Commentary on:
"The Imagery Debate Revisitied: A Computational perspective,"
by Janice I. Glasgow, in: *Computational Intelligence.*
Special issue on Computational Imagery,
Vol. 9, No. 4, November 1993

Varieties of Formalisms for Knowledge Representation

Aaron Sloman
School of Computer Science
The University of Birmingham
Birmingham, B15 2TT
England
A.Sloman@cs.bham.ac.uk
Phone: +44-21-414-3711

—

Abstract

Whilst I agree largely with Janice Glasgow's position paper, there are a number of relevant subtle and important issues that she does not address, concerning the variety of forms and techniques of representation available to intelligent agents, and issues concerned with different levels of description of the same agent, where that agent includes different virtual machines at different levels of abstraction. I shall also suggest ways of improving on her array-based representation by using a general network representation, though I do not know whether efficient implementations are possible.

1 Introduction

Janice Glasgow's position paper discusses debates concerned with the relative merits of using propositional and visual or spatial representations in a variety of contexts and for a variety of tasks. Her conclusion is that "there are advantages to extending and enhancing descriptive knowledge representation techniques to include functions for generating, transforming and inspecting spatial representations of images". She recommends the study of these additional representational forms and techniques not only for engineering tasks (i.e. designing intelligent artifacts) but also for the scientific task of explaining human abilities, especially abilities that appear to depend on our use of visual imagery.

In what follows I shall make a number of criticisms of her paper, but let me start by making it plain that I broadly sympathize with her position, and have been making closely related points in several papers and one book over the last 21 years (Sloman 1971; 1975; 1978; 1985; 1989; 1993). Indeed, I conjecture that a great deal of human mental ability depends crucially on sophisticated visual representational and processing capabilities that we share with many other animals with powerful visual systems, but which are still not understood at all. I shall enlarge on this below.

My criticisms of Glasgow's paper, therefore, are criticisms from a standpoint that is very similar to hers. The main points are as follows:

1. Whereas Glasgow contrasts two main forms of representation, propositional and depictive, I think there are *many* different forms of representation whose properties need to be explored and contrasted (Sloman 1985).

2. There are many ambiguities and problems that she does not mention, concerning what it *means* for an agent or mechanism to be using one or other type of representation, especially where different virtual machines in an implementation hierarchy exist within the individual concerned.

3. The use of rectangular arrays to represent spatial information seems to involve arbitrary and unfortunate restrictions that might be overcome by a more general network-based representation.

Ideally, points (1) and (2) need to be embedded in a full theory of the nature and role of representations in intelligent agents. There is no space for such a full discussion here, so I can merely hope to highlight the main points.

2 Requirements for a general theory of representation

The concept of representation is one that is widely understood intuitively, but is very hard to define in a precise way. This is partly because it covers a very wide range of cases, and also because its boundaries are to some extent indeterminate. Moreover, nothing is intrinsically a representation or not a representation: anything can be a representation for a behaving system with suitable capabilities.

A representation for a behaving system is an object or structure either within the system or in its environment that encodes information that is accessible or manipulable by the system, and usable by it in order to perform some function or achieve some goal.

The information encoded may be *control* information (e.g. programs or other structures that generate or modify behavior) or *factual* information about the agent or its environment. The factual information may be *particular*, that is, concerned with certain individual objects, agents, places, regions, times, etc. or *general*, that is, applicable to whole classes of individuals. General information may vary in its degree of specificity: the most specific information is applicable to or concerned with relatively small, restrictively defined classes, whereas less specific, or more general, information is concerned with wider classes.

The representing objects and structures may have very different kinds of ontological status. For example a representation may be an explicitly isolatable physical object, such as a flag on a ship, or it may be implicit in relations between objects (for instance using the direction between two objects that do not represent anything to convey a direction in which to move), or it may not be a physical object or relationship at all, but exist only in a *virtual* machine which is only indirectly implemented in physical objects. Computer data structures, such as arrays, symbol networks or lists of propositions are typically entities in such virtual machines, and operations on them are operations in a virtual machine, such as the Lisp or Prolog virtual machines.

These considerations are relevant to the position paper in several ways, which will now be discussed separately.

2.1 Complexity correspondences between virtual machines

Where representations exist in virtual machines, they may be implemented in lower level virtual machines which also make use of non-physical objects. Thus, a Lisp or Prolog virtual machine is typically implemented in something like a Vax virtual machine or a Sparc virtual machine where the lower level virtual machine may itself be implemented either directly in hardware, or only indirectly via microcode.

When physical structures directly or indirectly implement structures in an abstract virtual machine there is no need for the physical structures to be isomorphic with the abstract structures. For example a ‘sparse array’ in a virtual machine may have many more components than the corresponding underlying implementation. A sparse array is one in which the vast majority of components have some ‘default’ value (e.g. 0) so that all that need be represented explicitly (using clever indexing schemes) is the small subset of cells whose values differ from the default. A user (whether person or program) of an efficiently implemented sparse array of size 1,000,000 by 1,000,000 may be quite unable to tell the difference between an implementation using a million million cells and one using efficient sparse coding techniques. In either case the *virtual* machine structure would have the same number of cells with values, even though they differ enormously in the size of the underlying *physical* structures used. (A virtual sparse array could actually be too large to fit into the physical universe if implemented isomorphically.)

Conversely, the indexing structures and mechanisms required to make a virtual structure work efficiently can take up space that does not correspond to components of the structure itself. So in some cases the physical implementation may use significantly more components than the virtual structure.

Similar comments can be made about processing steps. There need be no systematic correspondence between numbers of steps in a high level virtual machine and numbers of physical operations implementing those steps. In a sparse array, or in an object-oriented hierarchy making use of inherited default values, it may be possible to change the contents of a huge number of cells in a virtual structure in memory simply by altering one cell in a lower level implementation machine.

Conversely a simple operation in a high level virtual machine may require many low level steps and ultimately a large number of physical steps. Thus for example, consider a network containing lists of words comprising sentences at its nodes, such as

[The hungry cat chased the frightened mouse]

A single step in a high level virtual machine, such as unlinking this list from the network, could correspond to a very large number of operations in a lower level implementation machine, especially if the network structure uses a rich index to facilitate searches of various kinds: removing one item may require a considerable number of alterations in the indexing structures that associate words and phrases with locations in the network where they are used.

This lack of correspondence between virtual structures and proceses and the low level physical structures and processes that implement them is one reason why the phrase “physical symbol system” introduced by Newell and Simon to describe AI models of cognition is a source of great confusion, and should be dropped by the AI research community.

For these reasons, the discussion of ease of programming and run time efficiency of visual or spatial representations needs to take account of the implications of using virtual machines that can be implemented in different ways: if what is conceptually simple actually requires enormous physical complexity at run time, then it may be practically useless.

A good example of this is the ease with which it is possible to program certain problems, e.g. constraint satisfaction problems, in Prolog. Unfortunately the clearest Prolog programs often take inordinately long to run because of the simple-minded depth first search strategy built into Prolog. Of course one can implement different strategies in a Prolog program, though often at the cost of ease of programming.

Glasgow claims (Section 2.3) that her symbolic arrays provide “succinct and holistic encoding as well as their provision for updating and change” and goes on to claim that “these advantages translate into computational efficiency". Similarly in Section 4.1.1 when discussing transitive inference she claims that it is possible to "read off" certain facts directly from a symbolic array representation. It is far from obvious exactly what is being claimed here. In order to support efficiency claims regarding the use of array-manipulating virtual machines in contrast with some other type of virtual machine (e.g. a logic machine), it is necessary to show both (a) that the complexity of processing on certain tasks is much reduced by using the array virtual machine and (b) that the complexity does not resurface in the implementation of the array machine in a lower level machine, as could easily happen, for example, if the indexing that facilitates rapid traversal across the arrays made use of list-searching in the implementation virtual machine. Claims of type (a) are trivial, since an arbitrarily complex operation can always be encoded in a single named function, defining a new virtual machine operation, which makes certain tasks very simple and efficient from the programmer's viewpoint.

2.2 Virtual machines in the brain

Second, even if both (a) and (b) have been established it still leaves open the possibility that the lower level machine uses propositional representations and inference techniques.

The question whether propositional or depictive representations and techniques are used by a particular machine, or the human brain, is ambiguous insofar as the answer may be different at different levels. For example in a machine using the spatial array processing virtual machine described by Glasgow something quite different will be going on at the machine code level. In principle, the lowest level machine might be a logic processing machine, in which case a non-propositional virtual machine could be implemented in a lower level propositional machine, a possibility not allowed for by Glasgow.

This point was made by Hayes (1974) in his attack on my paper (Sloman 1971). It led to a slightly revised version of the defense of analogical representations in (Sloman 1975), acknowledging that they could in principle be implemented in propositional structures, as long as the relevant indexing operations and neighborhood relationships are implemented efficiently.

2.3 The unreliability of introspection

Third, when human subjects report the use of visual or spatial images for solving problems or performing tasks, this needs to be taken with a pinch of salt. Although we clearly have *some* internal self-monitoring capabilities, these will have evolved to serve specific biological needs and need be no more accurate at telling us what's really going on in our minds, for scientific explanatory purposes, than visual perception is at telling us about the internal physical structure of objects in our environment. (What is really going on contrasts with the phenomenology of what is going on, which is what introspection reveals accurately, by definition.) Introspective reports can, at best, report some properties of a high level virtual machine, and even if this is some kind of spatial image processing machine, that does not rule out the possibility of an underlying implementation that is completely different, for instance using propositions and logical inferences. (There may be other evidence ruling that out.)

To what extent the human brain uses virtual machines, what sorts of virtual machines, and how many layers of implementation there are, are all still open questions. Introspection certainly suggests that we (sometimes) use images that we can manipulate, but this could merely report some abstract properties of a high level virtual machine, or it could be completely deceptive. It is clear that some of the states that people report as states in which they are inspecting inner pictures do not have the same properties as states in which real external pictures are inspected. For example, it may be easier to read items off the inner image traversing it in one direction than another direction. This suggests that an anisotropic internal representation is used (e.g. using linked lists), unlike computer arrays or real paintings.

The only thing that is certain is that human (and some other animal) brains have extremely powerful ways of using visual information that we do not understand at all, and which may not be capable of being represented efficiently on conventional computers. In fact, it is likely that we have not yet even learnt to ask the right questions.

3 Towards a generalized theory of representations

A full theory of representations, whether physical or virtual, would have to distinguish arbitrary and systematic representations. Human beings are certainly able, in appropriate contexts, to use anything to represent anything, for instance, using a pebble to represent a person. Although this is interesting and worth explaining, the more profound case is the use of a *system* of representations, of which a set of strings of words in a language is one case, and a set of networks for representing places and routes between them another. The use of hierarchies of arrays as described by Glasgow is yet another system.

There are (at least) two important implications of using a *system*, namely the ability to encode *novel* information, and the ability to solve problems by transforming representations in a principled way that generates new representations within the system. Such transformations can play a role in making inferences, forming plans, creating designs, and so on. (This goes against the frequent comment made by philosophers that the essence of representation or symbolism is arbitrariness.)

3.1 Syntax

A precondition for systematicity is that representations can vary, and the forms of permitted variation are defined by a *syntax*.

It is common to stress that what makes a representation pictorial is a type of correspondence between representing structures and what is represented. What is not so often appreciated, though Glasgow does mention this, is the importance of available transformations. In particular, what I have previously called Fregean, or applicative, representations permit substitution of sub-structures without changing the remainder of the representation, whereas this is not possible for all forms of representation: e.g. in a picture including a giraffe in a field, it is not normally possible to replace the giraffe with an elephant without affecting something else in the image, because they occlude different parts of the background. By contrast there's no background that shows through when a large sub-expression is replaced by a smaller one in an applicative representation, such as a logical formula or computer program.

There are several additional dimensions in which representational systems can differ regarding permitted modes of variation, including:

- whether continuous or only discrete transformations are permitted,
- whether transformations change the complexity of representations or not (for example a representation in the form of a fixed size vector of numerical values can be transformed only by changing the numbers, and this will not alter the complexity, whereas modifying a structural description or parse tree can change its complexity),
- whether new 'atomic' elements can be introduced or whether the basic symbols are fixed (e.g. natural languages contain an alphabet permitting new words to be formed when needed, as well as novel sentences composed of old words).

One reason why differences in kind of variability are important in intelligent agents, as pointed out in Sloman (1971), is that different representations that share metaphysical or epistemological adequacy (as defined by McCarthy and Hayes (1969)) can differ in heuristic adequacy because of differences in the ways that they constrain search spaces. A representation that does not allow irrelevant options to be constructed can generate a much smaller search space than one that does allow them. Thus, whereas a logical notation would allow a conjunction of assertions about **A** being close to **B**, **B** being close to **C** and **A** being a very long way away from **C**, a spatial representation with the usual metrical properties would not, and this could be an important constraint in some problems.

3.2 Semantics

A representational system includes not only syntax, but also semantics, i.e. what information is encoded, or what things are represented. It is often assumed that the essence of pictorial representation is isomorphism with the 'scene' represented, and Glasgow writes as if this is so, at least in the cases she considers. However, it is very important to note that in general this is not a requirement: for example 2-D images can represent 3-D configurations, with which they are not at all isomorphic, since they do not include relationships like depth or 'partially hidden' even though they represent them.

The complexity of non-isomorphic semantic correspondence between image and scene is familiar to those who work in computer vision. Local substructures in 2-D images are usually inherently ambiguous and can be interpreted only in terms of context, where the context also consists of fragments that are ambiguous. When this inherent ambiguity and mutual dependence is combined with noise, partial occultation of one object by another, and self-occultation of invisible parts of an object by visible parts, the resulting 'semantic' correspondences between images and what they depict can only be defined in terms of very complex constraint-optimizing rules that include reference to background knowledge of the relevant portion of the world. (Constraint satisfaction would suffice were it not for the noise typically found in natural images.) It may be impossible to specify precisely the correspondence between image and scene without introducing a collection of intermediate layers of representation such as is common in computer vision systems, especially where images are not produced by a simple projection mechanism, but include conventional elements as in cartoons and paintings (for more on this see (Sloman 1989)).

Typical attempts by logicians to generalize their semantic ideas to cope with pictures fail to take account of this complexity, and the same appears to be true of Glasgow's account (in Section 3.4) of depictive representations in terms of mappings between parts and relations.

Another example where algorithmic compositional semantics breaks down, of course, is natural language, where words and phrases are heavily context dependent and often a great deal of knowledge about the speaker, the culture and the specific circumstances of utterance may be required for disambiguation.

A full theory of semantics would not merely have to explain how different forms of representation work in conveying *factual* information about what is the case, but would also have to account for other kinds of information, e.g. *control* information. For example, a computer program embodies a great deal of control information about what to do when, and it seems very likely that many learnt action sequences including how to say things, how to perform routine tasks, how to do the eightsome reel and other dances, how to sing songs and how to play musical instruments are all encoded in formalisms whose primary semantic role is control not description.

Insofar as Glasgow is drawing attention to the need to consider different types of representational formalisms, the point is strengthened by drawing attention to the variety of types of semantics.

3.3 Pragmatics

That last point is related to the fact that information can be encoded for many different purposes. A full theory of representations needs to take account not only of syntax and semantics but also *pragmatics*, that is, the variety of purposes or functions for which the representations can be used. Without a study of the pragmatics of a system of representation (as used by a particular agent or class of agents) it will not be possible fully to appreciate the pros and cons of using that system rather than some other.

Glasgow refers to Marr (1982) who describes the 'quintessential fact of human vision – that it tells about shape and space and spatial arrangement' of objects in the 3-D environment. I have argued (Sloman 1989) that this is a grossly oversimplified view of the role of vision, since the outputs of a visual system can include control signals (e.g. posture control, saccadic control), information about 2-D relations in the optic array that are useful for fine control of behavior, such as sighting a gun or moving in a straight line, and, above all, information about a host of what Gibson (1979) called 'affordances', which includes not only geometrical properties of objects but also their causal powers, functional roles, and even the mental states of other agents.

An illustration of this point is provided by the differences between familiar ambiguous pictures which cause visual 'flips': the Necker cube flips between interpretations that differ only in the ways that Marr is concerned with, i.e. geometrical properties and relations. By contrast when the familiar "duck-rabbit" picture flips what changes is not information about "shape and space and spatial arrangement" but far more abstract information about what sort of animal is depicted, the functional roles of things depicted (the rabbit's ears, the duck's bill) and which way the animal is *facing*, an attribute that implicitly refers to the object as an agent capable of moving, perceiving things, and so on.

It seems at least possible that the visual system has some way of combining factual and control information so that during skilled performances of various kinds (e.g. leaping across rocks, sporting contests) it simultaneously provides information about what is in the environment and what to do about it.

Although Glasgow's paper is not primarily about vision, it is concerned with the power of visual representations, and, like Marr, apparently assumes that visual representations are concerned exclusively with the representation of information about spatial arrangement and motion. If, however, real biological visual systems are far more powerful and versatile than that, and can encode and manipulate a host of additional kinds of biologically relevant information, including *modal* information about what is and is not possible in the situation, as described in Sloman (1989), then perhaps that may help to explain why people believe they are using visual imagery when they think about all sorts of problems that are not inherently visual. For example, the difference between depth first and breadth first search is often best understood visually, though there is a non-visual abstract characterization of the difference in terms of operations on queues and operations on stacks. Similarly, many abstract mathematical ideas, including the concept of an unending sequence of numbers, are often thought of in visual terms, though real vision cannot cope with infinite sets. How do mathematicians visualize infinite sequences?

If the human visual mechanism does contain some very powerful representational apparatus, including transformation mechanisms, then perhaps people in AI and cognitive science who are interested in explaining or modeling intelligence would do well to study visual systems since they may provide far more powerful means of manipulating information for a variety of purposes than anything so far developed in computational vision experiments, including the experiments reported by Glasgow. Unfortunately, I do not think anyone understands much about how vision works, since existing models seem to be completely lacking in the generality, flexibility and speed that characterize human and (some) animal vision.

4 Glasgow's representations

Glasgow reports on computational experiments using representations based on hierarchies of rectangular arrays. It is very common for computer-based image processing and interpretation systems to make use of rectangular arrays because these map in a simple way onto the linear address-spaces typically provided by widely available general purpose computer architectures. Moreover, as her examples show, rectangular arrays can be used for solving certain simple kinds of problems involving spatial relationships, where all the directions involved fall into two orthogonal sets (left—right and up—down, or east—west and north—south).

The problem is that rectangular arrays quantize space in a particular way that may not be suitable for all tasks. Animal retinas demonstrate a wider variety of ways of dividing up 2-D space, including the use of approximately concentric rings of receptive fields whose diameters vary with the distance from the center. As Funt (1983) has shown this structure can be useful for certain operations involving rotation, expansion or contraction of images. Overlapping fields can also be useful for some purposes.

Both Funt's and Glasgow's representations attempt to map represented items onto locations on a 2-D surface, albeit at varying resolutions. Both representations have the restriction that each location in the representation has neighbors in only two orthogonal directions, along cartesian co-ordinates in the case of Glasgow, and along radial and tangential directions in the case of Funt. This is a serious restriction when dealing with structures requiring a richer set of neighborhood relationships. Both Glasgow's and Funt's representations seem to be special cases of a more general class of representations that preserve partial ordering information by mapping relationships like connectivity on a transport network, onto arcs in a network in the computer, allowing as many neighbors in as many directions as needed. For example a representation of a road network could use some nodes where there are two links to neighbors, some where there are three, some where there are four, and so on. An electronic communications network could use much higher node connectivity.

If a visual reasoning system were to use a generalized network rather than a regular array, it could combine spatial and other relationships, for example causal relationships such as support, containment, transfer of torque, pivoting, etc. Just as Glasgow's rectangular arrays are hierarchical, with some array cells occupied by higher resolution arrays, so too can a network be hierarchical with some nodes occupied by sub-nets giving more detail. This would support similar operations of zooming in and out, and using each level in the hierarchy as a sort of index to the more detailed information at lower levels. A more general representation would also permit cross links supporting paths that are frequently traversed between one portion of a detailed network and another. In particular, instead of being stuck with inaccurate relationships arising out of the use of a rectangular grid, precise direction and distance links could be inserted whenever they are required, adding important redundancy, as discussed in the next section.

Obviously, general operations for inserting, comparing, searching and merging such networks will be more complex than the operations on rectangular arrays described by Glasgow. Nevertheless, it should be possible in principle to implement generalized versions of all the operations described by Glasgow, e.g. generating networks, adding or deleting links or nodes, composing and superimposing networks (by merging common links and nodes), scanning, rotating, focusing and unfocusing attention (by moving up and down the hierarchy). In fact a generalized network would support a richer variety of scanning operations by allowing different sorts of links to be traversed in different scans, rather than efficiently supporting only horizontal and vertical scans. If certain types of links form chains without branches the nodes that they join can be given numerical 'coordinates', thereby permitting rapid transitive inferences of exactly the sort discussed by Glasgow (Section 4.1.1).

Rotation, however, could not be defined in general, except for subnets incorporating metrical information. However, Glasgow does not point out that rotating a sub-portion of a rectangular array by angles that are not multiples of 90 degrees is actually a complex operation that introduces distortions and loses information, which is one reason why Funt used concentric rings of cells. A more general representation could include frequently used 'rotation-pathways', thereby superimposing something like Funt's and Glasgow's representations.

Such a generalized network-manipulating mechanism would provide what Glasgow describes as 'succinct and holistic encoding' of important and frequently used operations by providing a suitable high level virtual machine in which these are simple operations, without the restrictions of her rectangular array representation.

The network mechanisms proposed here are not original. They are very similar to network representations that have frequently been used in AI and elsewhere. For example, Glasgow (Section 4.2) discusses the use of such networks by Winston in his work on learning structural descriptions. (For some reason she claims in Section 3.1 that Winston's networks and Minsky's frame systems use propositions to express relationships, which is surely a mistake, even though they can be translated into propositions, just as her own arrays can.) She reports that Winston had difficulties matching sub-networks within his network and then suggests that the use of her array representation can reduce the complexity of the matching process. There are two problems with this claim, namely that it is not obvious that arrays can reduce the problem significantly as far as cluttered real-life scenes are concerned, and secondly that insofar as arrays do help their structure can be simulated by a subset of links in a general purpose network, which is exactly what an experienced software engineer would do.

Making these generalized network operations truly efficient at the lowest implementation level might require the design of new computer architectures. There is some scope for the use of parallelism in network processing, just as there is with the array representation. In both cases, however, the fact that the hierarchy of resolutions is variable and can change dynamically means that there cannot be a fixed allocation of processors to 'locations' in the representation, which complicates the task of implementing such a system on highly parallel computers. Nevertheless, human visual processing copes with rich and complex, yet very rapidly changing, scene and image structures, which seems to be an existence proof of the possibility of some sort of mechanism that meets the need, even if we have no idea at present how it works!

5 The need for redundancy

One unobvious aspect of the way in which a spatial representation helps constrain problem solving is that it supports a certain kind of redundancy of information by making certain relations explicit (in the relevant virtual machine) that might otherwise have to be inferred. For example, an array representation allows rapid traversal from a cell to its neighbors, either by following pointers or by a simple computation of memory offsets at the machine level, depending on the implementation used for the array. As Glasgow herself points out (Section 2 and elsewhere) the array representation does not necessarily encode any information that cannot be encoded propositionally: it is not superior in expressive power. All it does is provide *rapid* access to the information via *conceptually simple* operations. She also admits that clever indexing schemes can provide rapid access for propositional representations by eliminating the need for lengthy list searches (compare Hayes (1974) and Sloman (1975)).

I believe that there is an important point that she has not noticed, namely that this requirement for redundant encoding in forms that support rapid access to relevant information is a *general* requirement for intelligence. This is illustrated by the fact that it is not possible for anyone who has merely understood Peano's five axioms for arithmetic, or some other equivalent set of basic arithmetical principles, to make sensible use of arithmetic. Instead it is necessary to *memorize* many particular consequences of these basic axioms in such a way that they can be rapidly retrieved when needed. This is the fundamental truth on which many 'old fashioned' educational practices are based, even if they are non-optimal for educational purposes because they often kill motivation. More 'radical' practices often throw out the baby with the bathwater because they do not replace the old function of building up a rich store of redundant but rapidly accessible information. (Compare the discussion of learning facts about the number sequence in Sloman (1978), Chapter 8.)

If the human brain provides general methods for storing information using efficient mechanisms for creating cross linkages and indexes that support rapid access by storing redundant information, then it could turn out that the phenomena that seem to be based on something like Glasgow's array representations are in fact based on the more general network manipulating mechanisms.

I do not know whether such general network mechanisms could provide a sufficiently fast and powerful information processing engine, or whether some totally different way of looking at the problem is needed.

6 Conclusion

Although I agree with much of the spirit of what is claimed in Janice Glasgow's position paper, I have had to find fault with it in part because of what it doesn't say. In particular the paper discusses propositional and spatial representations without pointing out that there is a huge variety of different forms of representation involving differences in syntax, semantics and pragmatics, and that intelligent agents may need to use many different forms for different purposes. Moreover, although claims are made about efficiency of the array-based representations these are justified only at the level of the array-processing virtual machine, leaving open questions about the total efficiency including the low level implementation. Furthermore, I have suggested that the use of arrays imposes arbitrary restrictions on the kinds of relationships that can be represented easily, whereas there is no reason to believe that human visual systems have these restrictions. I have suggested that we can generalize the advantages of rectangular arrays by using less constrained network representations, combined with embedded coordinate systems associated with certain classes of links that can be scanned, though it is not at all clear how efficiently the required general network-manipulating procedures can be implemented on conventional computer hardware.

I have also pointed out that the advantages claimed for symbolic spatial representations are in fact a special case of the advantages of representations that are highly redundant in order to provide efficient access paths. In a fuller discussion it would also be necessary to relate this point to the distinction between solving problems by manipulating *descriptions* of some class of entities, and solving them by *simulating the behavior of those entities*. Most of the contrasts pointed out by Glasgow amount to the contrast between making inferences using descriptions and simulating processes. Which strategy is better will depend on the kind of problem and what sorts of analytical techniques have been developed for manipulating descriptions.

Finally, although I agree with Glasgow that the manipulation of spatial and visual representations plays a very important role in human mental processes (including mental processes in congenitally blind individuals who, after all, still have most of the neural apparatus developed for visual purposes by evolution), nevertheless I suspect that we do not really understand at all how visual information (spatial and non-spatial) is encoded in the human brain in order to make these capabilities so generally and rapidly applicable. The apparent success in tracing the low level neural representation of some image features may have diverted attention from the larger issue of how the brain represents properties and relations of perceived objects out there in the world.

REFERENCES

FUNT, B. V. 1983. A parallel-process model of mental rotation. *Cognitive Science*, 7(1): 67-93.

GIBSON, J. J. 1979. The ecological approach to visual perception. Lawrence Erlbaum Associates, Hillsdale, NJ, reprinted 1986.

HAYES, P. J. 1974. Some problems and non-problems in representation theory. *In* Proceedings of the AISB Summer Conference, University of Sussex. *Reprinted in* Readings in knowledge representation. *Edited by* R.J. Brachman and H.J. Levesque. Morgan Kaufmann, Los Altos, CA, 1985, pp. 4-21.

MARR, D. 1982. Vision. W.H. Freeman, San Francisco, CA.

MCCARTHY, J. and P.J. HAYES. 1969. Some philosophical problems from the standpoint of artificial intelligence. *In* Machine Intelligence 4. *Edited by* B. Meltzer and D. Michie. Edinburgh University Press.

SLOMAN, A. 1971. Interactions between philosophy and A.I.: the role of intuition and non-logical reasoning in intelligence. *In* Proceedings of the Second International Joint Conference on Artificial Intelligence. *Reprinted in* Images, perception, and knowledge. *Edited by* J.M. Nicholas. Reidel, Dordrecht-Holland, 1977.

SLOMAN, A. 1975. Obituary notice: analogical representations. AISB Quarterly. *Reprinted as* Afterthoughts on analogical representations. *In* Readings in knowledge representation. *Edited by* R.J. Brachman and H.J. Levesque. Morgan Kaufmann, Los Altos, CA, 1985, pp. 432-439.

SLOMAN, A. 1978. The computer revolution in philosophy: philosophy of science and models of mind. Harvester Press and Humanities Press, now out of print.

SLOMAN, A. 1985. Why we need many knowledge representation formalisms. *In* Research and development in expert systems. *Edited by* M. Bramer. Cambridge University Press, pp. 163-183.

SLOMAN, A. 1989. On designing a visual system: towards a gibsonian computational model of vision. Journal of Experimental and Theoretical Artificial Intelligence, 1(4): 289-337.

SLOMAN, A, 1993. The role of representations in intelligent systems. *In* Proceedings Workshop on Hybrid Reasoning, International Joint Conference on AI, Chamberey, France. (To appear)