

THE COMPLEXITY OF INFORMATION-PROCESSING TASKS IN VISION

JOHN SYMONS

University of Texas, El Paso

(Forthcoming in Carlos Gershenson, Diederik Aerts, and Bruce Edmonds (eds.) *Philosophy and Complexity: Essays on Epistemology, Evolution, and Emergence*. Singapore: World Scientific)

The basic presupposition of cognitive science is that mental life is more complicated (or complex) than it appears; there is more to memory, attention, perception and the like, than meets the inner eye. In most mainstream cognitive science, the hidden workings of the mind are treated as information processing systems. In defense of their approach, cognitive scientists often point to the importance of articulating the complexity of the information processing processes underlying cognition and perception. The use of complexity in defense of mainstream cognitive science deserves detailed examination, because of its importance and because it is more philosophically problematic than has been recognized to date.

This paper focuses primarily on the use of complexity in the science of vision. The information-processing model gained supremacy, in large part, because of its success in the science of vision. Visual perception was one of the preeminent areas of investigation and served as the testing ground for the most important conceptual presuppositions of the discipline.

By the mid 1980s, it was commonly assumed that a satisfying explanation of vision should permit us to program a computer to see. From the computationalist perspective, a satisfying explanation of some aspect of mental life is one that allows us to understand the phenomenon in question as a series of steps articulated in ways that can easily be presented to our colleagues without the appearance of question-begging.

However, the attempt to articulate the mind's informational complexity raises a series of fundamental questions. In the cognitive science of vision, we can ask, for instance:

- (1) What is the relationship between the information-processing complexity that we hypothesize and the actual embodied act of seeing?
- (2) What predictive power do we gain via an articulation of the information-processing complexity of the visual system?
- (3) How can we correctly capture the information processing complexity of a particular system?
- (4) Is it always necessary to provide an information-processing account of a system in order to provide a proper explanation?

One can easily imagine explanatory contexts where the information-processing complexity that cognitive scientists hope to understand is simply not relevant. After all, most scientific explanations do not take an information-processing form and for the most part, it is not obvious why one should opt for an explanation of this kind. There might be some justification for an information-processing style analysis in engineering contexts, but even here, it is not clear that engineers would really need to know the information-processing properties of the solutions they develop. In addition to their emphasis on the importance of non-enthymatic, proof-like processes in explanations, the computationalist commitment to an information-processing accounts of psychological and perceptual processes results from a representationalist conception of mind. As we shall see, the notion of information-processing complexity serves as the intersection for a tangled set of basic questions in the philosophy of cognitive science.

This paper examines some of the arguments in support of the information processing approach to the complexity of visual perception. David Marr's defense of the approach serves as the focus of the analysis. Marr's posthumously published book *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information* (1982) presents the most prominent computational theory of visual perception. *Vision* provides an overview of work conducted by Marr and his colleagues at the Artificial Intelligence Laboratory at the Massachusetts Institute of Technology from about 1973 to the late 1970s. Written under the pressure of serious illness and completed after his death by friends and colleagues, it is widely regarded as the most important book in cognitive science. *Vision* outlines a set of methodological and theoretical principles that have helped to consolidate the self-image of cognitive science as a discipline. Marr's computational theory of the visual system is seen as providing a model for theoretically sophisticated investigation into the biological mechanisms responsible for mental life. So, for example, Steven Kosslyn claims that

Marr's view provided "... an enormous advance over much of the previous work in neuroscience, which tended to minimize theory, and over most of the previous work in artificial intelligence, which tended not to rest on careful analyses of a problem." (Kosslyn 1994,35) For the purposes of this paper, Marr's text provides is an influential example of the reasoning that shaped the prevailing view of complexity in the visual system.

1. Recognizing Complexity

Accounts of the visual system from the computational perspective almost invariably begin by contrasting their approach with earlier attempts to provide "direct" or "naïve realist" theories of perception. Central to this contrast is the notion of complexity. Computationalists see their work as uncovering the complexity of the information-processing task involved in visual perception that previous theorists had neglected. When scientists are introducing and defending their approaches to vision, they regularly begin by taking a stand on the notion of complexity. So, for example, Alan Yuille and Shimon Ullman begin their introduction to computational theories of low-level vision as follows:

Vision appears to be an immediate, effortless event. To see the surrounding environment, it seems we need only open our eyes and look around. However this subjective feeling disguises the immense sophistication of the human (or animal) visual system and the great complexity of the information processing tasks it is able to perform (Yuille and Ullman 1990, 5)

Computationalists understand their predecessors as having ignored the complex process underlying visual experience. Nicholas Wade and Michael Swanston also echo this emphasis on the underlying process when they write that: "The world around us seems concrete and immediate, and our ability to perceive it is easily taken for granted... Clearly, there must be some process that gives rise to visual experience." (Wade and Swanston 1991, 1) By contrast with their predecessors' neglect, they contend that the computational approach to perception as offers the best route to recognizing and untangling the information processing complexity of the visual system. In his introduction to cognitive science, Paul Thagard explicitly connects our recognition of complexity to the problem of crafting a computer program that sees:

For people with normal vision, seeing things seems automatic and

easy. You look at a room and immediately pick out the furniture and people in it. The complexity of vision becomes apparent, however, when you try to get a computer to do the same thing." (Thagard, 1996, 96)

Introductory discussions of the science of vision show a striking level of uniformity in their theoretical presuppositions with respect to complexity. They echo a long tradition in cognitive science and cognitive psychology which takes a two pronged approach to defending its approach to vision and to mental life more generally. First, cognitive scientists criticize what they claims is the naivety of their predecessors, who fail to see the complexity of the processes under consideration. Second, they argue that the correct approach to depicting and understanding the complexity of visual perception involves the use of an information-processing model.

As early as Ulric Neisser's 1967 introduction to cognitive psychology we read the following:

Even psychologists who ought to know better have acted as if they believed [in naïve realism] [...] The information reaching the eye is subjected to complex processes of analysis, extended in time. Visual experience results from some of these processes. [...] Perception is not a passive taking-in of stimuli, but an active process of synthesizing or constructing a visual figure. Such a complex constructive act must take a certain amount of time. (1967, 16)

Neisser and others argued that the time taken to complete a process is an indication of the level of complexity involved. It is important to note that the complexity which interests cognitive scientists is said to exist at the level of information-processing, rather than at the biological or behavioral levels. Given the emphasis on the informational level analysis, sources of evidence at the biological or even at the behavioral level tend to be marginal to their choice of explanation. Given the nature of our access to the putative underlying information processing, the evidential constraints governing articulations of informational complexity become somewhat problematic. The time that experimental subjects take to perform visually-based tasks serves as the primary empirical basis for the construction of theories of information-processing complexity. However, the problem of understanding what the informational complexity of a process amounts to and how it should be studied is underdetermined by timing evidence alone. Timing measurements alone leave room for a variety of explanations that are com-

patible with the evidence. Consequently, other, non-empirical constraints must play a role in one's decision to favor one articulation of information processing complexity over another. Philosophical presuppositions, play an especially prominent role in this kind of decision, as we shall see.

2. How Marr saw vision

Marr describes vision as a computational process wherein the brain constructs a three dimensional representation of the distal scene based solely on the illumination of the surface of the retina. According to Marr, the visual system consists of three (or four, if one counts the initial effect of light on the retina) hierarchically organized stages. At each stage a new aspect of the representation is constructed from information originally contained in the retinal image. For Marr, the visual system functions somewhat like an assembly line whose final product is a useful, finished model of the visual scene suitable for the organism's manipulation. Along the way, each stage in the visual system is explained in terms of the transformation of one representational structure into another according to relatively fixed rules. No new material or content is added to the representations as they undergo the sequence of computational transformations. Instead, increasingly specific features are extracted from information derived from the original retinal image, until the visual system generates a fully elaborated model of the visual scene. Models of this kind represent a data-driven approach to perception. Data-driven models of perception are generally contrasted with theories that view belief and other higher-level phenomena as playing a role in the construction of what it is that the organism perceives. A gestalt theory of perception would be an example of a top-down as opposed to a purely data-driven theory.

The specific claims involved in his description of how the visual system actually accomplishes these stages have proven less important to cognitive science than the methodological framework articulated in the opening and closing chapters of *Vision*. Following Marr's methodological example, traditional cognitive science has assumed that the first step in the investigation of mental life is the construction of an information-processing model. This construction involves determining a set of well-defined information-processing problems that the researcher believes the visual system must solve. Marr recommended that these problems should, at least initially, be studied in-themselves, apart from consideration of the underlying neuroscience.

Marr's suggested methodological tripartition in *Vision*; his distinction between computational, algorithmic and implementational levels of analysis in cognitive science is consonant with the philosophical functionalist intuition that, when it comes to matters of the mind, neuroscience can be safely ignored, since mental states can be approached at the computational (or functional) level.

The notion of information-processing complexity plays a central role in the approach and as we shall see, the claim that these systems are complex is motivated by the assumption that the visual system traffics in information bearing representations. By insulating the information-processing approach from the constraints of biological or behavioral evidence, cognitive scientists were free to craft a system of rules and representations that solved the problems of perception that they believed the visual system evolved to solve. However, because they are crafted in response to computational characterizations of the problems in question, such solutions do not necessarily help us to understand the mechanisms in actual biological vision systems.

3. Information processors vs. resonators

As mentioned above, the computational or information-processing view is almost always introduced by contrast with so-called *direct* theories of perception. Marr follows this practice when, early in *Vision*, he distinguishes his approach from J. J. Gibson's. Gibson, who is widely regarded as the most important modern proponent of the direct theory of perception, denied that perception is mediated by sense data or representations and sought instead to provide an ecological theory of vision that accounted for the dynamical and embodied characteristics of the senses.

Gibson recognized that visual perception is mediated by highly complex biological processes. His denial of what Susan Hurley called *the classical sandwich* - the distinction between perception, cognition and action- is generally overlooked in computationalist characterizations of so-called direct theories of perception. (Hurley 1998) Rather than tackling his criticism of the functional-level account, critics assume that the ecological theorists share their analysis of the complexity of vision at the functional level and then that they simply fail to explain that complexity properly. This is what motivates the computationalist emphasis on Gibson's sketchy and highly simplified accounts of the internal processes that subserve the senses. So, for example, Marr takes Gibson to task for 'seriously underestimat[ing] the

complexity of the information-processing problems involved in vision and the consequent subtlety that is necessary in approaching them.' (1982, 29)

Marr's criticism of Gibson focused specifically on the idea of the senses as organizing behavior in relation to ecological invariants. According to Marr, Gibson's approach was the product of his failure to understand information processing (1982, 29) and he wrote:

Although one can criticize certain shortcomings in the quality of Gibson's analysis, its major and in my, fatal shortcoming lies at a deeper level and results from a failure to realize two things. First the detection of physical invariants, like surface images, is exactly and precisely an information processing problem, in modern terminology. And second, he vastly underrated the sheer difficulty of such detection. (1982, 30)

Gibson would object to the idea that the brain does anything like computing the solutions to problems. For example in a passage cited by Marr, he writes "the function of the brain, when looped with its perceptual organs, is not to decode signals, nor to interpret messages, nor to accept images, nor to *organize* the sensory input or to *process* the data, in modern terminology." (as cited in Marr 1982, 29) Consequently Marr and Gibson differed fundamentally on the nature of brain function. For Gibson, the problem with an account like Marr's is that it relies on what he called "the muddle of representation" (1988, 279) and that it takes what Gibson called a "molecular" rather than a "molar" approach to perceptual systems. (1966, 52). The purpose of Gibsonian explanations, by contrast, is to show how "the receptive units combine and their inputs covary in systems and subsystems" (1966,52) Both Gibson and Marr are interested in understanding the function of vision. However, their views of what constitutes an explanation of this functional level are different. This functional level difference is not debated, and the computationalists, who assume that cognitive science is in the business of providing complex information-processing accounts, see Gibson as simply failing to understand the problem. In Marr's words: "Although some aspects of his thinking were on the right lines, he did not understand properly what information processing was, which led him to seriously underestimate the complexity of the information-processing problems involved in vision and the consequent subtlety that is necessary in approaching them." (1982,29)

From Marr's perspective, the important lacuna in Gibson's approach, was his failure to articulate the process involved in extracting informa-

tion about the environment from the flowing array of ambient energy. For Marr, it is not enough to say, as Gibson does, that the organism simply “resonates” with certain invariant features of its visual environment. Marr believed that he could explain, in information processing terms, how the organism extracted these invariant features in the first place. In Gibson’s terms, Marr can be understood as attempting to explain how an organism comes to “resonate” with some aspects of the environment rather than others. At its core, Marr’s project is an attempt to explain how we recognize things. For Gibson, by contrast, recognition happens via a set of biological and ecological mechanisms. These mechanisms can be described, but beyond that no further explanation is really necessary. The purpose of the functional level analysis according to Gibson is simply to *show* the interaction and covariation of what he called the “receptive units”. Marr viewed the appeal to mechanisms that resonate with certain features of their environment as missing precisely the information-processing strategies by which we recognize, to use his example, certain cat-like shapes as cats.

Denying representationalism, as Gibson did, does not entail a denial of the possibility of an explanation of recognition *per se*. Gibson’s criticism of representationalism is unequivocal, and yet it leaves space for explanations of how we might come to judge, for instance that some object that we directly perceive is a cat. He writes for instance:

Direct perception is what one gets from seeing Niagara Falls, say, as distinguished from seeing a picture of it. The latter kind of perception is *mediated*. So when I assert that perception of the environment is direct, I mean that it is not mediated by *retinal* pictures, *neural* pictures, or *mental* pictures. *Direct perception* is the activity of getting information from the ambient array of light. I call this a process of *information pickup* that involves the exploratory activity of looking around, getting around, and looking at things. This is quite different from the supposed activity of getting information from the inputs of the optic nerves, whatever they may prove to be. (1979, 147)

Gibson was also interested in providing an explanation for the capacity to recognize distinct objects as instances of a particular kind. However, Gibsonian accounts of how such recognition takes place, do not employ the representationalist strategies used by the computationalists. Rather than processing information, for Gibson, the organism is simply reacting to patterns of information in its environment. A more appropriate target for

Marr's frustrations and criticism would have been the handful of philosophers who argue that it is simply impossible to explain how we recognize cats as cats, tables as tables, etc. Marr's desire to explain how we can, for example, see a cat as a cat or a table as a table, provides the basic impetus for his idea that we should treat vision as a set of information-processing problems. Marr actually cites John Austin's discussion of what it means to say that we recognize or fail to recognize the *real shape* of an object. (1982, 29) Marr criticizes Austin's claim that there is no such thing as the real shape of a cat. Austin writes :

What is the real shape of ... a cat? Does its real shape change whenever it moves? If not, in what posture is its real shape on display? Furthermore, is its real shape such as to be fairly smooth outlines or must it be finely enough serrated to take account of each hair? It is pretty obvious that there is no answer to these questions—no rule according to which, no procedure by which, answers are to be determined. (Austin 1966, 67 as cited in Marr 1982, 31)

Contrary to the skepticism that Austin promotes, Marr shows how each of Austin's questions can be given an answer, perhaps not the kind of answer that would satisfy the philosophical skeptic, but certainly the kind that would satisfy ordinary scientific curiosity. Marr sees himself as providing precisely these answers in the fifth chapter of *Vision* where he provides an account of how the visual system generates representations of shape. The purpose of Marr's account in that chapter is to show how the visual system might provide some content for a judgment that some represented shape can be identified as an instance of some concept— —how, for example, some represented shape is the shape of a cat.

To conclude this section, we have seen that Marr's criticism of direct theories of perception is misdirected. Gibson was not suggesting that Marr's explanatory goals could not be achieved. Philosophers like Austin and perhaps occasionally Wittgenstein may have held such positions, however, as we have seen, where Gibson and Marr really differed was with respect to the representational approach to perception. The significant difference centers on the notion that the complexity of the information-processing task must be articulated in representational terms. Gibson certainly believed in some sort of mediating process that supported perception. What Gibson denied was the need for a system of representations to serve as mediators between perception, cognition and action. By contrast, the notion that the brain generates a series of representations is central to Marr's view.

4. Representation and Information Processing Complexity

The idea that brains traffic in representations was encouraged by renewed interest in the processes underlying behavior in psychology. While cognitive psychology as a discrete field of study only emerged in the early 1960's, the idea that recourse to representations and psychological processes is necessary to explain behavior was already in the air thanks to Tolman's cognitive mapping experiments in the 1930's. Theories of brain function that posited internal manipulation of symbolic structures seemed to solve a number of important problems for experimental psychology. They provided a model of the kind of internal processing that Tolman's cognitive mapping experiments were thought to require and, through the physical symbol system hypothesis, they offered a neat solution to the problem of the relationship between psychological and physical properties. Representations stored in a physical form in a computer could be understood as having both semantic and causal properties.

Relying on what had, by then, become a standard line of thought in cognitive science; Marr claims that we must have representations in the brain, because in order for an organism to know something about an object, it *must* be able to represent the object to itself. (1982, 3) Marr offers a clear definition of what he means by 'representation'. He writes: "A *representation* is a formal system for making explicit certain entities or types of information, together with a specification of how the system does this." (1982, 20). Marr's line of thought with respect to representation runs roughly as follows: in order to know we must have representations in the brain, and these representations are (or are part of) a formal system that makes certain kinds of information explicit along with a set of instructions or rules for interpreting these representations. In this paper, I will not provide an analysis of Marr's view of representation *per se*, instead I will focus on his unique and extremely influential computational account of the relationship *between* representations. This is where the notion of information-processing complexity plays an important role.

According to Marr's view, cognitive scientists should articulate information-processing problems into their component sub-problems and present an explanation that consists of a string of computable solutions to each of these problems. However, one unresolved issue faced by these explanations involves the criteria for judging whether they fit the mechanism under consideration. For instance, in biological systems a relatively

simple neural mechanism such as the edge detecting cells in the visual system may solve what computationalists treat as a complex information-processing task. In such cases, the neural mechanism can be interpreted as computing a range of alternative algorithms for the information-processing task. The challenge for the computationalist is to decide which of the set of possible solutions to chose.

In his defense of the dynamical systems approach to perception and cognition, Tim Van Gelder made the valuable point that simple systems can accomplish feats that can also be construed as incredibly complex information-processing tasks. (Van Gelder 1995,1991, Van Gelder, T. J., & Port, R. 1995) His example of such a system is Watt's governor. In the late 1700's James Watt invented a device for controlling the relationship between the speed of a steam engine and the pressure of the steam driving the machine. This device itself is simple, however, given computationalist inclinations, we could interpret it as solving a complex information-processing task. Careful analysis of the problem could, for example, lead us to write a computer program attached to a system of sensors (input) and valves (output) which could calculate the adjustments that are necessary to maintain a constant relationship between the pressure, speed, load etc. While nothing stops us from producing such a system, (let's call it Turing's governor) Watt's governor was a far simpler and more elegant solution.^a

How much would an articulation of this task in terms of a computer program – the Turing Governor - add to our understanding of what it is that the Watt Governor accomplishes? Practically speaking, a project to develop Turing's Governor would almost certainly have been an impediment to the creation of a technological solution. It is possible that such a solution may have been of some theoretical interest, but the chances of such theoretical advancement would hardly be sufficient to abandon Watt's solution

Theorists of direct perception, like Gibson and others, saw the senses along the lines that Watt took to the governor. The senses were to be

^aThe Watt governor is a spindle with two hinged arms. Connected to the engine's flywheel, the spindle turns at a rate that varies according to the demands on the engine (speed and load). As the spindle turns, the hinged arms are driven towards a horizontal position by centrifugal force. The arms are connected to valves that control the release of steam such that as the flywheel spins faster, the arms of the spindle move upwards, slowing the release of steam and thereby slowing the engine. If the flywheel slows, one gets the opposite effect, the arms descend and the valve opens again, thereby increasing engine speed. This simple device thereby establishes equilibrium between a variety of forces without any recourse to the complex information-processing techniques that would have been involved in our imaginary Turing governor.

understood in terms of the dynamic interaction between the organism and the environment. Sensory systems were collections of feedback loops that tended towards a kind of equilibrium between organism and environment. In the case of the visual system, these visuomotor loops obviously involved some kind of mediating biological process, but these processes were thought to be more like Watt's governor than my imaginary Turing Governor. For direct theorists of perception, a detailed presentation of the interaction of the organism and the environment was a sufficient explanation of the senses.

Computationalist criticisms of direct theories of perception attack this basic idea that the senses can be understood solely in terms of the interaction of their component parts. The study of cells, for example, was seen as failing to explain how an information-processing task, like recognizing a cat as a cat, could be accomplished. To compensate for this apparent deficiency, an additional level of explanation was introduced wherein the dynamic relationship between organisms and their environment could be analyzed into components that had no direct relationship to physical objects or to observable behavior. This level, the level of information-processing complexity, permits the development of theories of perception that incorporate the all-important notion of representation.

This complexity bears no direct relationship to the philosophical problem of computational complexity and has relatively little to do with Shannon and Weaver's formal definition of the notion of information. Instead, the standard strategy in cognitive science is to recommend that perception be broken down into a series of stages consisting of transformations over information bearing states or representations. As we have seen, articulations of this complexity may often have no correlate in physical reality. Consequently, it is often very difficult to know what it is that we are being told to recognize as complex.

5. Conclusion

The problem with Marr's argument for the importance of information-processing complexity is its implicit assumption that we can understand what the visual system does or what its function is, without knowing its physiological or anatomical form. Marr's basic assumption, understood in its strongest form, detaches accounts of this complexity from biological or other sources of evidence and precludes the possibility that any new discoveries regarding the structure or physiology of the brain or mechanism in question could change the way we understand its computational or func-

tional properties. This weakness is due to the lack of evidential constraints on the characterization of information processing complexity.

Marr *qua* scientist was a not dogmatic adherent to the autonomy of the functional level and would probably have been sufficiently skeptical in scientific practice not to stick to the strongest version of his methodological principle come what may. Nevertheless, Marr's methodology does recommend that scientific practices be governed by a principle that is equivalent to the belief that a suitably attuned philosopher or psychologist can determine psychological functions without concern for messy details of biology or ecology.

Historically, the bias in favor of the functional or computational level of investigation over the evidence of neuroscience did not originate with Marr. It dates back at least to the classic early works in cognitive psychology and is clearly articulated in Neisser's *Cognitive Psychology*. (1967) Again, the dismissal of biological evidence rested, almost from the start on the analogy between minds and computer programs, as Neisser writes:

The task of the psychologist trying to understand human cognition is analogous to that of a man trying to understand how a computer has been programmed. In particular, if the program seems to store and reuse information, he would like to know by what "routines" or "procedures" this is done. Given this purpose, he will not care much whether his computer stores information in magnetic cores or in thin films; he wants to understand the program not the "hardware." By the same token, it would not help the psychologist to know that memory is carried by RNA as opposed to some other medium. He wants to understand its utilization, not its incarnation. (1967, 6)

This radical denial of biology was not tempered by developments in neuroscience. For example, it is common to read more recent works in cognitive science arguing that:

Information processes can and must be studied without reference to the physics or biology of the system carrying them out... the algorithm for multiplication can be studied and understood in interesting ways without any knowledge of how it is implemented by neurons in the human brain... In general an important clarity of understanding is attained by separately considering an information process as a formal symbol manipulation process and as a physical

