DESIGN AND OPERATION OF BRAINS

K. N. Leibovic*

ABSTRACT. With the advance of information technology, the computer has become a model for the brain, replacing earlier ones, such as the telephone exchange and hydraulic machinery. But what kinds of computations does the brain perform? What we know about its structure and function imposes constraints. We can analyze the consequences and implications of these constraints to arrive at new insights. Thus we are led to a view of multivariable, broadly tuned single cells, overlapping computational areas and distributed representations of functions and operations.

Recent findings from imaging and neurobehavioral studies support the idea that brain programs or representations are dynamic and distributed and that brain-body interactions are inseparably linked. It is proposed here that the primary constraint on representations is that they be in brain-body correspondence with the appropriate task or operation and within that constraint there is considerable latitude in the ways these programs can be generated and regenerated. Based on these viewpoints, this paper presents an outline of principles and their illustration by quantitative treatments.

1 Introduction Computational theories of artificial intelligence, in the broad sense, even when not addressed to neuroscience, have drawn heavily on knowledge of the nervous system. Some such theories now lead an independent existence without any reference to biology. Yet, invariably, implementations simulate some degree of quasi-mental activity. It follows that a knowledge of neural function is relevant to biocomputation and artificial intelligence.

In trying to understand how the brain works we may look at its design and ask two questions: First, does it suggest a rationale in terms of function? Can we think of any reasons why this design should be superior to others? And second, what are the functional consequences of the particular design? What do the properties of the units and their interconnections imply for the ways they operate?

There have been numerous attempts to model the brain with the use of advanced mathematical techniques and, while interesting, it is doubtful whether they have given us much new insight. This may be due, in part, to the depth of our knowledge being insufficient. It is as if we were looking at a classical painting from a distance, where we can see the outlines of the composition, the balance of colors and of light and dark. But the essential content is still not within our grasp. Therefore, we need to ask some simple questions, and it is surprising how much new insight we can get based on what we know already. So, to begin with, let us outline some of the properties of neurons and networks. Starting with single cells, we know that different neurons are tuned to different stimuli and produce different responses. At one time it was thought that these responses are highly specific.

So one particular neuron might become active when grandmother appeared. It was the idea of grandmother cells, which we now believe to be untenable. In fact, neural tuning is quite broad. A visual cell responding vigorously to an edge oriented at 30° to the horizontal will respond with diminishing strength to progressive deviations from 30°. Moreover, the

same cell may give differential responses to parameters such as movement and color. The result is that neurons are multivariable and broadly tuned. This is a property which is not sufficiently appreciated in the literature.

At the network level, a striking feature of design is the impressive convergence and divergence of lines of communication between the various levels and stations of nervous systems. This is quite different from point to point transmission in man made systems. It raises a number of questions, including how detailed information might be represented. At first sight such a design may appear too diffuse to subserve accuracy of sensation or movement. We need to ask what distinguishes the operations in neural networks where many source cells converge onto a target cell and one source cell diverges to many target cells.

At a global level we find that information is processed and transmitted in streams or channels at the same time as there is a dense interconnectivity between different parts of the brain. There are distinguishable neural channels not only for the different senses and activations of different muscle groups, but also within each modality. For example, in vision there are channels for motion, color, shape from retina to cortex, and similarly in the control of movement. There is a dense interconnectivity through convergent, divergent and feedforward and feedback pathways, so that every neuron is ultimately connected to every other neuron in the brain. But at the same time there are channels and areas which are tuned to preferred parameters.

How can we make sense of these design features and what are their functional implications?

2   Neural Tuning To appreciate a rationale for broad tuning consider the case of color vision. The human fovea contains 3 types of cone, the red, green and blue cones. Each is tuned to a broad range of the 400nm visible spectrum with peak sensitivities at approximately 565, 535 and 440nm respectively. It is well known that the responses from these cones are combined to produce the sensations of a continuous range of colors with high discriminability. Suppose instead of this broadly tuned system we had photoreceptors which were tuned to a narrow wavelength range, analogously to grandmother cells. It has been estimated from psychophysical experiments that we can distinguish some 10^6 different colors. (see Boynton 1990). A US Bureau of Standards publication lists 7500 color names (Kelly & Judd 1976). Suppose we use the lower figure with 7,500 cones to cover the 400nm spectrum. Each cone would have to be tuned to a wavelength width of $4/75$nm $= 0.5\mathring{A}$ approximately. The cones are spaced some 2μm apart in the fovea. Then 7500 cones would occupy about 3% of the human fovea. This would create problems for the spatial resolution of colors. But, even worse, from the uncertainty principle, a 0.5Åtuning would imply a position uncertainty of about 800μm, extending over more than half the fovea. Clearly, broadly tuned cells are superior in performance.

As regards multivariable tuning, consider a circular receptive field with its associated target cell responding to the position and intensity of a spot of light. In principle, one could have one target cell computing each position and intensity, or one could have one target cell tuned only to position and another to intensity, or one could have two broadly tuned target cells, each responding to both position and intensity. Suppose the common receptive field contains k cells, each capable of j intensity responses. Then in the first case, the target cell would have to accommodate $k \times j$ responses. In the second case it would require $k + j$ and in the third case $2\sqrt{(k \times j)}$ responses for the two cells. (In the third case the total number of distinct responses of the pair is $\sqrt{(k \times j)} \times \sqrt{(k \times j)}$). For $k \times j \geq 4$, $k \times j$ and $k + j$ are both greater than or equal to $2\sqrt{(k \times j)}$. So the required response repertoire is smaller for broadly tuned, multivariable cells.

Let us now look at a specific model of broadly tuned, multivariable cells. When a spot of light moves across a diameter of the receptive field let the response of the cell increase to a peak in the center and then diminish again. For a given position let the response also increase with light intensity. Let two target cells A, B have overlapping receptive fields (Leibovic 1969). We can plot the responses of A vs. B for different positions and intensities of the spot of light. We obtain a series of distinct curves parametrized either by position or intensity from which both position and intensity can be uniquely determined. It illustrates how multivariable information can be encoded and suggests as a consequence that it requires the activities of groups of cells. The example has implications for the problem of representation to be discussed later.

We also find that this design is robust with respect to reliability. For, suppose there are three target cells with overlapping receptive fields. If one of them failed there would still be two for computing spot position and intensity. Instead, if there were separate target cells for position and intensity, one would need four cells to safeguard against failure, one extra one for the position cell and one for the intensity cell.

Thus, multivariable, broad tuning turns out to be economical in terms of the number of different kinds of units required for accurate processing and transmission of information as well as in terms of required response repertoire, and in terms of any redundancy subserving reliability.

3   Convergence and Divergence Successive transformations or computations are carried out between different levels in the nervous system. For example, if a target cell in the visual pathway is to compute speed and direction of movement it needs inputs from photoreceptors within its receptive field in the retina. This necessitates convergence, as a result of which there is a lack of information on the detailed shape of the image. To recover detail requires concurrent divergence and overlap of receptive fields as we saw in the previous section on neural tuning. To further pursue this, consider a linear system of target and source cells. Let the i-th target cell output y_i be a linear combination of the source cell outputs x_j, i.e.

$$(1) \qquad y_i = \sum_j a_{ij} x_j$$

where it is assumed that the x_j and y_i represent a single variable, e.g. the firing frequency of a neuron. Then, for $1 \leq j \leq N$ and $1 \leq i \leq T$, $T \geq N$, there can be a unique solution for the x_j under suitable conditions (Leibovic 1988). In other words, all the information contained in the source cells can be recovered. In eq. (1) the source cells are simply characterized by their outputs x_j. But suppose the x_j have a structure, so that successive segments of the output of x_j reflect different variables or aspects of the input. We can split the x_j into a number equal to the separate variables contained in the output of the j-th cell. For example,

$$(2) \qquad x_j = \sum_k b_{jk} x_{jk}$$

and then

$$(3) \qquad y_i = \sum_{j,k} a_{ij} b_{jk} x_{jk}$$

In this case the number of target cells must be larger than the number of source cells if all information is to be preserved. There are examples in the nervous system in which there

is more divergence than convergence and vice versa. Thus in the macaque visual pathway some 100×10^6 photoreceptors converge onto 10^6 ganglion cells and an approximately equal number of principal cells in the LGN, which then diverge to 50×10^6 cells in layer IV of V1 (Orban 1984). In the cerebellum 1 parallel fiber diverges to 300 Purkinje cells while at the same time more than 50,000 parallel fibers converge onto 1 Purkinje cell. (Eccles 1969).

Of course, the function of serial operations in the brain is not just to preserve point-wise information between different levels, but to compute meaningful parameters. The above considerations can serve as a guide in deciphering where detailed input information is preserved, where it is discarded and where the multivariable responses of source cells may be made explicit.

Reliability can also be achieved in various ways in convergent – divergent pathways (Leibovic 1993, 2000). For example, referring to the model of eq. (1), in a two layer net with divergence, let there be n cells in layer 1 diverging to $n + m$ in layer 2. The number of distinct subsets in layer 2 containing n target cells is

$$(4) \qquad\qquad\qquad\qquad T = \frac{(n + m)!}{n!m!}$$

Each of these sets can contain the complete information from layer 1 if the net is functioning correctly. But if one of the target sets fails, one could rely on a network which could take a majority vote. The appropriate m would depend on failure probability, but in any event this would be a much more economical system than the one considered in the classic paper of von Neumann (1956).

Thus convergence and divergence is necessary for computations in receptive fields and the preservation of information on the structure of activity patterns in groups of cells. It is an economical design which is well suited to ensuring reliability. As with multivariable, broadly tuned cells, a consequence of convergence and divergence is that information is distributed between levels.

4   Channels We have seen that successive computations are carried out through source cells converging onto target cells. We have also seen that there can be advantages to multi-variable tuning, in other words for computing more than one variable per target cell. This was the case in the discussion of response repertoire and reliability for position and inten-sity tuning. On the other hand, different computations have different network requirements. A case in point is illustrated by a model artificial vision system (Quesada-Arencibia *et al* 2000) which computes the size, position, speed and direction of a moving object. There is a channel for object size and position, and a channel for speed and direction which are calculated for each position of the moving object. In principle, the object position can be obtained from the motion channel. But the calculated movement trajectory then deviates appreciably from the correct one. More biological examples are readily at hand as follows.

To distinguish two points of light as separate there must be at least one unstimulated photoreceptor between two stimulated ones in the retina. But to distinguish their colors there must be at least two cones with different spectral sensitivities stimulated by each spot, with an intervening space of cones in between. Again, to detect retinal motion requires a larger receptive field than distinguishing between two spots of light (Leibovic 1988). Each of these functions requires different connectivities which presents a rationale for different form, color and motion channels. Such channels have been described in the visual pathway and they also occur in other sensory and motor modalities (see e.g. Maguire *et al* 1990). In vision, the different channels synapse in different cortical areas. Thus, the primary visual cortex V1 has subdivisions for color, motion and form parameters which undergo further processing in V2 (see Livingstone & Hubel 1987); color is represented in V4 and motion in

V5. There is also evidence for parietal and temporal processing streams for space (where) and form (what) respectively (Ungerleider & Mishkin 1982).

There are more than 30 cortical areas concerned with different aspects of vision. But it would be wrong to characterize them as devoted to a single variable. That would be repeating the mistake of grandmother cells. My proposition is that multivariable, broad tuning is a universal design feature of the nervous system and this applies not only to single cells, but also to functional groups of cells in sub-cortical nuclei as well as cortical and cerebellar subdivisions.

Thus a rationale for channels is to be sought in computational efficiency. A consequence is that categorical information is represented in distinguishable loci. But we must reject the notion that either channels or cortical areas are simply unimodal.

5 Problems of Representation A traditional view of the brain is that it consists of different channels and areas which deal with different stimulus and action parameters. Perception and recall are sub-served by serial processes from lower to higher levels in specialized areas of the brain up to association and integration areas, presumably in frontal cortex. Subsequently, instead of a unidirectional processing stream it has been proposed that there is an interaction between high level, specialized modules which receive their input from lower level analyzers. The high level interactions would be controlled by convergence zones (Damasio 1989) which bind or bring together the separate aspects of a percept in the specialized modules. Thus, the sight of motion and shape would result in the realization of a bird in flight. However, neither integration areas nor convergence zones have been identified in neural tissue, and there are some interesting recent and not so recent findings which suggest a different point of view as proposed here. I shall mention a few examples.

It is well known that left hemisphere damage to language centers at an early age leads to the development of language centers in the right hemisphere. It demonstrates the plasticity of the brain in development. With regard to the adult brain, there are patients with severe left hemisphere lesions who retain some semantic abilities. These are presumed to be due to residual or normally silent capabilities in the right hemisphere, which has traditionally been considered as nonlinguistic.

More recently, there have been discoveries regarding phantom limbs. After amputation of a limb the patient can feel pain and other sensations in the missing limb. It turns out that such sensations can also be evoked by stimulating specific areas on the patient's face. The face and hand areas are adjacent in the somatosensory cortex and some time after amputation the hand areas seem to be taken over by the face area as revealed by brain imaging. (Pons et al 1991, Ramachandran 1993). The interesting point, however, is that the face stimulation can produce amputated hand sensations very soon after the operation and before neural reorganization could have occurred. (Borsook *et al.* 1997). This suggests that the hand and face areas are not sealed off from each other but, in fact, are intrinsically multivariable, though hand sensation may, in general, be silent in the face area and vice versa.

Since the invention of brain imaging a whole new field of research has opened up. A remarkable discovery was that even in the performance of simple tasks, there is widespread activation of different areas of the brain. Of particular interest are some studies which are beginning to reveal the neural substrates of not only sensory and motor activations but of thought processes. I shall mention only a few.

There has been a lot of work as well as controversy on visual imagery and visual representation (see TINS Debate in TINS 1994). A case in point concerns the question: Which cortical areas are activated when a subject performs a visual task on the one hand or performs that task in his imagination on the other hand? Some have claimed that the areas

are identical (e.g. Kosslyn & Ochsner 1994) and include V1 and V2, which have retinotopic projections and are involved in analyzing local parameters of shape, color, motion. Others (e.g. Roland & Gulyas 1994) have maintained that only higher level visual areas are activated in imagery. There are a number of methodological problems which may explain the discrepant results and there is evidence that imagery and perception, not being identical, do not produce identical activations (see e.g. Kosslyn et al 1997). Yet it remains a fact that lower as well as higher visual areas are active in both conditions in at least some experimental paradigms. Thus, imagery, a reconstructive process, drawing on memory and attention, may use neural mechanisms which are traditionally associated with low level analysis of the visual scene: The mental reconstruction does not occur only in high level association areas or convergence zones.

Not only do imagery and performance reproduce similar neural activities, in addition imagery induces physical activation which is also part of the performance. For example, kinesthetic sensation is internally simulated during imagined movement (Naito *et al.* 2002), and eye movements accompany visual imagery (Brandt *et al.* 1989). Thus, the thought processes generating a particular mental imagery also contain a penumbra of associations involving the whole organism.

Bodily activity induces neural activity and vice versa. This is equally true at the molecular level: physical states turn on mental states (and v.v.) and induce hormonal activity. Hormones bind to DNA and affect gene expression, leading to the production of proteins which alter mental and physical states. The brain, on the view presented here is merely an organ which participates in a cycle of activities in support of the living organism. Any separation we may make between body and brain may be quite illusory.

6 Visual Representations are not Encoded in Fixed Retinotopic Tissue

We showed some time ago (Leibovic et al 1971) that visual percepts are not encoded in fixed retinotopically organized areas of the brain. In our experiment a subject is placed in a dark room without any visual space cues and has to set up a series of small, dim lights along a straight line, perpendicular to his line of sight and at several distances from himself. This is the well-known paradigm of the fronto-parallel lines. Instead of being straight, the lines set up by the subject are convex towards him at near and concave at far distances. It has been shown that binocular disparity is the dominant cue in this situation. Other cues, like accommodation and convergence play little, if any, role.

But if the cue is binocular disparity, then it should be possible to predict all the experimental fronto-parallel lines from just one of them at one distance from the subject. It turns out, however, that these theoretically derived fronto-parallel lines are quite different from the experimental ones. This implies firstly, that other than the purely retinotopic cues of binocular disparity are involved and secondly, that visual representations are not encoded in either retinotopically organized fixed cell assemblies nor in fixed patterns of activity in labile assemblies. For, if the representation resided in retinotopic, fixed cell assemblies then different retinotopic images should appear to be different; and if the representation was in a fixed activity pattern, then different retinotopic images should evoke different activity patterns and appear to be different. But this is not the case: the experimentally generated fronto-parallel lines are retinotopically different, but appear to be all the same.

7 Representations, Sensorimotor Activity and Thought Processes

Our results on the fronto-parallel lines have implications for functional imaging studies. Together they suggest that representations are not stored in some fixed groupings in higher brain centers or convergence zones. If they were, then only the high level centers or convergence zones would need to be activated to go through an imagined or recalled task. But, as the imaging

studies show, the low level areas are also recruited. True, one could make the argument that distributed memory residing partly in low level centers need to be accessed from convergence zones. But, as stated earlier, no such convergence zones have been identified. I suggest instead that something like a program or plan of widespread neural activation which depends on a sequence of functional operations is at work and is generated, or regenerated, by the appropriate input or cue (Leibovic 2000). Such input or cue need not come from any convergence zone, but may arise in any part of the brain. The high interconnectivity ensures spread of activation and interactions through re-entrant pathways. I propose that primary activity in a specialized area also contains a substrate of more or less silent activity for other specialized areas. Thus activity in one area primes activity in other areas. If these other areas have also received primary activation, the to and fro of information leads to a resultant co-activation which represents the appropriate percept. Thus, activation in areas specialized for form (what) and for position or movement (where) can lead to the percept of a bird in flight. There is no need of convergence zones for binding the component activities which, in themselves, have only partial or no meaning. The representation resides in the global activity arising naturally through coactivation and co-operative interactions. There is no distinction of high and low level areas, only of functional differences.

A percept then arises from the fit of distributed activity in relation to the constraints provided by the input or cue. As a consequence, like a perceptual or motor task which can be performed in more than one way, the corresponding neural activity can exhibit more than one pattern. The important point is that there should be a correspondence between neural activity and task.

A compelling example of such a correspondence is the discovery of mirror neurons in the monkey brain (Gallese et al. 1996). These are found to discharge when either the monkey manipulates an object or sees the experimenter performing a similar action. My interpretation of this is that, like the imaging results, it suggests that the working neurons are also the thinking neurons. Sensorimotor activity and thought have similar representations.

8   Conclusion In summary, neurons are not univalent grandmother cells, but are multivariable and broadly tuned. This makes for efficiency in terms of the number of units and the response repertoires required to encode information. As a consequence, information is carried in the simultaneous activities of populations of cells. Moreover, with this design any redundancy in the service of reliability is minimized.

I propose that multivariability and broad tuning is a general feature of design in the nervous system including, in particular, channels and cortical areas.

The views I have presented place the brain within the context of the body, as suggested by the examples on mirror neurons and imaging, where imagined motion engenders kinesthesia and visual imagery is accompanied by eye movements. We are thus led to a dynamic view of organic, functional operations and neural representations. All these are sub-served by broadly tuned, multivariable neurons and neural assemblies, interconnected through converging and diverging pathways and computing different parameters. It is an interdependent, highly adaptable, efficient, dynamic system which has evolved through natural selection.

References

[1] Boorsook B, Fishman S, Bacerra L et al (1997): Acute plasticity in human somatosensory cortex following amputation. Soc. Neurosci. Abstracts

[2] Boynton R M (1990): Human color perception (in K. N. Leibovic ed. Information Processing in the Nervous System, Springer Verlag, N.Y.)

[3] Brandt S A, Stark L, Hacisalihzade S et al (1989): Experimental evidence for scanpath eye movements during visual imagery. Proc 11th I EEE Conf. on Engineering, Medicine & Biology.

[4] Damasio AR (1989): The brain binds entities and events by multiregional activation from convergence zones. Neural Computation 1, 123-132.

[5] Eccles J C (1969): The dynamic loop hypothesis of movement control (in K N Leibovic ed. Information Processing in the Nervous System, Springer Verlag, N.Y.)

[6] Gallese V, Fadiga L, Fogassi L, Rizzollatti G (1996): Action recognition in the premotor cortex, Brain 119, 593-609.

[7] Kelly K L, Judd D B (1976): The ISCC-NBS method of designating colors and a dictionary of color names. NBS special public. 440, Washington DC

[8] Kosslyn S M, Ochsner K N (1994): In search of occipital activation during visual mental imagery. TINS, 17(7), 290-291.

[9] Kosslyn S M, Thompson W L, Alpert N M (1997): Neural systems shared by visual imagery and visual perception: a positron emission tomography study. Neuroimage 6(4), 320-334.

[10] Leibovic K N (1969): Some problems of information processing and models of the visual pathway. J. Theoret. Biol., 22, 62-79.

[11] Leibovic K N (1988): Parallel processing in nervous systems with converging and diverging transmission (in L M Ricciardi ed. Biomathematics and Related Computational Problems, Kluwer Academic Publishers).

[12] Leibovic K N (1993): Biophysics of neural computation (in J Mira, J Cabestamy, A Prieto eds. New Trends in Neural Computation, Springer Verlag, Berlin).

[13] Leibovic K N (2000): Computers, grandmother cells and dynamics of functional operations in the brain (in U Halici, K Leblebicioglu, V Atalay E Nalcaci eds. Brain-Machine Workshop 2000, Ankara, Turkey).

[14] Leibovic K N, Balslev E, Mathieson T A (1971): Binocular vision and pattern recognition. Kybernetik 8(1), 14-23.

[15] Livingstone M S, Hubel D H (1987): Psychological evidence for separate channels for the perception of form, color, movement and depth. J Neurosci 7(11), 3416-68.

[16] Maguire W, Weisstein N, Klymenko V (1990): From visual structure to perceptual function (in K N Leibovic ed. Science of Vision, Springer Verlag, N.Y.)

[17] Naito E, Kochiyama T, Kitada R et al (2002): Internally simulated movement sensations during motor imagery activate cortical motor areas and the cerebellum. The J. of Neurosci 22(9), 3683-91.

[18] Orban G A (1984): Neuronal Operations in the Visual Cortex. Springer Verlag, Berlin.

[19] Pons T , Garraghty P E, Ommaya A K, Kaas J H, Taub E, Mishkin M (1991): Massive reorganization of the primary somatosensory cortex after peripheral sensory deafferentation, Science 252: 1857-60

[20] Quesada-Arencibia A, Leibovic K N, Moreno-Diaz jr.R, Aleman-Flores M (2000): Visual analysis of size, location and motion by a multiple channel system (in V Halici, K Leblebicioglu, V Atalai, E Nalcaci eds. Brain-Machine Workshop 2000, Ankara, Turkey).

[21] Ramachandran V S (1993): Behavioral and MEG correlates of neural plasticity in the adult human brain. Proc.Nat.Acad.Sci (USA) 90,10413-20.

[22] Roland P E, Gulyas B (1994): Visual imagery and visual representation. TINS 17(7), 281-287.

[23] TINS Debate Visual imagery and visual representation (1994): TINS 17(7), 281-297.

[24] Ungerleider L G, Mishkin M (1982): Two cortical visual systems (in M A Goodal, R J W Mansfield eds. Analysis of Visual Behaviour. MIT Press, Cambridge.)

[25] von Neumann J. (1956): Probabilistic logic and the synthesis of reliable organisms from unreliable components (in C Shannon, J McCarthy eds. Automata Studies).

STATE UNIVERSITY OF N.Y. AT BUFFALO 105 HIGH PARK BOULEVARD 14226 BUFFALO NY U.S.A.