Introduction

The scientific mind does not so much provide the right answers as ask the right questions.

-Claude Levi-Strauss

The goal of this dissertation is to integrate learning and analogy-making. Although learning and analogy-making both have long histories as active areas of research in cognitive science, not enough attention has been given to the ways in which they may interact. To that end, this project focuses on developing a computer program, called Analogator, that learns to make analogies by seeing examples of many different analogy problems and their solutions. That is, *it learns to make analogies by analogy*. This approach stands in contrast to most existing computational models of analogy in which particular analogical mechanisms are assumed *a priori* to exist. Rather than assuming certain principles about analogy-making mechanisms, the goal of the Analogator project is to learn what it *means* to make an analogy. This unique notion is the focus of this dissertation.

In the following section, I discuss the assumptions that I have made – and not made – in the Analogator project.

1.1 Traditional Assumptions

Although the history of modeling analogy-making is only a few decades old, some well-entrenched traditions have developed. The traditional view of modeling analogy-making appeals to the intuition. A common assumption is that analogy-making begins with two structures – one representing a well-known *source* situation, and the other a lesser known *target* situation (see, for example, Gentner, 1983; Holyoak and Thagard, 1989a). These two structures are imagined to be similar to Minsky's frames (1975), or Schank and Abelson's scripts (1977). Analogy-making is assumed to be a search through the structures, matching analogous components. Data structures are assumed to be composed of structured parts, namely *objects*, *attributes*, and *relations*. There are assumptions about what types of parts can match, which parts to consider first, and which parts should be weighted most significant. An analogy is considered to have been made when a complete set of "mappings" between source and target components has been found.

Much research into analogy-making has focused on creating efficient searches through the data structures. Because these data structures are seen as being a core part of analogy-making, and analogy-making is viewed as a search through them, analogy-making (like much of high-level cognition) has been considered better handled by traditional symbolic artificial intelligence (AI) techniques. Briefly, here are the major assumptions made by most of the traditional models of analogy-making:

1. *Analogy-making begins with two structures.*

2. *Analogy-making is a search through the structures in an attempt to find analogous parts.*

3. *Syntax alone determines the similarity between any two objects, attributes, or relations.* For example, `LARGER-THAN` and `BIGGER-THAN` are not seen as any more similar than `BLACKER-THAN` and `WHITER-THAN`.

4. *For any two relations to be seen as analogous, they must exactly match in terms of their number of arguments, and types of arguments*. For example, `LARGER-THAN(radius, circle-1, circle-2)` would be seen as having nothing in common with `LARGER-THAN(circle-1, diameter, circle-2, meters)` even though they express similar relations.

5. *Relations, attributes, and objects are distinctly different things*. Because of this assumption a relation, such as `CIRCLE-OF(obj1, obj2, obj3, obj4, obj5)`, cannot be seen as analogous to a circle-shaped object.

6. *Context plays no part in making the analogy*. This is a common simplifying assumption.

7. *The result of making an analogy is the creation of a mapping between corresponding pieces of the two structures.*

We will examine these assumptions in detail in Chapter 6. We now turn to examine the assumptions made in the Analogator project.

1.2 Analogator's Assumptions

My goal was to construct a computational model with few assumptions so that analogy-making could "fall out" as naturally as possible. Although the Analogator model is not an implementation of assumed analogical mechanisms *per se*, it is not without assumptions. Analogator's assumptions are few and quite general. They are:

1. *Analogy-making should be the natural by-product of perceiving the similarity between two things.*

Some researchers have postulated that analogy-making is a special mode that we enter into when our reasoning system hits an impasse (see, for instance, Burstein, 1988). Yet, analogy-making has been shown to occur spontaneously (French, 1992). In addition, it has also been argued that analogy-making is a perceptual process (Mitchell, 1993). Therefore, it is a simpler explanation to posit analogy-making as a process of perception that is constantly in operation rather than appealing to more complex theories.[1] This assumption places analogy-making deep into our cognition, down at the core of our intelligence.

2. *Analogy-making should be linked to lower-level processes*.

Mitchell (1993) has shown that analogy-making is related to categorization and recognition. Likewise, I believe that a computational model of analogy-making should be connected to those same lower-level processes. A unification of low-level and high-level processes is a highly desirable trait of any cognitive model.

---

[1] Hofstadter (1995) makes a similar point.

3. *Perception should be seen as naturally breaking things into the* figure *and the* ground.

Things that we perceive, such as scenes, sentences, and sounds, fall into two pieces: the “figure” and the “ground” (see early gestalt psychologists Rubin, 1915, and Wertheimer, 1923, or more recent linguists Talmy, 1983, and Langacker, 1987). The figure of an image is defined to be the area that has our focus of attention, and the ground is everything else. Being able to attend to a voice at a noisy cocktail party is a related “figure-ground segregation” ability (at which most humans do quite well). Descriptive sentences are often described in figure-ground terms. In fact, everything that we perceive seems to be subjected to this division (See Figure 1-1).

I believe that seeing the figure and ground of a situation is also related to the notion of “gist extraction” (Hofstadter, 1995). Hofstadter believes that gist extraction should be considered a part of the analogy-making problem description; however, it is rarely included in a model.

4. *Analogy-making should always occur in context*.

All perception occurs in a context and affects the way in which we see things. In the analogical literature, specialized contexts are sometimes referred to as “goals” or “pragmatics” (for example, see Holyoak and Thagard, 1989a).

One may notice that these primary premises do not mention many notions that most researchers would insist be included in a description of a model of analogy-making. For instance: there is no mention of “relations”, “objects” or “attributes”; no mention of “rules of mapping”; no mention of “search”, “matches”, or “slippage.”

The goal can be stated succinctly: create a learning model that can perceive the figure and ground of a novel situation, given the context of another. Due in part to the principles' generality, the resulting model, as we will see, is also quite general, and many analogical problems can be expressed in a manner suitable for testing in Analogator's framework. The following section outlines specific types of analogy-making problems explored in this dissertation.

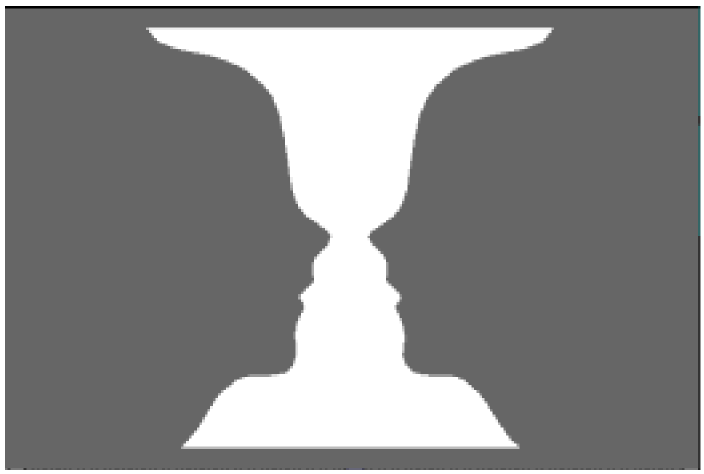

Figure 1-1. A classic example of the battle between figure and ground segmentation. Is this a picture of two faces or a vase?

1.3 Analogator's domains

The idea for Analogator's domain is based on Hofstadter and French's Tabletop project (Hofstadter *et al.*, 1995; French, 1992). To illustrate the basic idea, consider the two scenes in Figure 1-2. Quite simply, the goal is to examine the source scene and the element selected in it (as indicated by the pointing hand), and identify the "same" element in the target scene. The problem, of course, is determining what "same" means in a given situation. In any particular problem, there is no "right" answer. A discussion of possible answers is provided in the next section.

1.3.1 Type #1: Geometric-Spatial Analogies

The scenes of geometric shapes in Figure 1-2 represent the first domain explored. We will restrict problems in the geometric-spatial domain to be analogies composed of squares, circles, and triangles. In these samples, the colors will be restricted to black and white.

At this point, the reader is encouraged to examine the analogical problem of Figure 1-2 and consider the possible issues that come into play. In attempting to make an analogy between the two scenes in Figure 1-2, one might realize that there is not a white triangle in the target scene like that selected in the source scene. Therefore, a completely literal-minded approach will not work in this particular case. Also, there is not an object in the upper-right corner of the target (like that in the source), so considering just *absolute location* will not solve this problem. One possible solution is to realize that the selected object in the source differs from the other two objects in the source on the dimension of shape. Applying the rule "the object that differs from the other objects on the dimension of shape" would lead one to choose the black square in the target scene. Although that is not the only way to conceptualize the problem, it is one way that captures the basic elements of abstract analogy-making. Notice that in order to make the analogy, one must

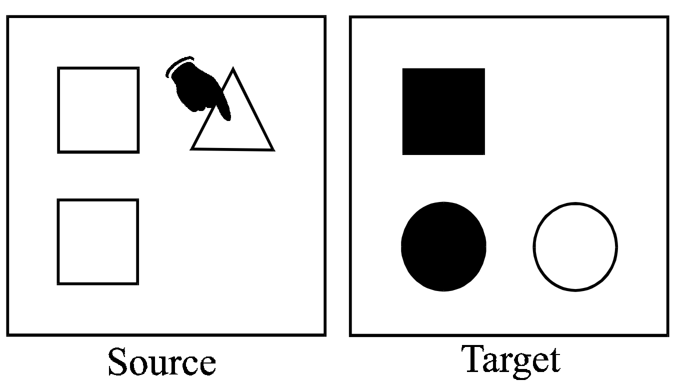

Figure 1-2. Sample #1: A geometric-spatial analogy problem. Which object in the target is the "same" as the object being pointed to in the source? (After French, 1992).

see how the selected object differs from the other objects. “Differs on the dimension of shape” is an abstract relation that is not directly represented by the scene, but is a perceived category created on-the-fly.

To further illustrate, consider Sample #2 (Figure 1-3). Notice the similarity between Sample #1 and Sample #2. In Sample #2, the white triangle in the source scene has been replaced with a black square. All of the other objects’ shapes, colors, and positions have remained the same. In trying to make an analogy with the scenes in Sample #2, one might recognize that there is a black square in the target scene, just like the one being pointed to in the source. If one put oneself in “superficial mode” this would be the preferred object to choose. But if one resists the superficial treatment of the problem, one might notice that the selected object differs from the other objects in the source on the dimension of color. Again, this is not a category represented directly in the source scene, but is created on-the-fly due to the specifics of this particular problem. Following this line of thinking, one might choose the white circle, as that is the object in the target scene that “differs in the same way” (i.e., “the object that differs on the dimension of color.”) After seeing this relationship, the mapping seems most natural to many people, even though black becomes white, and white becomes black.

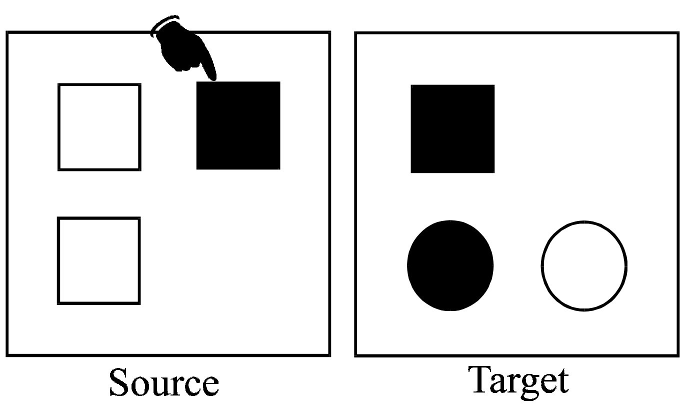

Figure 1-3. Sample #2: A variation of Sample #1. In this version, the black triangle has been changed into a black square. Notice how this change affects the choice of an analogous object.

As demonstrated, by altering an analogy problem just slightly, one might perceive it completely differently. That is, two analogy problems may be superficially very similar, but are perceived very differently. This distinction, related to *categorical perception* (Harnad, 1987), is the flip side of analogy-making (e.g., things may be superficially very different but are perceived the same).

In Samples #1 and #2, we did not consider any object's position in the way we perceived the problem or selected an analogous object. However, one might wish to do so. Consider Sample #3 (Figure 1-4). One is immediately struck by the similarity between the two sets of objects. However, imagine ignoring all object *attributes* (e.g., color and shape) for the moment, and focusing only on an object's position. The selected object is located in the bottom left-hand corner of the source scene. There is an object in the bottom left-hand corner of the target scene, and that could be considered to be the "analogous" object, albeit through a pretty superficial analogy. Thinking more abstractly, one might see the source and the target as mirror images of each other. In that way, the source scene is a backwards L, and the target scene is a normal L (which is a view supported by the objects' attributes). However, one might also see the target scene as being in the same configuration, rotated 90° counterclockwise. Perceiving the problem in this manner, one might choose the black square in the target scene.

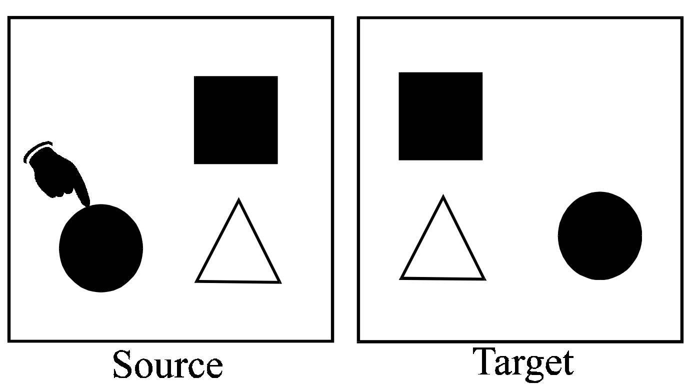

Figure 1-4. Sample #3: Position of objects in a scene may effect perceived similarity.

Finally, one could take into account both position and attributes when making an analogy. Consider Sample #4 (Figure 1-5). In this problem, one might notice that the selected object is one of a pair of same-shape objects. Likewise, the target also has two objects that have the same shape, and a third that does not. However, this does not uniquely identify which is the analogous object in the target scene. If one imagines the target as a mirror image of the source scene and sees the white squares as mapping onto the black circles, then one could pick between the black circles, choosing the one in the bottom right-hand corner.

With geometric-spatial analogy problems, we have seen that this simple domain exhibits many subtleties, requiring concepts such as “mirroring” and “rotation”, dimensions such as “color”, “shape”, and “position”, and more complex, on-the-fly categories such as “the object that differs on the dimension of shape.” We will now examine another spatial domain.

1.3.2 Type #2: Letter-Part Analogies

This domain has also been adapted from Hofstadter and colleagues, specifically from their Letter Spirit domain (McGraw, 1995; Hofstadter and McGraw, 1995). These

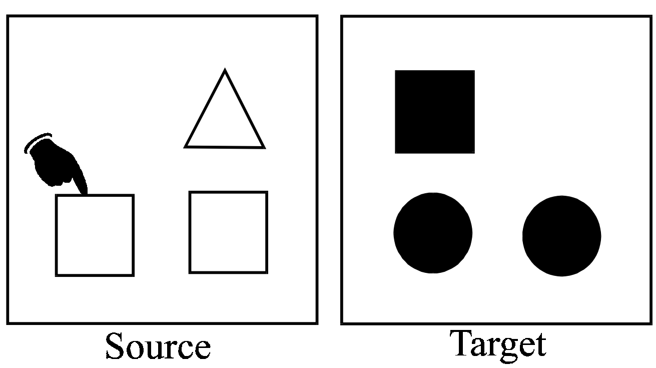

Figure 1-5. Sample #4: Creating on-the-fly categories allows objects to be grouped together in substructures.

problems are treated in a similar manner to those in the geometric-spatial domain. However, rather than being arbitrary objects in a spatial arrangement each source and target scene consists of a letter ‘a’. For example, consider the two a’s in Sample #5 (Figure 1-6.) As before, consider the selected portion as indicated by the pointing hand (the selected portion has also been colored gray). The question is: what is the corresponding part in the target scene?

In this example, one naturally sees the a’s in the two images, and finds the corresponding parts. In this example, most people would probably consider the topmost

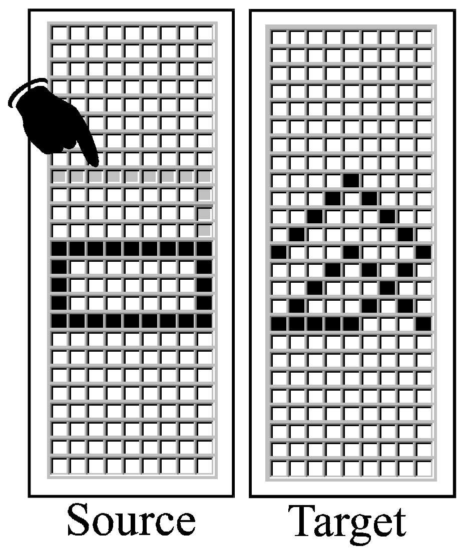

Figure 1-6. Sample #5: A letter-part analogy. Shown are two letter a’s. Considering the selected part in the source scene (pointed to and gray), what is the analogous part in the target scene?

upside-down V in the target letter to be the analogous part.

Consider the similar problem shown in Figure 1-7. The top of the target ‘a’ is quite a bit higher than that of Figure 1-6. Notice that the black squares (or *pixels* as I shall refer to them) that formed the analogous part of the target ‘a’ of Figure 1-6 are also a part of the target ‘a’ of Figure 1-7 (this is illustrated in Figure 1-8). This demonstrates that the same exact pixels may be used for different parts of different a’s. That is, analogous parts

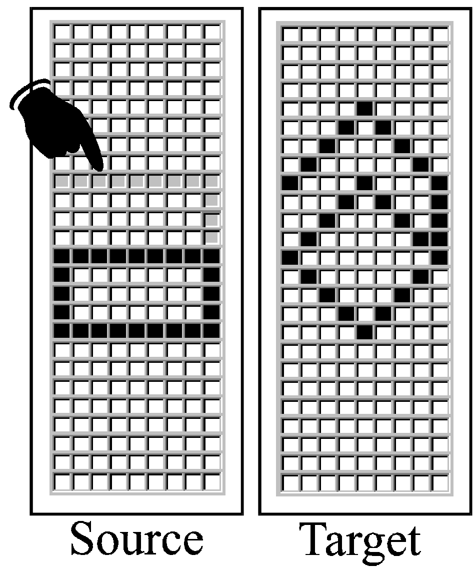

Figure 1-7. Sample #6: A similar letter-part analogy to that of Figure 1-6.

cannot be determined by local properties of a small set of pixels, but involve global relationships of the entire letter.

Like the geometric-spatial analogies, the letter-part analogies involve spatial relations. In this domain, there are just a couple of parts with specific (if not well-defined) relationships, but many instances of a single letterform. We will now examine the final domain.

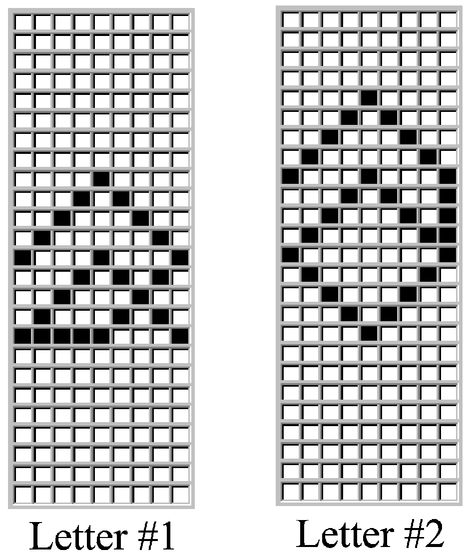

Figure 1-8. Comparison of two a's. Although Letter #1 and Letter #2 have many pixels in common, the parts that those pixels compose are different. A model based purely on "pixel statistics" would have trouble correctly identify the different letter-parts.

1.3.3 Type #3: Family Tree Relationship Analogies

The last domain considered in this dissertation is that of family trees, based on a domain created by Hinton (1986). Consider the two families in Sample #6 (Figure 1-9). The question in this domain is: Who in the target family is the analogous person to the person selected in the source? Although the representation of Figure 1-9 is given in a spatial format, I introduce this domain to test Analogator in purely syntactic domains. Therefore, this domain will be presented in terms of syntactic "facts", such as "Christopher is to Penelope as Roberto is to Maria". Given all of the relations represented schematically in Figure 1-9 as facts, can one still identify the corresponding person? Using the graphical version, it is easy to actually see which person is in the "same" place. However, simply using a set of facts to find the analogous person is a much harder task,

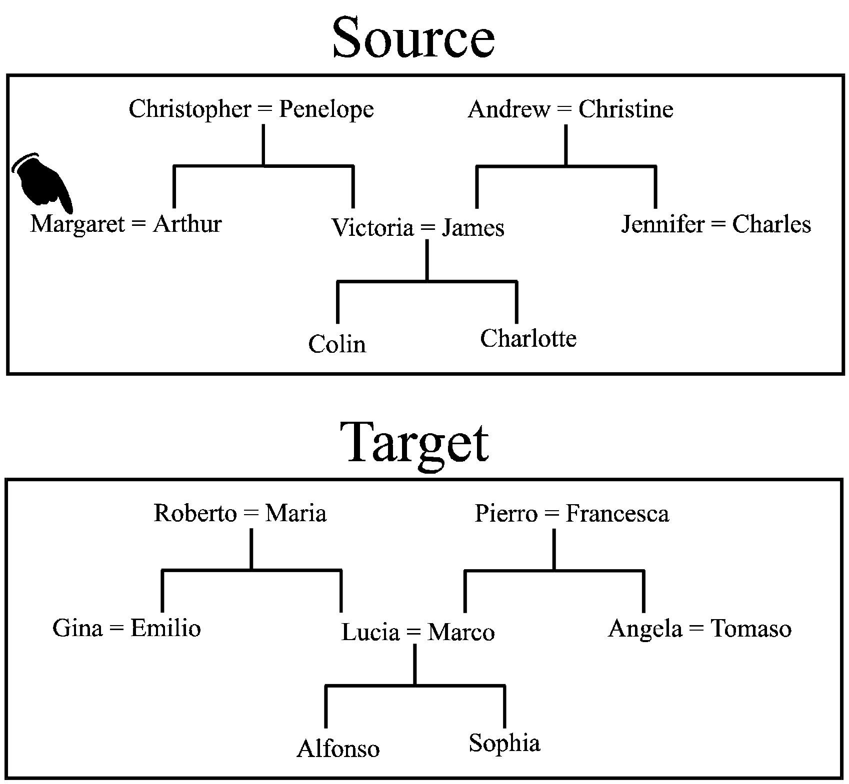

Figure 1-9. Sample #6: two isomorphic families. Who, in the lower family, can be seen as analogous to Margaret? (After Hinton, 1986).

even for humans. However, by aligning the structures from the two families, one can correctly identify corresponding people.

As demonstrated, the Analogator domains contain many of the subtleties of real-world, abstract analogies, yet are distilled down into just a few dimensions. Having examined Analogator's domains, we now turn to a brief overview of the Analogator model.

1.4 The Analogator model

Analogator has been implemented as a connectionist network. A connectionist network is a mathematical model based loosely on neurons. By sending simulated activation between simulated neurons and adjusting connections between them, a set of these neurons can learn to associate a given *input pattern* with a desired *output pattern*. Basically, Analogator takes analogy problems as input, and produces an appropriate answer as output. In addition, by generalizing, it can make appropriate analogies to novel problems.

Analogator learns to make analogies by seeing many example analogies; it is trained by having many instances of analogy problems given as input, and requiring the network to produce the desired output. During training, Analogator is told what the appropriate answer is for a given problem. The output is interpreted to be an indication of the object in the target scene that Analogator "computes" to be the most analogous.

In order to perform the types of analogies depicted in the previous section, a new network training method was created, called the *recurrent figure-ground associating procedure*. Specifically, it is an implementation of the assumptions described in Section 1.1. The recurrent figure-ground associating procedure will be described in detail in Chapter 4.

1.5 Summary

The research described in this dissertation is fundamentally different from other research on analogy-making; it asks very different questions, and explores very different issues. Unlike most computational models of analogy-making, my goal was not to build a model capable of making sophisticated analogies between abstract scenarios. Rather, my goal was to design a system capable of learning to make analogies by beginning with only low-level perceptual information – those data generally restricted for use in “simpler” tasks, such as categorization and recognition. Although most researchers would not deny that abstract analogy-making is related to these lower-lever processes, little has been done to attempt to bridge the gap between them.

The methods used to model these two types of processes – the categorization of low-level sensory data on the one hand, and analogy-making based on abstract, structured representations on the other – have traditionally been very different. Connectionist models have proved superior in low-level tasks (like categorization), while more abstract, cognitive processes (like analogy-making) have been tied to the classical techniques in artificial intelligence. Analogator was designed to explore the gap between these processes by attempting to explain both within a single framework.

1.6 Overview

Chapter 2 examines analogy-making, learning, and generalization. Chapter 3 provides a gentle introduction to the connectionist mechanisms used throughout this dissertation. Chapter 4 discusses the Analogator representation, network architecture and training procedure in depth. Chapter 5 describes a series of analogy-making experiments performed with the Analogator system. Chapter 6 compares and contrasts other computer models of analogy-making, and Chapter 7 summarizes and concludes.