Knowing About Diagrams

Robert K. Lindsay

University of Michigan
205 Zina Pitcher Place
Ann Arbor, Michigan 48109, USA
lindsay@umich.edu

Abstract

Understanding diagrams and using them in problem solving requires extensive knowledge about the properties of diagrams, what diagram elements denote, how their parts are distinguished and referenced, how they relate to linguistic statements, and so forth. This knowledge is most naturally represented linguistically. Nonetheless, diagrams or imaginal representations of them are used in substantive non-linguistic ways as part of the problem solving process. The interaction of linguistic and diagrammatic representations must be understood in order to construct a theory of diagrammatic reasoning. In this paper, an example is examined to illustrate some of the ways in which diagram manipulation may be used in geometric reasoning and to informally identify some of the knowledge necessary for such reasoning.

Introduction

Humans and other animals continually engage in interactions that require perceptual reasoning. For example, an experienced person can catch a thrown ball, cross a busy street, sight read a musical score, judge shape by touch, recognize a friend's face in a crowd, and so forth. Each of these activities requires complex and rapid computations, usually not accessible to conscious description. The neural mechanisms and processes that perform these tasks are general in the sense that they apply to an unbounded variety of similar situations. However, each application is specific to one situation, and explicit generalization is not needed and usually not available to the person: a ballplayer cannot precisely describe how he fields fly balls. Thus for many perceptual reasoning abilities, language is not required.

Other forms of human reasoning do involve explicit generalizations, often stated in a natural language or, on occasion, in a formal artificial language such as a predicate calculus. Reasoning with language about both special cases and generalizations is frequently at least in part done consciously.

Formal mathematical reasoning is usually seen as the quintessential form of linguistic reasoning. The in-principle mechanizability of proof-verification and production by linguistic methods, though qualified by certain undecidability issues, gives linguistic reasoning an aura of correctness, both as a model of correct reasoning and as a model of human cognitive processes as applied to mathematics. However, the lack of obvious procedures for the discovery of important new ideas or for inventing new mathematical tools, methods, and subfields of utility belies the completeness of this model as a description of how mathematicians understand and create mathematics. Further, the fact that mathematicians use both natural language explanations and non-linguistic notational devices in performing their work and communicating it to others suggests that a purely linguistic model of mathematical reasoning is at best incomplete.

Here I am focusing on a small piece of this extra-linguistic machinery of mathematical reasoning, namely the use of diagrams. Diagrammatic reasoning in this context employs perceptual reasoning in the service of establishing general statements by applying perception-like processing to diagrams (artifacts) or mental representations of them (mental images). One way to use diagrams for reasoning is to translate them into a formal predicate calculus and use deductive methods to make inferences. Recent work by several researchers has addressed the problem of translating simple diagrams to statements, but that research has not reduced the translation to a mechanical process for general cases. Note specifically work by Barker-Plummer & Bailin (1992), Chou (1988), Shin (1995), and Wang (1995). Another approach is to use a non-linguistic representation of diagrams in conjunction with methods of inference that manipulate those representations directly. For example, inference processes can be constructed from procedures that manipulate a diagram by constructing new elements of it and moving other elements about, and then reading off the ensuing changes that are then interpreted as inferences. Interpreted as a psychological model, the mind is assumed to run experiments, using the mind's eye and mind's hands to perform them. While this explanation appeals to many as more psychologically plausible than the manipulation of propositions, before the question of psychological validity can be addressed, the perceptual models must be more precisely specified. The first step is to show they are plausible "first order" theories, that is, that they are capable of reasoning at all.

One way to formalize diagrammatic perceptual reasoning of the above sort is to devise an inventory of diagram manipulation processes and demonstrate that these can be used as a programming language for manipulating diagrams in productive ways. For example, Furnas (1992) has devised the "bitpict" system that allows the specification of processes that transform small pixel arrays into other pixel arrays, and has demonstrated how some "programs" composed of such operations can be written to solve problems. Anderson & McCartney (1995, 1996) define a system of picture element manipulation processes that can be combined according to explicit rules into programs that solve problems strictly by manipulating pixel-arrays. My work is similar to those approaches in that it also performs computations on pixel arrays. However, the picture manipulation processes are *described* at a higher level of aggregation, namely at the level of the manipulations of geometric objects, such as "construct a line between two points" and "determine if two lines are parallel." These processes are then available to write programs that create and manipulate diagrams. One class of such programs can follow diagram manipulation steps that demonstrate certain geometric propositions and verify that the results support the conclusion (Lindsay, in press). In the present paper I am concerned with how such processes can be employed strategically to *devise* programs appropriate to a given reasoning task.

I have emphasized elsewhere – Lindsay (1996) – that an account of conscious geometric reasoning with diagrams must employ propositional as well as pictorial representations and manipulations. This would be true if only because the mathematical statements proved, demonstrated, or understood are propositional in form, usually universally quantified assertions (e.g., "The sum of the interior angles of *any* plane triangle is equal in measure to a straight angle"), and some connection must be established between the reasoning and the conclusion. More fundamentally, the very strategy of using diagrams in a substantive way requires the constant interaction of diagram elements and propositions, notably in the form of constraints imposed on the diagram ("This angle is – *always* – to remain a right angle no matter what other changes are made"). Indeed it is undeniable that mental experiments can be run according to a variety of assumptions, for example that the objects are rigid, or that they are plastic. That is, diagrammatic reasoning is "cognitively penetrable" (Pylyshyn, 1984), where what is "cognitive" is described propositionally.

However, propositional knowledge is used in many essential ways other than to state a conclusion and the premises underlying it. This essential knowledge includes generalizations about invariants, inventories of special cases, definitions of symmetry, knowledge of algebraic relations, generalizations of prior conclusions, and knowledge of problem solving strategies and heuristics. It is desirable to establish an inventory of the knowledge that, in addition to the knowledge of space built into the programming language, is needed to devise methods that can achieve human-like reasoning about geometry. Once this is done, that knowledge must be represented in ways that can interact with diagrammatic representations and be used to verify demonstrations, create demonstrations, and ultimately perhaps to discover geometric propositions that were not explicitly conveyed to the program.

I will illustrate by example how a "programming language" must be augmented to support the invention of demonstrations and the discovery of geometric relations. I will use one example to illustrate a range of diagrammatic reasoning strategies and the knowledge underlying them.

The Example

The example concerns the Quadrilateral Theorem (QT), informally: "The figure formed by connecting the midpoints of the sides of any quadrilateral is a parallelogram." This theorem, perhaps initially surprising, is true for convex as well as concave quadrilaterals.

Understanding by Proof

One way to understand this theorem is to understand its fully formal proof, by following the proof and checking its steps for validity one-by-one. Diagrams are not part of a fully formal proof. For most of us, a better way to understand this theorem is a "textbook proof." By textbook proof I mean the sort of quasi-formal argument favored in textbooks and classroom teaching. They are almost invariably accompanied by a diagram whose connection to the propositional argument is not formalized. Typically the diagram accompanying a textbook proof illustrates only one specific instance of the general proposition. Textbook proofs usually omit proving many properties that are true of the diagram; the bulk of a fully formal proof typically addresses these omissions from the textbook proof. Thus textbook proofs are not rigorous.

Understanding a textbook proof requires the ability to understand both natural language and diagrams. It also requires seeing how the conclusion of the theorem is based upon other theorems and lemmas that are previously understood, and so on back to the axioms of geometry that presumably are accepted as obvious, whatever might cause that. Devising a proof is a more difficult task than understanding one, and reflects an understanding of the theorem in an even deeper sense.

Reasoning by Diagram Manipulation

The following discussion illustrates a variety of ways that diagram observation and manipulation can be used to support reasoning. It will be understandable because the reader understands a number of unstated assumptions that need to be captured if a program is to be able to follow or generate such arguments.

The QT applies to "any" quadrilateral; it is a universal statement. The starting point for using a diagram typically is to choose one particular illustration of the concept or proposition in question. For the QT this means selecting four points on a sheet of paper and connecting them to form a quadrilateral, or doing something of that sort mentally. However, a single diagram is notoriously deficient in its ability to represent universal propositions.

One common way around this deficiency is to choose a particular case that has no special properties that may be responsible for the generalization. That is, one wants a representative instance. Care is taken so that the figure so selected is not one of a number of "special" quadrilaterals, those that have other properties, such as rhombuses or trapezoids. However, this presupposes knowledge about what properties are special in this sense. For the QT, each side and each angle of the chosen quadrilateral should be of different measure, and special angle measures, such as 0, 45, and 90 degrees, should also be avoided. No pairs of sides should be equal, parallel, or perpendicular.

Draw or imagine a representative quadrilateral, find the midpoints of its sides and connect them pairwise. Finally, examine the inscribed figure and note that it is a parallelogram. Since the original quadrilateral was chosen "arbitrarily" and had no properties known by previous experience to yield additional special features, one is encouraged to accept the generality of the proposition. As a further check one could try other quadrilaterals as starting points to further reduce the probability of having hit upon a special case. One could also attempt to construct a counterexample, since a single counterexample suffices to disprove a conjecture.

Deeper Understanding

Although the observation of several appropriate special cases is compelling, there is a deeper and more important sense of understanding that is aided by diagrams. One would like to know "why" the theorem is true.

One could modify the original quadrilateral by dragging one vertex and noting that the constructed figure remains a parallelogram. This not only provides a sequence of several examples, but shows how altering one property alters others in an exactly compensatory way.

Given that one has a specific instance of the appropriate diagram, say Figure 1, one could ask if an *arbitrary* change will *preserve* the parallelogram property. A given quadrilateral can be transformed into any other quadrilateral by moving its vertices to the location of the new quadrilateral vertices. The order of making these alterations is immaterial; only the end result matters.

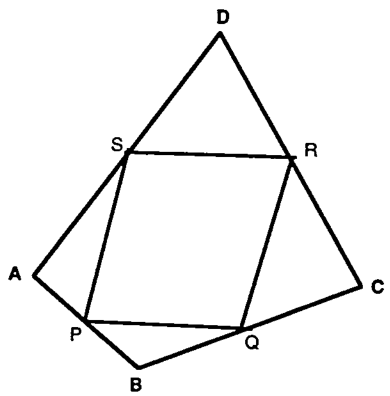

Figure 1

Furthermore, there is a "syntactic" symmetry to the situation (Gelernter, 1959, Lindsay, 1996) in the sense that, since the figure is arbitrary to begin with, if one could show that the movement of any one vertex preserved the property, one could show it true for the movement of any other vertex by a similar procedure. So we are encouraged to believe that no counterexample exists if we have one valid instance and can show that moving a single vertex preserves the midpoint-parallelogram property.

Note further that an arbitrary target movement of a vertex can be broken into any number of steps whose vector sum is the target movement. In particular, the target movement can be broken into just two steps in distinct directions that need not be orthogonal. This presumes path independence, that the relevant properties of a geometric figure are determined by its static configuration, independently of how it was constructed. This is a property of geometric figures, but not in general of physical situations involving energy transformations. If each of these sub-movements is property preserving, the total movement must be.

The next step might be to select two distinct directions such that movements in those directions can be easily seen to preserve the parallelogram-property. The vertical and horizontal directions of the paper are not *a priori* interesting choices since geometry is orientation independent; the only relevant directions are those relative to objects in the figure. These include the directions of the quadrilateral sides and the sides of the parallelogram. Closer observation shows that the parallelogram sides are also the directions of the (unconstructed) diagonals of the quadrilateral. That observation, triggered by the search for significant directions, may be valuable (in particular, it helps to understand the textbook proof given later).

What happens if vertex C of Figure 1 is moved in the direction of the orientation of segment PQ (from P toward Q), directly away from its opposite vertex A? What happens if vertex C of Figure 1 is moved in the direction of segment QR? We could actually make these movements and observe that the parallelogram-property is preserved, but can we show in a deeper sense why it must be preserved?

Let's examine more closely the relation between quadrilaterals and inscribed parallelograms. Clearly it is possible to inscribe non-parallelograms by choosing appropriate points on the quadrilateral's sides. Is it possible to inscribe a parallelogram within a quadrilateral without connecting midpoints? If we connect two non-midpoints P and Q with a line (See Figure 2) and attempt to construct a parallelogram with that as one side, the length and orientation of the opposite side of the inscribed parallelogram are determined. We can make a copy, P'Q', of the first line PQ and move it parallel to itself until it contacts one of the other two quadrilateral sides, hold that end on the contacted side, maintaining orientation and length, and slide P'Q' along the contacted side until it contacts the fourth side with P' at S and Q' at R. Completing the figure yields a parallelogram PQRS that does not connect the quadrilateral's midpoints. Thus the midpoints do not define the only inscribed parallelogram. R and S are uniquely determined once P and Q are chosen because there is at most one "slice" of the other half of the quadrilateral of that length and orientation because the sides converge to D. This is observed by diagram manipulation combined with an analysis of the available degrees of freedom. However, if PQ is too long (say UV) there may be no solutions, since there is no possible "slice" at that orientation.

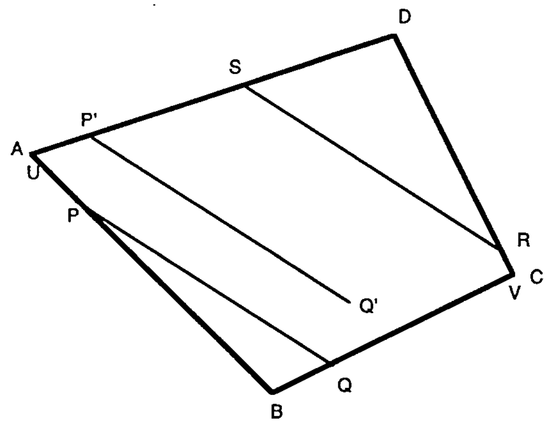

Figure 2

If we had picked P and Q as midpoints, would this procedure force R and S to be midpoints? Experiment seems to confirm this but does not establish it in general.

Here is another approach. Starting with an arbitrary parallelogram, let us circumscribe a quadrilateral whose sides are tangent to the parallelogram's vertices. Do the vertices of the parallelogram bisect the quadrilateral sides? This proves not to be true; an indefinite number of quadrilaterals can circumscribe a given parallelogram.

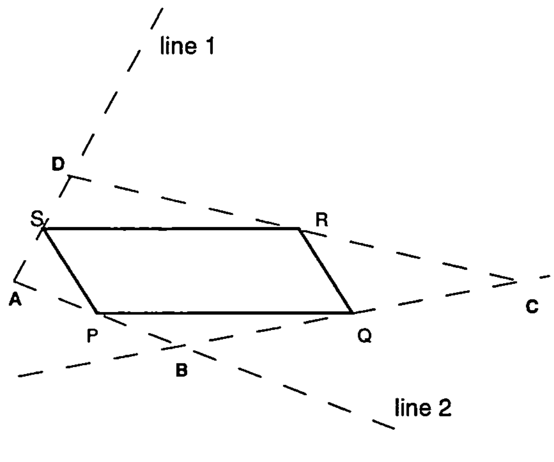

Figure 3

Next consider Figure 3. Starting with a parallelogram PQRS and an arbitrary point A outside it, the directions of two circumscribing lines (1 and 2) are uniquely determined. This is because two points uniquely determine a line. However, one quadrilateral vertex on each of line 1 and line 2 may be freely chosen anywhere on the line subject to the constraint that they lie on the same side of the extended opposite parallelogram side QR as does A. Choice of these two points B and D determines 3 of the vertices of the quadrilateral. But placing B and D also determines the directions of the remaining two quadrilateral sides since they must pass through Q and R, respectively. Their intersection, the fourth vertex C, is thus also determined. Since points B and D may be altered subject only to the constraint above, we again see that the midpoint property is not necessary. What if we force it to hold by choosing B and D accordingly, as in Figure 3? Then it will be observed that the quadrilateral is uniquely determined, and it satisfies the midpoint property. It is thus possible to construct a circumscribing quadrilateral obeying the parallelogram-property around an arbitrary parallelogram. We have also learned that the situation has only a few degrees of freedom: selecting one quadrilateral vertex fixes the midpoint quadrilateral for a given parallelogram.

The foregoing experiments have not led to a complete understanding of why the QT should hold, but have provided information about the quadrilateral-parallelogram relationship.

Another useful strategy is to start with the *simplest* special case, and verify the proposition for it. This is often relatively easy because special properties enforce symmetries; that's what makes them special. We may then be able to show that departures from the special case do not alter the proposition. Here we would start with the simplest, most symmetric quadrilateral, a square. We see that the inscribed midpoint polygon is also a square. In the previous case, difficult measurements were needed to verify a potential relation. However, in this symmetric case the measurements are simple because they follow from the observation of bilateral symmetry, something human perception is good at.

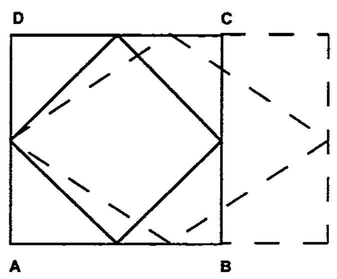

Figure 4

Now what would happen if we make the square into a rectangle by lengthening a pair of opposite sides equally? See Figure 4. It is readily seen that the inscribed square's vertices that lie on the elongated lines (AB and DC) move in the same direction and by the same amount, half of the elongation. The sides of the inscribed figure change orientation, but our bilateral symmetry detectors immediately see that opposite pairs change by the same amounts and the figure remains symmetric. We now have a parallelogram inscribed in a rectangle – the condition remains true. Again, *syntactic* symmetry detection shows us that the same will happen if we now stretch or shrink the rectangle in its other dimension.

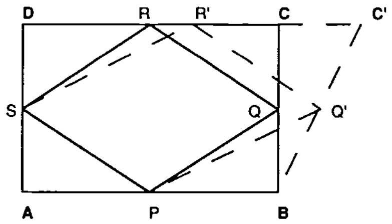

Figure 5

Symmetry is thus a compelling means to understand the theorem in the special cases of rectilinear, equal, orthogonal stretching of pairs of sides. Consider next the movement of a single vertex in a direction that violates symmetry by extending one side of the rectangle, say DC in Figure 5, and ask if the parallelogram-property is preserved. In this case the midpoint Q of the right side BC moves half the movement of the vertex, which is exactly how far the midpoint of the upper line moves. Seeing this is straightforward in this special, rectilinear case. This means that QR moves parallel to itself to Q'R', and thus remains parallel to its opposite side SP, which remained unchanged since its endpoints were not altered. Also, while the orientations of the other two sides change, they change equally because their moved endpoints move in the same direction by the same amount. Thus they remain parallel, and the parallelogram-property is maintained. Can this relation be generalized to arbitrary movements of C?

Move point C in any direction other than from D toward C, say in the direction of the diagonal PR. See Figure 6. If we can establish that the midpoints of the altered segments move in the same direction by the same amount we know that the orientation of the line connecting them (R'Q') has not changed. Examining a simpler case makes this easier to see. In Figure 7 we consider a segment with one fixed endpoint E. Moving the other endpoint Y in a given direction to Y' moves "every" point on the line in the same direction, but each point moves by a fraction of the endpoint movement proportional to its distance from the fixed point. Note that we are imagining the original segment and the new segment to have points in 1-1 correspondence. (Technically this is true, even though there is a non-denumerable set of points in the abstract idealization of a line.) However, it is sufficient to imagine the line as composed of a finite number of points, with the line restricted to passing though each of them. Imagine the line as a rubber band that stretches. Applying this to Figure 6 (EYY' maps onto BCC' and DCC') means that R and Q each move by half the distance that C moves away from D and B, in each respective direction, and all three points move in the same net direction, the P to R direction. Thus both ends of QR move by the same amount in the same direction and the line must maintain its orientation and length.

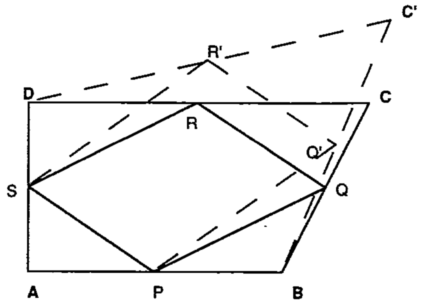

Figure 6

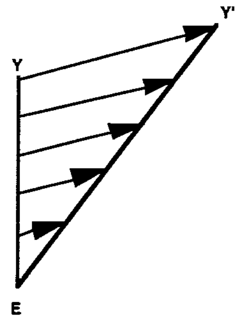

Figure 7

Recalling path independence we now see that an arbitrarily chosen amount and direction of movement of one vertex of the quadrilateral can be broken into two components one of which is parallel to one diagonal of the quadrilateral, the other of which is parallel to the other diagonal of the quadrilateral. (Further experiments would confirm that any two distinct directions would suffice.) Movements in diagonal directions are parallelogram-property preserving, so any movement of the vertex is parallelogram-property preserving. Finally this "same" argument applies to any vertex, *mutatis mutandis*, so any alteration of the quadrilateral is parallelogram-property preserving because it can be broken into a set of eight independent and parallelogram-property preserving steps. Since we can start with a square for which the theorem is clearly true by symmetry, we now see that the general theorem is true.

Extra credit: Is the QT true of non-planar quadrilaterals?

This is the sort of reasoning by experiment in which some people engage in lieu of proof-style reasoning when thinking about geometry, and which for them underlies the discovery and understanding of formal proofs. Clearly it is different from formal, linguistic processing, being a hybrid of propositional thinking and diagram (or image) manipulation in which the perception of the altered diagram plays a substantive role. I have suggested a few of the items of knowledge on which this reasoning draws. I will now collect an inventory of some of this knowledge, informally stated.

Knowledge-Based Understanding

Some of the knowledge needed for diagrammatic reasoning about geometry resides in the ability to construct and manipulate a diagram according to propositional specifications. This includes knowledge implicit in the diagram (or a mental image or other representation of it) that enforces the essential properties of space. In addition there is knowledge of basic concepts such as line segment, midpoint, quadrilateral, and so forth. All of this knowledge is already represented in my programmed system in ways that permit it to be employed appropriately. Thus the system can be told what experiments to perform, and then do them. In addition to performing constructions and constrained manipulations it can observe the effects of the changes and record them in its inventory of facts about the diagram.

The ability to *propose* the experiments of a demonstration must employ additional knowledge, such as the following.

Some of the properties of relevance to the theorems of plane geometry depend upon relative values and are independent of absolute values. For example, the shape of an object is independent of its absolute location in space. Similarly, some of the properties of relevance, such as shape again, are independent of orientation. Other properties are independent of scale as well, while some are not. The proficient geometer knows which invariances apply to which properties.

Many propositions of plane geometry are implicitly about rigid figures. In general, the properties of rigid figures are static properties in the sense that they can be determined by examination of the static configuration without knowledge of how the figure was constructed, e.g., in which order the components were drawn. This was referred to earlier as the assumption of path independence.

In the case of diagrams, one knows that the object of study is not the diagram itself, but the abstract idealization that it represents, where lines have no thickness and are meant to be perfectly straight. It is known that the imperfections of an actual diagram may yield incorrect results and one must know how to avoid such traps.

Other knowledge is provided by definitions. For example, the definition of a parallelogram as a four-sided figure with opposite sides parallel must be known in order to understand statements about parallelograms. Furthermore there is knowledge about the relations of classes of objects, such as the fact that all squares are rectangles. This knowledge may be used in conjunction with knowledge of the inheritance of properties concept to conclude that a property shown to be true of every rectangle is also true of every square, for example.

Overlying all of the foregoing knowledge is the knowledge of when each is true, for example, what properties are independent of scale or which ones are not. This knowledge perhaps can only be represented as explicit lists of facts.

There is also knowledge of logic. For example, a universal statement can be disproved by a single counterexample but generally cannot be proved by even a large number of positive instances. This knowledge is tempered in informal problem solving by other beliefs, some based on the concept of probability. For example, a larger number of positive examples increases support for a conclusion. If the possibilities can be factored into a finite number of cases, the generality can be concluded if it is shown to be true of each case.

Belief in a generalization is increased if it can be shown that it is true of any arbitrarily selected member of the class to which it applies. The definition of arbitrary selection is difficult to capture. Usually it means that no constraints are placed on the example to make it a "special" case. The constraints are those that apply to the properties of the figure class in question. Thus "any quadrilateral" should exclude special cases, such as degenerate cases (where two or more vertices are identical, for example). A special case is where any subset of elements of the figure are related in "special ways."

"Special ways" is an inventory of properties that are known, presumably from prior learning, to be important. For experienced geometers this includes properties of parallelness, perpendicularity, equal length, and equal angle measure. Special angle measures are 0, 90, 45, 30, and 60 degrees and their integer multiples. Although absolute lengths are arbitrary, relative lengths are not; special cases of length are those where one length is an integer multiple of another.

Many problems also require knowledge about how to reason about some equality and ordering relations. For example, the equality (of area, of length, etc.) relation is symmetric, transitive, and reflexive, the greater-than relation is anti-symmetric, transitive, and irreflexive. A geometer must know what these properties mean and how to use them to draw new conclusions from established facts.

Other knowledge takes the form of previous conclusions, for example previously proved theorems in the case of formal geometry. Having this knowledge available in turn requires knowledge of how to represent and recall generalizations and how to apply them to new cases. Thus to demonstrate that two specific triangles are congruent, one might rotate and translate them into superposition (ignoring mirror images which require flipping in 3-space). One might then construct or examine a demonstration of the side-angle-side congruency theorem. The demonstration might take the form of showing that specifying these properties of a triangle leads to a rigid figure, that is, one that cannot be altered in shape if those values are fixed. Having accepted this generalization it would be stored. It could then be used in the future to establish congruence without the perceptual reasoning steps of rotation and translation.

Other knowledge involves knowing strategies for problem solving, including knowledge of when they are useful. The use of symmetry is one example. If bilateral symmetry is detected, certain conclusions can be immediately drawn about the sameness of the relations among objects on opposite sides of the axis of symmetry. Syntactic symmetry is a different method. This involves detecting the similarity of procedures and discovering a mapping of variables that preserve the form and permits a similar conclusion. Degree-of-freedom analysis is another useful method for systematically applying a set of constraints to sequentially limit the possible properties of an object. When the object is a point and the property is its location, this is the method of loci. Beginning with a simple, highly-constrained example and relaxing the constraints is another frequently useful strategy. Another strategy is the analysis of a movement into a series of steps whose effects are easier to see. This was used in the QT example, and is of wide use.

Beyond Verification: Discovering Demonstrations

Most of the above knowledge is propositional, although it refers to diagrammatic properties and is only useful when it becomes procedural. Knowledge must be related to the available inventory of diagram processing functions. The above inventory is only a sample of the relevant knowledge needed, even if we restrict the topic to plane geometry. To invent demonstrations requires the selective use of this knowledge, in appropriate sequence, so that the demonstration will be relevant to the theorem.

One way to employ this additional strategic knowledge is to encode it in a representation analogous to a "script" (Schank & Abelson, 1977). For example, "When attempting to demonstrate a claim about an entire class of figures, select one according to its rules of non-arbitrariness, examine it to see that the alleged property holds; repeat with other cases." And "After demonstrating the property for one arbitrary instance of the class, alter that instance by altering in turn each of its components in ways that are arbitrary for that component." And "To demonstrate congruence of figures, attempt to rotate and translate the figures into coincidence."

This model of demonstration construction would yield a set of special cases (scripts), rather than a general procedure. The system would find the relevant script, if it knows one, and apply it, instantiated with the appropriate definitions of its parameters (e.g., "arbitrary" is interpreted in the context of the particular component). New models for other classes of demonstrations could be constructed *ad hoc*, or could perhaps be induced from a set of cases.

Understanding the Textbook Proof

A textbook proof of the QT is given in the following table (paraphrased from Fogiel, 1994, pages 66-67). Note that several steps of the QT textbook proof call upon previous definitions and theorems. The proof establishes that one pair of sides must be both parallel and of equal length, and the result follows by a previous theorem. The key idea is the construction of a quadrilateral diagonal and the use of theorems about similar triangles. See Figure 8.

Statements	Reasons
1. P, Q, R, and S are the respective midpoints of sides of AB, BC, CD, and AD of quadrilateral ABCD	1. Given
2.. SP is a midline of triangle ABD	2. A midline of a triangle is the line segment joining the midpoints of two sides of the triangle
e. SP parallel to DB	3. The midline of a triangle is parallel to the third side.
4. SP = 1/2 DB	4. The midline of a triangle is half as long as the third side of the triangle.
5. QR is a midline of triangle CDB	5. Definition of midline.
6. QR parallel to DB	6. The midline of a triangle is parallel to the third side of the triangle.
7. QR = 1/2 DB	7. The midline of a triangle is half as long as the third side of the triangle.
8. SP parallel to QR	8. If each of the two lines is parallel to a third line, then they are parallel to each other.
9. SP = QR	9. Transitivity property.
10. Quadrilateral PQRS is a parallelogram	10. A quadrilateral is a parallelogram if two of its sides are both congruent and parallel.

Notice that the textbook proof is no more rigorous than the perceptual proof. It is defined in terms of a specific diagram, and makes no explicit effort to show that this diagram is not a special case. It uses informal reference ("the third side" "each other" and so forth) and alludes to "reasons" whose application is vague ("by definition" "transitivity property" for example). Thus while the textbook proof has a surface appearance of succinctness, that is bought at the price of vagueness and incompleteness. Furthermore, the textbook proof relies on the very same knowledge underlying perceptual reasoning, but in an impoverished way, and in many ways is inferior to perceptual reasoning as a means of understanding. Rather than observing how components and properties of the figure interact, the proof examines the static figure. In contrast, the manipulation approach addresses the problem from a variety of ways, playing with different figures and noting how movements reveal the connections among properties in a perceptually related way. Clearly the two approaches are not exclusive alternatives, but are complementary and synergistic.

Neither the textbook proof nor the perceptual argument cited above is formal, and yet both are revealing and perhaps convincing. The reason in each case is that the result is related, through intermediate stages, to underlying properties that may be observed in a particular diagram that is accepted as representative. Ultimately, the credibility of each depends on the reader's acceptance of the validity of the arguments because they coincide with his perception and understanding of visually processed information. Each also requires knowledge that can only be understood propositionally, including knowledge of how the perceptual and the propositional are related.

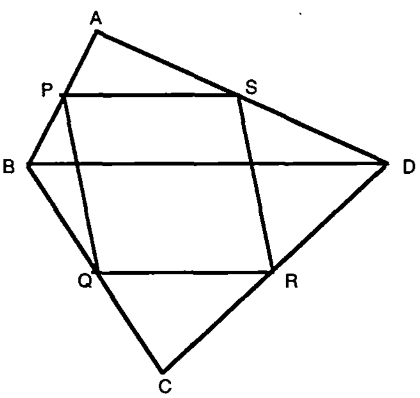

Figure 8

Acknowledgment

This material is based on work supported by the United States National Science Foundation under Grant No. IRI-9526942.

References

Anderson, M. and McCartney, R. 1995. Inter-diagrammatic reasoning. In *Proceedings of the International Joint Conference on Artificial Intelligence*, 878-884. San Mateo, CA: Morgan Kaufmann Publishers, Inc.

Anderson, M. and McCartney, R. 1996. Diagrammatic reasoning and cases. In *Proceedings of the Thirteenth National Conference on Artificial Intelligence*, 1004-1009. Menlo Park, CA: AAAI Press.

Barker-Plummer, D. and Bailin, S. C. 1992. Proofs and pictures: Proving the diamond lemma with the GROVER theorem proving system. In *Reasoning with Diagrammatic Representations. Technical Report SS-92-02*, 102-107. Menlo Park, CA: American Association for Artificial Intelligence.

Chou, S.-C. 1988. *Mechanical geometry theorem proving*. Dordrecht, Boston, Lancaster, Tokyo: D. Reidel Publishing Company.

Fogiel, M. ed. 1994. *The High School geometry tutor*. 2d edition. Piscataway, NJ: Research and Education Association.

Furnas, G. W. 1992. Reasoning with diagrams only. In *Reasoning with Diagrammatic Representations. Technical Report SS-92-02*, 118-123. Menlo Park, CA: American Association for Artificial Intelligence.

Gelernter, H. 1959. A note on syntactic symmetry and the manipulation of formal systems by machine. *Information and Control*, 2: 80-89.

Lindsay, R. K. 1996. Generalizing from diagrams. In *Cognitive and Computational Models of Spatial Reasoning*, 51-55. Menlo Park, CA: American Association for Artificial Intelligence.

Lindsay, R. K. in press. Using diagrams to understand geometry. *Computational Intelligence*, 14.

Pylyshyn, Z. W. 1984. *Computation and Cognition: Toward a Foundation for Cognitive Science*. Cambridge, MA: MIT Press.

Schank, R. and Abelson, R. 1977. *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Hillsdale, NJ: Erlbaum.

Shin, S.-J. 1995. *The logical status of diagrams*. Cambridge: Cambridge University Press.

Wang, D. 1995. *Studies on the formal semantics of pictures*. Ph.D. diss., Institute for Logic, Language, and Computation, University of Amsterdam.