can argue that the TV was a representation extrinsic to the observer, the retinal one was intrinsic to the observer, and the brain (as a holistic organism) is likely to use one type of representation only. I would counterargue, however, that the value unit type of representation from the TV screen to the retinal image has little to do with the brain's organization. It really has to do with the physical shape of the lens and the refractor properties of the light. Such value unit representations are, indeed, demonstrably used more in photographic devices than in neuronal systems, although it would be a mistake to categorically exclude a neuronal implementation. Rather than lamenting the difficulties of multidimensional or multimodal representations in terms of the value unit theory (duly listed in Ballard's paper) – the enormous number of units required for value unit representations (the so-called N^k problem – the storage problem of pictures if you think in terms of gathering intelligence by photographic means), the limited feasibility of storing all kinds of movements in "look-up tables" (if value unit representation is proposed for motor control), reservations about "grandmother cells," the evidence for multiple-type, not point-to-point, representations over cortices – a different, perhaps more constructive comment will be made below.

While value unit representation may yet be shown to exist in the CNS (not by an optical lens but by neuronal connectivities), the "alternative" type of representation demonstrably exists. Ballard lists this as "variable-encoding representation" but unfortunately dismisses it on the last page of his paper since he seems to feel that a clear either–or choice must be made. Accepting the fundamental assumption that the brain is organized hierarchically and represents sensorimotor invariants in a parallel manner, one may find that such invariants are difficult to pinpoint in visual systems. Indeed, colors, shapes, edges, directions, distances, etc. are all various invariants, but their sensorimotor character is not at all obvious. For example, how is color related to movement? One could possibly reach for the reddest apple, but such a sensorimotor act includes not only color vision but pattern recognition and stereoscopic vision as well as some higher-level decision making based on olfaction and personal preference of Macintosh over Golden Yellow.

In some other sensorimotor systems the invariants and their hierarchical representation are much simpler and more obvious. Consider gaze control, where a passive head movement (as a physical invariant) is expressed in the intrinsic system of coordinates of the vestibular semicircular canals and a corresponding compensatory gaze shift is generated by the motor apparatus of the head expressed in the coordinate system intrinsic to the neck muscles, or of the eyes themselves, expressed in the coordinate system intrinsic to the eye muscles. All such representations in fact utilize variable encoding, in the sense that each sensory or motor expression defines one point at one time in the multidimensional functional space. These are the six-dimensional semicircular canal space, the twelve-dimensional eye muscle space and the thirty-dimensional neck muscle space. These neurons are, in fact, frequency-coded elements; both the motoneuron firing and the primary vestibular neuron firing constitute an analogue expression of the physical invariant of head movement. Not having to "store" all possible head movements, but rather have to *adequately* respond to any actually occurring movements yields, as Ballard remarks, a hundredfold greater compactness of the gaze-stabilization neuronal mechanism.

Throughout the paper it is clear that Ballard is keenly aware of the fact that brain function (representing physical invariants) may actually occur by coordinate transformations, not necessarily by value unit representation. This is evident from the remark in Section 5.1: "Shape recognition is only one of several problems that require the computation of coordinate transformations," or from Figure 14, in which some of the necessary transformations are actually spelled out. Of course, sensorimotor transformations through neuronal networks of multi-dimensional vectors, expressed in coordinate systems intrinsic to the organism, are an approach (apparently unknown to Ballard) capable of providing not only conceptual guidance to the interpretation of brain function but also a powerful generalized vectorial (tensorial) formalism. This may be of significance, since no concept is likely to succeed (even if it is valid) if it is not coupled with a suitable mathematical formalism that can quantify and elaborate how the general concepts apply to specific neuronal mechanisms. In Ballard's article the need for mathematical formalism is evident (multidimensional vectors are often mentioned, albeit as ordered sets of binary digits, without realizing the underlying frames of reference that are intrinsic to the organism). Also, the need to elaborate general concepts into specific neuronal models of particular CNS subsystems is clearly expressed. Having done away with the "brain as computer" dogma and replaced it with the axiom that the CNS is a parallel system, the ground has been cleared by Ballard for the construction of an axiomatic, mathematical brain theory.

Computational neuroscience

Terrence J. Sejnowski
Biophysics Department, Johns Hopkins University, Baltimore, Md. 21218

It is generally believed that the connections between neurons are important for the computational capabilities of the cerebral cortex; Ballard attempts to take seriously the details of how one could use cortical connections to perform computations for vision. I would like to suggest that this concern for computational realizability – that is, concern for both the computational problem and its network implementation – represents a promising new direction for research that could contribute a new computational branch to neuroscience.

Marr (1982) was influential in emphasizing the importance of computational analysis in thinking about how the nervous system solves difficult problems in vision. However, he argued for the independence of the computational level and the implementation level, which encouraged the belief that it was unnecessary to take the details of neural circuitry too seriously: There might be many ways to implement a particular algorithm, and the details of how it is implemented in the brain might not be particularly revealing. Ballard argues the contrary, that the hardware in cerebral cortex may reveal much about the style of cortical computation and that neuroscience should be taken seriously by cognitive scientists and computer scientists interested in understanding how to solve the same difficult problems.

The emergence of simple parallel models exhibiting nontrivial computational capabilities may be of great importance for future research in neuroscience because they offer one of the few ways for neuroscientists to test qualitative ideas about the representation and processing of information in populations of neurons. Suppose that the responses of single neurons in an area were sensitive to features that could be important for computing, say, optical flow. Knowing the goal of the computation, one could design a parallel algorithm for implementing the computation and then test it with a wide range of inputs. The process of specifying and testing an algorithm often reveals unexamined assumptions and refines the original motivation for the model. If one successful algorithm is found, the computational feasibility of the original proposal is strengthened; to test whether some form of the algorithm is actually implemented in cortex would be much more difficult; ultimately, the performance of the algorithm has to be compared with psychophysical testing.

Models are used in many different ways and we should make clear the type of model under discussion. A connectionist model is not a model in the sense of reconstructing molecular and cellular detail. Rather, a connectionist model is a simplified,

stripped-down version of a real neural network similar to models in physics such as models of ferromagnetism that replace iron with a lattice of spins interacting with their nearest neighbors. This type of model is successful if it falls into the same equivalence class as the physical system; that is, if the qualitative phenomena (such as phase transitions) are the same for both the real system and the model system (Ma 1976).

What is needed to make even modest progress in computational neuroscience is a balance between constant awareness of the biological constraints and concentration on the concerns of specific computational problems. A good example of this approach applied to a problem at the subcortical level is the recently proposed neural network model of the integrator in the oculomotor system (Canon, Robinson & Shamma 1983), which is based on the classical model of lateral inhibition in the Limulus retina (Ratliff 1974). At the least, a network model should not be inconsistent with the known biological facts (Crick & Asanuma 1986); on the other hand, it may be premature to rule out possible mechanisms for which there is not yet evidence (Sejnowski 1986).

Some neuroscientists may feel uncomfortable because connectionist models do not seem to take into account much of the known cellular properties of neurons, such as the variety of membrane channels that have been found. What if the processing capabilities of cerebral cortex were to depend crucially on some of these properties? In this case it might not be possible to get networks of oversimplified model neurons to solve difficult computational problems, and it might be necessary to add new properties to the model neuron. The added capabilities would yield a better understanding of the roles played by these neural properties in processing information, and suggestions could emerge for useful properties that had not yet been observed (Shepherd, Brayton, Miller, Segev, Rinzel & Rall 1985). In a sense, the present models are guideposts for thinking about the computational capabilities of neural networks and benchmarks that set the standards for future models.

One of the key insights that have already emerged from studying one class of simple nonlinear networks with recurrent collaterals is that among the large number of possible states of a network only relatively few of these states, called attractors, are stable (Anderson 1983; Cohen & Grossberg 1983; Hinton & Sejnowski 1983; Hogg & Huberman 1984; Hopfield 1982; Hopfield & Tank 1985; Sejnowski 1976; Wilson & Cowan 1972). The existence of stable attractors is a feature that is likely to generalize to more complex networks. Objects and their relationships can be internally represented by these attractors, and the search for the best match between the world and the internal representation of the world by the dynamics of the network is much more powerful than previous template-matching procedures. This opens a large number of research problems, some of which are discussed in Ballard's article, such as the issue of local versus distributed representations and the binding problem. The identification and study of these issues in simple network models will greatly help us in understanding the principles that went into designing the cerebral cortex.

What does the cortex do?

Mriganka Sur
Section of Neuroanatomy, Yale University School of Medicine, New Haven, Conn. 06510

Ballard sets out with an important but ambitious goal: to describe the organization and connections of the visual areas of the cerebral cortex in terms of a parallel computational architecture. The underlying hypothesis – that the cortex computes collections in invariants at different levels of abstraction – has been suggested earlier but in different forms (e.g., Phillips, Zeki & Barlow 1984). I see nothing wrong with the hypothesis or with

the present effort to support it. One problem is simply that not enough is known yet to support or refute many of Ballard's conjectures.

Ballard introduces the term *value unit* to denote an elementary processing unit of cortex that represents a single parameter and that accesses or can be accessed by other units representing other, complementary parameters. The idea is not far from our current notion of a cortical column or module, that is, a set of neurons united by a common task. Modules are one solution to the problem of representing many variables on a two-dimensional surface such as the cortex. The dimensions of modules differ, depending on the type of module considered. Ocular-dominance columns in area 17 or V1 (Hubel & Wiesel 1977) and slowly and rapidly adapting bands in area 3b of somatosensory cortex (Sur, Wall & Kaas 1981) are strips of cortex on the order of $500 \mu m$ in width. Orientation "columns" appear to be an order of magnitude smaller. Such "minicolumns" (Mountcastle 1978) may be ubiquitous and may constitute the irreducible processing elements of cortex. A full set of columns must be present in prototypical blocks of cortex to analyze each elementary unit of the sensorium (Hubel & Wiesel 1974; Sur, Merzenich & Kaas 1980).

To return to Ballard: Although "value units" may be anchored in columns, they actually take us beyond columns, for we now have specific suggestions about organization and connectivity based on the indexing parameters in a given area. (Incidentally, all visual cortical areas studied so far contain retinotopic maps (Merzenich & Kaas 1980; Van Essen 1979), so that the primary and secondary indices in all areas are the coordinates of visual space.) Making columns, however, may only be a manifestation of the central task the cortex performs on its input: the activity-dependent weighting of synapses. By definition, the cortex must not be anchored too rigidly to a hard-wired anatomy. A single cell must be a part of many different circuits, and a synapse must change its weight and hence its role in information processing as a result of changes in activity down its parent fiber or in the postsynaptic cell. Evidence that the same cortex can serve different functions comes from experiments that show alterations in somatosensory cortical maps after peripheral nerves are cut or digits are amputated (Merzenich et al. 1983; Merzenich et al. 1984). Such manipulations cause dramatic shifts in the cortical loci of representation of the normally innervated digits and other parts of the hand. The implication is that cortical maps are dynamic and that alterability must be a dynamic process that continues normally, even in adulthood.

What else does the cortex do? Apart from dynamic weighting of synapses in a cortical area to make maps and modules and intracortical connections between areas to achieve functional diversity and specificity – Ballard's suggestion that corticocortical connections perform coordinate transformations for shape perception is perhaps the most challenging idea in the paper – cortical areas are also connected in highly parallel fashion with corresponding thalamic nuclei. Furthermore, many of the columnar features that are anatomically defined in the cortex have their substrate in the thalamus. Thus, the lateral geniculate nucleus contains neurons from the two eyes segregated into eye-specific laminae, and these project to appropriate ocular dominance columns. The ventroposterior nucleus of the thalamus contains modules of slowly and rapidly adapting neurons (Dykes et al. 1981), and these project to the appropriate type of column in area 3b or area 1 of somatosensory cortex. Cortical areas feed back heavily to appropriate thalamic nuclei. Clearly, the thalamus must be considered with the cortex in defining parallel information-processing schemes.

All in all, it is clear that we still have a way to go in specifying what the cortex does and how. Still, attempts such as Ballard's to cast experimental data into a coherent theoretical framework are valuable, if for no other reason than that such attempts persuade experimentalists like myself to think a little more about what we do and why.