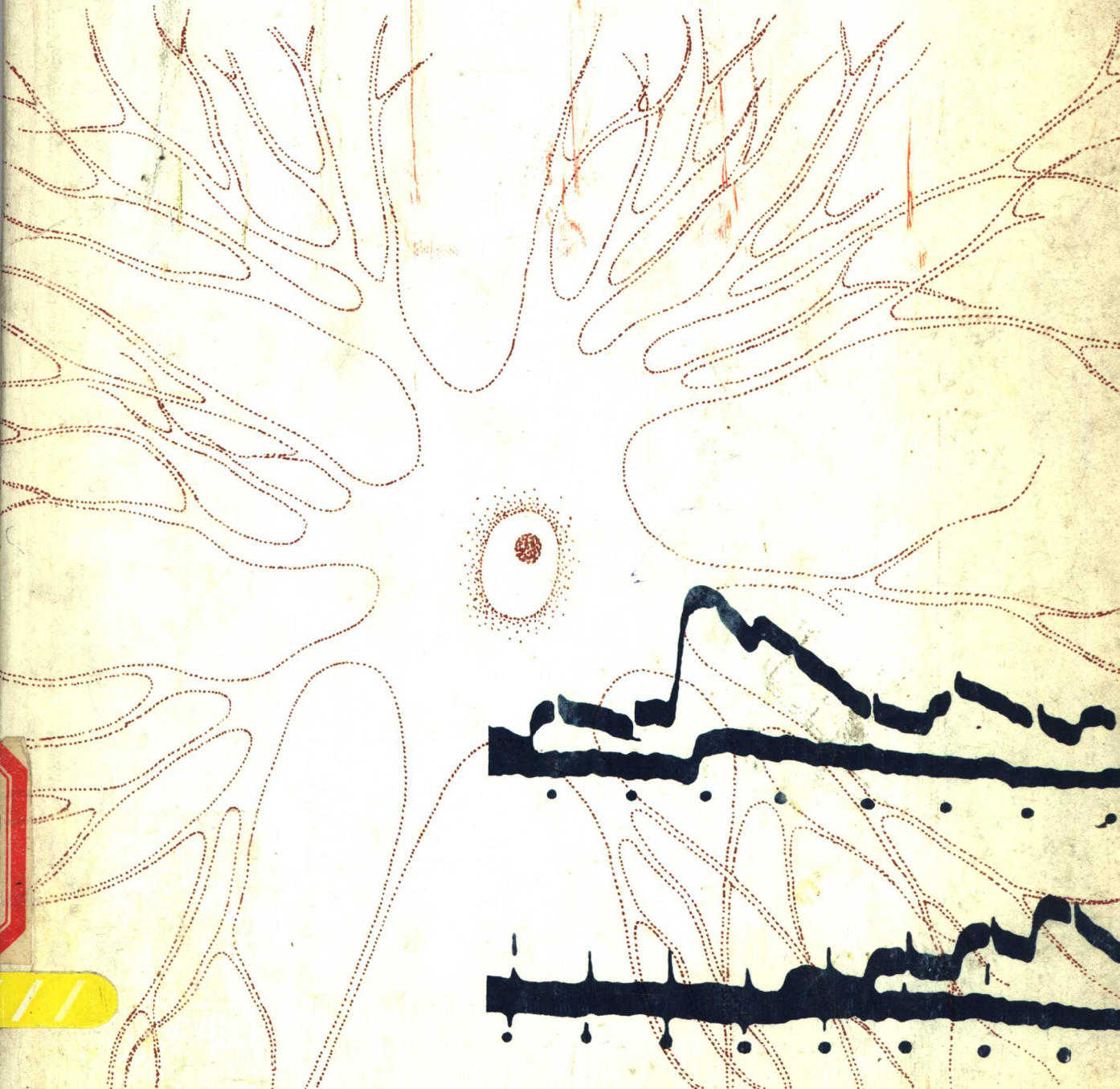
The Physical Principles of **Neuronal and Organismic Behavior**

Edited by
Michael Conrad + Magar E Magar
Center for Theoretical Studies,
Coral Gables, Florida

Gordon and Breach
Science Publishers



THE PHYSICAL PRINCIPLES OF NEURONAL AND ORGANISMIC BEHAVIOR

Edited by

MICHAEL CONRAD and MAGAR E. MAGAR
Center for Theoretical Studies, Coral Gables, Florida

Proceedings of the Conference on Physical Principles of Neuronal and Organismic Behavior, held at the Center for Theoretical Studies, Coral Gables, University of Miami.

GORDON AND BREACH SCIENCE PUBLISHERS

London New York Paris

Gordon and Breach, Science Publishers Ltd.
42 William IV Street
London W.C.2.

Editorial Office for the United States of America

Gordon and Breach, Science Publishers, Inc.
One Park Avenue
New York, N.Y. 10016

Editorial Office for France

Gordon and Breach
7–9 rue Emile Dubois
Paris 14e

These papers were first published in *The International Journal of Neuroscience*, Volume 2, Number 4/5, Volume 3, Numbers 1 to 6 and Volume 4, Number 1.

Dr. Michael Conrad is at present at the
University of California (Berkeley).

Library of Congress catalog card number applied for. ISBN 0 677 12290 X.

PREFACE

This volume contains a number of the papers delivered at the "Coral Gables Conference on the Physical Principles of Neuronal and Organismic Behavior", held on December 16–18, 1970. This was the second Coral Gables Biology Conference sponsored by the Center for Theoretical Studies. The first was on the physical principles of biological membranes.

In the past two decades great progress has been made in both molecular biology and the information sciences, as well as in neurophysiology and related areas. It is altogether likely that further progress in the neural sciences area will depend on a combination of these approaches.

The original intent in planning this conference was to focus on problems central to this program. Our feeling was that it would be most valuable to include diverse aspects of neural and behavioral science, but with the provison that these be represented by both experimental and theoretical work. Broadly speaking, papers in this volume examine the nervous system from the point of view of development, through theoretical and physiological studies of neurons and information processing, and from the molecular point of view. The last session was devoted to the evolutionary aspect of organism behavior, artificial intelligence, as well as methodological discussions.

Needless to say, the inherent limitations of a conference prevented us from including many important aspects of the subject. Indeed much of the discussion at the conference centered around which approaches are important, as well as on the applicability of the methodology of physics, the possibility for novel physics, and the proper relation of theory to experiment. These questions can only be answered by future developments.

We would like to acknowledge the Alfred P. Sloan Foundation for their financial support.

M. Conrad
M. Magar

CONFERENCE COMMITTEE

Behram Kursunoglu (Chairman)
Michael Conrad
Magar E. Magar
Norma Gayle Hagan

Conference sponsored by Alfred P. Sloan Foundation

CONFERENCE PROGRAM

Session I

Models of Embryological Processes
(Session Organizer: Michael Arbib)
Chairman: Francis O. Schmitt

Session II

Theoretical Models of Neurons and Neuron Interactions
Chairman: Jerome Wolken

Session III

Cortical Events as a Result of Diverse Stimulation
Chairman: Sidney Fox

Session IV

Molecular Aspect of the Nervous System
Chairman: Howard H. Pattee

Session V

Invariance, Constraint, Hierarchy and the Evolutionary Aspect of Organism Behavior
Chairman: David E. Green

CONTENTS

PART ONE
Models of Embryological Processes

ORGANIZATIONAL PRINCIPLES FOR EMBRYOLOGICAL AND NEUROPHYSIOLOGICAL PROCESSES†

MICHAEL A. ARBIB

Department of Computer and Information Science, University of Massachusetts, Amherst, Massachusetts 01002

The author argues for the development of organizational principles for multicellular assemblies to complement the approach of molecular biologists to cellular mechanisms, and notes that such development requires an inductive phase before detailed deductions can be useful. He notes the challenge posed by neural specificity to theorists of embryological development, and suggests 'layered somatotopy' as a useful organizing principle for brain theory.

This paper is dedicated to the memory of Donald M. Wilson whose research on insect flight and locomotion combined incisive experiment with theoretical insight. His untimely death in a 'rough river' boating expedition has saddened all of us who valued his friendship and had looked forward to enjoying his company and sharing his insights many times in years ahead.

1. *Approaches to Theoretical Biology*‡

There seem to be two main directions whereby the physicist can approach biology. In one, he tries to do conventional physics as long as he can, and gets into the biophysics of DNA, say, whence he finds himself slowly growing into molecular biology, and, hopefully, thence into the principles which govern the interactions of cells and the growth of those interactions. A second, more romantic, approach is to start from questions about overall function of body and mind, and try to find mechanisms in terms of which one could answer them, and so go down, for example, to explore what sort of brain functions we would need for intelligent behavior, and thence to the neural structures that could subserve those functions.

To point up this distinction between going 'up' from the basic biochemistry and physics and going 'down' from overall functional questions, we might well look at Computer and Information Science, where we see a very marked division of labour. On the one hand, we have the electrical engineer using solid-state physics in trying to push to the ultimate reduction of size, increase in speed of operation and flexibility of function for devices which are then to be built into computers. However, the computer scientist, having got those components, is not at all concerned with the actual physics involved. He wants to be guaranteed that the components have certain functions, but then his concern is to figure out how to put together large scale organizations in terms of those com-

† Read at Conference on Physical Principles of Neuronal and Organismic Behavior, December 16–18, 1970, Center for Theoretical Studies, University of Miami, Coral Gables, Florida, U.S.A.

‡ In this section I indulge in a long-winded discussion of some problems in the philosophy of science raised by the interactions between experimentalists and theoreticians at the conference. Some readers may wish to turn directly to Section 2.

ponent functions to get some overall sophisticated function. Thus, the component level—or the cellular level, to be more biological about it—is the meeting ground for two quite different approaches. To explain how cells 'work', and their capacities for interaction, is the task of the biophysicist and biochemist. On the other hand, understanding how to organize large collections of such components seems to require such approaches of Computer and Information Science as automata theory and computer simulation.

To make this distinction another way, we may recall those happy days a few years ago when having 'cracked the genetic code' many people talked (at least to the popular press!) as if all biological mysteries were solved, at least in essence, and some people were not joking when they talked of 'cracking the brain code'. Now that we have had some years to reflect upon the 'cracking' of the genetic code, we well realise that the transduction from DNA via RNA to amino acids does not explain all the wonders of cellular behavior, let alone the complex dynamics of embryology. No more will finding out in great detail the functional characteristics of neurons solve all the problems of neural organization.

By now, many hitherto mysterious properties of cells—the basic 'components' of all organisms—have been explained in terms of biochemistry and molecular biology, and many papers in these *Proceedings* document the power of such explanations. It would be foolish to try to belittle such achievements, and I do not wish to do so—but I do wish to caution the reader against the all too common mistake of being so dazzled by their success as to believe that biochemistry alone can unravel all the knotty problems of biology, and that the development of new theoretical approaches is unnecessary. I believe that some of the less subtle research on the chemistry of memory amply testifies to the dangers of such a view. We all agree that cells are living systems and that learning involves changes in the brain, and thus—in some sense—learning is a growth phenomenon. Thus it is hardly surprising that substances which block RNA synthesis—and thus cell growth—interfere with an animal's learning. But to go from this to making statements like 'Therefore, RNA is the memory molecule' is as useful as noting that cutting off the electricity supply disrupts the storage of information in a computer and deducing that 'Therefore, electrons are the building blocks of memory'. We have a theory of complex computer memory structures based on the properties of the switching and storage elements. It is irrelevant to this theory whether component properties are mediated by electrical, magnetic, hydraulic or chemical mechanisms. Similarly a theory of the brain will not be so much in terms of biochemistry as in terms of organizational principles for neurons. Biochemistry is irrelevant to such a theory of organization *per se*—but is vitally important in helping us understand the detailed properties of these components. In studying human perception, biochemistry may be of little relevance, while organizational principles predominate. In studying drug therapy, precisely the opposite balance may hold.

In the rest of this paper, I shall stress the search for organizational principles, but do *not* try to argue the superiority of this approach to that of the molecular biologist, but rather argue the complementarity of the two approaches. In fact, even in the sketchy presentation that so short a paper as this decrees, we shall often see the organizational approach—'If cells can do such-and-such then an array of them with certain properties will develop or process information in a way which is thus explained'—immediately raising complementary questions—'Is it physically possible for a cell to do such-and-such, and if so what biochemical mechanisms are involved?'

The failure to note this complementarity and instead argue for the superiority of one's own approach led to certain tensions at the conference itself,† perhaps expressed most noticeably in the desire of a few experimentalists to discredit all theoretical efforts which were not slavishly tied to experiment—forgetting that new concepts and paradigms are often required before experiments can be designed which extract really meaningful data. [Incidentally, there is a certain delicious irony in the chiding of some biologists against over-involvement in theory, for at other conferences, some mathematicians complain that over-involvement with physical intuition distracts from the cool algebraic beauty of purely formal deductions within an axiomatic system. However, one may suggest—somewhat presumptuously, perhaps—that it may be possible to converge eventually upon

† Here, and elsewhere in this section, I am building a 'straw-man' whose demolition will bolster my arguments. I hope that no conferee, noting a straw or two of his own, will make the hurtful mistake of believing that the whole assemblage is intended as an unkind reconstruction of his overall viewpoint.

some judicious balance of intuition and theory which will yield enough understanding to compensate for such chiding!]

The real cleavage which emerged at the conference was not, perhaps, that between experiment and theory, but rather that between science and technique. The scientist seeks for understanding and will use whatever blend of theory and experiment seems best to match his intellect to the task at hand, while the technician seeks to apply a technique he has mastered. Sometimes the technique is appropriate to the task, and the lucky technician may bask in the glow of scientific achievement. But all too often the technique is good for nothing but producing irrelevant papers—be they by the experimenter-technician pumping drugs to be assayed in cat after dying cat, or the theorist-technician using the mathematics of physics to churn out equation after tedious equation assessing the progress of a Hamiltonian which has little relevance to the organism it is posited to represent.

Perhaps with this distinction in mind, we can see one of the greatest pitfalls facing the physicist turning theoretical biologist. Too much of his education has involved his mastering long-polished mathematical techniques, and finding that wide classes of problems can be solved simply by 'plugging-in' these techniques. With this as background, it is all too easy to believe that he can solve the biologist's problems by the technician's approach (though he will think it is scientific) of 'plugging-in' these techniques to biological situations, little realizing that a great period of induction and experimentation (yes, even theorists must experiment—even if only with symbolic constructs) was required to match technique to problem. But theory *is* required in biology—as in any science where constructs become subtle enough to escape the domain of the immediately observable and where the depth of argument comes to exceed the usual grasp of common sense—and what remains to be determined is not whether there shall be theoretical biology, but rather what forms theoretical biology shall take. The theorist who can make a substantial contribution will probably be one who combines an intimate knowledge of the experimental data of some restricted problem in biology with a broad command of theoretical techniques, and uses the interaction between his reformulation and reconceptualization of the data and his reworking of the techniques to evolve genuinely new insights into that particular biological problem—only to find that those insights are valid elsewhere. There is no recipe for this.

2. *Neural Specificity and Theoretical Embryology*

We cannot understand the central nervous system unless we understand to what extent we are dealing with a genetically determined structure and to what extent we are dealing with a loosely specified structure which is to be shaped by adaptation. In fact, we now know that there is a great deal of specificity in the organism and in particular in the nervous system, and the papers of Ede, Schroeder and Laing (this volume) give some rather detailed ideas about the experimental and theoretical tools with which we can determine the mechanisms which give rise to that specificity.

Until perhaps 1940, a popular view of the nervous system was as an essentially random network. The idea was that if a creature, with adequate receptors and effectors, were put in some complicated environment and 'punished' when it did something 'wrong' and 'rewarded' when it did something 'right', then eventually the correct connections would be made to enable the organism to function effectively in its environment. Such a belief was based on rather gross observations upon humans with polio who had a flexor muscle, say, wither away and had had the remaining healthy extensor muscle cut in two and so sutured that half the muscle kept its old function of extension while the other half now had the opposite function of flexion. After extensive therapy, patients were able to adapt the muscle to its new use, so long as they were carrying out careful voluntary movements. By extension from these results, many people thought the nervous system was completely 'plastic'—i.e., that all connections could be—and in fact were—moulded by experience. We owe to Paul Weiss and Sperry and other workers the knowledge that plasticity is *not* unlimited, and that in fact there is a great deal of neuronal specificity—i.e. *genetics constrains many details of neuronal connections which cannot be changed by experience unless there exist specific brain structures to exploit that experience.*

On the other hand, a newborn baby has to be able to suck, to breathe, to excrete and so on. It cannot do many other things at birth, but has to be able to learn how to do them, and this cannot happen unless it has appropriate structures to implement learning. This point may seem obvious, but is so often lost sight of that it may pay to

belabour it with an obvious example. Think of tossing a coin repeatedly. Every time it comes up heads, spray it with Chanel No. 5, and every time it comes up tails spray it with stale cabbage juice. It hardly seems profound to doubt that the coin will eventually tend to come up heads rather than tails, but it may be helpful to explicate the grounds for our doubt.

Firstly, the coin does not have receptors which allow it to distinguish Chanel No. 5 from cabbage juice. Secondly, even if it could distinguish them, it has no inbuilt criteria to determine which is preferable. Thirdly, even if it could tell which was preferable it has no mechanism whereby it could make use of that knowledge to change its behaviour. Thus, in looking at the embryology of the nervous system we have to look for specificity, whether in direct sense of determining networks which will mediate innate behaviour patterns, or to provide the adaptational substrate to enable the organism to adapt its evolutionary heritage to the exigencies of its own environment. We have to understand how appropriate receptor and effector arrays can be structured, how basic drive mechanisms can be 'built into' the organism so that it can shape its behaviour on the basis of some evolutionarily determined criteria of biological usefulness or destructiveness, and we must understand—at least in mammals—the determination of a sufficiently rich cortical structure to allow sophisticated learning. To enhance the latter point by a striking contrast, we may recall Paul Weiss' [1941, for an overview] intriguing experiments in which the forelimbs were reversed in the larval stage. When the salamander grew to salamanderhood then, whenever it would see some food in front of it, the brain would send the appropriate command of 'advance', but unfortunately the neural circuitry in the brainstem which interpreted the command did not 'know' that the forelegs were back-to-front and so would send the sequence of muscular activation which would cause the forelimbs to try and make the animal scurry away from its food. No matter how long the animal was exposed to this unfortunate situation it could never learn what was wrong—or, at least, if it learned what was wrong, it could not do anything about it. Thus we see the necessity for adequate structure if learning is to ensue.

Notice that what we are talking about in the nervous system is not the development of individual organs per se, but rather the development of functional systems which involve the whole organism. The animal at birth has to be able to take tactile stimuli on the lips and go through the 'computation' required to convert this into a sucking reflex. If we look at animals such as the guinea pig in which the hindlimbs are more important than the forelimbs at birth we will find the uneven development of the spinal cord which insures that the hindlimbs are ready to function at birth. This is what the Russian physiologist Anokhin [1964] refers to as *systemogenesis*—we have to think of the nervous system not in terms of anatomically defined lumps of tissue, but rather in terms of an interacting overlapping collection of systems for carrying out biologically important functions. Thus, our task becomes even more complicated when we realize that it is not enough to look at one small part of the body or the nervous system and explain how it grows, but we have to explain the sort of synchrony which allows functioning systems of various kinds to be available at birth and at later stages of maturation. The models of Ede and Schroeder are at the simpler stage of studying morphogenesis of single organs—this seems to be a necessary way-station in the evolution of our models before we can tackle the synchrony problems of systemogenesis. At the moment, we look at one organ in a system and try to explain what sort of cellular interaction can give rise to its shaping. We may hope that, later on, when we understand this, we will have the intellectual apparatus in place to combine together our models of several systems to understand what sort of overall mechanisms allow coordination of their development.

Having established, in Section 1, the cellular level as an appropriate intermediate between the study of macromolecules and organelles by the biophysicist and biochemist, and the study of organismic control by the computer and information scientist and having now seen the interest of understanding embryological processes, let us briefly mention some of the mechanisms at the cellular level which shape the overall form of the organism, including that of its nervous system. (The reader will find an excellent overview of 'the forces that shape the embryo' in Trinkaus [1969].)

One mechanism whereby a tissue may change its form is that of the autonomous change in cell shape. For example it is now well-known that various microstructures may be synthesised within cells during characteristic changes of shape, and that their destruction impairs such changes. Thus cells seem able to elongate themselves by producing

microtubules aligned parallel to the axis of elongation. Again, cells seem able to constrict a portion of themselves by producing microfilaments which can then contract to provide the constriction by a sort of 'purse-string effect' [Baker and Schroeder, 1967]. Schroeder [1970, 1971] has combined such mechanisms to provide an elegant model of neurulation—the process whereby a plate of cells on the back of the embryo is formed into a trough which then rolls up into a tube running the length of the embryo to then disappear beneath the surface of the back and form the rudiments of the spinal cord and brain. A crude caricature of the mechanism is shown in Figure 1—the reader will find a more subtle and careful treatment in Schroeder's paper in this volume.

Another mechanism whereby a tissue may change its form involves the combined effects of cellular adhesiveness and cellular motility. Such a mechanism helps us understand situations in which the attachments of cells change over time, but where there seem to be important specificities in the ensuing pattern of cellular attachments. Gustafson and Wolpert [1967—for an exposition see also Wolpert and Gustafson, 1967] have given a masterly analysis of cellular movement and contact in sea urchin morphogenesis. Ede and Agerbak [1968] have been able to correlate changes in adhesiveness of cells (and the consequent change in their motility) in normal and *talpid*[3] mutant chick embryos with changes in the developing limb pattern in these embryos, while Ede and Law [1969] (see also Ede [1971]) have expressed this correlation in the specific form of a computer simulation of limb development.

While elegantly showing how changes in cell shape, motility or adhesiveness can provide mechanisms for morphogenesis—both in the nervous system and elsewhere—the above schemes do not make explicit how a cell 'knows' what contribution it is to make in the overall pattern. It is for this reason that other workers have developed the idea of 'positional information'. Here, the line of argument runs 'If the cell is to change appropriately it must have information about its position within the embryo (and perhaps it will need to consult a clock, too).' An early approach to such positional information was in gradient theory (e.g. Child [1941])—if a source of some metabolite were located at one end of the axis and a sink at the other, with a uniform gradient in between, then the concentration of metabolite in any cell would signal its position on the axis. Wolpert [1969] has suggested ways in which such a model needs refinement and elaboration, and Goodwin and Cohen [1969] have instantiated Wolpert's ideas in a model in which position is signalled by the phase differences between families of pulses propagating with different delays from cell to cell. By contrast, automata theorists have shown how cells may be formed into complex arrays without explicit 'addressing'. Rather, each cell is capable of a finite number of states, and at any time the cell changes state in a way dependent upon its previous state and that of its neighbors. For example, von Neumann [1966] exhibited a self-reproducing array with tens of thousands of components, but the cells were only capable of 29 states, and so could not 'know where they were'. Arbib [1967] has attempted to place this approach

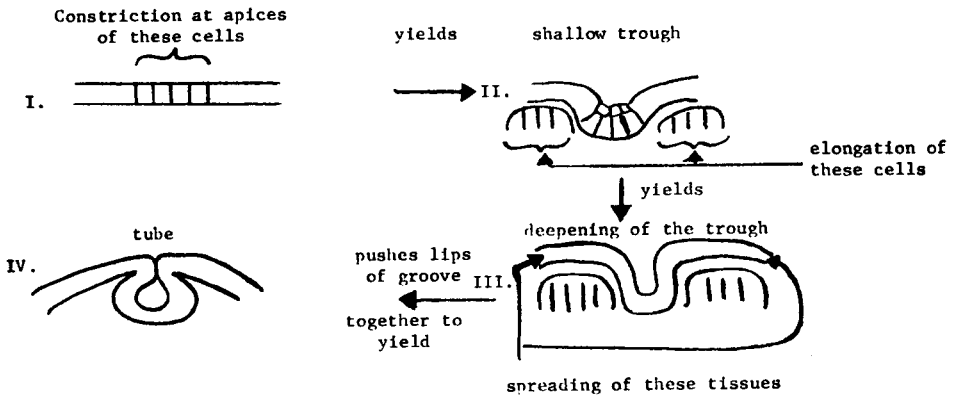


FIGURE 1 Dangerously oversimplified schematic of neurulation. The 4 stages are not chronological. Rather, each of the 3 transitions schematises a mechanism (there are others) found by Schroeder [1970] to play a role in forming the neural tube. All views are in cross-section.

in a more biological context. The work of Apter [1966] should also be mentioned here. Other authors have compared the change of state rules used by von Neumann and others to the rewriting rules employed by linguists to 'grow' a sentence via its grammatical description, and are now exploring the applicability of formal linguistics to theoretical embryology (Lindenmayer [1968], Laing [1969, 1971]).

In considering the specificity of cellular connections, we must not be misled by estimates that the amount of information in DNA is far less than that contained in the connections of the brain, which some have taken to imply that connections in the brain must be random. To see this, consider the following computer program which comprises four instructions:

1. Set n equal to zero.
2. Print out n.
3. Replace n by $n+1$
4. Return to the second instruction.

If you observe a computer executing this program, it will emit a stream of numbers which is endless—at least till you have exhausted the capacity of the computer. Arguments that a comparison of the number of DNA bases with the number of connections in the brain shows that the brain must be a random network are as naive as comparing the four instructions of the above program with the number of positive integers and concluding that the sequence of positive integers is a random sequence! In other words, one of the things we know from our study of programming computers to do clever things, is that our programs have loops within them which are hierarchically structured to provide for a great deal of economy in the way we specify processes. As a biological example of a plausible 'use' of such 'nested subroutines', we may cite the retina of the frog. The connections between the interneurons of the second layer in the retina and the ganglion cells which send their output down the optic tract to the brain have been schematized by Lettvin and Maturana as falling into two or three segregated layers. A plausible wiring scheme would then prescribe that certain types of axons from the interneurons terminated in one layer and so are highly likely to connect one level of the dendrites of the ganglion cells while other types of axons bearing different transforms of the visual input would terminate in the other layer thus hitting other parts of the ganglion cell dendrites. By this means, one can very simply specify how to get a retina that would function perfectly for the frog trying to snap flies in its world, without having to specify point-by-point interconnections. Hence, a sort of 'nested subroutine' approach could probably explain a great deal of the specificity of the nervous system without requiring an immense investment in genetic material. In the next section we shall see—among many other facts—that such economy of genetic prescription augurs well for economy of functional description when we come to describe organizational principles for neurophysiological processes.

3. *Implications for Brain Modelling*

Having gained some idea of the specificity of structure there is in the nervous system let us see the effect our views of brain structure will have upon our approach to modelling brain function.

In certain invertebrates, we may find that the function of the system we wish to explain is mediated by a rather small neural network and so we might actually hope to track down, by explicitly simulating the behaviour of say 100 or so neurons, all the details of their interaction, and so obtain a plausibly complete explanation of how a locust, say, walks or flies. (See, for example, the beautiful review of 'Insect Walking' by Donald Wilson [1966]).

When we turn to vertebrates, this strategy does not work, save in studies of peripheral circuits for muscle control, for there are just too many neurons. There are various strategies to take, depending on ideas about structure, as to how one might make a model. The physicist has one ready answer for how we might model a system with millions or even billions of neurons. From his study of gases he would suggest statistical mechanics [or—in technical terms—'average the hell out of it']. Unfortunately, such averaging may destroy the very parameters of interest to us if we want to explain linguistic behavior or coordinated motor behavior as in a frog snapping at a fly. On the other hand, if we want to understand how the cooperative behavior of many billions of neurons in the cortex gives rise to evoked potentials or electroencephalograms, then some sort of statistical mechanical approach may well be worth while. However, a straightforward statistical approach to the very large system will not do for more detailed structural questions about complex information processing in brains.

In this context, it may be worthwhile to contrast two types of statistical models. Winograd and

Cowan [1963] were concerned with the fact that in as large a system as the brain one has both the likelihood of not specifying completely all neural interconnections accurately by genetic parameters and also the likelihood of many 'malfunctions' of components during actual information processing by the system. They wished to design networks with enough redundancy to insure that the organism would not be too unreliable. Their strategy was to start from a very specific function they wanted a hypothetical 'nervous system' to undertake, and then provide ways in which they could transform this 'nervous system' into a new form which was sufficiently redundant that quite a lot of sloppiness in the 'wiring' and in the behavior of the 'neurons' would still give correct overall function. This strategy of starting from a specific structure for computing some function and finding ways of introducing redundancy to make it resistant to certain types of damage both in growth and function is radically different from the strategy Cowan [1969] took in his later work, in which he looked at interactions between thalamus and cortex only in terms of gross statistical parameters of their interconnection, and then asked if certain aspects such as cortical rhythms could be explained on this basis. In this case one only wants some crude parameters of overall system behavior such as the period of rhythm recorded in gross potentials, and so one can 'average out' a lot of detail by statistical mechanical techniques. But if one wants to look at the detailed state-dependent processing of inputs to get outputs then one has to impose far more structure, and study deterministic operation at a certain level.

Another approach to modelling a large system is that of compartment models. A brain modeller taking such an approach will not try to average over the complete system, but will look at the gross anatomy of the brain to subdivide the brain into various regions. He will thus try to simplify the problems of explaining one large region of the brain by breaking it down into a collection of interconnected 'black boxes' and see if by making multiple plausible guesses about those boxes and their interconnections he can put together a reasonably functional model of the overall system. It may then be easier to take those individual boxes with their plausible functions and try to model them back down to the cellular level than trying to do the whole thing directly. Perhaps one of the most interesting brain models of this kind is that of Kilmer, McCulloch and Blum [1969] on the reticular formation. They looked at the neuroanatomy and saw that the formation could be viewed as a series of 'modules' ascending the longitudinal axis. Each module could then be modeled as a whole, and then the simulacra could be interconnected to get the overall change-of-mode behavior which they posit to be exhibited by the reticular formation.

With this as background, I want to suggest eight principles which may help us understand how the human brain can control the complexities of a human's behavior.† [Their elaboration will appear in my forthcoming book 'The Metaphorical Brain' (Arbib [1972])]:

1. *Theory Must Be Action-Oriented*

One often talks as if human perception merely involved being able, when shown an object, to respond by naming it correctly. However, it is often more appropriate to say of an animal that it perceives its environment to the extent that it can interact appropriately with that environment. We can perceive a cat by naming it, true, but our perception may involve no conscious awareness of its being a cat per se, as when it jumps on our lap while we are reading and we simply classify it by the action we take as 'something-to-be-stroked' or 'something-to-be-pushed-off'. In computer jargon, then, we may say that perception of an object generally involves the gaining of access to 'programs' for controlling interaction with the object, rather than simply generating a 'label' for the object.

2. *Redundancy of Potential Command*

We were careful to characterize perception as a gaining of access to a program rather than the execution of a program—you perceive something and yet may still leave it alone. Thus in gaining access to the program, the system only gives it *potential* command, further processing being required to determine whether or not to act. A key question will thus be 'How is the central nervous system structured to allow coordinated action of the whole animal when different regions receive contradictory local information?' McCulloch suggested that the answer lay in the *Principle of Redundancy of Potential Command* which states,

† Many of the ideas in the rest of this section were developed in Arbib [1971] on the basis of work by Arbib Dev and Didday (see also, Arbib, Dev and Didday [1971] and Didday [1971a,b]).

essentially, that command should pass to the region with the most important information. He cited the example of the behavior of a World War I naval fleet controlled—at least temporarily—by the signals from whichever ship first sighted the enemy, the point being that this ship need not be the flagship, in which command normally resided.

McCulloch further suggested that this redundancy of potential command in vertebrates would find its clearest expression in the reticular formation of the brain stem (RF). Kilmer and McCulloch then made the following observations towards building a model of RF:

(i) They noted that at any one time an animal is in only one of some 20 or so gross *modes* of behavior—sleeping, eating, grooming, mating, urinating, for example—and posited that the main role of the core of the RF (or at least the role they sought to model) was to commit the organism to one of these modes.

(ii) They noted that anatomical data of the Scheibels [1958] suggested that RF need not be modelled neuron by neuron, but could instead be considered as a stack of 'poker chips', each containing tens of thousands of neurons, and each with its own nexus of sensory information.

(iii) They posited that each module ('poker chip') could decide which mode was most appropriate to its own nexus of information, and then asked, 'How can the modules be coupled so that, in real-time, a consensus can be reached as to the mode appropriate to the overall sensory input, despite conflicting mode indications from local inputs to different modules?'

This was the framework within which Kilmer, McCulloch and Blum [1969] designed and simulated the compartment model, called S-RETIC, which we have discussed above of a system to compute mode changes, comprising a column of modules which differed only in their input array, and which were interconnected in a way suggested by RF anatomy.

Another useful system for the study of redundancy of potential command is the frog, which is normally immobile, but will snap at any fly that comes into suitable range—'snapping' comprising a movement of the head (and, when necessary, the body) to aim at the fly and the rapid extension of the tongue to 'zap' the fly. The situation seems very simple in that the frog does not seem to recognize flies as such—rather it will snap at any wiggling object, but will not snap at a stationary (i.e. dead) fly. A frog confronted with two flies then presents us with a beautifully simple redundant command situation—normally the animal snaps at one of the flies, and so we have sought to model the brain mechanism that determines which of the flies will 'take command' of the frog.

Lettvin, Maturana, McCulloch and Pitts (1959) found that most ganglion cells of the frog's retina could be classified as being one of four types—such as 'moving spot detectors' and 'large moving object detectors'. Incidentally, none of these are found in the cat, whose retinal ganglion cells may better be characterized as 'contrast enhancement devices'. This ties in with our first action-oriented point of view—a frog with little visually-guided behavior beyond snapping at 'wiggles' and jumping away from 'enemies' has a retina which 'throws away' most aspects of the visual input not related to these features, whereas a cat, leading a subtle life (such as watching a mousehole intently, and only springing when the mouse pokes his head out far enough) cannot function with so specialized a retina.

To get to our third point, however, what we want to emphasize is the way in which the information from the four types of detectors is distributed in the brain. Their axons terminate (among other places) in a brain structure called the *tectum*, with the terminations forming four separate layers, one atop each other, with the properties that (a) different layers correspond to different types of detector; (b) each layer preserves the spatial relations between the original cells (i.e. there exists a direction along the layer corresponding to moving across the retina); (c) that terminations stacked above one another in the four layers come from ganglion cells with overlapping receptive fields. This is another dramatic case of the neural specificity we have discussed in Section 2 (see also Sperry [1951]). Such a relationship between two layers of cells—in this case the retina and any of the four tectal layers—is called somatotopic, from the Greek *soma* (body) and *topos* (place), since it preserves information about place on the body as we move from receptors to the central nervous system. We thus state our third principle, known to all neuroanatomists, which has been strangely neglected in brain theory:

3. *The Brain 'is' a layered Somatotopic Computer*

It should be noted that somatotopy may preserve