DISTINGUISHING VISUAL IMAGERY AND SPATIAL REASONING

ROBERT K. LINDSAY

Mental Health Research Institute, University of Michigan, Ann Arbor, MI 48109, U.S.A.

The imagery debate in its most recent revival inquires into the "psychological reality" of pictorial representations, asking whether images serve any cognitive function in humans, whether knowledge representation is analog or digital, whether images serve their functions in a stand-alone manner or require tacit knowledge, and so forth. However, these questions have not moved far enough away from the original phenomenological puzzle (how and why do I "see" things that are not there) because they are ill-specified and lack an encompassing theoretical context. For these reasons the debate has not been resolved by psychological experiment and never will be without more precise theories to test.

Janice Glasgow joins those who have proposed that we address instead questions of how images can be represented and used by computers. The issue then becomes whether specific depictive representations, whatever form they might take, have computational advantages over alternatives such as predicate logic with deduction. The judgments must be based on comparisons of total systems, including the access processes, not merely the notation used. Later the psychologists can make of these results what they will.

Of course, there has already been interest in artificial intelligence (AI) on several problems of spatial reasoning such as navigation. Nonetheless, Glasgow's emphasis on the computational questions and imagery is a minority one. Classical AI, particularly knowledge representation and problem solving, has been almost exclusively wedded to search-based forms of deduction and descriptive representations. Since I have also been a member of the minority, I applaud Glasgow's efforts, both her proselytizing and her research, and on the general issues she and I appear to be in substantial agreement. Here I will contrast my own research with hers to provide some perspective on our differences.

Psychology and philosophy have placed emphasis on the percept-like nature of imagery, particularly visual imagery. However, the cognitive roles of most sensory features such as color, pitch, and saltiness are relatively limited in inferential richness and structure and are probably amenable to conventional methods of analysis. However, a consortium of sensory-perceptual abilities evolved to provide efficient ways to move about in space and deal with physical objects in real-time, and these abilities have come to include in man the means to make inferences, especially predictions, about non-current events because such abilities have clear adaptive advantages that are central to our success. My approach adopts the position that the cognitive role of imagery is primarily a result of the fact that visual imagery and, in less obvious ways, tactile, proprioceptive, and auditory imagery, involve the representation and use of the structure of space, time and mechanics.

Thus, I have chosen to avoid problems of nonspatial sensory features, and instead study computational methods that use spatial information in reasoning about well-defined tasks, e.g., to understand the role of diagrammatic reasoning in addressing mathematical problems (Lindsay 1988; 1989; 1992). Similar problems have long been addressed in AI by others, from Gelernter (1959) to Larkin and Simon (1987) and Novak and Bulko (1990), but without explicitly attempting to capture the structure of space in a uniform representational scheme.

We have clear mathematical descriptions both of space and of the mechanics of rigid bodies acted upon by forces. The most straightforward ways to represent space are based on these descriptions. One way of doing this is illustrated by the work of Kaufman (1991),

who has constructed an axiomatization of spatial and temporal structure along lines familiar in formal logical analysis. Using these axioms and standard proof methods, one might then deductively prove, for example, that a string can be used to pull but not push an object. However, 30 years of AI research has demonstrated that predicate calculus with deduction may fail to meet efficiency criteria in nontrivial domains.

My approach is to represent geometric objects algebraically and define processes that manipulate the structures in ways that cannot violate the structural properties of space. I have long advocated this "natural models" approach in other contexts (Lindsay 1961; 1963; 1973), and see it as the basis of the power of many familiar but less general AI techniques such as inheritance hierarchies. With such representational systems, one alternative to inference by deduction is inference by simulation, the incremental manipulation of representations and the "observation" of the inferences that follow from the structural constraints. If the representations reflect the underlying algebraic structure of that which they model, computational advantages might be achieved. Cognition by simulation is an alternative to the axiom-deductive form of computation, and is an idea with a long, prescientific history as well as many contemporary advocates. In the context of imagery and spatial reasoning Funt (1977; 1980), as far as I know, was the first to explore it with a specific computational model.

Glasgow's work is in this spirit, but rather than explicitly attempting to capture the mathematical structure of space, she proposes representations that capture some spatial properties and some additional features of visual imagery as well. Thus, in one sense her work is more general and nearer to the usual issues debated in psychology than is mine, but in another sense it is less focused and systematic on the issue of space itself.

In Glasgow's model, there are two working representations. She refers to the occupancy array as "visual" and the nested symbolic array as "spatial," and intends that "the visual subsystem is concerned with the geometric properties of images and the spatial subsystem is concerned with understanding spatial configurations." Elsewhere Glasgow and Papadias (1992) state that the spatial (symbolic) array makes explicit "symmetry and adjacency" and "preserves spatial and topological relationships . . . but not necessarily relative sizes and distances." The contrast of geometric and spatial properties is puzzling. She apparently intends to separate visual features plus metric features on the "visual" side from topology on the "spatial." However, the symbolic array includes a limited metric notion of direction, without which it would be unable to make inferences based on the transitivity of its directional adjacency. On the other hand, the quantification of direction is too coarse for many important inferences, and it is not clear how it could handle symmetry, which is a geometric predicate, or orthogonality, even of the symbolic array's two preferred directions. Furthermore, the visual array must preserve topology as well as geometry. Thus, Glasgow's division does not cleanly separate spatial from nonspatial properties. She has chosen a different cut, and we might wonder why.

One argument she offers is by reference to neuropsychological evidence that there are two visual pathways, one for "spatial" information and the other for "visual" information. That distinction is based on work of Ungerleider and Mishkin (1982) and others. The two visual pathways have sometimes been labeled "what" and "where." In brief, subjects with damage only to the "where" pathway can identify objects but not point to them, subjects with damage only to the "what" pathway can point to objects but cannot identify them. It has become common also to refer to these pathways as "visual" and "spatial." This is misleading because both are visual; they are after all visual pathways.

Furthermore, both are spatial as well. Note that "spatial" normally means more than "location in my perceptual field," which is roughly the information apparently dependent on the "where" neural system. However, the recognition of shape by the "what" system

requires identifying the relative locations of parts of the perceived object, and this is no less spatial than identifying the relative locations of different objects in a larger scene. Indeed, both location and object shape are properties that are meaningless without knowledge of space.

The sets of contrasting behavioral tasks reported in the literature do not make a sufficiently clear computational distinction between the functions of the two pathways, but what examples there are (Farah *et al.* 1988) have not been shown to correspond to the Glasgow division. For example, without the occupancy array would the system be able to determine whether Tennessee is closer to Oregon or to Maine, a "where/spatial" pathway ability?

These relatively minor differences aside, our approaches have much in common. However, many interesting issues, such as how these methods combine with goal-directed problem solving, are still difficult, open questions.

REFERENCES

FARAH, M. J., K. M. HAMMOND, D. N. LEVINE, and R. CALVANIO. 1988. Visual and spatial mental imagery: dissociable systems of representation. Cognitive Psychology, 20:439–462.

FUNT, B. V. 1977. Whisper: a problem-solving system utilizing diagrams and a parallel processing retina. *In* Proceedings of the Fifth International Joint Conference on Artificial Intelligence (IJCAI-77), pp. 459–464.

FUNT, B. V. 1980. Problem solving with diagrammatic representations. Artificial Intelligence, 13(3):201–230.

GELERNTER, H. 1959. Realization of a geometry theorem proving machine. *In* Proceedings of the International Conference on Information Processing, UNESCO House, Paris, pp. 273–282.

GLASGOW, J., and D. PAPADIAS. 1992. Computational imagery. Cognitive Science, 16(3):355–394.

KAUFMAN, S. G. 1991. A formal theory of spatial reasoning. *In* Proceedings of the Second Conference on Knowledge Representation and Reasoning, pp. 347–356.

LARKIN, J., and H. A. SIMON. 1987. Why a diagram is (sometimes) worth ten thousand words. Cognitive Science, 11:65–100.

LINDSAY, R. K. 1961. Toward the development of a machine which comprehends. Doctoral dissertation, Carnegie-Mellon University, Pittsburgh, PA.

LINDSAY, R. K. 1963. Inferential memory as the basis of machines which understand natural language. *In* Computers and thought. *Edited by* E. A. Feigenbaum and J. Feldman. McGraw-Hill, New York, pp. 217–233.

LINDSAY, R. K. 1973. In defense of ad hoc systems. *In* Computer models of thought and language. *Edited by* R. Schank and K. Colby. Freeman, San Francisco, pp. 372–395.

LINDSAY, R. K. 1988. Images and inference. Cognition, 29:229–250.

LINDSAY, R. K. 1989. Qualitative geometric reasoning. *In* Proceedings of the Eleventh Annual Conference of the Cognitive Science Society, Ann Arbor, MI. Lawrence Erlbaum, Hillsdale, NJ, pp. 418–425.

LINDSAY, R. K. 1992. Diagrammatic reasoning by simulation. *In* Working Notes of the AAAI Spring Symposium on Reasoning with Diagrammatic Representations, Stanford University, Palo Alto, CA. *Edited by* N. Hari Narayanan, pp. 130–135.

NOVAK, G., and W. BULKO. 1990. Understanding natural language with diagrams. *In* Proceedings of the Eighth National Conference on Artificial Intelligence (AAAI-90), pp. 465–470.

UNGERLEIDER, L. G., and M. MISHKIN. 1982. Two cortical visual systems. *In* Analysis of visual behavior. *Edited by* D. J. Ingle, M. A. Goodale, and R. J. W. Mansfield. MIT Press, Cambridge, MA.