Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Visual Reasoning in Discovery, Instruction and Problem Solving

Permalink

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 16(0)

Author

Narayanan, N . Hari

Publication Date

1994

Visual Reasoning in Discovery, Instruction and Problem Solving

N. Hari Narayanan
Knowledge Systems Laboratory
701 Welch Road, Bldg. "C"
Stanford University
Palo Alto, CA 94304
`narayan@ksl.stanford.edu`
(Organizer)

Abstract

The symposium on "Visual Reasoning in Discovery, Instruction and Problem Solving" will consist of three talks focusing on the role of visual reasoning in higher level cognitive processes. As this is a newly emerging research area spanning cognitive science and artificial intelligence, the symposium is designed both to inform and to stimulate interest, discussion, and further enquiry. The speakers will consider visual reasoning in three areas: scientific reasoning and discovery, learning and instruction, and problem solving. The talks will show how several different types of data can contribute to a clearer understanding of processes, mechanisms, and strategies underlying visual reasoning. First, Nancy Nersessian will cover the discovery aspect by providing a historical view on visual representation in creative scientific reasoning. Next, Rogers Hall will cover the educational aspect by laying out a set of general educational questions concerning the role of representational forms, and discussing studies of how people coordinate representational resources while working on problems in different instructional and work settings. Finally, Mary Hegarty and Hari Narayanan will together cover the problem solving aspect from both experimental and computational perspectives.

Introduction

> I was about seventeen, when my thoughts turned seriously to invention. Then I observed to my delight that I could visualize with the greatest facility. I needed no models, drawings or experiments. I could picture them all as real in my mind. Thus I have been led unconsciously to evolve what I consider a new method of materializing inventive concepts and ideas. – *Nikola Tesla, Inventions, 1919*[1].

For the purposes of this symposium we define the term *visual reasoning* broadly, to cover thinking, reasoning, and solving problems using visual modality specific external representations (diagrams, pictures, CRT displays) and internal representations (mental imagery). According to this perspective, visual reasoning is not an activity confined only to gifted inventors like Nikola Tesla, but one that most of us engage in at one time or another. However, its apparent ease is somewhat deceptive because it hides complex interactions among perceptual, cognitive and deliberative processes that support this capability. Though visual reasoning involves processes of visual perception, cognition, comprehension, and deliberation, this symposium will focus on comprehension and deliberation. A clearer understanding of the nature of these processes and their interactions will help explicate an important and intriguing human capability and allow us to harness the power of visual reasoning in diverse applications including instructional technology, visualization in science and engineering, and multimedia systems.

Whereas research on mental imagery has had a long history in cognitive science, a research area on the use of visual reasoning in higher level cognitive processes (e.g., problem solving and scientific reasoning) has emerged relatively recently among cognitive scientists and artificial intelligence researchers. In fact, the recent past has witnessed a surge of research and related activities on computational and cognitive aspects of visual reasoning. This symposium is a response to this increased interest. Its aim is to highlight the significant role of visual reasoning in certain human activities. In particular, it will address the role visual reasoning in three areas: discovery, instruction and problem solving. First, we will examine insights that historical data provides on the role of visual reasoning in creative thought processes. Second, we will consider data from studies examining the use of different representational forms in instruction. Third, we will look at results from experimental investigations of people engaging in visual reasoning while solving the problem of understanding and reasoning about devices from diagrams and text. Finally, we will examine how computational models can clarify accounts of visual reasoning processes.

Visual Reasoning in Scientific Discovery

Visual Representation in Scientific Reasoning – A View from History

Nancy J. Nersessian
School of Literature, Communication, & Culture,
School of Psychology, and College of Computing
Georgia Institute of Technology
Atlanta GA 30332
`nancyn@cc.gatech.edu`

What are all those visual representations doing in the notebooks, diaries, letters, and publications of scientists? Are they "mere aids" to discovery, replaceable by linguistic/formulaic expressions, or do they play an ineliminable role? Why, especially, are they so prevalent during

---

[1] As quoted in (West, 1991).

periods of conceptual innovation and change? This talk will explore these and related questions from the perspective of cognitive philosophy and history of science.

Until quite recently visual representations were largely ignored by philosophers and historians of science. Positivist philosophers, concerned to justify scientific reasoning, saw visual representations as at best "mere aids" to the real reasoning process and at worst as impediments to reasoning, leading to false inferences. Reasoning is using deductive or inductive algorithms. One of the great achievements of this philosophy was Hilbert's axiomatization of Euclidean geometry: the pictures really are eliminable from the reasoning process. Further, philosophers of diverse persuasions have long argued against a "logic of scientific discovery", placing discovery within the province of historians and sociologists, not philosophers. Earlier historians under the positivist influence also tended to ignore the pictures and focus on the linguistic/formulaic dimensions of scientific discovery: Galileo's law of free fall, Newton's laws of motion, Maxwell's equations, the theory of evolution, the concept of "gene". Although some contemporary historians, under the sway of the sociology of knowledge, acknowledge the rhetorical and communicative role of the visual representations employed in science, scant attention has been paid to the question of their possible generative role.

I will survey some of the recent work by historians and philosophers on visual representation in the geological and biological sciences and discuss my research into its role in the development of electrodynamics. I propose that scientific discovery is inherently a problem-solving process that comprises such activities as articulating conceptual structures, mathematizing domains, and devising real-world and thought experiments. It is more properly termed "invention" or "construction" than "discovery". Further, something is not a discovery until it is understood and sanctioned by a community. Communicating discoveries to the members of a scientific community is an instructional process. The aim is to get others to understand, accept, and possibly further the discovery. Addressing the questions posed at the outset requires an interdisciplinary analysis that involves creating a working synthesis between research in philosophy and history of science and research in the cognitive sciences, in this case, research on visual representation in problem solving and learning (Nersessian, 1992, 1993, in press).

My analysis leads to a methodological hypothesis and a psychological hypothesis. The methodological hypothesis is that in many instances of scientific discovery, visual representations are not eliminable from the discovery process. They play a significant generational role. The psychological hypothesis is that visual representations employed in discovery processes act in support of internal reasoning via modeling processes. This hypothesis assumes that although the historical record provides only external representations – the inventors cannot be asked to give protocols or be the subjects of reaction time experiments – nevertheless, taken in conjunction with cognitive psychological research, these lend support to hypotheses about the nature of the internal phenomena they engage.

Visual Reasoning in Instruction

Making and Seeing Mathematics in and out of School

Rogers Hall
Division of Education in Mathematics, Science and
Technology
University of California
Berkeley, CA 94720
rhall@garnet.berkeley.edu

Research in cognitive science has begun to look outside the individual in an experimental setting to understand how discipline-specific forms of competence are acquired and used (Greeno, 1989, in press; Hutchins, 1990, to appear; Pea, 1993). My contribution to the symposium follows this shift in two ways: (1) shared use of interactional resources like talk, activity, and inscription take theoretical parity with mental representation in accounts of complex problem solving, and (2) schools and workplaces become preferred empirical sites for basic research in cognition. This talk draws on comparative studies of mathematical problem solving, in and out of school, to explore the development of heterogeneous representational practices. These include specific forms of visual reasoning.

In one line of work (Hall, 1990, 1993; Hall, Kibler, Wenger & Truxaw, 1989), people with different levels of achieved competence (students, graduate students, and teachers) are asked to solve typical school algebra problems. Across participants, surprisingly diverse representational forms (e.g., narratives, drawings, arrays, and expressions) are actively constructed, combined, and repurposed during problem solving. Many of these forms lie outside the sanctioned algebra curriculum (e.g., drawing snapshots and carrying out oral calculation as vehicles move towards each other), and their material features are differentially useful for inference, calculation, justification, or explanation.

A second line of work examines how mathematical problems emerge and are solved in the ongoing work of design teams, comparing middle school students and professional engineers who both use computer-aided design tools (Hall & Stevens, in press). In both settings, specific representational practices develop for modeling space and projecting its intended use. Diverse representational forms (e.g., screen or paper displays, narratives, gestural manipulation of displays and animation, and various types of calculation) are assembled together to make proposals, evaluate alternatives, or justifying existing design commitments. There are common features in the way that interactional resources like talk (turn structure and narrative elaboration), activity (gesture and kinesics), and inscription (drawing or tracing over the surface of existing displays) are assembled by students and professionals. Much of this interactional work is densely indexical and serves to hold their collective enterprise together.

However, there are also striking differences: (1) Representational practices for projecting views of space are more extensively integrated and show a higher degree of referential precision among engineers. This reflects both their facility for using multiple 2D displays to work on 3D structures and extensive organizational requirements on how an acceptable design can proceed. (2) Although design projects in school can be intensely engaging, they are generally short-lived and, as group activities, fit poorly into cycles of individual assessment that determine student trajectories into higher grade levels or out of school. (3) As an organizational undertaking, engineering work allows people to anticipate and influence the reception of their work by other parties. Relations of accountability are much less open for students in school.

The temporal trajectory of a "project" or developing a "professional identity", and forms of accountability that structure participation in both (Lave & Wenger, 1991), are serious problems for the design of learning environments intended to foster specific representational practices among learners in school or the workplace (Roschelle, 1990). We have a great deal to learn about the interactional character of cognition in collective activity, the role of different representational forms in this work, and how people learn to participate in these activities.

Visual Reasoning in Problem Solving

Mental Animation - Understanding and Reasoning about Devices from Text and Diagrams

Mary Hegarty
Department of Psychology
University of California
Santa Barbara, CA 93106
`hegarty@psych.ucsb.edu`
and
N. Hari Narayanan
Knowledge Systems Laboratory
701 Welch Road, Bldg. "C"
Stanford University
Palo Alto, CA 94304
`narayan@ksl.stanford.edu`

Scientific problems are often presented in mixed-mode format, i.e., as diagrams accompanied by text. Our research suggests that visual reasoning has a significant role to play in problem solving in these situations (e.g., Hegarty, 1992; Hegarty & Sims, 1994; Narayanan, Suwa & Motoda, 1994a). The process of solving a problem can be characterized in terms of two stages: (1) problem comprehension, or the process of generating an internal model from the information given, and (2) reasoning, or the process of generating a solution by making inferences from this model. We will consider the roles that diagrams play in each of these two stages.

Comprehension of mixed-mode input

When a mechanical problem is presented as a text accompanied by a diagram, the reasoning agent (human or machine) must build a unified internal representation of the problem from these two kinds of input. Consideration of how such a mental model might be formed raises several interesting questions, some of which have been investigated experimentally and computationally, and others which remain open.

First, we might ask how comprehension of a mixed-medium presentation differs from comprehension of either a text or diagram alone. We have found that either a text or diagram alone is sufficient for communicating the basic components and configuration of a device, but that readers who study mixed media descriptions construct better models of the kinematics of a device (Hegarty & Just, 1993) and are better able to solve problems concerning the functioning of the device (Mayer, 1989).

Second, we can ask how people coordinate their processing of a text and diagram while reading a problem description. Preliminary research, based on students' eye fixations while reading text and diagrams, suggests that the text plays an important role in directing processing of the diagram (Hegarty & Just, 1989, 1993). On the basis of this observation, we can make some prescriptions about how the text in a mixed-medium presentation might be written in order to facilitate comprehension: (1) the elements of the diagram should be labelled clearly, (2) the accompanying text should use these labels consistently and repeatedly, and (3) the text should include explicit direction to all the important information in the diagram.

Third, we might ask what functions diagram inspection serves in comprehension of mixed-media presentations. The eye-fixation data suggests different kinds of diagram inspections that can be classified as *local inspections* for encoding relations among a small number of adjacent components, and *global inspections* for integrating relations among many components. Local inspections are particularly text directed and help form an initial internal representation of information in the text. Global inspections serve to reactivate part of an existing internal representation and integrate new information from the diagram with this internal representation.

Finally, we ask what kinds of prior knowledge might be required for correct comprehension of a mixed-medium presentation. We suggest that two types of knowledge are involved: (1) knowledge of diagramming conventions and (2) background knowledge of the domain.

Visual reasoning during problem solving

The processing of the problem display does not end at the comprehension stage of problem solving. For example, mental visualization based on the problem display can facilitate problem solving in various ways (Antonietti, 1991). We examine the use of diagrams in the solution of a type of mechanical problem in which solvers are shown a static diagram of a mechanical system and asked to infer the motion (kinematics) of the components of the machine when it is in operation. We refer to this process as "mental animation" (Hegarty, 1992).

How are diagrams used during problem solving? Eye-fixation and protocol studies (Hegarty, 1992; Hegarty & Sims, 1994; Narayanan, Suwa & Motoda, 1994a; Schwartz & Black, 1992) reveal that a diagram serves as an external memory in two ways. First, it allows the problem solver to reactivate his or her representation of the configuration of the system to be mentally animated. Second, an internal representation of the diagram supports a visualization of the kinematics of the system. We postulate that unlike the static external diagram, its internal representation is a dynamic entity that is actively modified during this visualization. Our hypothesis implies that as problem solving proceeds there will sometimes be a mismatch between the external diagram (which remains unchanged) and the internal diagrammatic representation which has been modified. We will discuss how this mismatch affects reasoning.

How do diagrams influence reasoning? First, diagrams support reasoning by allowing the indexing of information by spatial location and making explicit information about the topological and geometric relations among the components of the problem (Larkin & Simon, 1987). Second, configurations depicted in a diagram may directly cue applicable problem solving knowledge, thereby reducing problem space search (Koedinger & Anderson, 1990). Third, diagrams facilitate problem solving by acting as external memory as described above. Fourth, we propose that in mental animation problems, the connectivity and spatial adjacency of components in a diagram guides the reasoning process along the perceived direction of causality (Hegarty, 1992; Narayanan, Suwa & Motoda, 1994a).

Computational modeling of visual reasoning from diagrams

The representational and facilitative roles of diagrams in reasoning have also been addressed by artificial intelligence researchers (Narayanan, 1992), who have developed schemes for machine representation of diagrams and imagery (e.g., Glasgow & Papadias, 1992) and reasoning with such representations. These computational modeling efforts will be the focus of the last part of our talk.

We will discuss four recent systems that have attempted to model diagrammatic reasoning: BEATRIX a system that parses mixed mode input to generate a unified internal model (Novak & Bulko, 1993); BH a system currently under development that hypothesizes behaviors of simple mechanical devices from labeled schematic diagrams (Narayanan, Suwa & Motoda, 1994b); POLYA a geometric theorem proving system that uses diagrams for proof planning (McDougal, 1993); and REDRAW - a system that solves elementary structural analysis problems from civil engineering using both diagrams and domain knowledge (Tessler, Iwasaki & Law, 1993). These four systems illustrate various aspects of the interaction between conceptual and diagrammatic aspects of problem solving. BEATRIX shows how referents in text and diagram may be matched, thus establishing co-reference, while parsing the diagram and text in order to generate a unified internal model. POLYA shows how information from the diagram can be used to index and trigger appropriate geometry proof plans. BH shows how a simulation corresponding to mental animation can support inference by facilitating the detection of component interactions. REDRAW shows how conceptual knowledge can guide the selective retrieval of information from the diagram.

The general lesson to be drawn from an examination of these systems is that computational models of visual reasoning must take a hybrid or heterogeneous approach toward representation and reasoning. Reasoning procedures must draw on the conceptual knowledge in addition to effectively utilizing diagrammatic information. This requires a combination of traditional AI methods (rule-based, deductive, constraint-based, case-based, etc.) and techniques for information retrieval from, and manipulation of, diagrammatic representations.

In summary, solving problems with mixed-mode input involves a complex interplay among text processing, diagram inspection and prior knowledge. Further delineating the specific roles played by each of these poses an exciting challenge to cognitive and computer scientists.

Conclusion

Despite a surge of research and related activities, research on cognitive analysis and computational modeling of visual/diagrammatic reasoning is still in its infancy. It is an area brimming with many interesting open questions. Most importantly, results from different avenues of research on visual reasoning have the potential to advance our understanding of creative reasoning processes underlying discovery, to facilitate the design of effective instructional materials and environments, and to help develop computer-based problem solving systems that utilize multi-modal representations and media for both reasoning and communicating with humans. Therefore, the primary goals of this symposium are not just to inform but also to stimulate interest, discussion, and further research.

References

Antonietti, A. (1991). Why does mental visualization facilitate problem-solving? In R. H. Logie & M. Denis (Eds.), *Mental Images in Human Cognition* (pp. 211–227). Amsterdam: Elsevier Science Publishers.

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

Greeno, J. G. (1989). Situations, mental models, and generative knowledge. In D. Klahr & K. Kotovsky (Eds.), *Complex Information Processing: The Impact of Herbert A. Simon* (pp. 285–318). Hillsdale, NJ: Lawrence Erlbaum Associates.

Greeno, J. G. (in press). Understanding concepts in activity. In C. Weaver, C. R. Fletcher, & S. Mannes (Eds.), *Discourse Comprehension: Essays in Honor of Walter Kintsch.*

Hall R.P. (1990). Making Mathematics on Paper: Constructing Representations of Stories about Related Linear Functions (Tech. Rep. 90-17). Doctoral dissertation. Irvine: University of California, Department of Information and Computer Science. Also appears as Monograph 90-0002, Institute for Research on Learning.

Hall, R. (1993). Representation as shared activity: situated cognition and Dewey's cartography of experience. Presented at *the 1992 AERA meetings*, San Francisco. Technical Report, Institute for Research on Learning.

Hall, R., Kibler, D., Wenger, E., & Truxaw, C. (1989). Exploring the episodic structure of algebra story problem solving. *Cognition and Instruction*, 6(3), 223–283.

Hall, R. & Stevens, R. (in press). Making space: a comparison of mathematical work in school and professional design practices. To appear in S.L. Star (Ed.), *The Cultures of Computing*. London: Basil Blackwell.

Hegarty, M. (1992). Mental animation: inferring motion from static displays of mechanical systems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(5), 1084–1102.

Hegarty, M. & Just, M. A. (1989). Understanding machines from text and diagrams. In H. Mandl & J. R. Levin (Eds.), *Knowledge Acquisition from Text and Pictures*. Amsterdam: Elsevier Science Publishers.

Hegarty, M. & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language*, 32, 717–742.

Hegarty, M. & Sims, V. K. (1994). Individual differences in mental animation during mechanical reasoning. *Memory and Cognition*, in press.

Hutchins, E. (1990). The technology of team navigation. In J. Galegher, R.E. Kraut, & C. Egido (Eds.), *Intellectual Teamwork* (pp. 191–220). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hutchins, E. (to appear). *Cognition in the Wild*. Cambridge, MA: The MIT Press.

Koedinger, K. R. & Anderson, J. R. (1990). Abstract planning and perceptual chunks: elements of expertise in geometry. *Cognitive Science*, 14, 511–550.

Larkin, J. & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11, 65–99.

Lave, J. & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge: Cambridge University Press.

Mayer, R. E. (1989). Models for understanding. *Review of Educational Research*, 59, 43–64.

McDougal, T. F. (1993). Using case-based reasoning and situated activity to write geometry proofs. In *Proceedings of the 16th Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Narayanan, N. H. (Ed.). (1992). *Working Notes of the AAAI Spring Symposium on Reasoning with Diagrammatic Representations*. To appear as a Technical Report. Menlo Park, CA: AAAI Press.

Narayanan, N. H., Suwa, M., & Motoda, H. (1994a). A study of diagrammatic reasoning from verbal and gestural data. In *Proceedings of the 16th Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Narayanan, N. H., Suwa, M., & Motoda, H. (1994b). How things appear to work: predicting behaviors from device diagrams. In *Proceedings of the 12th National Conference on Artificial Intelligence*. Menlo Park, CA: AAAI Press.

Nersessian, N. J. (1992). How do scientists think? Capturing the dynamics of conceptual change in science. In R. Giere (Ed.), *Cognitive Models of Science*. Minneapolis, MN: University of Minnesota Press.

Nersessian, N. J. (1993). In the theoretician's laboratory: thought experimenting as mental modeling. In D. Hull, M. Forbes, & K. Okruhlik (Eds.), *PSA 1992, Volume 2* (pp. 291–301). East Lansing, MI: PSA.

Nersessian, N. J. (in press). Abstraction via generic modeling in concept formation in science. In N. Cartwright & M. Jones (Eds.), *Idealization in Science*. Amsterdam: Editions Rodophi.

Novak, G. S. & Bulko, W. C. (1993). Diagrams and text as computer input. *Journal of Visual Languages and Computing*, 4, 161–175.

Pea, R. D. (1993). Practices of distributed intelligence and designs for education. In G. Salomon (Ed.), *Distributed Cognition*. New York: Cambridge University Press.

Roschelle, J. (1990). Designing for conversations. In Working Notes of the AAAI Symposium on Knowledge-based Environments for Learning and Teaching.

Schwartz, D. L & Black, J. B. (1992). Shuttling between depictive models and abstract rules: induction and fall back (Teachers College Technical Report CCT92-2).

Tessler, S., Iwasaki, Y., & Law, K. (1993). Qualitative structural analysis using diagrammatic reasoning. In *Proceedings of the 7th International Workshop on Qualitative Reasoning about Physical Systems*.

West, T. G. (1991). *In The Mind's Eye: Visual Thinkers, Gifted People with Learning Difficulties, Computer Images, and the Ironies of Creativity*. Buffalo, NY: Prometheus Books.