# From Stimulus To Symbol: The Economy Of Biological Computation

Heinz Von Foerster

BCL #135

Man's heritage is of two different kinds. One has been accumulated through perhaps two billion years of evolution and is encoded in the molecular structure of his genetic make-up. The other has been built-up during approximately one million years of communication and is encoded in the symbolic structure of his knowledge.

While man evolved as a result of interplay between genetic mutability and environmental selectivity, his self-made symbols evolved as a result of interplay between his flexibility in expressing and his sensitivity in distinguishing. This observation links these two evolutionary processes in a not too obvious way, and gives rise to the formidable problem of demonstrating this link by tracing structure and function of the symbols he uses back to the cellular organization of his body.

It is clear that we are today still far from a solution to this problem. First, we do not yet possess a consistent comprehension of structure and function of our symbols, to wit, the Cyclopean efforts by various linguistic schools to establish a concise language for dealing with language; second, our knowledge of the cellular organization of the body is still meager, despite the incredible amount of knowledge accumulated over the past decades. As a matter of fact, it is indeed doubtful whether with presently available conceptual tools this problem can be solved at all. These tools, however, will permit us to get an insight into the magnitude of this problem.

An approach which considers symbolization in the framework suggested by the formulation of this problem does have the advantage that it can tie together evidences accumulated in a variety of fields. Moreover, within the framework suggested here it becomes impossible to talk about symbols in a static, ontological way and not consider the dynamic evolution of symbolic presentation. Likewise, it becomes impossible to separate a symbol from its symbolizer, his sensory motor and mental capabilities and constraints. And further, it becomes impossible to separate symbol and symbolizer from his environment which we have to populate with other symbolizers in order that symbolization makes any sense at all.

The following is an attempt to establish clues for the understanding of potentialities and limits of symbolization through the understanding of variety and constraints in the maker and user of symbols and in his environment.

The argument will be presented in three steps. First, the concept of "environment" and the relation "environment-environmentee" will be discussed. The second step will be to briefly sketch some basic principles and some hypotheses of the processes that permit internal representations of environmental features. Third, modes of projecting externally these internal representations will lead to the consideration of possibilities of interaction by symbolization.

## Environment: An Analysis

Evolution, like memory, is an irreversible process. The man who once knew a datum, but has forgotten it now, is different from the man who never knew it. Irreversibility in evolution permits one to picture this process in the form of a tree with divergent branch points only. Fig. 1 is such a representation of evolutionary differentiation in vertebrates over the last 500 million years. Time runs from bottom to top and the number of different species at any time within each branch is indicated by the width of this branch. A subspecies among mammals called *homo sapiens*, including its entire temporal extension, occupies in this graph but a tiny speck of space in the upper right corner of the mammalian branch, number 8.

It is perhaps easy to see that this graph represents paleontological estimates of only those species that were sufficiently stable to leave detectable traces. All instable mutants escape detection, and thus cannot be accounted for. In other words, this graph is essentially a picture of the success story of living forms. This observation permits us to look at this representation in a slightly different way, namely, to consider each point in a branch as being an instant at which a crucial problem is presented to a particular species. If it solves this problem the point will be retained and moves upward an ever so slight amount. If not, the point will be removed, *i.e.*, the species is eliminated. It is clear that the crucial problem referred to here is how to survive, and it is also clear that this crucial problem is posed by the properties of the particular environment which is in interaction with elements of this species or its mutants.

Fig. 1. *Evolution of vertebrates over the last 500 million years. Time runs from bottom to top. Width of branches corresponds to approximate abundance of different species within the branch (class). 1. Jawless fishes. 2. Cartilage fishes. 3. Placoderms. 4. Bony fishes. 5. Amphibians. 6. Birds. 7. Reptiles. 8. Mammals.*

From this viewpoint "environment" is seen in a twofold way: as a set of properties of the physical world that act upon an organism; and also as an accumulation of successful solutions to the problem of selecting such conditions in the physical world which are at least survivable. In this discussion "environment" will always carry this relative notion as "environment of...," where environment and the organism associated with it will be duals to each other in the sense that a particular organism $O$ implies its particular environment $E(O)$, and vice versa, that a particular environment $E$ implies its appropriate organism $O(E)$.

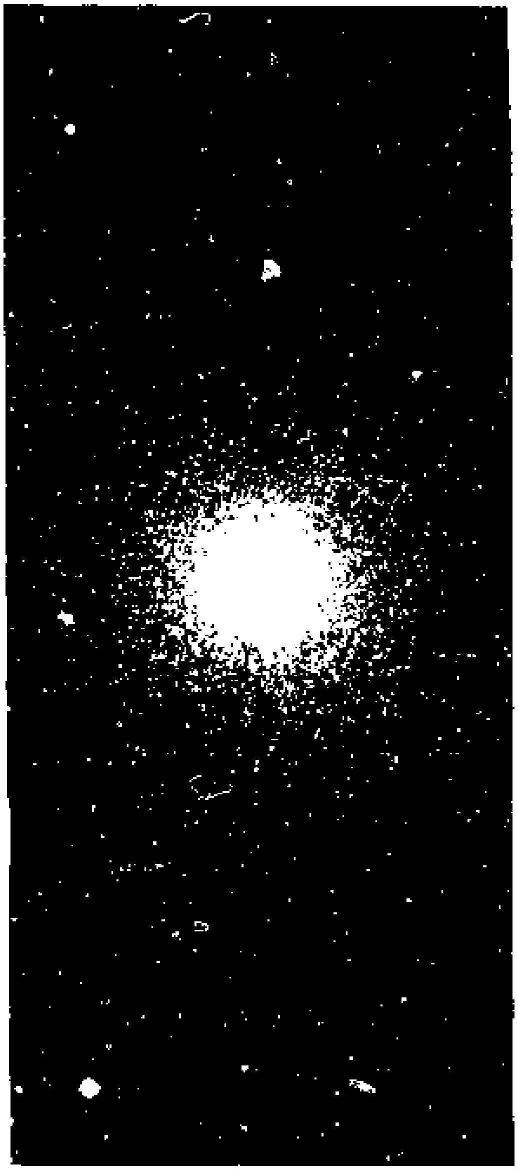


Fig. 2. *Globular star cluster NGC 5272 in the constellation* Canes Venatici, *Mount Wilson and Palomar Observatories.*

By carving out from the physical universe just that portion $E(O)$ which is "meaningful" for this organism $O$, one has carved out a portion that is necessarily of compatible complexity with that of the organism. An organism that tolerates a variation of temperature of, say, thirty degrees Fahrenheit around a certain mean, cannot "dare" to move into places where temperatures vary beyond this tolerance.

This statement can be expressed differently. An organism that is matched to its environment possesses in some way or another an internal representation of the order and the regularities of this environment. How this internal representation within the cellular architecture of living systems is achieved will be taken up later in this paper.

At this point the concept of "order" needs further clarification. Intuitively one would associate order with the relation of parts in a whole. But what are parts? Again, intuitively, parts emerge as "separabilia," because the relation among their components is of higher order than that of the parts of the whole. Although this definition is circular, it points in the right direction, for it relates order to the strength of constraints that control the interaction of elements which comprise the whole. These constraints manifest themselves in the structures they produce. The globular star cluster (Fig. 2) has simple spherical symmetry, because the weak gravitational forces that hold the approximately 100,000 elements of this system in statistical equilibrium have themselves radial symmetry. Of course, much more sophisticated structures are obtained if this constraints are more numerous and stronger. The volume *Structure* of this present series[1] abounds with beautiful examples from nature and art, where either strong molecular forces (*e.g.*, the paper by Cyril Stanley Smith) or the application of strong principles of construction (*e.g.*, the paper by R. Buckminster Fuller) generate structures of great intricacy and sophistication. Here only one shall be given, the almost inexhaustible variety of hexagonal symmetries in snow crystals (Fig. 3). The growth mechanism of these crystals is subjected to a major constraint, namely the triangular shape of the water molecule $H_2O$ which has two hydrogen atoms attached to the big oxygen atom at angles which are close to either 30°or 60°. This slight deviation from the condition that would produce equilateral shapes introduces a certain amount of "freedom" for the molecules to attach themselves to each other, which in turn allows for the large variability within this constraint. Note that in spite of the great difference in the individual shapes of these crystals, no difficulty arises in recognizing these forms at a glance as snow crystals. This suggests that the cognitive apparatus that "figures out" — or computes — the answer to the question "What is this?" is the one thing that is common to all these shapes, and this is the constraint in their growth mechanism. The name we give to this constraint is simply "snow crystal."

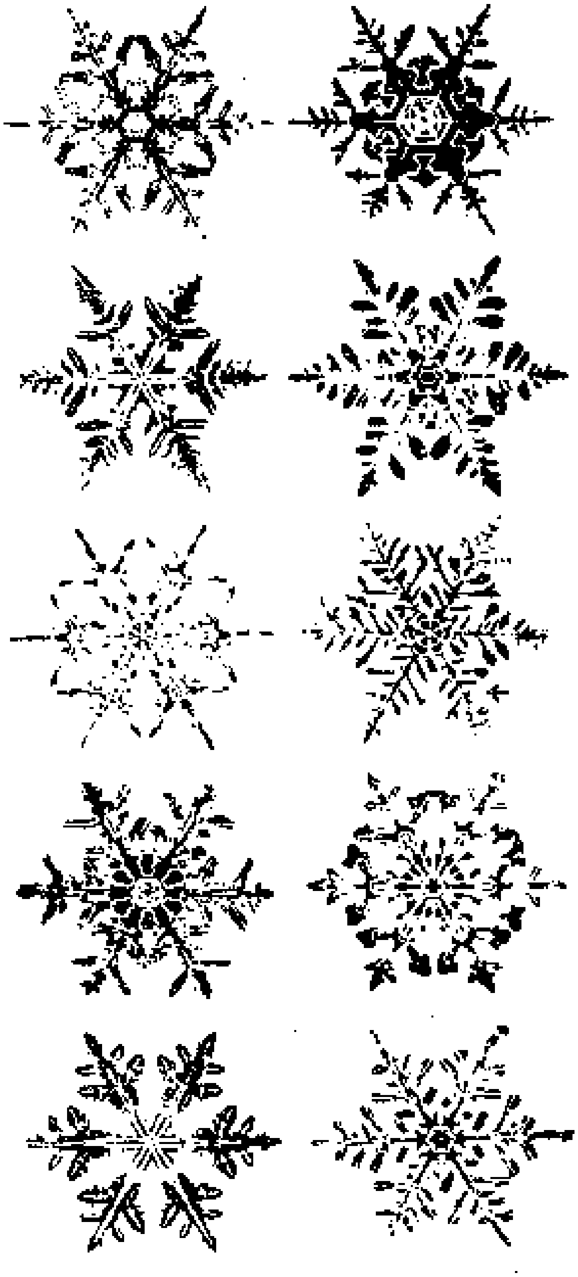


Fig. 3. *Snow crystals.*

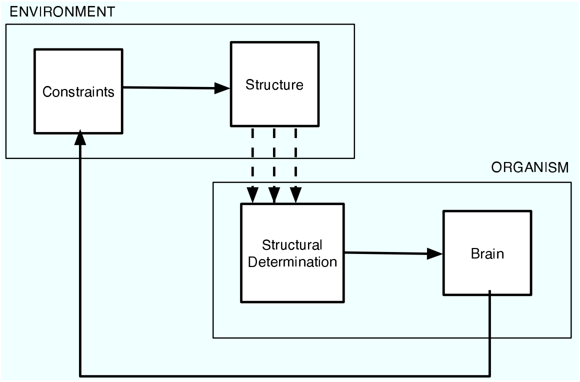
In the temporal domain order is again generated by the constraints of the "Laws of Nature" which, on the macroscopic scale of direct observation, control the chain of events. Chaos would permit transitions from any state to any other state, mountains transforming themselves into flying pink elephants, pink elephants turning into yellow goo, etc. Not only are organisms impossible in this world, for by definition, there is no law that holds the organism together, but also this world is indescribable, for description requires names, and names refer to the "invariabilia" — the constraints — in the environment.

One clue of how to compute these constraints from the apparent structure of the environment is suggested by the preceding examples. Structure in space was determined by a law in the growth mechanism that permitted attachment of new neighbor elements only at particular points; structure in time was determined by a law in the transition process that permitted only a particular event to be neighbor to an existing one. In other words, spatiotemporal order is generated by constraints that control spatiotemporal neighborhood relationships. Hence, if these can be "sized up," the constraints can be evaluated.

If chaos permits every event to appear with equal probability, order emerges from chaos when certain transitions of events become more probable than others. Certainty of an event following another creates a perfect, deterministic universe, and the problem of how to survive in such a deterministic universe is reduced to finding the constraints that govern the transitions from one event to the next. Clearly, the simplest of all such deterministic universes is the one where no transitions take pace, *i.e.*, where everything is at motionless and uniform tranquility. Hence, the oceans, where temperature variations, changes in the concentration of chemicals, destructive forces, etc., are kept at a minimum, were the cradle for life.

The dual interdependence of organism-environment permits a dual interpretation of the tree of evolution (Fig. 1). Instead of interpreting points on this graph as *species of organisms*, one may interpret them as *species of environments*. Thus viewed, this chart represents the evolution of environments which were successively carved out of the physical universe. These environments evolved from simple, almost deterministic ones, to extremely complex ones, where large numbers of constraints regulate the flow of events. An environmental subspecies among mammalian environments, called "E (*homo sapiens*)," occupies in this graph a small speck of space in the upper right corner of branch number 8. Hence, its dual, "*homo sapiens*(E)," sees "his universe" as a result of two billion years of environmental evolution, which step by step carved out from the physical universe an ever increasing number of constraints of all those in this universe that are computable within the limits of the evolving organism.

The diagram shown here below sketches the circular flow of information in the system environment-organism. In the environment constraints generate structure. Structural information is received by the organism which parses this information on to the brain which, in turn, computes the constraints. These are finally tested against the environment by the actions of the organism.



With the emergence of self-reflection and consciousness in higher organisms a peculiar complication arises. A self-reflecting subject may insist that introspection does not permit him to decide whether the world as he sees it is “real,” or just a phantasmagory, a dream, an illusion of his fancy. A decision in this dilemma is important in this discussion, since, if the latter alternative should hold true, no problems as to how organisms represent internally the features of their environment would arise, for all environmental features would be just internal affairs in the first place.

In which sense reality indeed exists for a self-reflecting organism will become clear by the argument that defeats the solipsistic hypothesis.[2] This argument proceeds by *reductio ad absurdum* of the thesis: “This world is only in my imagination; the only reality is the imagining ‘I.’ ”

Assume for the moment that the gentleman in the bowler hat in Fig. 4 insists that he is the sole reality, while everything else appears only in his imagination. However, he cannot deny that his imaginary universe is populated with apparitions that are not unlike himself. Hence he has to grant them the privilege, that they themselves may insist that they are the sole reality and everything else is only a concoction of their imaginations. On the other hand, they cannot deny that their fantasies are populated by apparitions that are not unlike themselves, one of which may be *he*, the gentleman with the bowler hat.

With this, the circle of contradiction is closed, for if one assumes to be the sole reality, it turns out he is the imagination of someone else who, in turn, insists that *he* is the sole reality.

The resolution of this paradox establishes the reality of environment through evidence of a second observer. Reality is that which can be witnessed: hence, rests on knowledge that can be shared, that is, “together-knowledge,” or *con-scientia*.

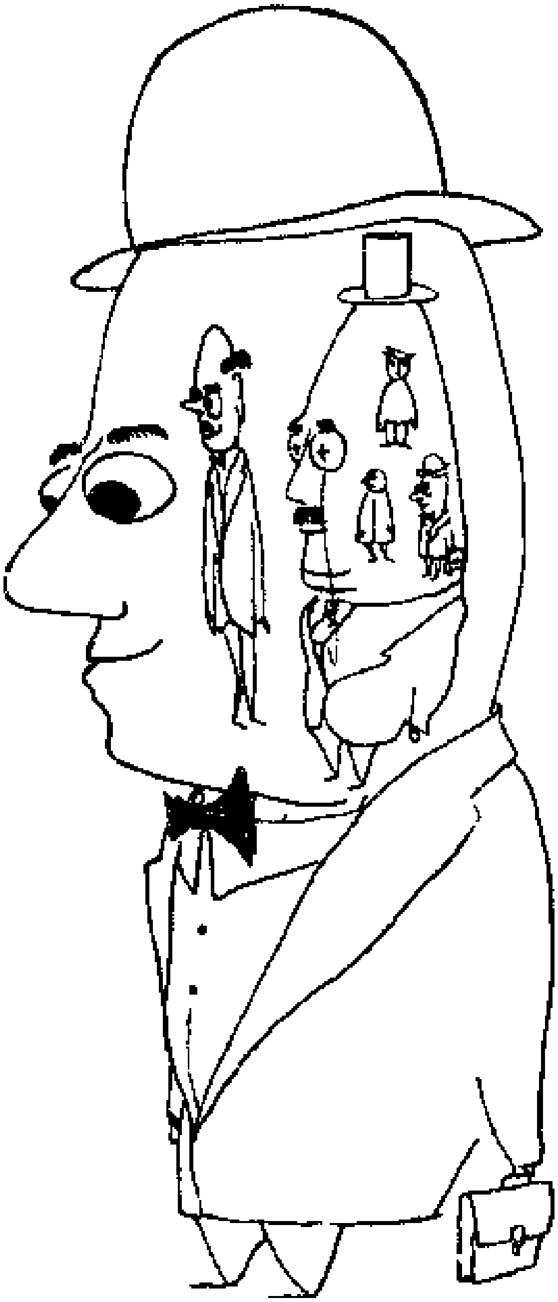


Fig. 4. Reductio ad absurdum *of the solipsistic hypothesis. The hominid apparitions of the gentleman with the bowler had have the gentleman with the bowler hat as apparition. Picture by Gordon Pask.*

## Internal Representation Of Environment: A Physiology

Distributed over the surface of multicellular organisms are highly differentiated cells that establish the interface between the proceedings of the external world and the representations of these proceedings within the organism. To some variables in the physical universe these cells, called sensory receptors, have become specifically sensitive: for example, cells sensitive to changes in pressure are insensitive to, say, the changes in the concentration of sodium chloride in the water surrounding the organism, etc., etc., and vice versa.

Sensitivity of a receptor cell to a specified perturbation is observed by its response in the form of a short electric discharge, which, after it has been initiated at the surface, travels into the interior of the organism along a thin fiber, the axon, which protrudes from the cell.

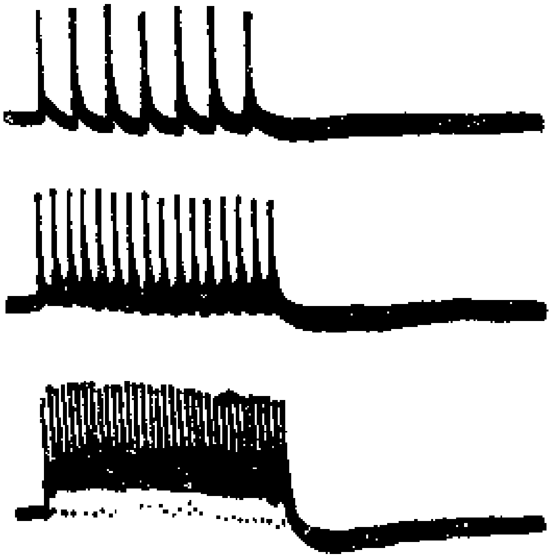


Fig. 5. *Electrical pulse activity measured with a microprobe on the axon of a tactile sensor neuron under different pressures. High frequency corresponds to high pressure.*

The approximate duration of this discharge is several thousandths of a second and its magnitude always about one-tenth of a volt, irrespective of the intensity of the perturbation. A prolonged perturbation produces a sequence of discharges the frequency of which corresponds approximately to the logarithm of the intensity of the perturbation. A series of such pulse sequences measured with small electrodes in the axon of a tactile sensor is shown in Fig. 5. In engineering language the encoding of an intensity into frequency of a signal is called frequency modulation, or FM, and it may be noted that all sensory information — irrespective of sensory modality — is coded into this common language.

If a perturbation is permanently applied, the interval between pulses slowly increases until the sensor fires at a low frequency — called the resting rate — which is independent of the intensity of the permanent perturbation. This phenomenon, "habituation," is one example of computational economy in living organisms, for a property of the universe that does not change in space or time can safely be ignored. Air has no smell. It is the change of things to which an organism must be alerted.

A specific perturbation that elicits responses of a sensory receptor is called stimulus. Stimulus and receptor are duals in the same sense as are environment and organism. Consequently, a tree of the evolution of sensory receptors could be drawn which, at the same time, would show the successive acquisition of specified properties of the physical universe that are selectively filtered out from the rest of the universe.

In the higher animals the most intricately developed sensory system is that of their visual organs. Distributed over the human retina are 180 million sensory receptors of essentially two kinds, the rods and the cones. Rods respond to brightness in general and are more concentrated on the periphery, while cones respond to brightness modified by a variety of pigments and are more concentrated in the central part of the retina, the fovea. The fovea, by proper accommodation of the crystalline lens, has the lion's share in transducing the information contained in the inverted image focused on the retina.[3] The concentration of sensors in the fovea is very high indeed. An area on the retina of the magnitude of the small, black, circular spot that indicated termination of the previous sentence contains approximately 20,000 cones and rods. The projected image of this spot, when looked at under normal reading conditions, is "seen" by about 200 cells. Since each cell distinguishes about 60 levels of brightness, the number of images distinguishable by this small ensemble of 200 cells is exactly $(60)^{200}$, or approximately $10^{1556}$. This is a meta-astronomical number which, if printed out on this page, spreads over 13 lines.

It is clear that this overwhelming mass of information is neither useful nor desirable, for an organism has to act; and to act requires making a decision on the available information, which in this case is so large that it would take eons of eons to initiate action, even if the evidence were scanned at lightning speed. Moreover, any accidental distortion of the image — may it be ever so slight — caused, say, by light scattering in the vitreous humor, by optical aberrations in the lens, such as achromatism, astigmatism, temporary failure of single receptors, etc., etc., would pass as evidence with equal weight and be admitted in the decision-making operation.

What, then, protects the brain from overflow of information?

A first clue was discovered by counting the fibers in the optic tract that is the bundle of nerves which connect the eye with the brain. Here one counts only one million fibers, a reduction by 1/180 compared with the number of sensors. Why this waste on the sensory level, or why this redundancy? Is all this tremendous sensory information just discarded? One has to look and to measure in order to answer these questions.

The anatomy of the post-retinal neural structures is known over many decades; the knowledge of its functions emerges only slowly with advances in electronics and the refinement of micro-electrodes that permit penetration of single fibers *in vivo* and thus permit the recording of their activity under controlled conditions of illumination.

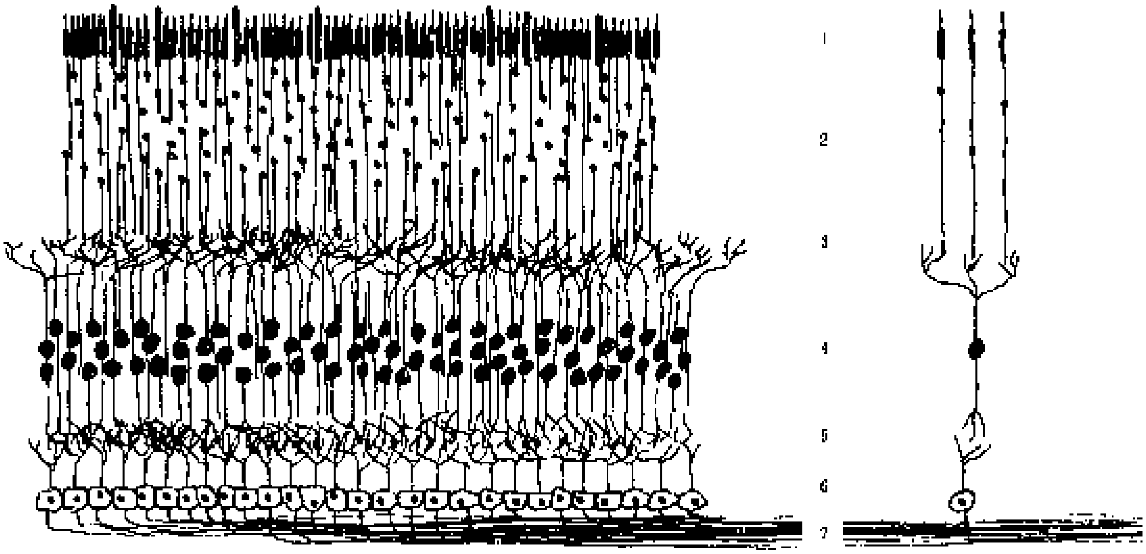


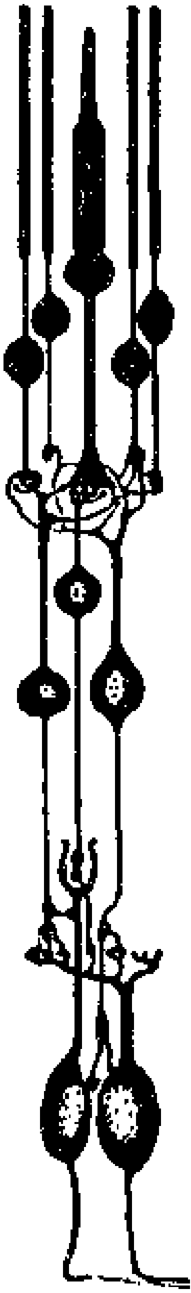
Fig. 6. *Semi-schematic drawing of the post-retinal neural network. 1. Rods and cones. 2. Nuclei of cones and rods. 3. Interaction between sensors and bipolars. 4. Bipolar cells. 5. Interaction bipolars and ganglion cells. 6. Ganglion cells. 7. Optic nerve.*

Fig. 6 shows a semi-schematic sketch of the multi-layered post-retinal neural network that connects sensors with the fibers of the optic tract. Rods, and a few cones, with their associated cell bodies containing the nucleus, comprise layers 1 and 2, the light-sensitive nerve-endings in 1, the nuclei in 2. Their axons descend into layer 3 where contacts are established with fibers emerging from the nuclei of a second layer of cells, the “bipolars,” in layer 4. Their axons, in turn, connect in layer 5 with branch-like ramifications, the “dendrites,” emerging from cells of a third kind, the first ganglion cells in layer 6, which send their axons into deeper regions of the brain, making up the fibers of the optic tract, layer 7.

Two features of this network should be noted. First, that only a few sensors within a spatial neighborhood contribute to one ganglion cell, as can be seen more clearly in Fig. 7, which shows an elementary net of four rods, one cone, three bipolars and two ganglion cells, drawn directly from microscopic observation.[4] Second, that the signal pattern generated at the cones and rods may be modified only in two places, namely in layers 3 and 5 where cells in different layers connect, and thus may act on their successors according to rules of signal transmission from neuron to neuron and according to the local connection scheme.

The mechanisms that determine the response of a successor cell when stimulated by the activity of its predecessor at the place of their junction — the synapse — are still today not clear. Nevertheless, it is clear that two types of interaction can take place, excitation and inhibition. An excitatory synapse will transmit to the successor the oncoming discharge, while an inhibitory synapse will cancel the trigger action of another excitatory synapse. Fig. 8 suggests a symbolic representation of these two synaptic functions. The triangular figure represents the successor neuron with its axon extending downward. Axons from predecessor neurons forming excitatory synapses are indicated as knobs, those with inhibitory synapses as loops.

This observation of the two kinds of signal transmission suffices to see neural interaction in a new light, for it suggests the possibility of seeing the function of a neuron in the form of a logical operation, the affirmative corresponding to excitation, negation corresponding to inhibition. Hence, a network of synapting neurons can be regarded as a system that computes certain logical functions depending upon the type and structure of the connections.



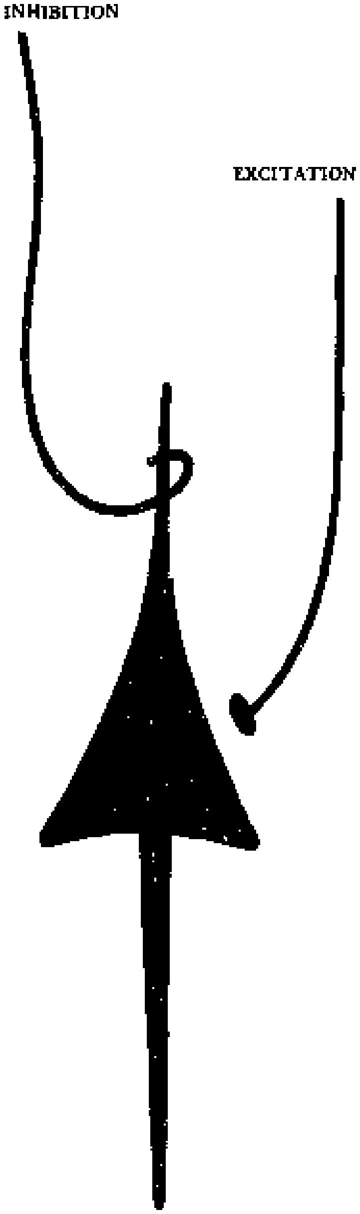


Fig. 7. *Elementary net composed of four rods, one cone, three bipolars and two ganglion cells. Reproduced from S. L. Polyak,* The Vertebrate Visual System, *Chicago, The University of Chicago Press (1957).*

Fig. 8. *Symbolic representation of excitation (knob) and inhibition (loop) of a neuron.*

To see clearly the significance of this observation, an idealized two-layer neural network is drawn in Fig. 9. The first layer consists of "rods," each of which acts upon precisely three neurons in the second "computing" layer. Two fibers with excitatory synapses connect with the neuron just below, while two other fibers with inhibitory synapses connect with its left- and right-hand neighbor. This we shall call an elementary net. It repeats itself periodically over the entire strip, which is thought to extend far out to both sides of the figure.

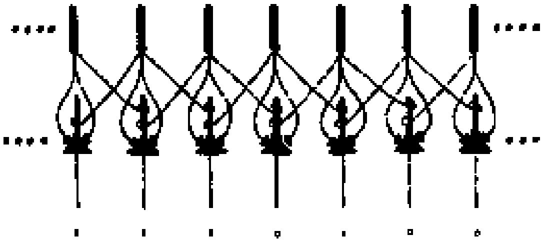


Fig. 9. *Periodic network of idealized neurons incorporating lateral inhibition.*

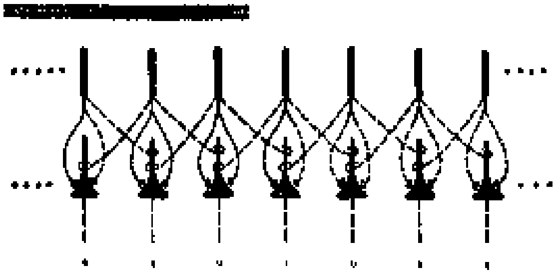


Fig. 10. *Periodic network with later inhibition computing the property "edge."*

What does this net compute? Assume that all sensors are uniformly illuminated. An arbitrary neuron in the computer layer receives from its corresponding sensor immediately above two excitatory stimuli which are, however, canceled by the two inhibitory stimuli descending from the immediate neighbors of its corresponding sensor. Due to the perfect cancellation of the two "yeses" and the two "noes," the net result is no response at all. Since this is true for all other neurons in the computer layer, the whole net remains silent, independent of the intensity of light projected on the sensors. One property of this scheme is now apparent: the net is insensitive to a uniform light distribution.

What happens if a perturbation is introduced in the light path? Fig. 10 illustrates this situation. Again, under regions of uniform darkness or uniform illumination the computer cells do not respond. However, the neuron at the fringe between darkness and light receives no inhibitory signal from the sensor in the shade, double excitation overrides single inhibition and the cell fires. Due to the periodicity of the elementary net, this property, namely, the presence of an edge, will be computed independent of the position of this edge and independent of the level of over-all illumination. Hence, such a network may be called an "edge detector," which when the same principle is extended into two dimensions, may be called a "contour detector."

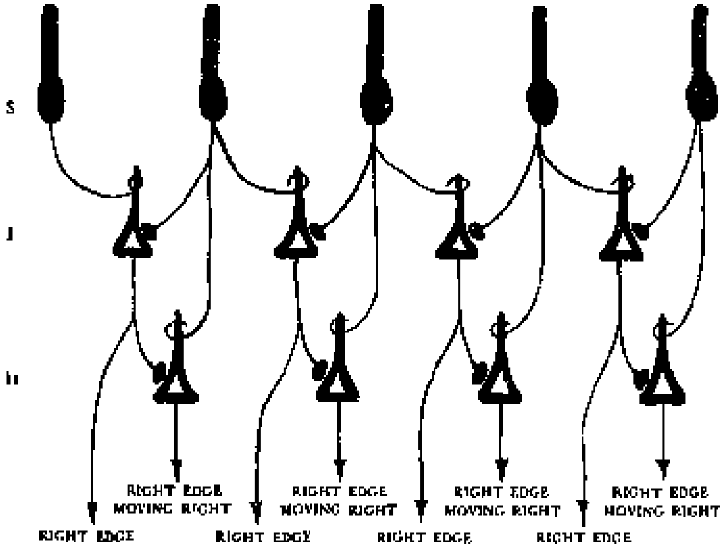


Fig. 11. *Anisotropic periodic network computing the property "right edge" and the property "right edge moving right."*

Other connections will compute other properties in the visual field. Fig. 11 shows a periodic net with two computer layers, 1 and 2, where layer 2 utilizes results computed by layer 1. Inspection of the connection scheme may easily show that layer 1 computes the presence of a right edge (light right, dark left); while layer 2, utilizing the synaptic delay in elements of layer 1, computes a right edge moving right. Of course, no responses are obtained for left edges (light left, dark right), left edges moving left or right, and right edges moving left.

These examples are intended to show that owing to the basic computational properties of the neuron, parallel, periodic arrays of elementary networks are capable of extracting a variety of useful "invariants" in an otherwise complex environment. The theory that connects structure and function of such networks with the invariants they compute is fully developed. Given any universal property to be computed, the appropriate network to carry out this computation can be synthesized.[5]

To establish similar correlations in actual physiological nerve nets is infinitely more difficult. Nevertheless, during the last couple of years in a series of brilliant experiments[6] the computation of invariants by post-retinal networks in some vertebrates (frog, pigeon) has been demonstrated. The experimental procedure consists of observing responses of single fibers in the optic tract elicited by the presentation of various visual stimuli to the retina of an anesthetized animal. These observations show indeed that certain fibers respond only if the appropriate invariant is present in their receptor field. Some of these invariants are:

1. Local sharp edges and contrast.
2. The curvature of edge of a dark object.
3. The movement of edges.
4. Local dimmings produced by movement or rapid general darkening.
5. Vertical edges only.
6. Vertical edges moving.
7. Vertical edges moving right (left) only.
   (etc.)

These abstracts are still on a primitive level, but it is the way in which they are computed that invites further comments. Although only those operations of the perceptive apparatus have been described which are an immediate consequence of the stimulation of sensors, some basic principles are now visible which underlie the translation of environmental features into representations of these features within the cellular architecture of the organism. Perhaps the most fundamental principle involved in this translation is the correspondence between the *neighborhood relationships* that determine environmental structures, and the *neighborhood logics* that are incorporated into neural connectivity which determine the "whether" and "where" of certain environmental properties.

This suggests two levels of computation. First, computation on the grand scale of evolutionary differentiation which incorporates the environmental constraints into the structure of those networks which on the second level, compute within the limits of their structure spatiotemporal quantities of useful universal parameters. Clearly, the first level refers to the species, the second to the specimen. It is on the first level that the notion of "Platonic ideas" arises, for they refer to the fabric without which experience cannot be gathered.

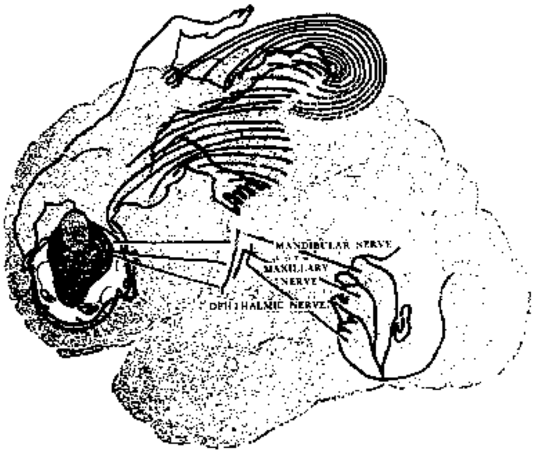


Fig. 12. *Topological mapping of the sensation of touch into the outer layers of the brain "homunculus."*

The importance of distributed operations that can be carried out on a distributed stimulus is further emphasized by a careful preservation of neighborhood relationships even after the original stimulus has been relayed over many cascades of computational layers into the deeper regions of the brain. Fig. 12 shows a topological mapping — that is, a mapping which preserves neighborhoods — of our body with respect to the sensation of touch into the appropriate cortical regions. This "homunculus" is obtained by registering with microelectrodes those regions in the brain which become active when certain regions of the body are stimulated.[7] Such a "signal representation" must not necessarily conform with original proportions, as seen by the emphasis of organs that convey most of the tactile information. The importance is the preservation of neighborhoods which permit further computations of tactile abstracts.

The reliance on neighborhood relationships can cause peculiar breakdowns of the perceptive apparatus when presented, for instance, with a triple-pronged fork with only two branches (Fig. 13). Although in all details (neighborhoods) this figure seems right, as a whole it represents an impossible object.

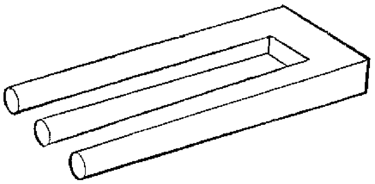


Fig. 13. *Triple-pronged fork with only two branches.*

Similar difficulties arise when the visual system is confronted with unusual projections which do not allow quick reconstruction of the unprojected image. Erhard Schön's anamorphosis (Fig. 15 [next page]) seems to picture a somewhat peculiar landscape, but "actually" it portrays the three Emperors, Charles V, Ferdinand I, Francis I, and Pope Paul III. Faces of these personalities, including their names, can easily be recovered by looking at this engraving under a gazing angle from the left.[8]

Since all sensory modalities translate stimuli into the universal language of electric pulse activity, invariants computed by different senses may be compared on higher levels of neural activity. Since it is on this level where we have to search for the origin of symbolization, this point may be illustrated by an example.

A hypothetical anthropologist visits a fictitious tribe whose members use symbolic representations, two of which are shown in Fig. 14. One is referred to as "Ooboo," the other one as "Itratzky." It is significant that no further information is required to identify these symbols.

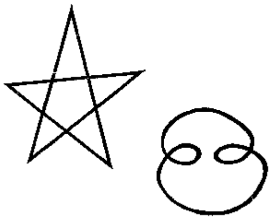


Fig. 14. *"Ooboo" and "Itratzky."*

In the light of the preceding discussion it may indeed be argued that in this case the pattern of neural activity, which represents the visual stimulus configuration, is homologous to that generated by configurations of the auditory stimulus. This argument is going in the right direction, but it fails to cope with a strange situation, namely, that earlier experience and learning is not involved in this spontaneous identification process.

Since associations gained from experience are excluded, one must assume that this audio-visual correspondence rests upon the fabric without which experience cannot be gained. The structure of this fabric must permit some cross-talk between the senses, not only in terms of associations, but also in terms of integration. If this structure permits the ear to witness what the eye sees and the eye to witness what the ear hears then there is "together-knowledge," there is *con-scientia*.

Fig. 15. *Anamorphosis by Erhard Schön, 1535.*

Fig. 16. Nature Teaching Nature*, Allegory from* Scrutinum Chymicum *of Michael Maier, 1587.*

## Symbolization: A Synthesis

To survive is to anticipate correctly environmental events. The logical canon of anticipation is inductive inference, that is, the method of finding, under given evidence $E$, the hypothesis $H$ which is highly confirmed by $E$ and is suitable for a certain purpose. This is computation of invariants within the limits of insufficient information, and follows the principles of invariant computations as before, only on a higher level. Knowledge is the sum total of these hypotheses (invariants, laws, regulations) and is accumulated on three levels. First, on the molecular level in the genetic structure which tests the viability of its hypotheses, the mutations, through the vehicle of the developed organism; second, on the level of the individual organism through adaptation and learning: and third, on the social level through symbolic communication which cumulatively passes information on from generation to generation.

Since these are evolutionary processes, and hence irreversible, error would accumulate with knowledge, were it not for a preventative mechanism: death. With death, all registers are cleared and untaught offspring can freshly go on learning. This mechanism works on the first and second levels, but not on the third.

To cumulatively acquire knowledge by passing it on through generations, it must be communicated in symbols and not in signs. This separates man from beast. Communication among social insects is carried out through unalterable signs which are linked to the genetic make-up of the species. While signs refer to objects and percepts, and serve to modify actions and manipulations, symbols refer to concepts and ideas and serve to initiate and facilitate computation.

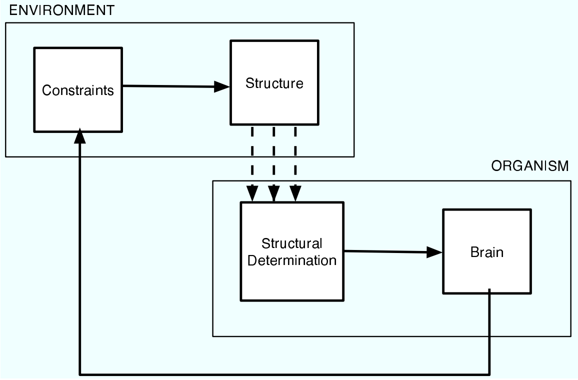
Since the ultimate relation between symbols and environmental entities is cascaded over the relations symbol/concept and concept/environment, it is in its logical structure very complicated indeed. This gives rise to breakdowns that manifest themselves on various levels of semantic morbidity.

Symbols share with concepts and ideas the property that they do not possess the properties of the entities they represent. The concept of roses "smells" as much, or as little, as the concept of jumping "jumps." The concept of a square is not quadratic. If this point is missed, a number would be just so many fingers and a square with area 2 would have non-existing sides.

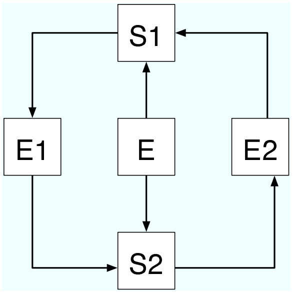
Since symbols refer to concepts and ideas, they too may not have the properties they represent. The symbol of a square may not be quadratic, as can be clearly seen by the string of peculiarly shaped little marks on this paper that have just been used to refer to this geometrical figure. This was, of course, well understood when mystical experience was to be coded into symbols. Michael Maier's allegory, entitled *Nature Teaching Nature*, from the fifty allegories of his *Scrutinum Chymicum* of 1587[9] is here presented for contemplation (Fig. 16). It may be noted that no commentary — except the title — accompanies these pictures.

What, then, determines the form of a symbol; is it an arbitrary convention, or does it convey its meaning by its shape? Again, ontologically this question cannot be resolved. One has to look into the ontogenesis of symbolic presentations.

We here repeat the diagram seen earlier which represents the information flow between a single organism and its environment:



Since symbolization requires at least two interacting subjects who are immersed in an environment that is common to both, we must extend this diagram to admit a second subject. This is done here below:



Subjects $S_1$ and $S_2$ are coupled to their common environment $E$. In contrast to the first diagram in which the organism is faced only with an environment with given constraints, now each of these subjects is confronted with the additional complication of seeing his environment populated with at least one other subject that generates events in the environment $E$. Hence $S_2$ sees, in addition to the events generated by $E$, those generated by $S_1$, and since these take place in $E$, they shall be labeled $E_1$; conversely, subject $S_1$ sees in addition to events generated by $E$ those generated by $S_2$ which will be called $E_2$. Thus, in spite of the fact that both $S_1$ and $S_2$ are immersed in the same environment $E$, each of these subjects sees a different environment, namely, $S_1$ has to cope with $(E,E_2)$ and $S_2$ with $(E,E_1)$. In other words, this situation is asymmetrical regarding the two subjects, with $E$ being the only symmetrical part.

Assume that $E_1$ and $E_2$ are initial attempts by $S_1$ and $S_2$ to communicate environmental features to each other. It is clear that these attempts will fail unless — and this is the decisive point — both subjects succeed in eventually converging to like representation for like universal features. This process may be expressed symbolically:

$$\begin{tikzcd}
E_1 \arrow[dr] & & E_2 \arrow[dl] \\
& E_0 &
\end{tikzcd}$$

The arrows indicate the convergence process, and $E_0$ stands for the final universal "language" spoken by both subjects. At this point the initial asymmetry ceases to exist and both subjects see the same environment $(E, E_0)$.

As in all evolutionary systems, the outcome of this process cannot be predicted in the usual sense, because the goal which establishes equilibrium is not directly visible in the final equilibrial state which is a communicable symbol, while the goal is communicability.

Symbols must not necessarily have the shape of the objects they ultimately refer to, yet within that freedom there are constraints working in the evolution of symbolic representation which confine their development within reasonable limits. One of these constraints is dictated by the tools with which these symbols are generated, the other one is their syntactical structure.

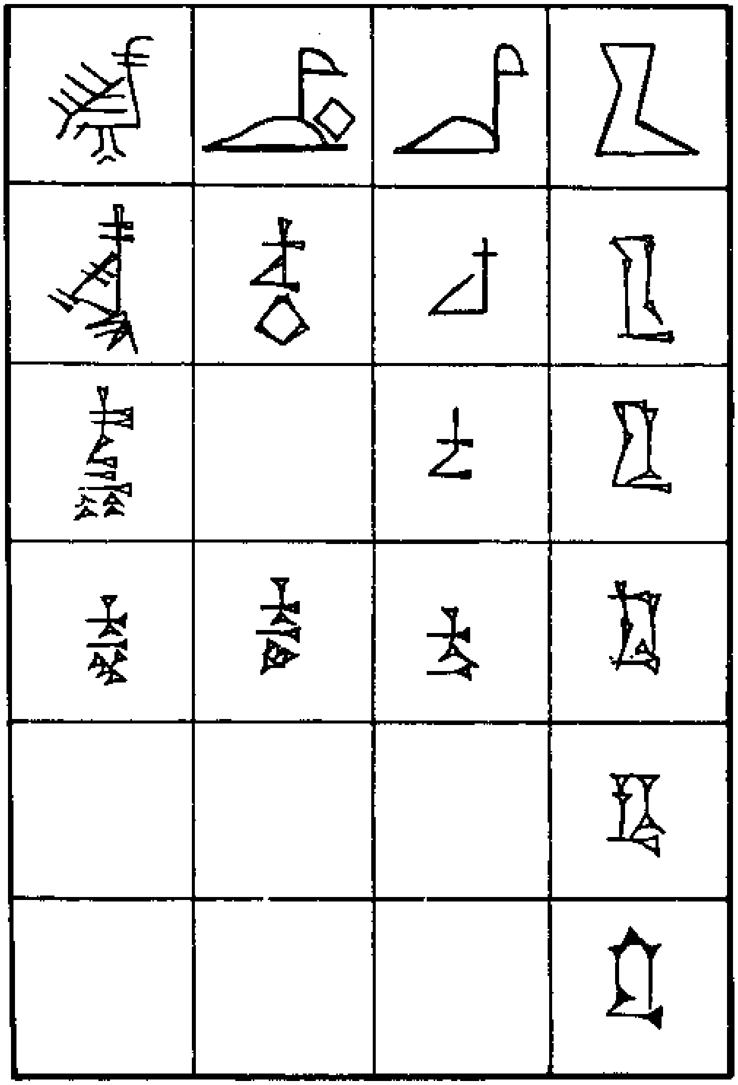


Fig. 17. *Formalization of pictograms through constraints imposed by writing tools. This development is estimated to have taken place in Mesopotamia during a period from the fourth to the second millennium B.C.*

An example of the first kind of constraints operating on the development of written symbols is given in Fig. 17, which shows the development of highly stylized symbolic forms from initially representational pictograms.[10] This transition is believed to have taken place in the two millennia of Sumerian cultural activity between 4000 and 2000 B.C. As one goes down the rows it is clearly seen how the constraints imposed by the writing tools — a stylus with triangular cross-section pressed into soft clay — strongly modify the early pictograms given in the top row. It may be interesting to note that simultaneously with this departure from structural representation goes an increase in the possibility to add modifiers to the original meaning. While the pictogram at the top of the right-hand column indeed says "foot," after two thousand years of stylization (bottom row) it may stand for "walking," "running," "delivering a message," or other "foot-connected" actions if associated with appropriate modifiers. Nevertheless, in some instances it seems to be possible to see behind the form of later symbols the shape of earlier pictorial representations.

The other kind of constraint is a structural one and does not show itself in an obvious way, for symbols carry rules of connectivity and not so much rules of entity. Symbols may be compared to atoms which react to particular atoms only to form the molecular compounds, but are inert to all others. Take, for instance, these "molecular" sentences:

> "Socrates is identical."
>
> "4 + 4 = purple."

The disturbing thing about these is that they are neither true nor false; they are nonsensical. The connection rules of the symbols have been violated in these examples. "Identical" sets up a relation between two entities. "Socrates is identical with Plato" is a sentence that makes sense although it happens to be a false proposition. The compound "4 + 4 = " requires a number to follow. Putting "6" at the end is a good guess, but "purple" is an operator with an entirely different structure. This indicates that somehow structure is still preserved in symbolical discourse, but in a syntactical and not in a representational sense. The language of symbols has, so to speak, its own logical grammar.[11] Uniqueness in symbolic expressions is established in a way similar to that of a jigsaw puzzle in which pieces can be put together in one, and only one way. It is the sometime far-extending neighborhood relationship among the pieces — the symbols — that puts them into place,

It is clear that the constraints expressed in the neighborhood relationships of symbols reflect constraints in the environment. For instance, a sentence that refers to two particular persons must employ two proper names. To establish connection rules among symbols of speech is the linguistic problem. One of the most primitive connectivities among words is the probability of their succession. With the following two examples the emergence of order by tightening the constraints of succession will be demonstrated. These examples are random sequences of words generated by a chance device which, however, takes into account the various probabilities by which a particular English word follows a number of precursors. In the first example the number of precursor words is two:

> . . . THE HEAD AND IN FRONTAL ATTACK ON AN ENGLISH WRITER THAT THE CHARACTER OF THIS POINT IS THEREFORE ANOTHER METHOD FOR THE LETTERS THAT THE TIME OF WHO EVER TOLD THE PROBLEM FOR AN UNEXPECTED . . .

In the second example the constraints are tightened by extending the neighborhood relationship up to four words:

> . . . HOUSE TO ASK FOR IS TO EARN OUR LIVING BY WORKING TOWARDS A GOAL FOR HIS TEAM IN OLD NEW YORK WAS A WONDERFUL PLACE WASN'T IT EVEN PLEASANT TO TALK ABOUT AND LAUGH HARD WHEN HE TELLS LIES HE SHOULD NOT TELL ME THE REASON WHY YOU ARE IS EVIDENT . . .

Symbols are no proxy for their objects.[12] There are two morbid states of the mind, magical thinking and schizophrenia, in which this distinction is erased. In both cases symbol and object become indistinguishable. In purpose-oriented Jou Jou and in Voodoo the identity of symbol with object is used to manipulate the world by manipulating the symbol. In schizophrenia symbol and object are freely interchanged to produce peculiar hierarchies of identities. In order to comprehend in depth the modality of this affliction, a short passage of the extensive description of the case of a six-year-old boy by the name of Walter ( = water) is given here:[13]

*Late in November, 1936, presumably because he had heard a rumor about a child killed in an accident in an elevator there, he became terrified when taken to a department store. He trembled, cried, vomited and remained "hysterical" for two days during which time he made little jerking movements of his body and shoulders and said scarcely a word. The following day, Dr. Hamill was for the first time able to make out that he failed to distinguish between himself (Walter) and water. Walter shifted to water, thence to Deanna Durbin who played in "Rainbow on the River" and so to water again. Being water, he felt he could not be drowned, but might be imprisoned* *in the radiator. On hearing the knocking of water in the radiator, he said, "elevator just came up and gave the kid a knock" and again, "they are killing the kid," which terrified him because he was the kid. Then followed "the telephone burnt and got water after Suzy burnt."*

*(Dr.: "Where does water come from?") "I come from the show." (Dr.: "You thought water and Walter were the same thing.") "My father used to take me across the river." (Dr.: "And he called you Walter?") "And got drowned. I do not live on Springfield. Bad boys drink water. They do not drink milk. Good boys live on Springfield. I used to live on Springfield-Mississippi River."*

It may be speculated that evolution did not weed out mental diseases that afflict proper use of symbols because the survival value of the ability to symbolize is so enormous that occasional morbid deviations of this ability in individuals and in whole cultures could still be tolerated. The enormous advantage of organisms that are able to manipulate symbols over those who can only react to signs is that all logical operations have not to be acted out, they can be computed. It is obvious that this saves considerable amounts of energy. But the really crucial point here is that errors in reasoning are not necessarily lethal.

The recognition of the fact that information is a precious commodity and can be processed by manipulating symbols gave rise to the quick emergence of the fast and large electronic computer systems. These systems manipulate symbols only and do not know objects. The laws of algebra and logic are incorporated in their structure. Hence, they cannot err by confusing modality as does a schizophrenic, nor can they err in syntax and generate nonsense. The only error they can make is confusing true with false and false with true.

The human retina with its associated, genetically structured networks may be compared to these computer systems from a purely quantitative point of view, namely, by the sheer amount of information that is processed. The retina, with its 180 million sensors which operate in parallel at millisecond intervals, performs equivalently to a modem digital computer system that occupies 800 square feet of floor space and uses 4 tons of highly sophisticated electronics. In comparison, the retina's extensions are 2 square inches by 4/1000 of an inch, and it weighs approximately 100 milligrams. This may be taken as an indication of the economy of biological computation.

## Notes

[1] *Structure in Art and in Science* (Gyorgy Kepes, editor), New York, George Braziller (1965).

[2] Self-Organizing Systems and Their Environments," in *Self-Organizing Systems* (M. C. Yovits and Scott Cameron, editors), New York, Pergamon Press (1960), pp. 31–50.

[3] R. Held, "Object and Effigy," in *Structure in Art and in Science…*, p. 42–54.

[4] S. L. Polyak, *The Vertebrate Visual System*, Chicago, The University of Chicago Press (1957).

[5] W. Pitts and W. S. McCulloch, "How We Know Universals: The Perceptions of Auditory and Visual Form," in *Bulletin of Mathematical Biophysics*, vol. 9 (1947), pp. 127–147; H. Von Foerster, "Structural Models of Functional Interaction," in *Information Processing in the Nervous System* (R. W. Gerard and J. W. Duyff, editors), Amsterdam, Excerpta Medica Foundation (1963). pp. 370–383.

[6] J. Y. Lettvin, H. R. Maturana, W. S. McCulloch, and W. Pitts, "What the Frog's Eye Tells the Frog's Brain," in *Proceedings of the Institute of Radio Engineers*, vol. 47 (1959) pp. 1940–1951; H. R. Maturana, "Functional Organization of the Pigeon Retina," in *Information Processing…* , pp. 170–178.

[7] F. A. Mettler, *Neuroanatomy*, St. Louis, The C. V. Mosby Co. (1958), p. 432.

[8] J. Baltrusaitis, *Anamorphoses*, Paris, Oliver Perrin (1955).

[9] R. Caillois, *Au Coeur du Fantastique*, Paris, Gallimard (1965).

[10] O. Neugebauer, *Vorgriechische Mathematik*, Berlin, Springer (!934), pp. 40–78.

[11] L. Wittgenstein, *Tractatus Logico Philosophicus*, New York, Humanities Publications (1961).

[12] S. Langer, *Philosophy in a New Key*, New York, The New American Library (1962), p. 61.

[13] L. J. Mednna and W. S. McCulloch, "The Modern Concept of Schizophrenia," in *Symposium on Neuropsychiatric Diseases*, Philadelphia, W. B. Saunders CO. (1945), pp. 147–164.

This article is in part based on work sponsored by the Air Force Office of Scientific Research under Grant 7-64 and by the National Institute of Health under Grant GM-10718.