

Logical Structure of Environment and Its Internal Representation

Heinz Von Foerster

BCL #96, 1962

It is for the cultural anthropologists to explain the great popularity of a recent publication¹ that bases man's evolutionary success on his alleged intrinsic habit to kill and to destroy. Since this belief is contrary to the thesis on which my paper today rests, I hope you will forgive me if I open my remarks with an apparently unpopular confession: in contrast to the tenets of some self-appointed apostles of the glories of annihilation and destruction, I believe that love, not hatred and mutual killing, is the mainspring of man's cultural and spiritual evolution.

Clearly you will not let me get away with such an outrageous assertion, without sufficient evidence to support my proposition. To this end permit me to translate this assertion into somewhat more scientific terms. This can be done most advantageously in terms of a fast developing, fascinating branch of mathematics, namely "game theory." This theory deals, among other game situations, with decision-making under insufficient information, determining those moves of a player which would give him maximum gain with minimum risk. In the particular case where a group of players have to play against a not completely determinable opponent — as man in his struggle against the partly unpredictable forces of nature — it can be easily shown that a "coalition structure" is much stronger than a "competitive structure." By "stronger" I mean that the pay-off per element is higher; and my assertion that a coalition structure is stronger than a competitive structure refers to the fact that two elements jointly can do things which the two elements separately can never achieve. In other words, by joining, each element gets more out of the deal than if it remains single. As trivial examples of advantages of coalitions, may I point to man's social build up during historical time, his vigorous urbanization in recent centuries, and his extensive development of the means of mass communication. Although this observation is almost trivial there are two facets which are — in my eyes — of considerable importance.

First, a coalition is a much more sophisticated structure than a competition, because it requires the possibility of the elements to communicate with each other.

As you probably know, all social animals — bees, ants, or animals that live in herds — constantly exchange denotative information about food, danger, and individual states of mind such as anger, submissiveness, etc. If I am not mistaken, von Frisch's² analysis of communication among bees allows for a bee's vocabulary of about 200 "words." I could give you a host of fascinating examples of information exchange in animals. And it is quite obvious that those poor creatures doomed incommunicado have to resort to a rather poor competitive game. Since evolution is cashing in at even the slightest edge of an advantage, it is clear that evolution fosters communication.

I. The Logical Structure of "Environment."

The second facet of the advantage of a coalition to which I want to draw your attention is that this structure is an example of the old saying that "the whole is more than the sum of its parts." Although we seem to understand very well what this means, this statement has been attacked by positivists and operationalists time and again, rightly so, I think because the way it stands it clearly is nonsense. Two and two are both parts of four, but $2 + 2 = 4$ and not a tiny bit more or less. However, if what we want to say by this statement is properly formulated, a most profound principle is defined. It is the principle of super-additive compositions for elements making up a system. Let me first give a precise formulation of this principle in rather abstract terms, and later illustrate the application of this principle to pertinent concrete situations. What we really want to say is: "A measure of the sum of the parts is larger than the sum of the measure of the parts." This statement can be formulated in even more precise, mathematical terms. Consider F to be a measure function. If you recall that the symbol " ζ " stands for "left side larger than right side," the above statement can be written in the following way:

$$F(a + b) > F(a) + F(b).$$

In order to make this highly symbolical expression more tangible, let me suggest a simple example which

may evoke old high school memories. Take for the moment as an example of a measure function the operation “squaring;” that is

$$F(x) = (x)^2.$$

Squaring all F-expressions in our first equation we obtain:

$$F(a+b) = (a+b)^2 = a^2 + b^2 + 2ab.$$

and

$$F(a) = a^2, F(b) = b^2.$$

Putting the results back in form of our first equation, we obtain the undeniable truth that indeed:

$$a^2 + b^2 + 2ab > a^2 + b^2.$$

The margin which makes the left hand side of this inequality larger than the right hand side is, of course, the product $2ab$. This provides us with an important clue. The product $2ab$ is nothing else but the measure of the interaction of the two parts a and b , namely the interaction of a with b and b with a . Hence, by taking the mutual interaction of elements in a system into consideration, the system as a whole indeed represents a more valuable entity than the mere sum of its independent parts. That a coalition is such a structure, where the individual elements interact for the benefit of the system as a whole, and hence for the advantage of each element comprising the system is, I believe, now reasonably clear.

These so called “non-linear composition rules” allow in a non-trivial way the description of systems composed of interacting elements. Take, for instance, a colony of about a hundred million flatworms of the genus *planaria*. Each of these creatures has about one hundred nerve cells. Thus, all together they have about ten billion nerve cells. The human brain also has about ten billion nerve cells. Why don’t these hundred million *planariae* represent the intelligence of a human brain? With this short course on super-additive composition rules, you are certainly now in a position to answer this puzzle. It is because brain cells are in a state of perpetual interaction, constantly co-ordinating, abstracting, and sifting pertinent information for the system as a whole. Poor *planariae* cannot do it; add a couple of million *planariae* to our colony, and nothing changes in the structure of this colony. They do not interact. If they interact, they interact by competing for a limited supply of food. At this point, it is high time to justify my lengthy elaboration on the concepts of non-linear composition rules. You may rightly ask what all this has to do with “Environment,” the theme of this conference, and in particular with the title of my paper, “Logical Structure of Environment and Its Internal Representation.” I hope that I will

be able to show in a moment that the concept of super-additivity will be crucial in the appreciation of environment, because, as some of you may have sensed already, it is the interaction part of the totality “Environment–Environmentee” that is today of central interest to us, and not merely the enumeration of independent events or entities. This, however, will be evident after we have discovered the peculiar logical modality that is implicit in the concept “Environment.”

I sincerely hope that I do not offend the kind sponsors of this conference if I first stipulate that it is nonsensical to talk about an environment *per se*. Stop to think for a moment what you would answer if I asked you, “What is the environment?” Most probably you would retort with a counter question, “Environment of what?” When we use the concept “environment” we tacitly assume an “environmentee” or that which is surrounded, or acted upon, by the environment. A cockroach who stands in line with me at the box office to *La Dolce Vita*, although he occupies almost the identical spatio-temporal neighborhood as I do, certainly has an entirely different environment. He may be looking forward to relishing a batch of old noodles which are, alas, protected by my foot, while I may be looking forward to Anita Eckberg’s antics. Thus, if we are able to specify the environmentee, then at the same time, we have defined its environment. Clearly, this is a highly subjective affair, as you can see from the cockroach example. On the other hand, it might be conceivable that, by appropriate specification and enumeration of certain objective features, one may establish an appropriate environmentee. However, this is rarely done, as we have seen at this conference where the environmentee was always implicitly or explicitly given.

From all this it seems apparent that the concept “environment” belongs to a class of concepts called “dyadic relations” because they involve two entities, say, A and B , where — in a loose sense — A implies B and B implies A . If this would be interpreted in a strict logical sense, the identity of A and B could be inferred. The sense in which this inference is meaningful will emerge after I have discussed the modes of internal representations of the structure of one’s environment. Presently, however, I would like to show that the logical structure of the concept “environment” is indeed more sophisticated than anything that can be described by a dyadic relationship. In carrying out my proof of the insufficiency of the assumption that the relationship “environment–environmentee” is a dyadic relation I shall use a popular logical artifice called *reductio ad absurdum*³.

Assume that a relationship “environment–environmentee” will indeed describe adequately the situation under consideration. Under these circumstances it can be argued that the whole environment does not exist at all in reality, but that it is the sole product of the environmentee’s imagination. In other words, we may insist

that introspection does not permit us to decide whether the world as we see it is “real,” or just the result of our dreams, an illusion of our fantasy. However, by looking around in our dream-world, which perchance may be this conference, we cannot deny that our imaginary universe is populated with apparitions that are in many respects similar to us. Consequently, we have to grant them the privilege to insist that they are the sole reality and everything else is only a concoction of their imagination. On the other hand, they cannot deny that their fantasies will be populated by apparitions — and some of them may be we!

With this we have reduced our original assertion to an absurdity, because if I assume I am the sole reality, it turns out that I am the imagination of somebody else, who in turn assumes that he is the sole reality. Of course, this paradox is easily resolved by postulating the reality of our environment in which we all happily thrive.

I hope that the crucial point of my argument has become sufficiently clear, namely that, for the establishment of the logical structure of the concept “environment” there must be at least two elements observing this environment, and they must be sufficiently alike in order to serve as mutual witnesses for any objective event. In other words, that which can be witnessed can claim to be real and to be part of our environment. Again, to put it differently, only knowledge that can be shared belongs to our environment. Thus, it turns out that the logical structure of “environment” is that of a “triadic relation” because it involves the relationship of three entities: an observer, *A*; a witness, *A**; and that which is witnessed, *B*. Environment can be called “together-knowledge” which in English undoubtedly sounds awkward. However, in Latin there is a splendid expression for exactly this term, namely *conscentia* from which, of course, our word “consciousness” is derived. It probably was the triadic logical structure of this concept that gave all philosophers over the last three thousand years difficulty when they tried to resort to a simple true-false, two-valued, Aristotelian logic, where at least a three-valued logic is required⁴. You may at this point argue that consciousness is, by all means, a single man’s affair and that you do not need a witness in order to be conscious. This seems, at first glance, to be true. However, closer inspection shows that our consciousness is produced by the “together-knowledge” of the rapport of our different sensory modalities. The ear is witness to what the eye sees; the eye is witness to what touch reports; and so on. In order that these senses may compare their various experiences, it is necessary that they translate these into the same language. I shall discuss this point in greater detail in a moment when I turn to the third portion of my paper “The Internal Representation of the Environment’s Structure.” Presently, however, I still owe you a clear description of what is supposed to be represented.

This leads us to the question: what is representable at all?

II. Environment as Structure.

A fact that never loses fascination for me is that there is not only a universe, but that there also are elements which are capable of observing this universe. That life with its intricate order and complex control mechanisms could ever evolve is due to the fact that this universe is a highly ordered affair and not a sequence of completely random events. Life can only exist in a cosmos — in the Greek sense of an ordered universe — and not in a chaos. It is quite clear that for anything to be anything, it has to remain unchanged for a while, it has to be an “invariant.” It is not necessary that this refers to a shape, or volume or a particular configuration; it may also refer to a chain of events that follow each other. In a complete chaos everything can happen. In a world with some order, not everything happens that could happen. From a purely logical point of view, there is nothing which would prevent this package of cigarettes I just took out of my pocket from pumping itself up into an elephant who slowly becomes airborne and flies out of this tent while singing the *Marseillaise* in a high pitched soprano. However, I venture to say that this does not happen; or — to be a bit more cautious — that the probability for this chain of events is “vanishingly small.” The nice thing about our universe is that there are invariants as, for instance, this package of cigarettes which I can put back in my pocket again, or its predictable trajectory if I would toss it in a particular direction. Thus, order presents itself in the form of constraints on a system that is not free to do what it would do if it were not subjected to these constraints. It is, therefore, not an accident that we denote these regularities of our universe as the “Laws of Nature”; and it is only due to their invariance that we are able to name them. Since “naming” is clearly a form of representation, it emerges that in order for something to be representable it must have some order. In a chaos, nothing can be named. Hence, whatever is representable in my environment must have some intrinsic order. Reversing this statement: my environment is defined by the kind and the amount of order I can discover.

Permit me to elaborate on this point, because if I succeed in showing that it is the kind of constraints that are determining the particular kind of emerging order, I shall have an easy task in showing that the corresponding process in the internal representation of the environmental order is neural inhibition. Let me first demonstrate by a simple example, what I mean by the emergence of order by adding constraints to an initially chaotic — or random — system. Consider for the moment a gambling situation in which a “fair” die is repeatedly tossed. As you all know, the outcome of a particular toss is completely

independent of the outcome of a previous toss. In other words, if face 4 has just come up and one picks up the die to toss it again, one is again completely uncertain about the next outcome, it can be 1, 2, 3, 4, 5, or 6, with equal probabilities. Let's consider for the moment an extraordinarily simple "universe" that is represented by this die. There are just six states in this universe S_1, S_2, S_3, S_4, S_5 and S_6 , namely the six faces of the die. The "behavior" of this universe is described if we are able to state the probabilities for this universe to go from any one of its six possible states into any other one of these six possible states. A representation of this situation is most easily done in the form of a quadratic matrix as shown in Table I, which lists on the left-hand side all present possible states as [illegible?] all possible subsequent states and inserts [missing line(s)?]

		Subsequent State					
		S_1	S_2	S_3	S_4	S_5	S_6
Present State	S_1	1/6	1/6	1/6	1/6	1/6	1/6
	S_2	1/6	1/6	1/6	1/6	1/6	1/6
	S_3	1/6	1/6	1/6	1/6	1/6	1/6
	S_4	1/6	1/6	1/6	1/6	1/6	1/6
	S_5	1/6	1/6	1/6	1/6	1/6	1/6
	S_6	1/6	1/6	1/6	1/6	1/6	1/6

TABLE I: Transition probability matrix for a completely undetermined (free) system (unloaded die).

In the particular case of our "fair" die, this matrix will be filled everywhere with equal probabilities of 1/6, because — as we have seen before — all outcomes are equiprobable and independent of past outcomes. I now propose to get a bit more zing into this boring game by slightly manipulating this die. Let's drill a hole into the die, and fill it with a viscous goo into which we insert a heavy piece of metal. Of course, we do not forget to seal up the hole so that this die looks like any other harmless die. However, its performance will be quite different from before. If we now allow the die to rest after each toss, the piece of metal will slowly sink and will load the die so it tends to fall on the same face. With this manipulation our previously unpredictable simple "universe" has become more predictable, because, if, say, face 4 has come up, we will be right in most cases if we predict that face 4 will come up. This situation can similarly be expressed in form of a transition probability matrix which may look as in Table II if we do not let the die rest too long between tosses.

		Subsequent State					
		S_1	S_2	S_3	S_4	S_5	S_6
Present State	S_1	1/4	1/6	1/6	1/6	1/6	1/12
	S_2	1/6	1/4	1/6	1/6	1/12	1/6
	S_3	1/6	1/6	1/4	1/12	1/6	1/6
	S_4	1/6	1/6	1/12	1/4	1/6	1/6
	S_5	1/6	1/12	1/6	1/6	1/4	1/6
	S_6	1/12	1/6	1/6	1/6	1/6	1/4

TABLE II: Transition probability matrix for a partially determined (constraint) system (loaded die).

[begins in mid-sentence] tribution where certain state transitions are more probable than others. In other words, the system is no longer free to go into any possible states with equal chance, but has to yield certain constraints within its structure. This, I believe, will become quite apparent if we let our die sit and wait between tosses until the piece of metal has settled on the ground. Under these circumstances the die will come up in subsequent tosses with overwhelming probability on one and the same side and all other state transitions will be "vanishingly small".

The corresponding matrix is shown in Table III, which now contains probabilities being either zero or unity depending upon the state transitions being impossible or being certain.

		Subsequent State					
		S_1	S_2	S_3	S_4	S_5	S_6
Present State	S_1	1	0	0	0	0	0
	S_2	0	1	0	0	0	0
	S_3	0	0	1	0	0	0
	S_4	0	0	0	1	0	0
	S_5	0	0	0	0	1	0
	S_6	0	0	0	0	0	1

TABLE III: Transition probability matrix for a completely determined system (Law of Nature).

In this case the constraints in our simple universe are so strong that it is impossible for the system to do anything else but to "observe the laws of nature" which are intrinsic in the structure of this system.

I realize that I may have used an over-simplified example to demonstrate the difficult concepts of determinacy and indeterminacy. Nevertheless I hope that the methods of describing these situations have become palatable. My excuse for elaborating on these ideas, perhaps unduly, is that in my next step I would like to show that small constraints on an otherwise random system represent themselves in the form of remarkably developed structures. This I shall do with the aid of three examples. Two of these are taken from Nature with the constraints fixed; the third example is an artificial system in which the experimenter is in a position to manipulate the transition probabilities.

Fig. 1 shows a collection of twelve patterns all displaying a beautiful hexagonal symmetry. Of course, you recognize these stars immediately as snow crystals⁵. Why is this recognition carried out so quickly on such a relatively large set of different objects? The reason is that the growth mechanism of these snow crystals is subjected to a major constraint, and it is this constraint which is "sized up" immediately. The growth mechanism of snow crystal is determined by the triangular shape of the water molecule⁶ having two small hydrogen atoms attached to the big oxygen atom in angles which are close to either 30° or 60°. This, on the other hand, accounts for a certain "freedom" in attaching themselves to each other, which in turn allows for the large variability within this constraint which we recognize as a unifying principle in

the construction of these shapes. We have a name for this constraint; we call it “snow-crystal”. However, this microscopic constraint would not become apparent to us if it were not applied over and over again. Since such a snow flake contains approximately a billion billion H_2O molecules, this constraint has been in operation in exactly the same number of times.

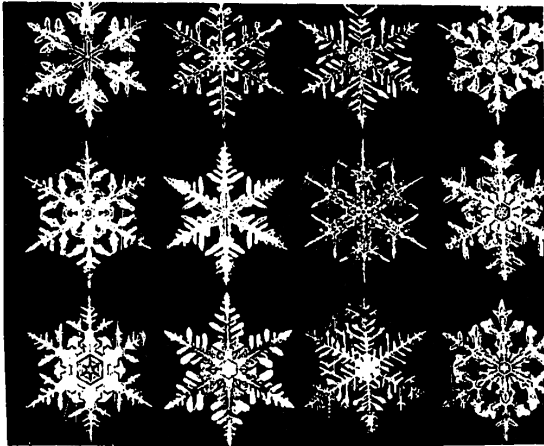


Fig. 1. Snow Flakes.

My second example from Nature refers to a case where she anxiously has to guard against even the slightest variations occurring. I am referring to the “genes”, those macro-molecules which govern the stability of the hereditary traits from generation to generation. It is their particular molecular configuration which determines the building program of the organism in its macroscopic appearance. But a program is nothing else but a set of commands: “do this; do that ...” which in other words means: “don’t do this; don’t do that ...” Again we have words for the various sets of these invariable constraints as, for instance, “elephant”, “mouse”, “hippopotamus”, etc. We may even recognize the fine-structure of these constraints by referring to certain family traits: “Oh, he is a typical Jones.”

In these two examples you may argue that the constraints I am talking about are rather strong, because they are produced by the strong electro-static forces which bind the atoms together to form the various molecules. Hence one should not be too surprised if this results in a variety of remarkable structures. However, one should not forget that all these molecular configurations are subjected to a strong random thermal agitation which tends to destroy all order and symmetry. Let me show now in a final example the emergence of structure in a system where we have the transition probabilities from state to state, so to say, at our fingertips, because it is only necessary to turn a knob which will automatically regulate these probabilities.

In Fig. 2a you see the first 23 steps of a luminous spot produced by an electron beam on a television screen:

The motion of this spot upwards or downwards, and from left to right, or from right to left, is “determined” by a fast electronic “coin tosser” which makes about twenty thousand random decisions in one second. Clearly there are only four states, namely a unit step into any one of the four directions. The transition probability matrix for these states looks exactly like that in Table I, with, of course, only 4×4 entries, all filled with a probability of $1/4$. We now apply to our electronic coin tosser the constraint that it should become slightly “loaded” in the sense that it should (a) give slightly higher probabilities to switching the spot into another direction by reducing its probability to continue its movement in either the horizontal or the vertical direction, and (b) give slightly higher probabilities for switching the spot in a clockwise direction. In Fig. 2b you see the result of the first 200 steps of this spot after the uniform probability distribution has been shifted only a few percent in the desired direction. Since this pattern shows only 16 steps, it is clear that the spot must have run several times through this configuration. In other words, such a pattern shows a certain amount of stability produced by the constraints inherent in the motion of this spot. Another feature may be noteworthy, and that is that adjacent squares are being circumnavigated in opposite directions as indicated by the arrows pointing in the direction of motion.

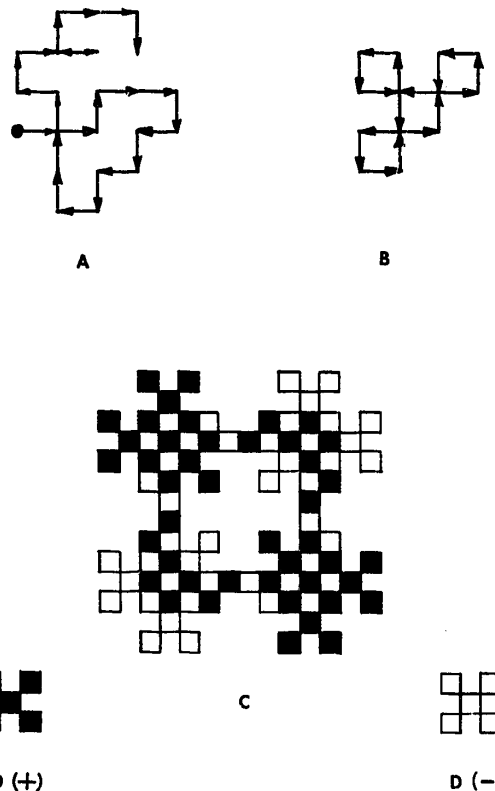


Fig. 2. Order emerging from applying constraints to 2 random system. (a) Random walk. (b) Single spot with slight turning constraint. (c) Four interacting spots, (d) Apparent “building blocks”.

A highly sophisticated configuration is shown in Fig. 2c which presents the results of the first 200,000 steps of the motion of four spots each of which has the same constraints as our spot before, with the additional constraint that they all interact weakly with each other in the sense that they “repel” each other when they come too close (the transition probability for turning away from each other is slightly increased when near), and that they “attract” each other when they go too far (the transition probability for turning toward each other is increased when apart). Clockwise circumnavigated squares are painted black. Since there are only 256 steps visible in this pattern, it is clear that some of the steps must have been repeated several thousand times. Hence, this pattern has reasonable stability.

A student of this structure, who does not know how it is created, will come to the conclusion that this “molecule” is built of two kinds of “atoms”, one black (+) and one white (–), with shapes as suggested in Fig. 2d, which obey a law of nature that forces them to bind into higher structures such that opposite signs attract. We may smile at the naivete of this natural scientist who discovers these “laws”, because we know that this whole pattern is generated by only four spots zooming around like mad in an almost random fashion. However, we should not forget that the accent lies on the almost. That is the crux of my thesis: Small constraints are sufficient to produce considerably ordered structures. Hence, the discoveries of our natural scientist are not so naive after all; he only put his knowledge into a different language. The two descriptions are equivalent.

Before I turn to my concluding discussion of the internal representation of environmental order, let me briefly summarize the train of thought that brought us to this point. First we established that the saying “the whole is more than the sum of its parts” is meaningful only if the parts are interacting in the form of a super-additive composition. Second, the logical structure of environment revealed itself as a triple interaction between at least two observers and that which can be communicably observed; this we called the environment. Finally, we have just attempted to show that order, not chaos, is that which can be communicated; and order is the result of constraints.

I believe, we are now properly prepared to explore the modes of internal presentation of environmental order because we know what to look for. We must look for how environmental constraints mirror themselves in the internal structure of the observing systems.

III. The Internal Representation of the Environment’s Structure.

It is a generally accepted notion that all information about our environment reaches us through the activity of

our sense organs. The sense organs act as intermediaries which translate the various physical modalities and their variations (light, color, sound, heat, etc.) into some appropriate activity of the nervous system. The result of this translation has usually been called “the sense data”. Until a couple of dozen years ago, it was believed that these sense data were quickly relayed to the brain which, by some of its miraculous powers, put these data bit by bit together and reconstructed a true replica of one’s environment. In other words, the brain was thought to be populated by little men, who themselves could see, hear, feel, etc., and who informed us about the state of affairs. If one asked about the brain of these little men, one was either “out”, or their mental activity was again explained by little men in the heads of these little men, and so on *ad infinitum*.

The remarkable advances in experimental neurophysiology in the last two decades have shown us that this picture of the information processing in the nervous system is not only unsatisfactory, but also false. The most important lesson those advances taught us is that the sense organs deliver, in the form of symbols, highly processed and reduced information to the brain, which in turn performs “non-numerical” computations on these symbols. The results of these computations become in part apparent in the form of the activity of the organism.

In view of the extraordinary complexity of the functioning of the nervous system, I hope that you will forgive me if, in the limited time of this lecture, I confine myself to a few points which, I hope, will stimulate you sufficiently to look for more information in this fascinating field^{7, 8, 9}.

In order to leave you with at least an inkling of how and what is computed by the simultaneous interaction of (at times) more than ten thousand parallel nerve fibers, let me briefly describe the nerve cell, the neuron, the elementary component of this intricate computer.

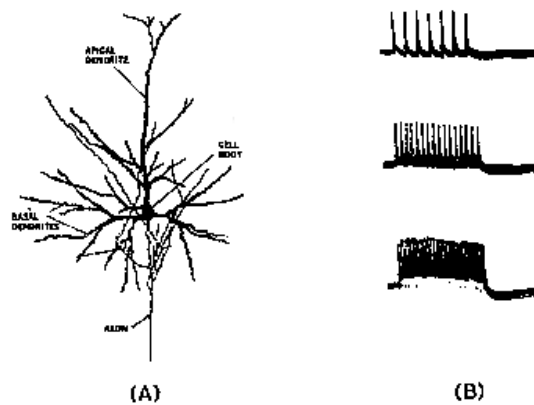


Fig. 3. Form and function of a cortical neuron. (a) Cell body, dendrites and axon. (b) Electrical activity measured with a microprobe on the axon of a sensory input neuron (Frequency Modulation).

A typical cortical nerve cell (as those of which you may find several billions in a cat's brain) is shown enlarged about 10,000 times in Fig. 3a. Like any other cell, the neuron has a cell body, called the *soma*; which houses the nucleus and is probably responsible for the neuron's metabolic activity. The same membrane which envelopes the soma also forms the tubular sheaths around the many ramifications extending from the soma. There are two kinds. One kind, the dendrites, are seen branching off in all directions in a tree-like fashion. The other, the axons, are smoothly surfaced and rather straight. They extend downward, bifurcating further below on many points, exhibiting a more regular, somewhat perpendicular pattern. The diameter of the axons may vary from a few microns to hundreds of microns; its length in a neuron may range from a few millimeters to a meter or more. Most of these axons terminate on other neurons, establishing two different kinds of connection as sketched in Fig. 4a. One is a direct attachment to the soma of the other cell by formation of an "end bulb", the other one is a somewhat haphazard ascending intertwining with the dendritic ramifications of the target neuron. These different connections fulfill different functions, as we shall see in a moment.

If one penetrates with a micro-probe into the enclosing membrane at any point of the neuron, one finds a change of the electric potential of somewhat less than a tenth of a volt, which indicates that a whole structure in its rest state can be considered as a charged, distributed, electric battery.

If at the soma this electrical potential is momentarily perturbed beyond a certain threshold value, the neuron will "fire", *i.e.* the perturbation will travel along the axon in form of an electric pulse and will pass on this perturbation to its connectees, which, in turn, will respond to this perturbation. You may note the following interesting details in connection with this interaction process. First, the magnitude of the electric pulse is independent of the intensity of the perturbation. However, a prolonged perturbation in sensory cells will produce a chain of pulses with a repetition frequency approximately proportional to the logarithm of the intensity of the stimulus. Fig. 3b shows three tracks of the pulse-activity of a single neuron as measured with an electric probe. While in this example the stimulus varied according to the ratios 2:4:16, the corresponding frequency ratios are only 1:2:4. This is a very elegant method indeed to compress a large amount of information into a relatively narrow band. The translation of outside information into pulse intervals of various frequencies is universal. That is, all sensory modalities will code their information into this pulse language, which is not only understood but also spoken by all other nerve cells. Indeed, it is the only form of communication between neurons and it is the only way in which all our experiences, thoughts, feelings and ideas are represented: through sequences of electric pulses.

I hasten to add that this does not mean that I belittle this situation. I did not say "merely sequences of electric pulses". I just wanted to point out that all mental activity is represented by this electric activity. In a similar way a physicist might point out that all matter is composed of about a hundred different atoms. It is undeniable that atoms as well as pulse sequences are respectable, sophisticated entities.

Another detail that is worth mentioning is that the magnitude of the traveling pulse does not diminish while it travels along the axon, even when a bifurcation is reached, and from then on two pulses travel along their tracks. This is, of course, due to the neuron's important property of acting as a distributed, charged battery, which supplies the necessary energy at any point along the line. Consequently, several neurons operating in series may act as an impressive signal amplifier.

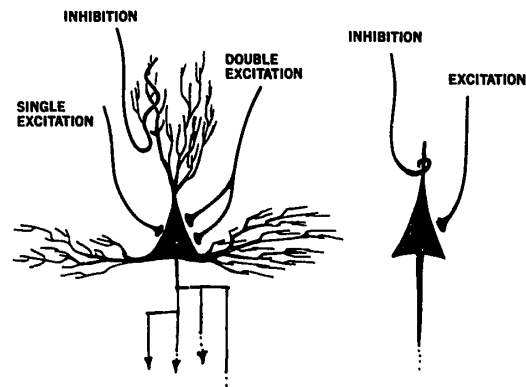


Fig. 4. Schematic (a), and symbolic (b), representation of a single neuron.

However, I still owe you a description of how these neurons exchange and process information. In Fig. 4a we indicated two different kinds of connections. Their functions can now be discussed. Connections via end bulbs provide units of excitation while axons terminating in the dendritic ramifications will inhibit the firing tendency of the target neuron¹⁰. In other words, if two pulses arrive almost simultaneously, the one over an end bulb and the other one over the dendrites, their action will cancel out and nothing will happen. If, however, at a particular instant the only active input is a single end bulb and the threshold of the cell is zero, this will suffice to trigger the neuron and it will transmit a pulse via its output axon to other neurons in the connection field. These, in turn, may or may not fire, according to the type of connections the descending axons have established.

It is not difficult to show that, with these types of connections, symbolized on an "idealized" neuron in Fig. 4b (namely excitation [yes], inhibition [no], or excitation [true], inhibition [false]) a single neuron is in a position to compute various "logical functions", and small nets of only three neurons are capable of computing all logical

functions of the form “A and B”, “A or B”, “if A so B”, “A equivalent B”, “A or not B”, etc. where A and B are two active afferent axons¹¹.

Although the recognition of these computations in the central nervous system is of the greatest importance from an epistemological point of view, I believe that at this conference you may be more interested in knowing how some of the spatial order in our environment is computed and represented by nervous activity.

Before I show (with the aid of a simple example) the principles of computation that are in operation, let me give you a clue as to what is computed. May I take you back for a moment to the point where I discussed our environment’s principles of order. As you remember from the examples I suggested, order was always established by producing constraints that acted either on temporal or on spatial neighborhoods. The trajectory of a flying object is defined at each moment by the direction, speed, air-resistance, etc. of this object. The object cannot go zig-zag. It must obey the laws that escort it, so to speak, during its flight. The snow crystal grows into its hexagonal shape because at each point there exists a constraint that allows a water molecule to attach itself only in a certain way. Hence, we see that all these constraints produce particular neighborhood configurations. In these configurations the constraints manifest themselves. Thus, a system which wants to compute these constraints and “sees” only the manifestations of these constraints has to carry out an analysis of the neighborhood relations.

This observation suggests that neural networks composed of neurons whose neighborhood relationship (or neighborhood logic) is compatible with the external neighborhood relations may just “read out” the prevailing structure. This situation is perhaps most easily demonstrated with particularly simple networks (so called “periodic nets”) which are characterized by a periodic repetition of one and the same connection pattern, hence by a repetition of one and the same computational operation. The study of these periodic structures proved to be most rewarding and led to important clues in our understanding of the problem of “cognition”.

Fig. 5 gives an example of such a periodic one-dimensional net. A series of photosensitive elements (say, photocells) are connected to a series of corresponding idealized neurons so that the left and right hand neighbor neuron is singly inhibited, while the associated neuron is doubly excited. This connection scheme repeats itself periodically over the entire strip. With all thresholds equal and slightly above zero, clearly, the neural net will not respond if all photosensitive elements are uniformly illuminated (Fig. 5a), whether the illumination is strong or faint, because the double inhibition converging on each neuron from sensory elements to the left and right will cancel the double excitation coming from its corresponding photocell.

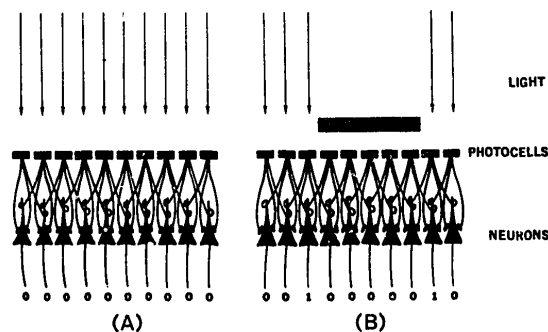


Fig. 5. Periodic net which computes the invariant “EDGE,” independent of location, strength of illumination and size of object.

An obvious effect of this particular interconnectivity is its insensitivity to variations of light, despite the fact that the “sensory layer” is composed of highly sensitive light receptors. You may ask now, why all this effort to make a light sensitive organ insensitive to light? However, you will see in a moment the interesting feature of this net, if an obstruction is now placed into the light path (Fig. 5b). The edge of this obstruction will be detected at once, because the only neuron which will now respond is the one on the edge of the obstruction, receiving insufficient inhibition from only one of its neighbor photocells in the light, while the other one is in the shade and silent. In other words, this net “computes” the invariant “edge”, independent of its location and independent of the strength of illumination. The efferent fibers of this network will be active only if edges are present in the visual field of this simple, one-dimensional “retina”.

Although the scheme is admittedly simple, this network can detect a property which cannot be detected by the nervous system built into us. Consider the simple topological fact that any finite, one-dimensional obstruction must have two edges. If N objects obstruct the light path to our edge detector, $2N$ neurons will be active and their total output divided by two gives exactly the number of objects in the visual field of that network. In other words, this strip sees each different number of objects as a different entity, say “seven-ness”, “twenty-ness”, etc., as we see different electromagnetic frequencies as different colors, “red-ness”, “green-ness”, etc.

We are today in possession of a general theory of neighborhood logics which compute an almost infinite variety of abstracts (as, e.g., straightness, curvature, topological connectedness, motion of shapes, etc.) in the visual field; chord and timbre independent of pitch, voicing and variation of frequencies typical for definition of spoken phonemes in auditory perception, and many more. It is most significant that these networks are not only on paper or [illegible] enterprising engineers¹², but have been [established to] exist in the sensory apparatus of cold and [illegible] animals. We know, for instance,

that post [illegible] neural networks in the frog's eye compute precisely four different invariants¹³: straight edges; changes of light; curvature of a dark object; and movement of edges. In the Platonic sense, these are the frog's "universals". It has recently been established that the pigeon has precisely six such property detectors in its visual apparatus¹⁴.

I hope that with these examples I have made sufficiently clear where to look for the representation of environmental structure within the observing organism. As these examples suggest, the constraints that produce the structure in the environment are mirrored by the constraints in the modes of interaction of neighboring neurons; that is, in their connection schemes.

I realize that it is not easy to envision this correspondence between neural connection structure and environmental order. Perhaps an obvious analogy may help. A superficial inspection of the score of Beethoven's Fifth Symphony and the microscopic wiggles in the grooves of a record of Beethoven's Fifth Symphony will not reveal their structural identity. However, when played by an orchestra and a record player respectively, their similarity cannot be denied. This shows that to any spot in the score there must correspond a set of wiggles on the record, and vice versa, otherwise the results obtained by transforming these representations into still another one, namely into patterns of sound, could not have so much resemblance.

At this point I could, in all fairness, conclude this presentation. First I have discussed the logical structure of "environment". Second I presented some approaches to the definition of our environment's structure; finally I attempted to give some clues to the understanding of the internal representation of environmental structures. However, I feel obliged to draw your attention to a difficulty in my presentation which I have not yet overcome. This difficulty I have posed myself by insisting that the reality of environment is guaranteed only by communication of environmental features between two observers. How do they communicate?

The first reaction to this question is that it borders on ridicule, because the answer seems to be obvious: by language, of course! But this, unfortunately, is too simple a way out, because — as I will show in a moment — it is impossible to learn some one else's language unless teacher and student have like internal representations of environmental order.

In order to prove my assertion of the impossibility to learn a language by just applying language I only have to show that language is a closed system. By this I mean that any word in a language is again defined by words in the same language.

If all explaining, and to-be-explained, words are unknown, they must remain unknown. A dictionary of a single language is a typical example of circular defini-

tions in language. I give an instance which, with amusement, may be extended to any other case. Suppose I do not know the meaning of "after". I consult a dictionary which explains "after" with the aid of "later", "subsequent", "succeeding" and "consecutively". Suppose I do not know the meaning of these. Consequently I look them up in my dictionary and again I do not know them, and so forth. The result of this search is given in matrix form in Table IV which lists for the American Collegiate Dictionary the left hand side the source words, on the top the target words, and indicates with an "X" at intersections of rows and columns the presence of a reference. Clearly, if I do not know any of the six terms, I shall never know what "after" means. How to escape this dilemma? It has been suggested that denoting by pointing is a meta-language, that breaks through the closure of language. Alas, this meta-language is also closed, as told in a charming story by Margaret Mead, the anthropologist. She wanted to learn quickly a basic vocabulary of a people whose language was unknown, and she pointed at various things with questioning sounds and gestures. However hard she tried, she always got the same sounds as answers. Later she learned that these people were helpful, concise, and understanding; the sound was their word for "pointing finger".

	AFTER	LATER	SUBSEQUENT	SUCCEEDING	CONSECUTIVELY	FOLLOWING
AFTER		X	X	X	X	
LATER	X				X	X
SUBSEQUENT						X
SUCCEEDING	X		X	X		X
CONSECUTIVELY						X
FOLLOWING	X					

TABLE IV: Illustration to the assertion that language is a closed system.

It is my personal belief that we find the clue for a resolution of this dilemma in the like representation of environment in the two elements who are separated by their skins, but alike in their structure. When they realize and utilize this insight then *A* knows what *A** knows, because *A* identifies himself with *A** and we have the equality

I — THOU.

This brings me back to my introductory remarks where I pointed out that cooperative interaction is super-additive. Clearly, identification is the strongest coalition — and its most subtle manifestation is love.

With this I have completed the circle of my presentation which, after all, turned out to be nothing else but a paraphrase of a relatively old story¹⁵: And Adam said,

“This is now bone of my bones and flesh of my flesh . . . Therefore shall a man leave his father and his mother, and shall cleave onto his wife: and they shall be one flesh”.

Notes

¹Robert Ardrey: *African Genesis*. Antheneum, New York 1961; 380 pp.

²K. von Frisch: *Bees*. Cornell University Press, Ithaca, New York, 1950; p. 53 ff.

³H. Von Foerster: “On Self-organizing Systems and Their Environments”. In *Self Organizing Systems*, M. C. Yovits and S. Cameron (ed), Pergamon Press, New York 1960; p. 31–50.

⁴G. Gunther: “Cybernetic Ontology and Transjunctonal Operations”. In *Self-Organizing Systems*, M. C. Yovits, G. T. Jacobi and G. D. Goldstein (ed), Spartan Books, Washington, D. C. 1962; p. 313–392.

⁵H. Weyl: *Symmetry*, Princeton University Press, Princeton, N. J., 1952; p. 63.

⁶P. Debye: *Polar Molecules*. Dover, New York, 1949; p. 73 ff.

⁷H. Von Foerster: “Perception of Form in Biological and Man Made

Systems”. In *Proc. I.D.E.A. Symp.* 1962, E. J. Zagorski (ed), University of Illinois, Dept. Fine and Appl. Arts, Urbana, Illinois, 1962; p. 10–37.

⁸W. R. Ashby: *Design for a Brain*. Chapman and Hall, London 1950; 286 pp.

⁹J. C. Eccles: *The Neurophysiological Basis of Mind*. Clarendon Press, Oxford 1952; 314 pp.

¹⁰Recent advances in the neurophysiology of the synapse indicate that the inhibitory mechanism is not restricted to dendritic interaction. Presynaptic inhibition from axon to axon has been established by Eccles. The chemistry of the inhibitory process is very complex and is as yet not completely understood.

¹¹H. Von Foerster: “Perception of Form in Biological and Man Made Systems”. In *Proc. I.D.E.A. Symp.* 1962, E. J. Zagorski (ed), University of Illinois, Dept. Fine and Appl. Arts, Urbana, Illinois, 1962; p. 10–37.

¹²P. Weston: “Photocell Field Counts Random Objects”. *Electronics*, September 22, 1961; p. 44–46.

¹³J. Y. Lettvin, H. R. Maturana, W. S. McCulloch and W. Pitts, “What the Frog’s Eye tells the Frog’s Brain”. *Proc. IRE* 47. 1940–1951 (1950).

¹⁴H. R. Maturana, Personal communication 1962.

¹⁵Genesis 2; 23–24.