Diagrammatic Representation and Reasoning: Some Distinctions

B. Chandrasekaran
Department of Computer & Information Science
The Ohio State University
Columbus, OH 43210
Email: chandra@cis.ohio-state.edu

1. Introduction

When I co-edited the book, "Diagrammatic Reasoning," [5] I often puzzled over the fact that rather different sorts of things were dubbed diagrammatic reasoning by different authors. I then spent sometime trying to categorize the different types of use of diagrams and the different tasks that they served. This short article is the result of that study. The distinctions I make here apply not simply to so-called diagrammatic reasoning, but the larger set of activities that have been called spatial and visual reasoning.

2. Representation

Let us first consider the notion of representation itself. One of the striking things about the current work in cognitive neuroscience is the degree to which it is impossible for the researchers to avoid talking about information being represented, projected, fed back etc. Information about objects is said to be represented in this area of the brain, about shape in that area, and so on. Thus, anti-representationalism is not a practical option any longer: information and representation are an inevitable part of the needed vocabulary. On the other hand, buying into information being represented does not necessarily imply having to buy into "information processing" or "representation processing." A mechanical coin-sorter that operates by passing the coins through levers and holes of different sizes can be analyzed in terms of its using information about weights and diameters of the coins, but saying that it performs information processing does not agree with everyone's intuitions. More directly relevant to our topic, Paul Churchland has described neural structures that are involved in the visuo-motor coordination of a frog's catching of its prey, a fly. He shows that, in a suitably transformed space, the solution to the control problem can be expressed as a linear relationship between the (transformed values of) the location of the prey and a motor control parameter. In the frog's neural structure, visual information about the prey is represented in one neural layer, and the motor control information in another layer directly below it. The two layers are so connected that the prey location information directly sets the motor control parameter. Again, information of various types is clearly represented, but whether the best way to describe what happens to it as "processing" is unclear.

At least among AI people, talk of information representation soon leads to talk of representation processing which soon leads to talk of reasoning. Conversely, this also sets up a counter-rhetoric of "no representation" [1]. With reference to our topic of interest, not all spatial representation in the brain is to be taken as grist for some spatial reasoning mill, if reasoning to be taken as being performed explicitly. On the other hand, knowledge and inference are appropriate terms if only a Knowledge Level account [2] is intended.

3. Visual versus Spatial versus Diagrammatic

While many people make distinctions between "visual," "spatial," and "diagrammatic" representations, there does not seem to universal agreement on the meaning of these terms. Glasgow [3] distinguishes between "visual" and "spatial" by using the former to refer to uninterpreted pixel-like representations and the latter to refer to representations in which objects and their mutual spatial relations are indicated. But many other people seem to use "visual" to denote representations, interpreted or not, arising from the visual modality, while "spatial" is used to denote knowledge of space, its occupants and their mutual spatial relations, leaving open which modality or modalities – visual, auditor, kinesthetic or haptic – that supplied the information. Diagrams and visual representations in general can also have color and texture as part of the representational repertoire, while spatial representations do not seem to involve these properties. Motion is potentially part of all three: diagrams can use animation, and visual and spatial representations can involve sequences of images. There seems to be much more of a general agreement, at least on an intuitive level, on the notion of diagrammatic representation. It is taken to be a form of spatial representation, explicitly constructed and intended to be visually processed, containing elements that have a conventional semantics, displaying the spatial relations among the elements. Diagrammatic representations may also have elements, such as labels or other annotations, intended to be interpreted as non-spatial, and elements with visual properties, such as color and texture, that are used as part of the representational vocabulary. Iwasaki [16] proposes that one of their important properties is that they represent *abstractions*, i.e., they abstract out some of the information and represent other information. Of course, diagrams also have details that are not to be taken seriously. The conventionality in element semantics provides clues to the consumer of the representation about which aspects of the diagram are to be taken seriously and which are to be taken as incidental to the representation.

4. Reasoning, Predicate Extraction and Projection, and Prediction by Simulation

The word "reasoning" is often used as a generic equivalent of computation – "geometric reasoning" or "spatial reasoning" programs often turn out to be programs that perform computations involving geometric or spatial information. There is, however, a narrower use of the term "reasoning." In this sense of reasoning, the agent starts with some given *assertions*, and makes additional assertions as inferences, using rules of inference. For example, when Barwise and Etchemendy [4] speak of diagrammatic reasoning, they intend for the diagrams to be assertions of some facts or hypotheses, and the goal is to arrive at inferences that satisfy the standard notions of valid inference.

Going back to the motivation for this paper, a perusal of the literature led me to the conclusion that that at least three distinct types of use of diagrammatic information are generally included in the term "diagrammatic reasoning." Generalizing this to use of spatial information in human or machine cognition, the three types of spatial information processing are predicate extraction and projection, reasoning, and simulation. (These three types are in addition to algorithms that use spatial information to compute something of interest and that have no special standing as model of any part of spatial cognition.)

4.1. Predicate Extraction and Projection

Given a spatial representation, a number of facts corresponding to the representation can be asserted; e.g., object A is in the picture, object A is in front of object B, A taller than B, and so on. The set of spatial predicates is open-ended, and domain- and task-dependent. For example, in this session at the workshop, the paper by Epstein and Gelfand discusses a game-playing program that constructs various predicates given a board representation. These predicates are intended to capture potentially relevant facts about the configuration. However, by definition, all the spatial predicates must be computable from the spatial information in the representation. Many so-called spatial or diagrammatic reasoning programs in fact focus on predicate extraction, i.e., computing a given set of predicates from a spatial representation. There is a complementary process, *predicate projection*, which has not received much attention in the literature. Given a set of spatial predicates pertaining to a situation, this projection process creates a spatial or diagrammatic representation, or modifies an existing one, so as to match the assertions. What spatial predicates are useful for what tasks and how to compute them efficiently are reviewed in some detail in Chapter 6 of [6].

4.2. Reasoning

As mentioned, spatial reasoning is making trustworthy inferences with spatial or diagrammatic representations. A simple model of spatial reasoning is one in which the agent starts with a spatial representation, extracts some predicates from it, and, using other axioms and rules of inference, proceeds to make inferences. Stated this way, there is nothing to distinguish the actual process of spatial reasoning from any other form of reasoning, except for the facts that the problem gets its start from the spatial representation and generation of initial predicates requires use of a predicate extraction module.

This is what has often puzzled many people who have not been enthused about special claims for diagrammatic reasoning. However, in fact, there are a number of properties of spatial reasoning that justify special treatment. Spatial representations are not of general quantified propositions, but of concrete instances. Nevertheless, we seem, in some cases, to be able to make correct inferences about general situations. Studying the conditions under which such general inferences are supported is a major concern in the logic of spatial reasoning. Second, spatial reasoning often involves special spatially sensitive axioms and rules of inference (such as the transitivity of *right-of*, or *inside-of* predicates). Study of spatial reasoning is an opportunity to focus on such modality-specific rules of inference. Thirdly, often, spatial reasoning does not simply involve just extracting predicates and then proceeding with reasoning as usual, but instead involves projection onto spatial representations and extraction of new predicates. That is, the spatial representation is inextricably involved in the whole process of reasoning. A useful model is one where the agent's cognitive state is represented as having two components, a spatial one and a corresponding predicate representation, with the two components matching one another. There are now two modules that operate on the side, a predicate extraction module, as before, and a predicate projection module that constructs or modifies the spatial representation to match the predicate component. The famous "Behold" proof of the Pythagorean Theorem (see Chapter 1 of [7] for a version of this proof) is just a series of diagrams, with no linguistically represented assertions along the way. But even in these instances the reader is assumed to be making implicit inferences in her head as she examines the diagrams.

A major concern in diagrammatic reasoning is the conditions under which such inferences using diagrams are reliable. Conversely, the creation and use of such representation to solve reasoning problems that are not intrinsically spatial (such as reasoning about sets) are also matters of intense interest. Barwise and Etchemendy [4] note that it has generally been assumed by logicians that diagrams are a mere heuristic aid, and "real proofs" have to be purely symbolic in character. A major focus of their work is showing that valid diagrammatic proof systems may indeed be constructed.

Because spatial representations are specific instances rather than quantified propositions, they often provide a convenient representation for *models* of propositions. The research of Johnson-Laird (presented at this workshop) shows that this property of spatial representation is useful in human non-specialist syllogistic reasoning. Given a syllogistic reasoning problem, human reasoners model the various propositions, represent them spatially, and try to see if counterexamples can be constructed to the proposition to be proved. Glasgow's work, presented in this session, discusses the use of spatial representations as models.

4.3. Simulation

Many spatial "reasoning" programs or cognitive models do not reason in the sense of moving from cognitive state to cognitive state by processing truth-values of assertions. Instead, they make use of simulation to generate the next state. The problems for which such simulation is appropriate are generally *prediction* problems: given a spatial situation and some proposed actions on, or interaction between, the elements in it, what will be the spatial representation corresponding to the new situation? The simulation is supposed to mimic the spatial transformations involved in the transition from initial state to final state. The simulation might additionally mimic any physical laws involved in effecting the changes. In AI, the work of Funt [8], DeCuypers, at al [9], and Gardin and Meltzer [10] are examples of this kind of spatial simulation. DeCuypers' program computes the results of diffusion of gas in a room, and Gardin and Metzler's program similarly performs analog computations to predict the behaviors of strings and other objects in three-space under various forces. Typically, the computation is implemented by some kind of array model, in which each processor performs some computation and passes on the results to its neighbors in the array. Glasgow [3] also uses a spatial array to perform spatial reasoning. The simulation approaches are often characterized as analog processes for obvious reasons.

In contrast to AI, where there is a certain freedom in choosing computational approaches to get the job done, there are more constraints in explaining human spatial information processing abilities. The scope of such analog computational processes in humans is among the most hotly contested of issues. A good deal of the debate on the nature of imagery – propositional versus imagistic representations – can be traced to this issue. The mental rotation experiments of Shepard, Cooper and colleagues [11] have often been interpreted as implying the existence of internal "analog" processes in humans for performing rotation and translation. Kosslyn [12] proposes the existence of some analog processes in human spatial reasoning. The alternative to the hypothesis of such analog spatial information-processing is that rotation, translation and other such spatial reasoning processes are actually solved by reasoning involving spatial predicates and predicate projections.

5. Control of Reasoning

Earlier when I discussed reasoning I focused just on spatial predicates. However, a more realistic model is one where the state of the agent is multi-modal, i.e., information in several perceptual modes and also in the conceptual mode is present at each state. For example, if the agent were thinking of an apple, her cognitive state would have a visual representation of the shape, texture and color of the apple. In addition, she might have knowledge about apples relevant to the context, say, their prices and where to buy them. She might also have an anticipation of its taste. Representational changes in one mode typically give rise to corresponding changes in other modalities, giving the multi-modal representation some degree of inter-modal coherence. Problem-solving is driven by the most relevant information in any domain. The reasoning might be driven by spatial predicates at one moment and by conceptual predicates at the next. Each move forward results in a new state and in updating of representations in all modalities. Koedinger and Anderson's geometry program [13] uses a representation in which both conceptual and diagrammatic information is present at each state. Narayanan and Chandrasekaran [14] investigate the control issues involved in the use of such multi-modal representations. Epstein and Gelfand [17] discuss a game-playing program which involves the use of such multi-modal representation – their program's game state has both spatial and propositional content.

6. Summary

The task of formalizing the logic of spatial reasoning has just begun, since we do not yet have accounts of all forms of valid spatial reasoning. How and when diagrams are effective in aiding reasoning in general remains relatively poorly understood. Goel [15] argues that, contrary to intuitions, diagrams can not only be vague, but designers and other problem solvers actually exploit this vagueness for effective problem solving. The degree to which human spatial cognition employs analog processes for performing prediction remains controversial, but in AI special purpose architectures that perform analog simulation for prediction remain promising. It seems to me that the emerging trend towards reformulating many AI problems in the context of robotic behavior, where the robots have sensory, action and reasoning abilities, is very important for progress in spatial reasoning. Broadening the notion of internal images from the visual modality to an integrated multi-modal representation that includes conceptual information as well and using this internal multi-modal "image" as the basic encoding of experience promises to open up new ways of thinking about learning.

7. Acknowledgments

Parts of this paper were included in a session report at the NSF Workshop on Visual Cognition and Decision-Making in the Spatial Domain, May 15-17, 1997. I thank the organizers for the invitation. I also thank Michael Anderson and Hari Narayanan for their comments on an earlier draft of this paper, though I haven't fully incorporated their suggestions in this version. I thank John and Susan Josephson for discussions on this topic.

Research that forms the basis for this paper was supported by DARPA, Order number: D594, monitored by Office of Naval Research, Grant #: N00014-96-1-0701.

References

1. Brooks, R.A., Intelligence without representation,. *Artificial Intelligence*, 1986. **47**(1-3): p. 139-159.
2. Newell, A., The Knowledge Level, in *AI Magazine*. 1981. p. 1-19.
3. Glasgow, J.I. and D. Papadias, Computational Imagery. *Cognitive Science*, 1992. **3**: p. 355-394.
4. Barwise, J. and J. Etchemendy, Visual Information and Valid Reasoning, in *Visualization in Mathematics*, W. Zimmerman, Editor. 1990,

Mathematical Association of America: Washington D. C.

5. Glasgow, J., N.H. Narayanan, and B. Chandrasekaran, eds. *Diagrammatic Reasoning: Cognitive and Computational Perspectives*, 1995, The MIT Press: Cambridge, MA. 780 + xxvii.

6. Davis, E., *Representations of Commonsense Knowledge.* 1990, San Mateo, CA: Morgan Kaufmann Publishers, Inc.

7. Allwein, G. and J. Barwise, eds. *Logical Reasoning with Diagrams.* . 1996, Oxford University Press: New York, NY.

8. Funt, B.V., Problem solving with diagrammatic representations. *Artificial Intelligence*, 1980. **13**: p. 201-230.

9. DeCuyper, J., D. Keymeulen, and L. Steels, A hybrid architecture for modeling liquid behavior, in [5], p. 731-752.

10. Gardin, F. and B. Metzer, Analogical Representations of Naive Physics. *Artificial Intelligence*, 1989. **38**: p. 139-159.

11. Shepard, R.N. and L.A. Cooper, *Mental Images and Their Transformations.* 1982, Cambridge, MA: MIT Press.

12. Kosslyn, S.M., *Images and Mind.* 1980, Cambridge, MA: Harvard University Press.

13. Koedinger, K.R. and J.R. Anderson, Abstract Planning and Perceptual Chunks: Elements of Expertise in Geometry, *Cognitive Science*, 1990. **14**: p. 511-550.

14. Narayanan, N.H. and B. Chandrasekaran. Reasoning visually about spatial interactions. *Proceedings of Twelfth International Joint Conference on Artificial Intelligence.* 1991. Sydney, Australia. San Francisco: Morgan Kaufmann.

15. Goel, V., *Sketches of Thought.* 1995, Cambridge, MA: The MIT Press/A Bradford Book.

16. Iwasaki, Y. 1997. Session Report, NSF Workshop on Visual Cognition and Decision-Making in the Spatial Domain, May 15-17, 1997. This material will be part of the Workshop Report to appear, edited by S. Epstein, J. Gelfand, and M. Marefat.

17. Epsteni, S. L, J. Gelfand, and J. Lesniak, Pattern-based learning ad spatially-oriented concept formation in a multi-agent, decision-making expert. *Cognitive Science*, to appear.