CONVOLUTION AND MATRIX SYSTEMS AS CONTENT ADDRESSABLE DISTRIBUTED BRAIN PROCESSES IN PERCEPTION AND MEMORY

Karl H. Pribram
Neuropsychology Laboratories
Stanford University
Stanford, California
USA

Data regarding the modifyability of receptive field properties of visual system neurons were presented to indicate that both convolutional and matrix models of cognitive processing could reflect real mechanisms operating in the brain. Both of these types of models invoke representations which are distributed and content-addressable. The convolutional model is adept at processing serial order effects; the matrix model does better with categorizing. Serial ordering has been shown related to the Frontolimbic portions of the forebrain, while categorization is related to the systems of posterior convexity.

The curious theory of Bain and others that each idea is lodged in a separate ganglion cell (is) an hypothesis impossible both psychologically and physiologically (Ebbinghaus (1964), p. 63 footnote).

The proper way in which to think of this causation (the indirect strengthening of associations) would become clear only by the introduction of physiological conceptions which must first be constructed or at least remodeled (Ebbinghaus (1964), p. 122).

INTRODUCTION

In tackling the problem of "the introduction of physiological conceptions which must first be constructed or at least remodeled," I proposed the following in the summation session of the International Congress of Psychology held in Festival Hall, London, England:

One of the facts of the brain is the peculiar finding, made much of by Lashley, that poking holes into it does not seem to bother recognition or recall. A patient blind in half or more of his visual field can yet identify and recognize whatever he knew before. He does this with the remaining brain tissue which must contain sufficient information to
allow reconstruction of the whole. Experimental analysis of this clinical observation has established beyond doubt that: memory storage in the brain is distributed and redundant within the system involved.

Until recently there was no easily conceivable model of a distributed redundant memory which also fitted the anatomical construction of the brain. The puzzle was that the brain is wired in a very specific fashion — not at all like a random net — but the psychological functions of this structure demanded flexibility which only a distributed mechanisms could provide.

The advent of holographic photography and the science of optical information processing has changed all this. Holography depends on precise and specific relationships to attain distributed and redundant storage. This is effected by coding not only the intensity but especially the spatial phase relationships, the spatial neighbourhood interactions among simultaneously occurring events. Most holograms, and I believe neural holograms are of this type, result when waveforms initiated by a source reach their destination by two or more routes so that patterns of interfering wavefronts become established.

Holograms have marvellous properties. A small piece of holographic film still allow reconstruction of the whole image of the object photographed just as a small piece of brain will allow recognition to occur. When two or more objects have been photographed, transillumination of the film with light from one of the objects will give rise to ghost images of the others — true associative recall which can act as an active filter in information processing systems. Further, the useful storage capacity of holograms is phenomenal: already 100 million bits of memory have been stored in the space of one cubic centi-metre. This capacity results from the fact that holographic patterns can be overlayed and each retrieved without affecting the others when the specific input which generated the pattern occurs.

I believe that holography, i.e., optical information processing and memory storage, will influence psychology during the 1970s much as computer technology has influenced our work during the sixties. The digital, sequential computer gave rise to ways of handling the programming of symbolic codes. The analogue, spatial hologram will give rise to ways of handling what I shall call co-ordinate coding. Co-ordinate transformations are powerful tools for content addressible parallel information processing, giving instantaneous cross-correlations when desired, and so speeding pattern recognition to orders of time compatible with those observed in biological organisms. (Pribram (1969b), p. 60).

The predictions made at the Congress have taken a bit longer to come about than I had hoped and expected. But by the mid-1980s, models based on holographic-like processes are begin-
ning to become serious contenders for journal space and for discussion at conferences. This essay will report some of the more influential of those articles in the context of research from my own laboratory.

The work of my laboratory which has bearing on the issue of holographic-like models of memory has centered on the distinction between the functions of the frontolimbic formations of the forebrain and those of the cortical convexity. This essay will trace the impact of this research on two types of mathematical models of distributed representations of memory currently used in cognitive science.

Murdock has recently reviewed the evidence which distinguishes convolution and matrix theories of associative memory. He points out that whereas the matrix model as developed by Pike has the advantage of simplicity in obtaining explicit expressions and to some extent in storage capacity, the convolution-correlation model is more powerful in many other respects such as the handling of serial order information.

This presentation will address two related problems which stem from Murdock's review. One is general and the other specific. The general issue concerns the relationship of mathematical models to the "reality" of brain function. The specific issue is the one which concerns Murdock, i.e., which is the better of the two models to handle cognitive processing? With regard to the general issue, I will develop the theme that the mathematics does in fact mirror the realities of brain function; with regard to the specific issue I will describe experiments that indicate that both the matrix and the convolutional models have their place. These experiments will be taken up first since they are also the basis for the assertion regarding the relation of the mathematics to brain function.

MODIFICATIONS OF RECEPTIVE FIELD ORGANIZATION

The particular experiments from which I want to generalize were performed on the receptive field organization of single neurons in the lateral geniculate nucleus and the primary visual cortex of cats and monkeys (Spinelli and Pribram, Lassonde, Ptito and Pribram). The receptive field of a neuron is that part of its dendritic field which is functionally excited by a particular sensory input. In the visual system the input is presented monocularly to a subject whose gaze is fixed either by having the subject fixate a point in the environment or by temporarily paralyzing the eye muscles. Maps are then constructed which delineate the locations in the environment in which a small light will excite or inhibit the cell's output which is measured by recording with microelectrodes and accumulating the number of impulses generated by the cell while the light is in a particular position. In our experiments the
light was displayed as a small moving dot on a constrasting background. The location and motion of the dot were computer controlled and thus the computer could sum in a matrix of bins which represented the range over which the dot was moved, the number of impulses generated for each position of the dot since it "knew" where the dot was located.

Figure 1

Three-dimensional display of visual receptive field: The Mexican Hat function.
The results for the lateral geniculate nucleus are shown in Figures 1 and 2. The three dimensional representation, Figure 1, is usually called the Mexican hat function for obvious reasons. The brim of the hat represents the spontaneous background of impulsive activity of the neuron. The crown of the hat represents the excitation of the cell by the dot of light when it is located at the center of the receptive field. Where the crown meets the brim there is a depression which indicates that the output of the cell had been inhibited.

This center-surround organization, first described at the optic nerve level by Kuffler (1953) is more clearly shown in Fig. 2 which is a cross section of the hat parallel to the brim. The cut is made two standard deviations above the background activity which constitutes the brim. The inhibitory surround has been shown (e.g., Creutzfeldt, Kuhnt and Benevento (1974) for cortical cells) to be due to hyperpolarizing activity in a lateral network of "local circuit neurons" (Rakic (1976)) which are essentially axonless and therefore do not generate nerve impulses.

It is this inhibitors surround which can be augmented or diminished by electrical excitation in other parts of the forebrain. Stimulation of the frontal cortex or the head of the caudate nucleus diminishes the inhibitory surround, as indicate by F in Figure 2; stimulation of the posterior intrinsic (association) cortex, specifically in this case, the inferotemporal portion of this cortex, or of another of the basal ganglia, the putamen, produces an augmentation of the inhibitory surround as indicated by IT in this figure.

The results for cortical neurons are somewhat more complex since their receptive fields are elongated as first described by Hubel and Wiesel (1968). However, some of these receptive fields (classified as "simple" by Hubel and Wiesel) demonstrate inhibitory flanks and others (classified as complex), while failing to show this type of internal organization, can nonetheless change their overall functionally active size. Despite these complexities the effects of visual cortical cell receptive fields of electrical excitation of frontal cortex and caudate on the one hand and of posterior cortex and putamen on the other essentially parallel those obtained from the lateral geniculate nucleus.

Dendritic fields overlap to a considerable extent. Thus when the excitatory portion of the receptive fields become enlarged, the dendritic fields essentially merge into a more or less continuous functional field. By contrast, when the excitatory portion of the receptive fields shrinks each neuron becomes functionally isolated from its neighbor.

This modifiability of the primary visual system in the direction of greater separation or confluence among channels was supported by testing the effects of the same electrical stimulations on the recovery cycles of the system as recorded.
Visual-receptive field maps show how information flowing through the primary visual pathway is altered by stimulation elsewhere in the brain. Map a is the normal response of a cell in the geniculate nucleus when a light source is moved through a raster-like pattern. Map b shows how the field is contracted by stimulation of the inferior temporal cortex. Map c shows the expansion produced by stimulation of the frontal cortex. Map d is a final control taken 55 minutes after recording a. (From Spinelli and Pribram (1967)).
The change produced by cortical stimulation in recovery of a response in an afferent channel. Cortical stimulation of 8--10 Hertz was maintained continuously for several months. Control stimulations were performed on the parietal cortex. Records were made immediately after the onset of stimulation and weekly thereafter. The initial recovery functions and those obtained after one month are shown. Vertical bars represent actual variability of the records obtained in each group of four monkeys. The amplitudes of electrical responses evoked in primary visual cortex constitute the observed data. (From Spinelli and Pribram (1967)).
with small macroelectrodes. Figure 3 graphs the effects obtained: frontal stimulations produce a slowing of recovery while posterior stimulations result in a more rapid recovery as compared with an unstimulated baseline. Slow recovery indicates that the system is acting in unison; rapid recovery that the system is "multiplexed", i.e., that its channels are separated and not encumbered by a more extensively interconnected system with consequent greater inertia.

INTERPRETATION OF EXPERIMENTAL RESULTS

With respect to basing a mathematical processing model on these data it is clear that frontal brain stimulation drives the visual system towards a more continuous mode of operation while posterior stimulation drives the system toward a more discrete mode. This suggests that the convolution-correlation model is more appropriate when the focus of brain activity shifts forward and that the matrix model is more appropriate when the focus of brain activity lies more posteriorly. To test this interpretation we need to relate the known behavioral functions of the frontal and posterior portions of the brain to the known advantages of the two types of models.

Convolution-correlation mathematics have been used to model sensory-motor and perceptual-motor learning and skills. Thus Licklider (1959), Uttal (1975) and Reichardt (1978) developed temporal and spatial autocorrelation models to account for their results of experiments on optimotor and perceptual performances. Cooper (1984) and Kohonen ((1972), (1977) have used similar models to describe a variety of properties both perceptual and cognitive. Thus, e.g., Cooper has developed a model based on the effects of monocular deprivation on the responsiveness of neurons in the visual cortex and has made successful predictions of outcomes of experiments inspired by the model. Our own efforts (Pribram and Carlton (1985a, b)) have used this type of model to tease apart, imaging as a function of convolving the various stage of processing in the primary visual system, from object perception which depends on correlations among patterns in which centers of symmetry are determined by operations performed in the superior colliculus and the visuomotor system.

None of these perceptual and motor skills depend on functions which can be ascribed to the frontal part of the brain. Nor are they related to the inferotemporal cortex and the posterior intrinsic "association" systems of which the inferotemporal cortex is a part. What is suggested by these successful models is that the convolution-correlation approach is the more appropriate for fundamental perceptual and motor performance and that the matrix model, the more appropriate for certain specialized cognitive operations.
But, as noted, certain aspects of cognitive processing are more receptive to the convolutional-correlational approach. The modelling by these techniques of serial position effects is especially fruitful. It is therefore interesting that Milner has shown that patients with frontal lobe lesions are deficient when "temporal tagging" is necessary to solve a problem (1974). In my work with monkeys a similar results was obtained (Pribram and Tubbs (1967), Tubbs (1969), Pribram, Plotkin, Anderson and Leong (1977)). Frontal lobe resections drastically impair the ability of monkeys to perform a delayed alternation task. In this task the monkey has to alternate from one trial to the next, his choices between two identical boxes in order to obtain a reward. Trials are separated by equal time intervals during which a screen is interposed between the monkey and the boxes, which allows baiting the appropriate box by the experimenter out of sight of the monkey. The screen has been shown to act as a distractor (Malmo (1942)) suggesting that the defective performance of the frontal lobe lesioned subjects is due to their susceptibility to pro- and retroactive interference, a suggestion that has been supported by a number of outer findings (see Pribram (1961) and (1973) for review).

As shown in Figure 4, when interference is minimized by making the intertrial interval asymmetric, frontal lobe lesioned monkeys perform as well as their unoperated control subjects. This result, in addition to that obtained on receptive field structure and that on recovery cycles, supports the suggestion that, as a guide for exploring the neural mechanisms involved in cognitive processing of this type, a convolutional model is preferable to a matrix model.

Where then does the matrix model have the advantage? An essential difference between the convolutional and matrix models is that in the convolutional model critical operations are performed on the inner products of its vector while in the matrix models such operations utilize the outer products of vectors. Murdock states the issue as follows:

... the basic issue seems to be as follows. I would suggest that an association can be represented as a convolution, information is stored in a common memory vector, and correlation is the retrieval operation. Pike would suggest that an association is the outer product of two vectors, information is stored in a memory matrix or set of matrices, and vector-matrix premultiplication and postmultiplication is the retrieval operation ... ((1985), p. 132).

As treated by both Pike and Murdock, memory is of a piece. This is probably so at the deepest level of processing, and at this level, as already noted, I agree with Murdock that the convolutional approach "is not quite ready to be abandoned in favor of a matrix system" (ibid). But all memory is not associative in structure. Categorization and hierarchy characterize referential (semantic) systems and their
The modified alternation task which could be mastered readily by monkeys with part of their frontal cortex removed. The brain-damaged monkeys had been unable to solve the standard left-right alternation task even when the interval between trials was only five seconds. The task was then modified so that the intervals between trials described the pattern R 5 sec. L 15 sec. R 5 sec. L 15 sec. R 5 sec. L 15 sec. When this change was made, brain-damaged monkeys performed about as well as normal monkeys, as shown. Performance curves of frontal (upper) and control (lower) groups. Errors are the number made each day before a monkey achieved 40 successful trials. Bars indicate the range of errors made by different monkeys. Data for the 15th day show the result when all the trials were again separated by equal intervals of five seconds. (From Pribram and Tubbs (1967)).
operations are served by the posterior intrinsic "association"
mechanisms of the brain (for review of the neuropsychological
evidence on patients, see Warrington and McCarthy (1983)).

The task which is used in animal research to test for this
type of memory is a sensory discrimination problem in which
the subject must identify a cue on the basis of its consistent
reinforcement history. Thus a monkey is trained to choose a
box whose lid has painted on it a square rather than a plus
by virtue of the fact that a peanut is always found in the
box under the square and never in the one under the plus. Frontal
lesions have no effect on performing this task. Infero-
temporal lesions do (and have no effect on delayed alternation).
Resections of other parts of the posterior intrinsic "association"
cortex affect discrimination learning and performance
in other sense modalities: anterior temporal for taste;
superior temporal for hearing; parietal for somatosensory.
In all cases the defective performance is unique to the parti-
cular modality --- the resection does not affect performance
in any other modality. There is thus a multiple dissociation
between identification of cues in the various sense modalities
and also between these and delay-type tasks such as the delayed
alternation (for review see Pribram (1954), (1969a), (1972),
(1984)).

My tentative suggestion is that the matrix model is also not
yet ready to be abandoned. This model is clearly variable in
the hands of Anderson and his colleagues when applied to
learning and performance of discrimination type tasks (see
Anderson, Silverstein, Ritz and Jones (1977) for review). When-
ever categorical and hierarchical processes are involved,
storage as outer products of vectors and retrieval by post-
multiplication may be more appropriate than storage by
association in a common vector produced by convolving inner
products.

This line of reasoning leads to the suggestion that reference
--- semantic --- memory is best represented by a matrix model
and that the convolutional-correlation model be reserved for
some other type of memory. Tulving has differentiated "episodic"
from semantic memory by a variety of tasks (see Tulving (1972),
(1985) for review). My inclination is to now proceed to find
out if indeed the convolution-correlation model will more
effectively model all aspects of episodic memory than will a
matrix model. This inclination is furthered by the fact that
one central characteristic of episodic memory is its preser-
vation of some sort of place keeping, time tagging, and serial
ordering.
DO COMPUTATIONAL MODELS REFLECT REAL OPERATIONS IN THE BRAIN?

If the stance taken in the preceding section has merit, a good case can be made for a match between models of the sort described and the brain mechanisms they are wont to describe. What is necessary, however, is to select from the variety of possibilities that model which will most effectively describe and predict data at the neural as well as at the behavioral level. This means that closer attention must be given to the differences in processes served by the several systems of the brain. As described above, there is a difference between the operations of the sensory and the motor systems and between these and their associated intrinsic systems. Further, there is a difference in processing by the frontal and the posterior intrinsic systems. It would be surprising indeed if the identical modelling approach would hold for the various processes served by these different brain systems.

In my own work I have emphasized the difference between processing in the space-time domain and processing in the Fourier transform domain which is sometimes referred to as spectral, sometimes as frequency (in the mathematical, not the ordinary time dependent sense), sometimes as holographic. Gabor, the discoverer of the mathematics of holography, described in detail the relationship of this mathematics to that used in quantum physics (1946). Gabor elementary functions have been shown to characterize the receptive field properties of cortical neurons in the visual system (Pollen and Ronner (1981), Marjella (1980), Burgess, Wagner, Jennings and Barlow (1981)). These functions are composed of a Gaussian envelope which limits the extent of an otherwise infinite Fourier component in the transform domain.

Campbell and Robson and their colleagues (see Blakemore and Campbell (1969), Campbell (1974) and Campbell and Robson (1968) for review), and DeValois and his coworkers (see review by DeValois and DeValois (1980), DeValois, Albrecht and Thorell (1978a,b)) have shown that under certain experimental conditions (the drifting of gratings) the receptive fields of neurons in the visual cortex are each "tuned" to approximately an octave of "spatial frequency", the entire ensemble of neurons thus making up an overlapping "sounding board" of resonant Fourier-like elements. The visual system analyzes (and reconstructs) spatiotemporal patterns, much as the auditory system analyzes (and reconstructs) temporospatial.

The experiments described in the initial section of this presentation indicate that lateral inhibition may be responsible for the Gaussian limit on the otherwise infinite Fourier component. The Gabor elementary function can thus be modified in the direction of the Fourier domain or in the direction of the Gaussian. When pushed towards its Fourier aspects a convolutional model is appropriate since the resonant frequencies (represented by a vector in the model) can be specified. When the receptive (i.e., functional dendritic) field is pushed toward its Gaussian, an impulse
function representing a specific space-time coordinate in a matrix is the more appropriate model. Vectors then indicate the space-time pattern of impulse functions. Movshen, Tolhurst and Thompson (1978a,b,c) have, in principle, analyzed the composition of complex receptive fields into subfields which display excitatory and inhibitory properties by such an approach.

Many scientists have claimed that because the Fourier relationship is invertible (i.e., one can reconstruct the space-time order from the Fourier by simply reapplying the transform a second time) that the models are simply mathematical or computational devices by which we gain access to one or another aspect of a hidden reality. This view is similar to Bohr's complimentarity view (1934) in quantum physics and the view of critical philosophers (e.g., Feigel (1960)) with regard to the mind/brain issue. This multiple aspects view is shared by Mackay who describes it as mind talk and brain talk.

I have countered these multiple aspects approaches to the mind/brain issue by proposing a multiple realization approach (Pribram (1971a,b), (1985)). In my presentation here I have extended this approach to suggest that various precise theories in the form of mathematical models are not to be treated as merely alternative views treating different aspects of the some global cognitive/neural relationship but that the different models based on cognitive and neural data reflect the realities of processing when made sufficiently specific with regard to the brain systems they are wont to describe.

CONCLUSION

Data regarding the modifyability of receptive field properties of visual system neurons were presented to indicate that both convolutional and matrix models of cognitive processing could reflect real mechanisms operating in the brain. Both of these types of models invoke representations which are distributed and content-addressable. The convolutional model is adept at processing serial order effects; the matrix model does better with categorizing. Serial ordering has been shown related to the frontolimbic portions of the forebrain, while categorization is related to the systems of the posterior convexity. Electrical stimulation of frontolimbic structures results in a greater coherence among the activity of dendritic receptive field processes in the visual system --- coherence which is better approximated by convolutional models. Electrical stimulation of systems of the cerebral convexity results in functionally separating the activities of these various receptive field channels from one another --- separation which is better approximated by matrix models. There is thus every reason to utilize both convolutional and matrix models, and to infer that each reflects a "real" brain process.


(33) Pribram, K.H., A further experimental analysis of the behavioral deficit that follows injury to the primate frontal cortex, Exp. Neurol. 3 (1961) 432-466.


(36) Pribram, K.H., The realization of mind, Synthèse 22 (1971) 313-322. (a)


(43) Pribram, K.H. and Carlton, E., Object Perception (1985) In preparation. (b)


