Title
Qualitative Geometric Reasoning

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 11(0)

Author
Lindsay, Robert K.

Publication Date
1989

Peer reviewed

QUALITATIVE GEOMETRIC REASONING

Robert K. Lindsay

Artificial Intelligence Laboratory
University of Michigan

ABSTRACT

This paper addresses an old and fundamental problem, the role of visual imagery in cognition. While the problem has a long history in philosophy and psychology, it has had less attention explicitly directed toward it in artificial intelligence research on *reasoning* (as opposed to machine vision and graphics). This paper addresses the question of what it means for a cognitive agent to think directly in visual images (*depictions*), and how such abilities might be formalized and accomplished with computer hardware. It is argued that the identification of imagery with non-propositional or with non-digital representations is incorrect; rather, the quality of imagery that gives it a special character is that it employs non-deductive inference, and this may well be achieved by descriptive, digital representations. Furthermore, research on knowledge representation within AI suggests approaches to the classical problems of imagistic thinking. A program is described for translating propositionally stated geometric assertions into diagrammatic representations and employing a constraint-propagating procedure to manipulate the representations, thereby making inferences and testing conjectures.

INTRODUCTION

Much of the imagery debate within psychology still grapples with several puzzles that have been long with us: If knowledge is represented as depictions which must be "observed", how can an endless regress of homunculi within homunculi be avoided? Since imagery lacks the detail and vividness of perception, and is not restricted by actual observation of the world but may be produced "at will" even to represent non-veridical possibilities, how can it be of use in reasoning? Since anything that can be represented can be represented descriptionally to any desired amount of detail, and since descriptions can be dealt with computationally in straightforward ways, why bother with depictional representation at all? Since depictions must be specific, how can they be used to reach general conclusions? The debate between Kosslyn (1980; Kosslyn & Pomerantz, 1977; Kosslyn et al., 1979) and Pylyshyn (1973, 1980, 1981) testifies to the continuing concern these problems cause.

Research in artificial intelligence has illuminated many of the issues that have stimulated cognitive research in general, and its emphasis on explicitly defined computational models offers certain insights and techniques that can apply productively to the imagery problem. For example, since a programmed model requires the specification of the processes that construct and use its representations (in this paper these are called *construction and retrieval processes*), the homunculus problem is addressed and solved with each AI implementation. But the major relevant insight of AI is that, while some sort of general calculus for reasoning may be possible, when one considers the computational complexity and efficiency issues that must be addressed in the construction of a real-time intelligence, it becomes clear that not all in-principle adequate representations are equivalent. Specifically, a representation that captures the geometric properties of space would be able to make inferences (and, as we will propose here, conjectures as well) by constraint satisfaction methods that avoid the combinatorial problems of deductive proof. Such methods, then, are a tool that addresses the frame problem (how knowledge should be updated when some facts change) by letting "the side effects take care of themselves (Haugeland, 1985, page 229)".

Anderson (1978) has argued that it cannot ultimately be decided whether human thinking is imagistic or propositional, since it is always possible to impose an imagistic interpretation on a propositional description, and vice versa. However, two informationally equivalent representations may nonetheless not be computationally equivalent (as pointed out by Larkin and Simon, 1987). Thus specific models can succeed or fail to match empirical facts, exhibit varying amounts of explanatory power, or be computationally and biologically more or less plausible. Theories must be judged not by direct verification of their underlying assumptions, but by the usual standards of empirical test of their predictions of observable behavior and their coherence with other established theory and fact. See Hayes-Roth, 1979. Consequently, this paper is concerned with accounts of the *functional* properties of representations rather than with issues of the nature of the base representation. I believe that imagery is indeed functionally different from logic-based systems, though not because logic is propositional. I have argued elsewhere (Lindsay, 1988) that there is a difference between knowledge representations that support inference by deduction, that is, by the use of a proof procedure such as that of first-order logic, and those that support inference without deduction, such as by heuristic search and/or constraint mechanisms inherent in the representation. In other words, it is not the distinction between propositional and non-propositional representations that is at issue, but the distinction between representations based on a logico-deductive formalism and those that have inherent structure, including but not limited to schema systems. This is an idea that has also been advanced by others in somewhat different terms, including Palmer (1978) and Dretske (1981).

Palmer's analysis is particularly illuminating. He defines a hierarchy of three types of isomorphism between a representation and that which it represents. Physical isomorphisms preserve information by virtue of representing relations that are *identical* to the relations represented. Thus a physical model of a natural terrain preserves the spatial relations of the represented terrain with the very same relations, including for example elevation, but on a different scale. Functional isomorphisms, on the other hand, preserve information by representing relations that have the same algebraic structure as the relations represented. Thus the elevations of a natural terrain may be represented as colors on a map of the terrain, provided the colors are interpreted appropriately (as an ordered set) and mapped so as to preserve the order of the physical elevations of the terrain. Thus physical models are a proper subset of functional models. Palmer introduces a class of isomorphisms between physical and functional, which he calls natural isomorphisms. In a natural isomorphism, the representation of preserved relations need not be by means of *identical* relations, as in physical isomorphism, hence not all natural isomorphisms are physical isomorphisms. On the other hand, not just any algebraically equivalent set of relations qualifies. In a natural isomorphism, the representing constructs have *inherent constraints* (Palmer's term); that is, there is additional structure imposed on the representing objects that limits the ways in which they may relate. If these inherent constraints preserve the relations of the represented world, we have a natural isomorphism.

Palmer identifies natural (including physical) isomorphisms with analog (including pictorial) representations, and functional but non-natural isomorphisms with propositional representations. Propositional representations are thus less restricted, as we normally suppose, because the structure of the representing world is extrinsic to it, that is it may be arbitrarily imposed, say in the form of rules of deduction. However, analog (including pictorial) representations employ representations that have inherent (non-arbitrary, unalterable) structure ('inherent constraints') that allow us to do away with deduction rules. This limits their applicability, but at the same time increases their power by reducing the computational complexity of inference (and easing the frame problem). A similar definition of "analog" was suggested by Dretske. Note that this use does not identify *analog* with *continuous* (non-digital) representations.

The concepts of *simulation* and *constraint satisfaction* which derive from artificial intelligence research provide means to realize these ideas. A program that simulates the behavior of a physical, geometric, or abstract system may serve the role of a natural isomorphism. A cognitive agent or program makes inferences about the external situation by running its model, whose inherent constraints mirror the structural and functional constraints of the situation. With appropriate implementations, the costly search required by proof procedure methods is avoided. The cost of using such representations is a loss of the generality of, say, a logic based system which both permits arbitrary descriptions and provides a proof procedure. A second cost is a loss of generality due to the need for a variety of representations for a variety of problems. This cost may be partially offset by the use of analogy, such as when non-geometric problems are translated into geometric terms.

The essence of this dissolution of the imagery problem is present in many knowledge representation methods that are not overtly depictional. For example, inheritance hierarchies (Touretzky, 1986) exhibit the essence of depictional representation simply because certain inferences (that property P is true of each member of any subset of a class to which P is appropriately attributed) follow "automatically" from the addition of information to the representation. Such inferences are inherent in the construction and use of the representation, without explicitly employing a proof procedure. Hierarchies are, of course, readily and universally represented propositionally.

Inheritance of properties, however, is a limited and specific form of reasoning. Reasoning with visual images is a more general and ubiquitous ability for which *cognitive* (as well as perceptual) mechanisms are presumably already in place in any organism with vision. Visual reasoning can be accomplished with a representation whose construction and retrieval processes embody the inherent geometric and/or physical constraints of space and mechanics. Once this is understood, the research issues no longer turn on whether the representation is propositional or not, nor whether it is analog or digital. The issue is how to represent geometric and physical systems in ways that are *computationally efficient* and can interface appropriately with other forms of knowledge representation. Should the representations look like or be readily translatable into a conventional descriptive format, such as frames, they remain, through the fact of their specialization, basically different in kind from a general, arbitrarily structurable formalism such as predicate calculus. They are "natural isomorphisms", which combine the virtues of a physical model, whose behavior is forced to obey physical laws, with the virtues of an abstract representation, which can be recorded and processed by digital computers or brains.

GEOMETRIC REASONING

The bulk of artificial intelligence research on reasoning has not directly addressed imagistic reasoning in spite of the seeming centrality, along with verbal reasoning, of this mode of thought as suggested by the long-standing interest in psychology, the common introspective impression of the layman, and extensive anecdotal evidence from literature on the history of scientific discovery. The special case of reasoning about two dimensional geometric objects has been addressed both in the cognitive modelling and AI literature, however. Space does not permit discussion and comparison with the present work except to acknowledge related research by Gelernter (1959), Novak (1977), Funt (1977, 1981), Kosslyn (1980), Forbus (1983), Anderson et al. (1985), Larkin & Simon (1987), and Koedinger & Anderson (in press). Some of this research has employed deductive methods augmented by *ad hoc* coding of diagrammatic information, while some of it has employed non-deductive representations more centrally, although for circumscribed tasks. The present work is an attempt to abstract and generalize the insights of these authors.

It is now widely understood that any knowledge representation must specify, in addition to the underlying format of the recorded data, the procedures for adding and recalling information from the store. As noted, I call these *construction processes* and *retrieval processes* respectively; I refer to the passive record as the *representation-proper*. Previous work on models of imagery have employed several representations-proper. None of these is "purely depictional" in any intuitively clear way, but as I have argued, that is not the issue.

Predicate representations: Individual constants are single letters, such as A, B, C, and D. Predicates introduce objects, for example, POINT(A), POINT(B), POINT(C), POINT(D), SEGMENT(POINT(A), POINT(C)), ANGLE(POINT(A), POINT(D), POINT(B)) and QUADRANGLE(POINT(A), POINT(D), POINT(B), POINT(C)). Metric relations may be introduced with the use of a numerical equality predicate '=' and numbers [LENGTH(SEGMENT(B,C)) = 10 and MEASURE(ANGLE(D, A, C)) = 90] or with additional predicates such as RIGHTANGLE(D, A, C). Non ratio-scale positional relations are expressed with predicates such as BELOW(D, C) and LONGER(SEGMENT(A, C), SEGMENT(A, D)). In general, arbitrary descriptions may be constructed including diagrammatically impossible descriptions such as BELOW(D, C) & ABOVE(D, C). Descriptions may also be incomplete and inaccurate.

Schema (Frame) representations: Schemas are structured propositional knowledge that relate several predicates and impose restrictions on variable types. For example, a schema for triangle includes the fact that it is composed of three vertices (which must be point names) and three sides (which must be segments). Point schemas and segment schemas in turn contain information about names, locations, and imposed constraints (such as rigidity of a segment or fixedness of a location); some of this information may be unknown at any given time. For example, a segment could be recorded in a schema with this format:

SEGMENT SCHEMA

NAME	NIL
END1	A
END2	C
SEGMENTPOINTS	({POINT}#61666 {POINT}#63333 . . . .)
MARKING	BLACK
SOURCE PREMISE	INPUT-1: "QUADRANGLE(A, D, B, C)"
LENGTH	10
LENGTHSTATUS	FIXED
BEARING	30
BEARINGSTATUS	ARBITRARY

Coordinate representations: The representation-proper consists of a set of marked points, each of which is assigned numerical x,y coordinates. Some of the marked points may have associated names. Objects exist by virtue of their coordinates satisfying certain defining numerical relations. For example, a segment exists if all of the coordinates "between" (algebraically defined) its endpoints correspond to marked points. The definition of a triangle is more complex. Metric relations are represented implicitly, and are retrieved by computations. For example, the length of a segment is computed by applying the Euclidian distance function to the coordinates of its endpoints. Ordinal metric relations such as "to the right of" are defined by appropriate computations on coordinates.

Pixel Array representations: This representation is similar to Cartesian coordinate representations except that it is digitized: coordinates must be, say, integers or rational decimals of limited precision. Each pixel may be one of a finite number of values to denote a grey level or a color. Pixels are referenced by indices, and relative locations may be determined by arithmetic on indices. In more elaborate pixel array representations, the "grain" of the array may be stratified under program control so that some or all of the representation may be "exploded" or "compressed" to obtain varying levels of detail.

Pixel Network representations: Elements of the representation-proper are interpreted as points each of which is connected to eight neighbors: in the north, northeast, east, southeast, south, southwest, west, and northwest directions. An element may have an associated name and a marking. Objects such as segments are represented implicitly by marking a connected, appropriately straight set of elements. Metric relations such as length are retrieved by determining the cardinality of the set of marked points. Ordinal metric relations are implicit in the topology and must be retrieved or tested by searching of the connections and applying appropriate definitions (e.g., B is right of A if there is a path from A to B that goes through more steps in the three "right" directions (northeast, east, southeast) than in the three "left" directions.

I have implemented a system for the representation of simple geometric elements such as points, lines, and triangles in two-dimensional space. At the heart of the system is a fixed set of construction and retrieval processes. For example, there is a process that constructs a line segment between two given points (a construction process) and a process that determines whether or not two line segments are of the same length (a retrieval process). The representation is a pixel array/network that combines the features of both of these representation types in a single structure, plus a structure of inter-related schemas for geometric objects (point, line segment, triangle, and so forth) present in a depiction. The pixel array is finite in extent, and has a certain inherent grain or minimum resolution both in distance and direction. The distance resolution can be set to coarser values and this will affect the behavior of the construction and retrieval processes; thus the test of segment length checks equality to within the currently operative resolution, and the construction of a line segment will exhibit a corresponding compliance. The frames relate the geometric objects to the pixel array, and record propositionally specified or derived descriptions of the objects, such as that a particular line segment is required to be rigid. Predicate representations are used as input (only). The information thus given to the system is used by the construction processes to build the three components of the representation system, thereby describing a specific instance of that which was described in the predicate input.

Although lines in a depiction have specific lengths and polygons have specific areas, only qualitative judgments about these metric quantities are known to the processes. Thus, a line may be judged longer, shorter, or equal in length to another, but differences in lengths cannot be compared unless a construction can be made (with these processes) that produces new line segments of the appropriate lengths to be compared.

In the human mind, imagery and verbal thought cohabit, interact, and complement one another. Therefore, it is important to see how such interactions could come about. Although the system cannot produce representations through visual perception, it can do so from descriptions provided either externally or by its own higher order goal-driven processes. In addition, the programs illustrate how images can be used both to produce and to test conjectures about what is possible, and thus have the ability to run *Gedanken* experiments in the service of problem solving. This requires the interaction of the representation with the more conventional search methods of AI as discussed later.

QUALITATIVE GEOMETRIC REASONING

Certain inferences are straightforward in the system. For example, if a triangle is constructed with two equal angles, the inference that it has two equal sides can be made by "observation", using the above cited retrieval process; this is an inference but not a deduction. More interesting is the use of the system to produce and test conjectures. This depends on the architecture of higher processes that set appropriate tasks. In my model, all of these problem solving methods work through a set of "visual routines" that employ the construction and retrieval processes. These visual routines manipulate existing depictions by propagating small perturbations of point positions or line lengths. This, in effect, simulates "in the mind's eye" what would happen if real geometric objects were manipulated physically.

These processes are called qualitative geometric reasoning processes, on analogy to work by de Kleer & Brown (1984), Forbus (1984), and others (see Bobrow, 1985) on qualitative physical reasoning methods. In that work, composite physical structures, such as an electronic circuit or a pressure regulating valve, are modelled from an inventory of components, each of which behaves according to equations that relate its inputs and outputs. A program reasons about the composite model by observing its behavior while small qualitative perturbations of physical quantities are propagated through the model, obeying the componential equations. In my model of qualitative geometric reasoning, small perturbations of positions (suggested by the higher processes in service of testing a conjecture or attempting to make a construction in accordance with an externally supplied propositional description) are propagated while obeying the inherent constraints of two-space (as implemented by the construction and retrieval processes) and any arbitrary, externally specified constraints on the depiction (as recorded in the frame structures). Thus the representation determines whether a proposed alteration is possible, and if it is possible, what follows from it.

To illustrate, consider Figure 1, interpreted as a shortest path problem. A representation of the left side of Figure 1 is constructed by the program from descriptive statements about points and lines, including that all line segments are rigid (and hence can be rotated but not stretched or compressed). External "pulls" are imposed on points A and B in the directions indicated by the arrows. This manipulation is external in the sense that it is imposed by the "higher processes" alluded to above, rather than by the geometric representation system itself. A step-wise simulation of the effects of these forces causes a gradual transformation to the representation illustrated to the right. At this configuration, no further motion is possible without violation of the rigidity constraints. This effectively demonstrates that the shortest path from A to B is through D, not C. While proof of this would require relatively little search of path combinations in this simple example, the manipulation provides a number of other inferences at the same time, such as that C remains above D, and the resulting figure is a triangle. On the other hand, while the same computational procedure could in principle handle networks of dozens or hundreds of nodes to compute shortest path, we note that people are unable to do so without the aid of a physical string model or some such external aid. The explanation of this within the terms of the present model comes from the amplification of errors by the qualitative, low resolution representation. Note further that the computations involved in this experiment were time consuming because they were carried out serially, point by point. However, the same computations lend themselves to distributed processing. If each pixel element in the representation had its own associated processor to do the vector calculations, the processes would be more rapid.

TESTING OF GEOMETRIC CONJECTURES

The testing of conjectures requires an exploration of "all possible" alterations of a given type. This involves higher processes that control search, backtracking, and enumeration of possibilities. For example, it is well-known (to us) that two triangles with corresponding sides equal in length must be congruent. A non-deductive "visual proof" of this theorem amounts to showing that at most one triangle can be constructed from three line segments of fixed length. The higher processes set the task of constructing one such triangle. This would prove impossible if the lengths chosen do not satisfy the triangle inequality; otherwise, the construction processes succeed in constructing a triangle. Doing so requires making specific choices for locations of vertices. The motivation of these choices is determined by the higher processes, and these embody important strategic assumptions. In the present case, choices are made to produce depictions of a size appropriate to the current resolution and at locations and orientations that avoid ambiguities and potential interference during subsequent manipulations of the depiction. Once a triangle is constructed, the higher processes construct the second triangle, with sides of the same lengths as those chosen for the first triangle. The triangles of course prove to be congruent. Then to complete the task of conjecture testing, modifications of the second triangle are proposed to "see" if another triangle can be constructed that differs from the first. *Any* proposed modification of the position of a point will force the modification of other positions in order that lengths remain unchanged, and it will be observed that angles thus do not change (the triangle can only rotate or translate rigidly). *Any* is here stressed to emphasize that the higher processes must be able to propose all possible modifications. This is not a trivial matter since it would be foolish to consider all possible triangles that could be marked on the array: it must be possible for the higher processes to recognize equivalence classes of proposed perturbations in order to avoid exhaustive search and yet conclude that a triangle is uniquely determined by three sides. It will be necessary to integrate the depictional representation with a search architecture to enable the model of geometric reasoning to be fully goal-directed.

REFERENCES

Anderson, J. R. (1978) Arguments concerning representations for mental imagery. *Psychological Review*, **85**, 249-277.

Anderson, J. R., Boyle, C. F., and Yost, G. (1985) The geometry tutor. In *Proc. of the International Joint Conference on Artificial Intelligence-85* held in Los Angeles. Los Altos, CA: Morgan-Kaufmann.

Bobrow, D. G. (1985) *Qualitative reasoning about physical systems.* Cambridge, MA: Bradford Books.

de Kleer, J. and Brown, J. S. (1984) A qualitative physics based on confluences. In D. G. Bobrow (Ed.) *Qualitative reasoning about physical systems.* Cambridge, MA: Bradford Books, 7-83.

Dretske, F. I. (1981) *Knowledge and the Flow of Information.* Cambridge, MA: MIT Press.

Forbus, K. (1983) Qualitative reasoning about space and motion. In D. Gentner and A. L. Stevens (Eds.) *Mental Models.* Hillsdale, NJ: Lawrence Erlbaum, 53-73.

Forbus, K. (1984) Qualitative process theory. In D. G. Bobrow (Ed.) *Qualitative reasoning about physical systems.* Cambridge, MA: Bradford Books, 85-168.

Funt, B. V. (1977) Whisper: A problem-solving system utilizing diagrams and a parallel processing retina. *Proc. of the 5th International Joint Conference on Artificial Intelligence, IJCAI-77*, held at MIT, August 1977. Pittsburgh: Carnegie-Mellon University, 459-464.

Funt, B. V. (1981) Multi-processor rotation and comparison of objects. *Proc. of the 7th International Joint Conference on Artificial Intelligence, IJCAI-81*, held at the University of British Columbia, August 1981. Menlo Park: AAAI, 218-220.

Gelernter, H. (1959) Realization of a geometry theorem proving machine. *Proc. International Conference on Information Processing*, pp. 273-282. Paris: UNESCO. Reprinted in E. A. Feigenbaum and J. Feldman, eds. (1963), *Computers and Thought.* New York: McGraw-Hill, 134-152.

Haugeland, J. (1985) *Artificial Intelligence: The Very Idea.* Cambridge, MA: MIT Press.

Hayes-Roth, F. (1979) Distinguishing theories of representation: A critique of Anderson's "Arguments concerning mental imagery". *Psychological Review,* **86**: 376-382.

Koedinger, K. R. and Anderson, J. R. (in press) Abstract planning and perceptual chunks: Elements of expertise in geometry. *Cognitive Science.*

Kosslyn, S. M. (1980) *Images and Mind.* Cambridge, Mass.: Harvard University Press.

Kosslyn, S. M., Pinker, S., Smith, G. E., and Shwartz, S. P. (1979) On the demystification of mental imagery. In Block (1981) *Imagery*. Cambridge, Mass.: MIT Press, 131-150.

Kosslyn, S. M. and Pomerantz, J. R. (1977) Imagery, propositions, and the form of internal representations. *Cognitive Psychology*, **9**:52-76.

Larkin, J. and Simon, H. A. (1987) Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, **11**, 65-100.

Novak, G. S. (1977) Representations of knowledge in a program for solving physics problems. *Proc. of the 5th International Joint Conference on Artificial Intelligence, IJCAI-77*, held at the Massachusetts Institute of Technology, August 1977. Menlo Park: AAAI, 286-291.

Lindsay, R. K. (1988) Images and inference. *Cognition*, **29**: 229-250.

Palmer, S. (1978) Aspects of representation. In E. Rosch and B. B. Lloyd (Eds.) *Computing and Categorization*. Hillsdale, NJ: Erlbaum. 259-303.

Pylyshyn, Z. W. (1973) What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin,* **80:** 1-24.

Pylyshyn, Z. W. (1980) Computation and cognition: Issues in the foundations of cognitive science. *The Behavioral and Brain Sciences*, **3**: 111-133.

Pylyshyn, Z. W. (1981) The imagery debate: Analog media versus tacit knowledge. *Psychological Review*, **87**. In Block (1981) *Imagery*. Cambridge, Mass.: MIT Press, 151-206.

Touretzky, D. S. (1986) *The mathematics of inheritance systems*. Los Altos, CA: Morgan-Kaufmann.

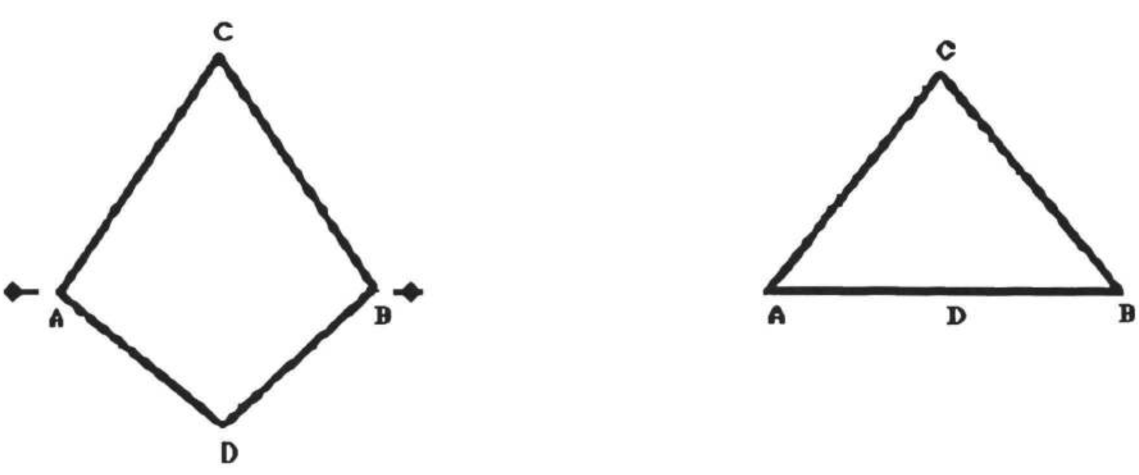

Figure 1
SHORTEST PATH PROBLEM