Computational Mental Imagery, and Visual Mechanisms for Maintaining a Goal-Subgoal Hierarchy

Maithilee Kunda MKUNDA@GATECH.EDU
School of Interactive Computing, Georgia Institute of Technology, Atlanta, GA 30308 USA

Abstract

Mental imagery is one important and powerful mechanism of human cognition that has been little explored in AI research. I present a computational framework for mental imagery which specifies the representational and inferential primitives needed to form an imagery-based reasoning system. I then present a computational model of mental imagery that uses a novel recursive memory structure that I call Recursive Visual Memory (RVM), which provides a purely visual mechanism for maintaining a nested sequence of reasoning goals and subgoals. I tested this model on the same set of reasoning problems attempted by Evans' classic ANALOGY program and show that, even without any prior geometric concepts such as shapes or lines, the use of RVM enables successful problem solving. I discuss several disciplines of AI that are commonly conflated with mental imagery and show why mental imagery is distinct. I close by considering broader impacts of this work, which include the construction of creative AI systems as well as improved understanding of individual differences in human cognition.

1. Introduction

While the existence of visual mental imagery was vigorously debated for much of the late 20th century (Kosslyn, Thompson, & Ganis, 2006; Pylyshyn, 2002), findings from neuroscience support the idea that mental imagery is a genuine form of mental representation in humans. Visual mental representations can be described as being analogical or iconic in nature, in that symbols in these representations bear some structured relationship to their referents (Nersessian, 2008).

In neurobiological terms, these kinds of analogical visual mental representations are those that are instantiated in brain regions containing retinotopically mapped neurons, which are found in the visual cortex as well as in other parietal, temporal, and even frontal cortical areas (Silver & Kastner, 2009). Visual mental imagery can be thought of as one particular type of visual mental representation that involves top-down neural activations in visual brain regions in the *absence* of corresponding external visual stimuli (Kosslyn, Thompson, Kim, & Alpert, 1995; Slotnick, Thompson, & Kosslyn, 2005); this activation in the absence of perception nevertheless resembles activation during initial perceptual processing of the stimuli under consideration (O'Craven & Kanwisher, 2000). Furthermore, this type of neural activity plays a functional role in performing certain cognitive tasks; if the activity is artificially suppressed, task performance decreases (Kosslyn et al., 1999).

Verbal representations, on the other hand, based on language or language-like elements, can be described as being propositional in nature, as symbols in these representations can stand for anything. In other words, mappings between symbols and referents are arbitrary. Propositional representations are extremely powerful in the kinds of inference they can support, e.g. the *physical symbol system hypothesis* of Newell and Simon proposed that such representations, realized in a physical system, are both necessary and sufficient for achieving general intelligence (Newell & Simon, 1976).

In contrast to this view of intelligence as being propositionally driven, many narrative accounts about human cognition, found both in informal anecdotes as well as in detailed epistemological analyses, maintain that visual representations can contribute directly to high-level reasoning through the use of visual mental imagery. These accounts emphasize the functional role played by visual mental imagery in complex tasks such as creativity, analogy, conceptual inference, and mental modeling, across numerous domains including engineering design (Ferguson, 1994), mathematics (Giaquinto, 2007), and scientific discovery (Nersessian, 2008). In addition, many individuals diagnosed with certain neurodevelopmental conditions such as autism (Grandin, 2006; Hurlburt, Happé, & Frith, 1994) and dyslexia (West, 1997) have introspectively observed that visual mental imagery seems to be atypically prominent in their cognitive processing (Kunda & Goel, 2011).

Many AI systems to date rely on propositional representations for high-level reasoning. Even for tasks that may seem suited to visual representations, such as diagram understanding or visuospatial reasoning, AI systems often convert visual inputs into propositional representations before the reasoning steps take place, e.g. (Chandrasekaran, Banerjee, Kurup, & Lele, 2011; Davies, Goel, & Nersessian, 2009; Larkin & Simon, 1987). This is perhaps partially due to this lack of specificity in the cognitive science of mental imagery, and perhaps also partially due to the power and ubiquity of symbolic computation on modern digital machines. There have been, of course, some examples in the AI literature that do address visual-imagery-based problem solving, such as:

- Hunt (1974) proposed a theoretical account of visual problem solving on matrix reasoning problems in which he argued that symbolic and imagistic approaches must both be explored as complementary aspects of dual-process accounts of such problem solving; however, no working AI system was built as part of this inquiry.
- Funt (1980) presents a computational system that reasons visually about problems in a blocks-world task domain, using imagery-like operations such as translation and rotation.
- Glasgow and Papadias (1992) present a computational, though not ostensibly cognitive, model of imagery-based reasoning that uses 3D occupancy grid arrays as a form of imagistic representation; their system supports operations such as rotation, translation, and zooming, as well as shifting of attention from one part of the representation to another.
- Schwartz and Black (1996) constructed a computational model that simulates reasoning over gear-rotation problems using different types of mental representations, including performing mental rotations of imagistic representations.

- Tabachneck-Schijf and colleagues (1997) present a computational model that combines visual and verbal reasoning pathways for diagrammatic reasoning in the context of graph understanding.
- Other examples have come from domains including map understanding (Schlieder & Berendt, 1998), reasoning about motion (Croft & Thagard, 2002), reasoning about conversational interactions over a set of objects (Roy, Hsiao, & Mavridis, 2004), spatial reasoning (Ragni, Knauff, & Nebel, 2005), robot navigation (Stein, 1991), and imagery-based reasoning within a cognitive architecture (Lathrop & Laird, 2007).

While most of these models have adopted a symbolic view of cognitive processing, Mel (1986) proposed a connectionist view of imagery operations.

These efforts present models of computational mental imagery in various task domains, with varying attention paid to the plausibility of these models in terms of mental imagery in human cognition. In previous work, colleagues and I developed a model of computational mental imagery (Kunda, McGreggor, & Goel, 2013) intended to get at basic human problem solving processes by looking at the task domain of Raven's Progressive Matrices, a standardized test that represents the best single psychometric measure of general intelligence that has yet been developed (Raven, Raven, & Court, 2003; Snow, Kyllonen, & Marshalek, 1984).

The purpose of this paper is 1) to extend this previous work by spelling out a computational framework for mental imagery that ties together many of the operations used in previous AI models, and that have been observed in studies of human cognition, and 2) to propose a new mechanism, recursive visual memory (RVM), for imagery-based image segmentation that also serves as a way to define and pursue goals and subgoals related to visual problem solving. The domain is that of geometric analogy—a fitting place to start to follow in the footsteps of one of AI's pioneers, Evans (1968).

First, I describe a framework for computational mental imagery, including a description of reasoning "primitives" that should be supported by mental imagery reasoning systems. Next, I present a description of recursive visual memory (RVM) and how it expands the capabilities of a previously developed computational model of mental imagery (Kunda et al., 2013). I close with a discussion of how mental imagery is (and isn't) related to other areas of investigation within AI, and finally with some remarks about the broader implications of this work.

2. Computational Mental Imagery

What is mental imagery? At its essence, "imagery" refers to the idea of representations as analogical—in correspondence with the represented world—as opposed to propositional—of arbitrary relationship to the represented world (Nersessian, 2008). Propositional representations are processed in purely syntactic ways, whereas images can be processed according to the rules that govern their structure. For example, a picture of a cat, since it obeys structural rules having to do with spatial relationships among its components, can be meaningfully manipulated in spatial ways, e.g. by stretching it, turning it upside down, etc. While imagery can technically take place in any modality, in this paper the focus is specifically on visual imagery.

One view of human mental imagery is that it is a form of analogical mental model in which the referents of imagery are modeled after physical objects in the world, and thus the operations of imagery are grounded in the affordances of manipulating (or observing manipulations of) such objects (Schwartz & Black, 1996); this view is consistent with ecological views of the development of cognitive capabilities (Shepard, 1984). The manipulations of a 3D rigid body (i.e. objects observed in the world) projected into a 2D plane (i.e. projected through human visual perception) correspond to imagery transformations such as translation, rotation, scaling, etc., and the combination of projections of multiple such bodies can be represented using set operations on points or other elements in the 2D image plane. Both of these classes of mental imagery operations have been observed in human studies.

2.1 Operations of Human Mental Imagery

The most well-known early experiments looking at operational properties of mental imagery are the mental rotation experiments of the early 1970s (Cooper & Shepard, 1973; Shepard & Metzler, 1971). Many variations on these experiments have been performed, with a recent plethora of studies that include neuroimaging measures (Zacks, 2008), but the basic findings have been that the reaction time taken by participants to mentally rotate an object is proportional to the angular degree through which the object must be rotated. This finding is replicated in situations when two stimuli are both visible for comparison (Shepard & Metzler, 1971) as well as when a single visible stimulus is compared to another representation from memory (Cooper & Shepard, 1973). There has also been evidence for scaling/zooming (Larsen & Bundesen, 1978) as well as translation or scanning (Larsen & Bundesen, 1998).

Set operations in mental imagery can include manipulations such as union, intersection, and complement. These manipulations correspond generally to the combination of elements of mental images in various ways, including: subtraction, in which participants were asked to subtract a visually presented shape from a remembered image in order to derive and identify a new mental image (Brandimonte, Hitch, & Bishop, 1992b), and combination, in which participants were asked to perform a similar identification after combining a new image element with a remembered image (Brandimonte, Hitch, & Bishop, 1992a; Finke, Pinker, & Farah, 1989). Additional studies have found evidence, in some cases, of verbal coding interfering with this type of image transformation (Brandimonte et al., 1992b; Brandimonte, Hitch, & Bishop, 1992c) and in other cases, of an interaction between seemingly pictorial depictive representations and non-pictorial descriptive representations (Hitch, Brandimonte, & Walker, 1995).

In a paper looking at mental rotation in autism, another experiment was conducted to examine image combination (Soulières, Zeffiro, Girard, & Mottron, 2011). Participants first inspected and memorized an array of visually presented letters and numbers, and then were briefly shown a circular segment with a portion of one character inside it. Then, upon looking at the segmented circle alone, the task was to determine which segment would contain a greater visual proportion of the original character. This experiment can be thought of as requiring an operation akin to intersection, in visualizing which portion of the character falls into each segment of the circle, as if the two were overlaid, and also a comparison of visual similarity in terms of which character portion embodies a greater visual area.

2.2 Computational Primitives of Mental Imagery

Thus, we define the following as computational primitives for visual mental imagery. Notice that while this forms a coherent mathematical framework, there is no one "correct" formulation of mental imagery. The unified mathematical framework illustrates that, at least in computational terms, the following are equivalent in some sense. For the purpose of building AI systems, it becomes a design choice.

1. Set **S** of visual elements with relations that are isomorphic to those of the 2D plane. This excludes diagrammatic representations in which entities have verbal labels, as in these representations, the "elements" are not visual. This also excludes propositional representations, which do not contain an inherent system of relations among knowledge elements. There are many forms that these visual elements could take, including:
    a. Points. Pixels are one example, but note that this scheme does not need to be restricted to rectilinear arrangements of points. Consider the cells of the visual cortex as an alternative example.
    b. Corners, edges, lines
    c. Shapes
2. A functionally complete collection of combination operations over these elements, such as {intersection, complement}.
3. Geometric operations over connected subsets of these elements. We define these in various classes corresponding to the classes of physical manipulations that they correspond to:
    a. Translations
    b. Similitude transformations
    c. Affine transformations
    d. Shape deformations

One additional class of possible operations is the set of colorimetric transformations over visual elements. This property could certainly be included in a system of mental imagery, but is not included in this framework for the following reason: a system of mental imagery can be complete without color transformations, but a system of mental imagery having only color transformations is not. In other words, the critical property of mental imagery is having systematic transformations that preserve spatial relationships among related subsets of planar elements.

3. Recursive Visual Memory (RVM)

In this paper, I build on a previously developed computational model of mental imagery called the Affine and Set Transformation Induction (ASTI) model. I present the model as used for the domain of geometric analogies, in the form A : B :: C : ?, like Evans' early AI work (Evans, 1968).

While this model was originally developed in prior work in a different problem domain, for matrix reasoning problems (Kunda et al., 2013), I present here a new version of the model that includes a new capability to segment images in a strategic, meaningful way, using purely visual representations and transformations to drive the segmentation mechanism. The model maintains this decomposition in an imagery hierarchy, akin to having a goal-subgoal hierarchy that can be expanded or contracted by adjusting working memory constraints. This type of goal hierarchy has been explored in a previous model of solving matrix reasoning problems, and in fact was found to be one critical variable contributing to successful problem solving performance on difficult problems, but this model used hand-coded propositional representations of the various shapes and features in each problem (Carpenter, Just, & Shell, 1990).

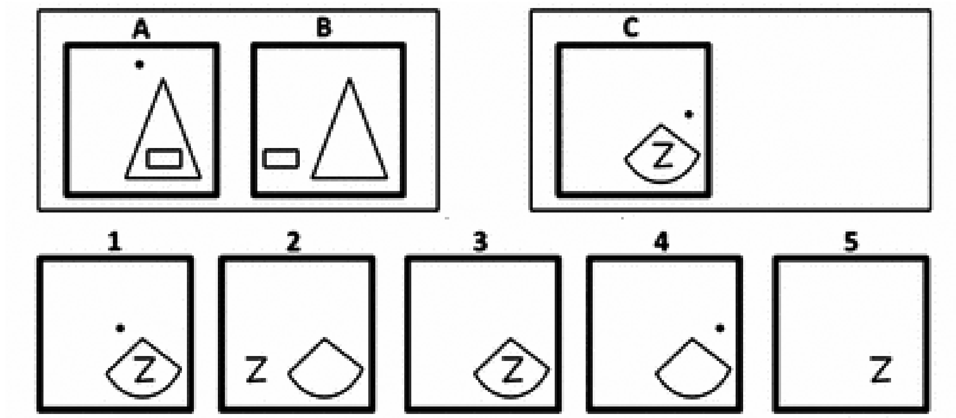

Figure 1. Example geometric analogy problem; image obtained from (Lovett, Tomai, Forbus, & Usher, 2009, p. 1225).

The central representation used by ASTI is at the point or pixel level, defined as binary features (i.e. black or white pixels). The primitive imagery operations supported by the model include:

1. Translation
2. Discrete rectilinear reflections and rotations
3. Set operations of union, intersection, and subtraction, defined over binary pixels as OR, AND, and NAND, respectively. Table 1 gives the truth table for these basic pixel operations.

Table 1. Truth table for set operations over binary (black/white) pixels.

p	q	$p \cup q$	$p \cap q$	$p - q$	$q - p$
0	**0**	0	0	0	0
0	**1**	1	0	0	1
1	**0**	1	0	1	0
1	**1**	1	1	0	0

There is one higher level operation also: calculating visual similarity. The ASTI model's conceptualization of visual similarity follows template-based similarity matching in which two images are compared according to the amount of visual overlap between them, as one is moved around to various locations relative to the other. At each relative offset, the similarity is computed as the Jaccard coefficient. Thus, for any two images $\boldsymbol{U}$ and $\boldsymbol{V}$, and for all possible offsets $\boldsymbol{(x, y)}$ between the two images, similarity is calculated as:

$$\mathrm{sim}(\boldsymbol{U},\boldsymbol{V}) = \max_{x,y}\left(\frac{\sum_{i,j}\left(\boldsymbol{U}_{i+x,j+y} \cap \boldsymbol{V}_{i,j}\right)}{\sum_{i,j}\left(\boldsymbol{U}_{i+x,j+y} \cup \boldsymbol{V}_{i,j}\right)}\right)$$

One clear simplification in this similarity mechanism is that the model performs exhaustive search across all possible offsets $\boldsymbol{(x, y)}$ between the two images. Humans certainly do not do this, nor would this be feasible for efficient artificial agents. While the heuristics used to guide this search are undoubtedly an interesting aspect of the overall cognitive system, the focus of my model is on the content of reasoning and not the temporal processing that unfolds. Thus, these questions of heuristics, efficiency, and reaction time are outside the scope of this paper.

The model takes as input eight individual images: the three images that constitute the analogy and the five answer choices, as shown in Fig. 1. The model then computes a cascading set of comparisons between six pairs of images: the first pair is A:B, and the rest are between image C and each answer choice: C:N1, C:N2, C:N3, C:N4, and C:N5.

The basic comparison mechanism in ASTI is to take two images and exhaustively test a set of transforms contained in memory to see which transform best accounts for the differences between the two images. The transforms contained in ASTI are: 1) identity, 2) rotate90, 3) rotate180, 4) rotate270, 5) flip, 6) rotate90flip, 7) rotate180flip, 8) rotate270flip, 9) add, and 10) subtract.

Testing transforms #1 through #8 is straightforward. For any pair of images X and Y, first X is manipulated according to the given transform t to produce t(X), and then the similarity metric given above is used to compare t(X) to Y.

For transforms #9 and #10, the process is slightly more complicated. For transform #9, addition, what we are trying to detect is whether Y represents a situation in which something has been added to X, and the similarity value generated from the comparison should reflect whether Y represents a "pure" addition relative to X. To obtain this estimate, we first define a dummy image Z = Y – X. Then, to obtain the desired comparison, we compare X – Z and Y – Z. If Y is a strict superset of X, then this comparison yield the maximum similarity value of 1.0. If X is a strict superset of Y, then this comparison yields the minimum similarity value of 0.0. The subtraction transform is defined is a straightforward extension of this approach.

Using this basic comparison mechanism, ASTI creates a cascading set of comparisons in order to solve each analogy problem. Given a pair of images X and Y, ASTI first computes a basic comparison as described above. Then, ASTI removes all the pixels from each image that have been completely explained by the best-fit transformation between these images. The result is two new images, which can then be compared in the same fashion. This process can continue until all pixels have been explained. This recursive process is what I call "recursive visual memory" (RVM).

Operationally, the process halts when one or the other of the images is completely blank, at which point the very next best-fit transform will be a pure addition or subtraction. Alternately, this can occur when the two remaining images are perfectly identical, after which both images will be completely blank. (In practice, a threshold is imposed on the number of possible recursions, as little improvement happens after a handful of recursions in this problem domain.) Fig. 2 gives detailed pseudocode for this entire process of solving a geometric analogy problem.

```
SolveAnalogy
Input: images A, B, C, N_1 through N_5
Output: number of answer choice (1 through 5)
    T = {identity, all rotations and reflections, AddTransform, SubtractTransform};
    data_AB = RecursiveCompare(A,B,T);
    for i=1:5
       data_i = RecursiveCompare(C,N_i,T);
       score_i = number of best transforms shared between data_AB and data_i;
    return i with max score_i over all i;

RecursiveCompare
Input: images X and Y, set of transforms T
Output: data_XY
    data_XY = m x n array, where m is maxRecursions and n is number of transforms in
    T;
    i = 1;
    while i < maxRecursions
       for each t in T
           sim_t = Compare(t(X),Y);
           data_XY[i][t] = sim_t;
       t_max = t with max sim_t over all t in T;
       newX = t_max(X) - Y;
       newY = Y - t_max(X);
       X = newX, Y = newY;
       if X or Y are blank, then halt loop;
       else i++;
    return data_XY;

Compare
Input: images X and Y
Output: similarity
    for each possible overlay (i,j) of X and Y
       similarity_ij = X∩Y_ij / X∪Y_ij;
    return max similarity over all (i,j);

AddTransform
Input: images X and Y
Output:  X_2 and Y_2 for comparison
    Compare(X,Y);
    (i,j) = position of maximum similarity;
    Z = Y_ij - X;
    X_2 = X - Z;
    Y_2 = Y - Z;

SubtractTransform
Input: images X and Y
Output:  X_2 and Y_2 for comparison
    Compare(X,Y);
    (i,j) = position of maximum similarity;
    Z = X - Y_ij;
    X_2 = X - Z;
    Y_2 = Y - Z;
```

Figure 2. Pseudocode for analogy problem solving in ASTI model.

Fig. 3 illustrates the contents of the RVM for the example problem shown in Fig. 1 above. Note that this figure only shows the RVM for the initial A:B image pair; similar cascades of images are computed for image C and each of the five answer choices.

Once this cascade has been computed, the A:B cascade is compared for similarity to the C:answer cascade for each possible answer choice. Similarity is computed as a weighted combination of the type of transform used to maximize similarity at each step in the cascade and the actual similarity value computed at each step.

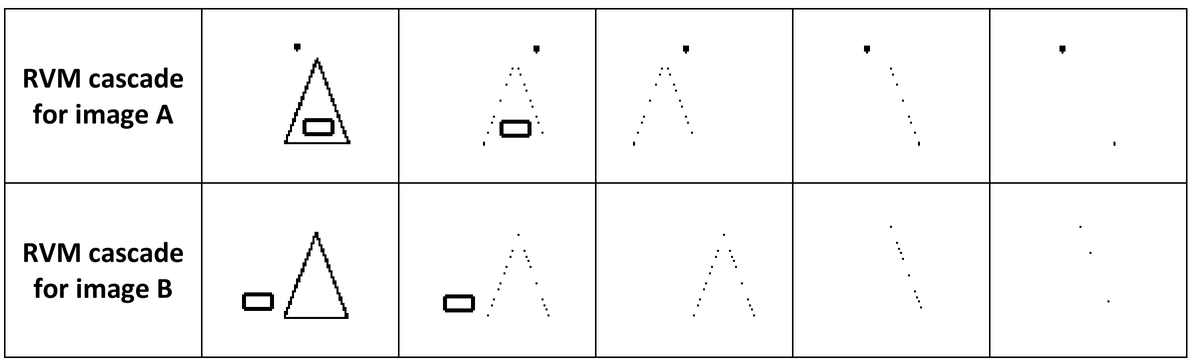

Figure 3. Illustration of contents of RVM for images A and B for the example problem in Fig. 1.

4. Experimental Results

We tested the ASTI model against the twenty cases presented for Evans' ANALOGY program (Evans, 1968). We obtained the actual problem images from Lovett et al. (2009).

Fig. 4 shows the number of problems solved correctly for two different levels of "working memory" in the model, defined as the number of RVM recursions that are allowed. As this figure illustrates, as the size of working memory increases from 1 (no image segmentation) to 5, the model is able to solve double the number of problems originally solved. Note that the current ASTI implementation is entirely deterministic, which is why the results show no variability.

The overall performance of the ASTI model reaches only about 50% on these visual analogy problems. Evans' model solved 18 of the 20 problems, and more recent attempts have produced computational models that have solved all 20 of the problems (Lovett et al., 2009). However, these previous models have relied on human hand-coding of problem inputs for segmentation of each problem into distinct elements. The ASTI model performs this segmentation automatically, using the RVM procedure as a mechanism for both segmenting and storing pieces of the problem, and then reasoning over them in a sequential fashion.

In particular, the ASTI model does not use any preliminary notions of shape, continuity, closure, etc. in performing its RVM-based segmentation. It looks purely at visual similarity and a set of mental-imagery-based visual transformations to obtain this segmentation. Thus, the fact that the ASTI model can solve 10 out of 20 problems is not intended as a point of direct competition with other models, but rather as a statement about the sufficiency of the set of visual mechanisms embodied by the ASTI model in solving these 10 problems.

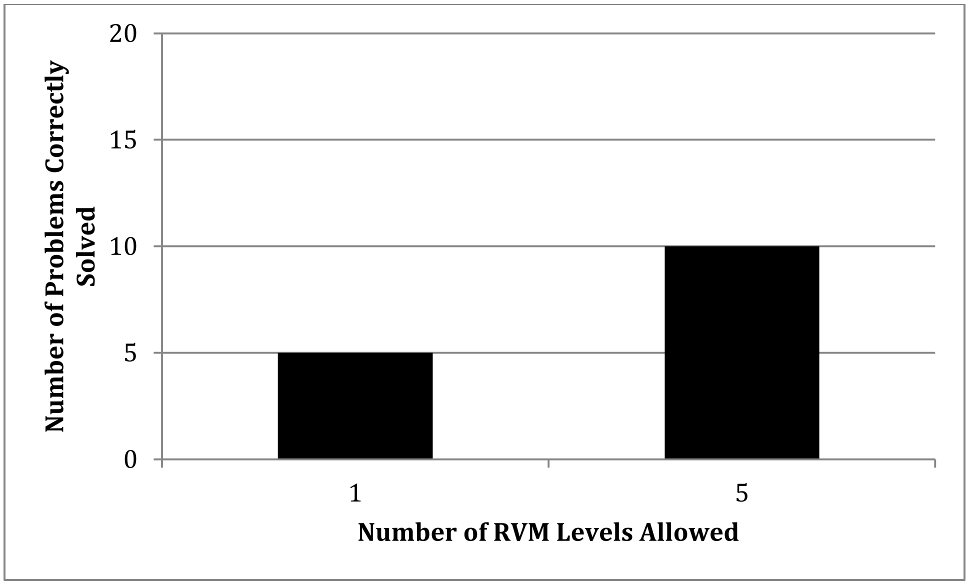

Figure 4. Number of analogy problems correctly solved by ASTI as a function of the size of ASTI's recursive visual memory (RVM) store.

5. Discussion

The main contributions of this work are twofold. First, the results of the ASTI model on Miller's test of geometric analogies closes the loop on a long-standing classic in AI, Evans' 1968 ANALOGY program, and illustrate how computational mental imagery can solve problems using an imagery hierarchy to segment the problem inputs into smaller subproblems. This is a novel capability for computational imagery systems, and an important step towards solving increasingly more complex sorts of problems.

Second, the framework for organizing primitives of mental imagery is a valuable contribution for thinking about how to build mental imagery in AI systems, and also how to relate computational imagery to imagery in human or nonhuman animals. There is often much confusion related to mental imagery, in both the cognitive science as well as the AI literature. In the spirit of the classic AI paper that examined relationships between representational commitments and intelligent capabilities (Brooks, 1991), we use the remainder of the discussion to elaborate on what differentiates mental imagery from other paradigms of cognitive processing—in other words, what computational mental imagery is not. (Of course, most complex intelligent processes probably involve many if not all of these kinds of capabilities; we present this list mainly to clarify the difference in focus between our work and other bodies of strongly-related yet differently-focused AI research.)

5.1 It isn't computer vision

Computer vision uses pixel based representations that are spatially organized and a plethora of spatially grounded operations on these pixels, including the operations discussed in our framework, and many, many more. However, the goals of computer vision are different. Computer vision is fundamentally about image understanding and is much more tied to questions of perception and perceptual inference, and in particular mapping visual inputs onto propositional outputs (such as category labels).

5.2 It isn't computer graphics

Computer graphics is closer; it involves the deliberate creation and manipulation of visual representations to create new ones. The biggest difference there is that in computer graphics, the intended user of the graphics is a human, and so the process of manipulating images is separated from the process of reasoning using those images. Mental imagery is a unified process with image manipulation and reasoning happening in concert. Furthermore, imagery requires that the manipulations are themselves instantiated visually, which is not a requirement for computer graphics.

5.3 It isn't gestalt perception

Gestalt perception involves the application of top-down heuristics, often defined by intuitive notions of how 3D bodies appear in the world and the physical constraints on them that we observe. However, the study of gestalt information processing has fundamentally been about automatic top-down effects on perception, and not about the deliberate top-down manipulation of mental representations. Certainly gestalt perception plays a role in the initial perceptual creation of representations that feed our imagery banks, and likely the operations performed within mental imagery as well, but the two are distinct processes.

5.4 It isn't qualitative reasoning

Qualitative reasoning often uses visual and spatial relationships. While these are certainly important, the point here is that mental imagery is quantitative, and these quantitative representations can support very powerful forms of inference. When (and how) quantitative representations get "thrown over the wall" to form qualitative representations, and vice versa, are important open questions for AI and cognitive research. Any *complete* account of visual reasoning in human-like intelligent systems will need to include both.

6. Conclusion

We envision that continued research on computational mental imagery will lead to new paradigms in AI and will also help unlock new avenues of inquiry into human cognition. While mental imagery plays a role in a vastly diverse range of intelligent capabilities, we pinpoint two areas—1) mental disorders and 2) creativity—that are currently of high interest to AI and cognitive science research communities, and for which this type of research will have far-reaching impacts and real-world extensions.

Recent studies of autism have raised the possibility that for some individuals on the autism spectrum, mental imagery actually dominates cognitive processing, perhaps as a compensatory mechanism for early impairments in the neurodevelopmental building blocks of language systems (Just, Keller, Malave, Kana, & Varma, 2012; Kana, Keller, Cherkassky, Minshew, & Just, 2006; Kunda & Goel, 2011). Temple Grandin is an individual who describes how she “thinks in pictures,” and how this gives her strengths in some areas but difficulties in others (Grandin, 2006). Building computational models of mental imagery not only clarifies the information processing properties of these tasks but can also lend insight human behavior to aid in the diagnosis, assessment, and understanding of cognitive conditions like autism, an approach now being dubbed “computational psychiatry” (Montague, Dolan, Friston, & Dayan, 2012).

Regarding creativity, there is much recent interest in developing truly creative AI systems, from art to science to design. Theories of human creativity often include mental imagery as a critical component, and many case studies of specific creative actors or acts have involved or even revolved entirely around mental imagery (Ferguson, 1994; Giaquinto, 2007). Spelling out a theory of computational imagery is the first step towards incorporating it into more and more high-level cognitive tasks, until we get to those that can truly be considered acts of creativity.

Acknowledgments

I would like to thank A. Goel and K. McGreggor for critical discussions of this work, as well as the reviewers for their many helpful suggestions.

References

Brandimonte, M. A., Hitch, G. J., & Bishop, D. V. (1992a). Influence of short-term memory codes on visual image processing: evidence from image transformation tasks. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *18*(1), 157–165.

Brandimonte, M. A., Hitch, G. J., & Bishop, D. V. (1992b). Manipulation of visual mental images in children and adults. *Journal of Experimental Child Psychology*, *53*(3), 300–312.

Brandimonte, M. A., Hitch, G. J., & Bishop, D. V. (1992c). Verbal recoding of visual stimuli impairs mental image transformations. *Memory & Cognition*, *20*(449.).

Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, *47*(1), 139–159.

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: a theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, *97*(3), 404.

Chandrasekaran, B., Banerjee, B., Kurup, U., & Lele, O. (2011). Augmenting Cognitive Architectures to Support Diagrammatic Imagination. *Topics in Cognitive Science*, *3*(4), 760–777.

Cooper, L. A., & Shepard, R. N. (1973). Chronometric studies of the rotation of mental images. In W. G. Chase (Ed.), *Visual Information Processing* (pp. 75–176). New York: Academic Press.

Croft, D., & Thagard. (2002). *Dynamic imagery: A computational model of motion and visual analogy*. (L. Magnani & N. Nersessian, Eds.). New York: Kluwer Academic/Plenum Publishers.

Davies, J., Goel, A. K., & Nersessian, N. J. (2009). A computational model of visual analogies in design. *Cognitive Systems Research*, *10*(3), 204–215.

Evans, T. (1968). A heuristic program to solve geometric analogy problems. In M. Minsky (Ed.), *Semantic information processing*. Cambridge, MA: MIT Press.

Ferguson, E. S. (1994). *Engineering and the Mind's Eye*. Cambridge, Mass.: The MIT Press.

Finke, R. A., Pinker, S., & Farah, M. J. (1989). Reinterpreting Visual Patterns in Mental Imagery. *Cognitive Science*, *13*(1), 51–78.

Funt, B. V. (1980). Problem-solving with diagrammatic representations. *Artificial Intelligence*, *13*(3), 201–230.

Giaquinto, M. (2007). *Visual Thinking in Mathematics: An Epistemological Study*. Oxford University Press.

Glasgow, J., & Papadias, D. (1992). Computational Imagery. *Cognitive Science*, *16*(3), 355–394.

Grandin, T. (2006). *Thinking in pictures*. New York, NY: Vintage Books.

Hitch, G. J., Brandimonte, M. A., & Walker, P. (1995). Two types of representation in visual memory: Evidence from the effects of stimulus contrast on image combination. *Memory & Cognition*, *23*(2), 147–154.

Hunt, E. (1974). Quote the Raven? Nevermore. In *Knowledge and cognition* (pp. ix, 321). Oxford, England: Lawrence Erlbaum.

Hurlburt, R. T., Happé, F., & Frith, U. (1994). Sampling the form of inner experience in three adults with Asperger syndrome. *Psychological Medicine*, *24*(2), 385–396.

Just, M. A., Keller, T. A., Malave, V. L., Kana, R. K., & Varma, S. (2012). Autism as a neural systems disorder: A theory of frontal-posterior underconnectivity. *Neuroscience & Biobehavioral Reviews*, *36*(4), 1292–1313.

Kana, R., Keller, T., Cherkassky, V., Minshew, N., & Just, M. (2006). Sentence comprehension in autism: Thinking in pictures with decreased functional connectivity. *Brain*, *129*, 2484–2493.

Kosslyn, S. M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J. P., L, W., … Alpert. (1999). The Role of Area 17 in Visual Imagery: Convergent Evidence from PET and rTMS. *Science*, *284*(5411), 167–170.

Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). *The case for mental imagery*. Oxford University Press. Retrieved from http://psycnet.apa.org/psycinfo/2006-10314-000

Kosslyn, S. M., Thompson, W. L., Kim, I. J., & Alpert, N. M. (1995). Topographical representations of mental images in primary visual cortex. *Nature*, *378*(6556), 496–498.

Kunda, M., & Goel, A. K. (2011). Thinking in pictures as a cognitive account of autism. *Journal of Autism and Developmental Disorders*, *41*(9), 1157–1177.

Kunda, M., McGreggor, K., & Goel, A. K. (2013). A computational model for solving problems from the Raven's Progressive Matrices intelligence test using iconic visual representations. *Cognitive Systems Research*, *22–23*, 47–66.

Larkin, J. H., & Simon, H. A. (1987). Why a Diagram is (Sometimes) Worth Ten Thousand Words. *Cognitive Science*, *11*(1), 65–100.

Larsen, A., & Bundesen, C. (1978). Size scaling in visual pattern recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(1), 1–20.

Larsen, A., & Bundesen, C. (1998). Effects of spatial separation in visual pattern matching: Evidence on the role of mental translation. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 719–731.

Lathrop, S. D., & Laird, J. E. (2007). Towards incorporating visual imagery into a cognitive architecture. In *Proceedings of the eighth international conference on cognitive modeling* (p. 25).

Lovett, A., Tomai, E., Forbus, K., & Usher, J. (2009). Solving Geometric Analogy Problems Through Two-Stage Analogical Mapping. *Cognitive Science*, *33*(7), 1192–1231.

Mel, B. W. (1986). A connectionist learning model for 3-d mental rotation, zoom, and pan. In *Proceedings of the Eighth Annual Conference of the Cognitive Science Society* (pp. 562–71).

Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. *Trends in Cognitive Sciences*, *16*(1), 72–80.

Nersessian, N. (2008). *Creating scientific concepts*. MIT Press.

Newell, A., & Simon, H. A. (1976). Computer Science As Empirical Inquiry: Symbols and Search. *Commun. ACM*, *19*(3), 113–126.

O'Craven, K. M., & Kanwisher, N. (2000). Mental Imagery of Faces and Places Activates Corresponding Stimulus-Specific Brain Regions. *Journal of Cognitive Neuroscience*, *12*(6), 1013–1023.

Pylyshyn, Z. W. (2002). Mental imagery: In search of a theory. *Behavioral and Brain Sciences*, *25*(02), 157–182.

Ragni, M., Knauff, M., & Nebel, B. (2005). A computational model for spatial reasoning with mental models. In *Proceedings of the 27th annual cognitive science conference* (pp. 1064–1070).

Raven, J., Raven, J. C., & Court, J. H. (2003). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. Pearson.

Roy, D., Hsiao, K., & Mavridis, N. (2004). Mental imagery for a conversational robot. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, *34*(3), 1374–1383.

Schlieder, C., & Berendt, B. (1998). Mental model construction in spatial reasoning: A comparison of two computational theories. *Mind Modelling: A Cognitive Science Approach to Reasoning, Learning and Discovery*, 133–162.

Schwartz, D. L., & Black, J. B. (1996). Analog Imagery in Mental Model Reasoning: Depictive Models. *Cognitive Psychology*, *30*(2), 154–219.

Shepard, R. N. (1984). Ecological constraints on internal representation: resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review*, *91*(4), 417–447.

Shepard, R. N., & Metzler, J. (1971). Mental Rotation of Three-Dimensional Objects. *Science*, *171*(3972), 701–703.

Silver, M. A., & Kastner, S. (2009). Topographic maps in human frontal and parietal cortex. *Trends in Cognitive Sciences*, *13*(11), 488–495.

Slotnick, S. D., Thompson, W. L., & Kosslyn, S. M. (2005). Visual mental imagery induces retinotopically organized activation of early visual areas. *Cerebral Cortex (New York, N.Y.: 1991)*, *15*(10), 1570–1583.

Snow, R. E., Kyllonen, P. C., & Marshalek, B. (1984). The topography of ability and learning correlations. *Advances in the Psychology of Human Intelligence*, *2*, 47–103.

Soulières, I., Zeffiro, T. A., Girard, M. L., & Mottron, L. (2011). Enhanced mental image mapping in autism. *Neuropsychologia*, *49*(5), 848–857.

Stein, L. A. (1991). Imagination and Situated Cognition. *MIT AI Memo*, *No. 27*.

Tabachneck-Schijf, H. J. M., Leonardo, A. M., & Simon, H. A. (1997). CaMeRa: A computational model of multiple representations. *Cognitive Science*, *21*(3), 305–350.

West, T. G. (1997). *In the Mind's Eye: Visual Thinkers, Gifted People with Learning Difficulties, Computer Imaging, and the Ironies of Creativity*. Prometheus Books.

Zacks, J. M. (2008). Neuroimaging studies of mental rotation: a meta-analysis and review. *Journal of Cognitive Neuroscience*, *20*(1), 1–19.