

Principles of Self-Organization – In a Socio-Managerial Context

H. von Foerster

One Eden West Road, Pescadero, CA 94060, USA

0. Opening

I have to confess that when I first received the kind invitation from Dr. Probst to participate in a meeting entitled "Management and Self-Organization in Social Systems" I was not quite clear about my role in such a meeting. I am not a stranger to the notion of Self-Organization; but when I considered it in the context of management and, moreover, in the environment of a Hochschule für Wirtschafts- und Sozialwissenschaften, I felt lost. I understand so little about management that already in grade school my teachers complained that this boy is unmanageable. In fact, I had to look "management" up in my dictionary (1). Here I found that it is derived from ... "constraining the movement of hands", having the same root as "to manacle", that is, putting someone into handcuffs : I was prepared to decline this invitation.

Fortunately not much later the organizers of this meeting sent me a paper by Messrs. Malik and Probst entitled "Evolutionary Management" (2), apparently with the idea of giving me a clue of what this meeting would be about. There are two mottos that initiate this paper. Since after I read them I knew I would accept the invitation, I shall read them also to you. The first is a quote by Peter Drucker who, like me, grew up in Vienna, and whose parents happened to be good friends with mine:

"The only things that evolve by themselves in an organization are disorder, friction, and malperformance ..."

That is not a bad start for a paper that addresses itself to self-organization in management. The second motto is again by a Viennese, the Nobel laureate Friedrich von Hayek, who participated in a conference on principles of our topic I had organized almost a quarter century ago. Here is his quote:

"... the only possibility of transcending the capacity of individual minds is to rely on those super-personal 'self-organizing' forces which create spontaneous order".

With these two mutually annihilating quotations the organizers of this meeting had me almost hooked, but succeeded completely after I had read the entire article. There were four points that were very much to my liking:

- (i) Hierarchies are inappropriate skeletons for a managerial structure;
- (ii) The importance of flexibility and adaptation;
- (iii) Limited control of, and knowledge in, the system;
- (iiii) And finally, the last line of this article which reads:

"As managers we have to ... learn to be what we really are: not doers and commanders, but catalysts and cultivators of a self-organizing system in an evolving context."

I found myself very close to this sentiment, an affinity with a point I once made at the end of one of my papers (3). I called it an "ethical imperative":

"Act always so as to increase the number of choices!"

My general impression was that the two authors were in search of an epistemology; an epistemology that takes account of the situation in which the manager is himself an element of the system he is managing.

A decade or two ago nobody in his right state of mind would have dared to consider this problem, or even to formulate it that way. And if one would have done so, all experts would have had the times of their life to show that this self-inclusion is the root of all paradox. If mildmannered they would have referred to the barber in the village who shaves all who do not shave themselves (clearly, those who shave themselves need not to be shaved). So far so good. But should the barber shave himself? Of course not, for he shaves only those who do not shave themselves. Apparently, he is not to shave himself. But then ... etc. If it is a learned expert he may cite Bertrand Russell's victory over the paradoxical "set of all sets that do not contain themselves as elements" (with the unanswerable question: does this set contain itself as element, or does it not?). This victory was celebrated as the "theory of types" in which this liberal gentleman simply forbade self-inclusion on logical grounds (a proposition must be either true or else false; here, however, these propositions are true when apprehended as false, and false when apprehended as true).

Fortunately, today the situation is quite different, thanks to the pioneering work of three gentleman. One is Gotthard Günther, a philosopher, now professor at the University of Hamburg, who developed a most fascinating multi-valued logical system (4), quite different from those of Tarsky, Quine, Turquette, and others. Then, there is Lars Löfgren, a logician in Lund, Sweden, who introduced the notion of "autology" (5), that is, concepts that can be applied to themselves, and in some cases need themselves to come into being. I shall dwell on these points in a moment. Finally, we have the work by Francisco Varela, who sits right here, who, as you all know, expanded G. Spencer-Brown's Calculus of Indication to become a Calculus of Self-Indication (6).

My plan for this paper is to build upon these ideas, and in attempt to maximize my usefulness to this meeting, I shall present my points complementary to those made by Malik and Probst in their paper (reprinted in this volume):

- (i) First, I shall expand on the notion of autology;
- (ii) Second, I shall give a brief account of a rather general interpretation of the concept of computation, and its (conceptual) realization in form of "machines", because I need this concept for the next point I wish to make, namely,
- (iii) Recursive Computations.
- (iiii) Finally, I shall make use of all that by talking about self-organization in the socio-managerial context.

1. Autology

I wish to contemplate the manager who considers himself a member of the organization he manages. If he takes this consideration seriously, he has to apply his managerial perceptions and acts to himself, to his own perceptions and acts. Management, clearly, is an autological concept. In some other context, such concepts are referred to as "second-order concepts".

To get a feeling for the peculiar logical properties that distinguish autologies from other concepts, I invite you to participate in the experiment suggested in Figure 1. Kindly follow the instructions as given in the caption of this Figure, and do not give up until indeed the black spot has completely disappeared. This phenomenon is usually referred to as the "blind spot" in our visual field, and physiologists have a straightforward explanation for this phenomenon (Figure 2). There is a place on our retina where there are no receptor cells, neither rods nor cones. This place is called the "disc", and it is there where the optic nerve leaves the eye ball. Of course, the black spot can not be seen when one is forced to project the spot on the disk when keeping the asterisk focused on the fovea.



Figure 1: Hold paper with right hand. Close left eye. Fixate asterisk. Move paper to and fro along the line of vision. Watch black spot disappear (at eye-paper distance between 12 and 14 inches). Keep asterisk fixated and move paper slowly parallel to itself up, down, left, right, or in circles: black spot remains invisible.

This explanation seems to take care of these affairs, and we could turn to other matters. However, I would like to make two comments here, one regarding the blind spot phenomenon itself, the other about this explanation.

What apparently is surprising in this experiment is its demonstration of the incompleteness of our visual field, an incompleteness of which we are totally unaware under normal conditions. If one were to stress now the autological nature of visual perception or, as a matter of fact, of perception in general, one may say that we don't see that we don't see!

This suggests that the problem here is not not-seeing, the problem is not seeing that one is not seeing. This is a problem of the second order, and it is graciously overlooked in the orthodox explanation above. Hence, not seeing the problem is the blind spot phenomenon all over again, only now on the cognitive level.

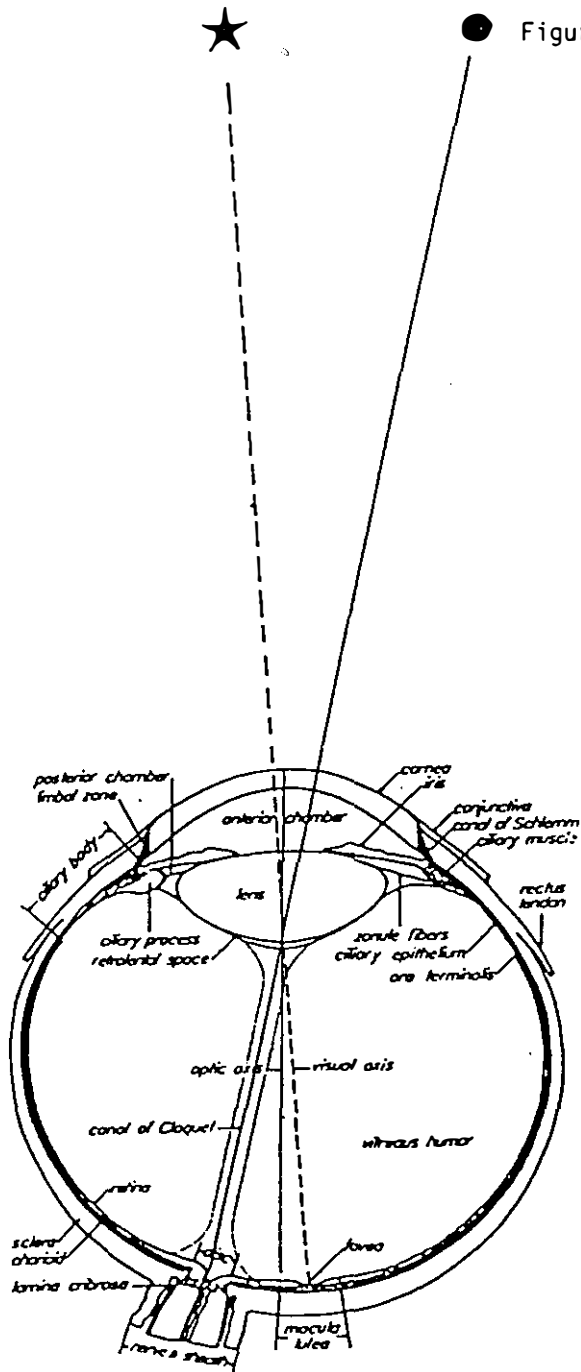
My strategy of introducing second-order concepts containing negatives was to show at once their unusual logical structure, for here double negation does not yield affirmation: not not-seeing does not imply seeing.

I shall now turn to examples of these concepts with an affirmative logical skeleton, again to draw your attention to the different "logical types" as Gregory Bateson may have said, of notions that are embedded in their own domain.

Let me begin with "purpose". If taken as a first-order concept one may speak of something "having a purpose". However, taken on its second-order level we may ask "what is the purpose of 'purpose'?", that is, to ask why introduce the notion of purpose in the first place. Of course, the answer here is straightforward, namely, to avoid contemplating variable and unpredictable trajectories by attending to a more or less invariant state of affairs; the "goal", the "end", *telos*. However, by paying attention to the autological nature of "purpose", our gaze is shifted from "something", the observed, to "somebody", i.e., the one who uses this term, that is, the observer (7).

Next, I turn to language: "What is language? Or better, what is "language"? Whatever is asked here, it is language we need for an answer; and, of course, we need language to ask that question on language. Hence, if we did not know the answer,

Figure 2: Horizontal section of the right human eye, showing locus of projections.



how could we have asked the question in the first place? and if indeed we did not know it, what will an answer be like that answers itself?" (8). The semantic loop I am stressing here suggests a logical constraint in a possible definition of "language", namely, its autological nature. That is, for any referential communicative conduct to be "language", it must contain reference to its communicative conduct (i.e., a language must be able to express the notion of "language" or, as Humberto Maturana is fond of saying, language must be able to refer to its referring, must be able "to point to the pointing"). Of course, the ultimate teaser in this context is Ludwig Wittgenstein's question (9): "What is a question?", and I will leave it to you to tackle it.

As a final example, I shall now deal with the autological nature of the central topic of our meeting, namely "organization". Let me again go through the shift from a first-order to a second-order interpretation of this concept. We take the corresponding transitive verb "to organize", then we stipulate a world in which the organizer and his organization are as fundamentally separated from one another as are the active and the passive forms; it is the world of organizing the other, it is the world of the injunction:

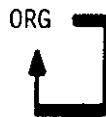
"Thou shalt ..."

On the other hand, if we contemplate the organization of an organization so that the one slips into the other, i.e., "self-organization", we stipulate a world where the actor acts ultimately on himself, for he is included in his organization: it is the world of organizing oneself, it is the world of the injunction:

"I shall ..."

From this it appears to be clear that shifting from first to second-order interpretations has as one of its consequences a shift in the epistemological foundations of ethics. The novelty appears in the latter case, where for the first time one may begin to see the ethical epistemologist becoming accountable for his own epistemology.

I hope that with all these examples of autology, and most explicitly in that of self-organization, my position not to yield to the Russellian escape route into meta-domains (e.g., "meta-languages", etc.) has become apparent. May be the essential feature of those concepts that can be applied to themselves, namely, "closure", has become apparent as well. Perhaps the following symbolization of, say, an organization that applies its competence to itself



suggest "closure" even more persuasively.

Moreover, those of you who are familiar with the formal development of this argument may recognize in the "recursive pointer" Francisco Varela's mark for the autonomous state



which he introduced almost ten years ago in his seminal paper on a calculus of self-reference (6).

While at first one would think that the introduction of closure adds richness to the arguments, it does in fact do the opposite. It removes one degree of freedom. This is so, for whatever we may consider the "end" in any domain, it must coincide with the "beginning", otherwise the system is not closed. Since this is a crucial point, as you will see in a moment, let me demonstrate this on two examples.

The first I will take from physics, from the early days of wave mechanics. As you may remember, some experiments with elementary particles, electrons in particular, suggested that they could be interpreted as the particles behaving like waves, aug-

menting each other when crests meet crests and valleys meet valleys; but annihilating each other when crests meet valleys. If this is so, de Broglie argued, electrons orbiting the nucleus in an atom would always annihilate themselves, unless they would move in orbits that are integer multiples of their wavelength (see Figure 3), only then crests would meet crests, and valleys valleys; that is, the end of a wave train must be its beginning.

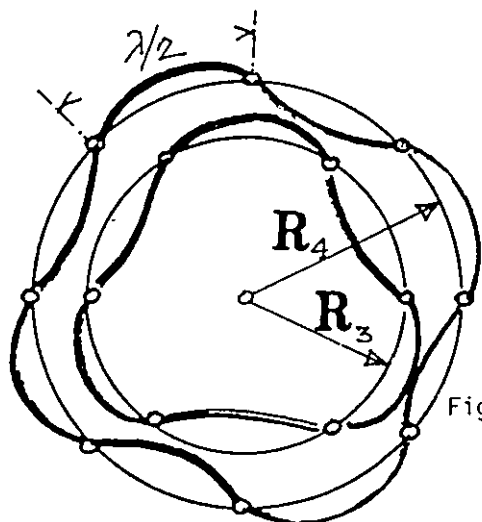


Figure 3: Stable electron orbits along "Eigen-Radii" corresponding to circumferences of multiples of wavelength λ :

$$R_3 = 3\lambda/2\pi; \quad R_4 = 4\lambda/2\pi.$$

With this condition to be fulfilled, it is clear that only certain orbits can exist, they are "quantum jumps" apart, and it was the confirmation of de Broglie's hypothesis through quantum physics that brought him the Nobel Prize.

Please note again from the argument or from Figure 3 that the condition of closure, i.e. the end fitting the beginning, carves out from the infinite possibilities these electrons could move around their nuclei, a set of discrete solutions whose values fulfill the desired condition.

These values are called "Eigen-values" ("self-values"), first so called around the turn of the century by the mathematician David Hilbert in connection with solutions of problems with similar logical structure.

My second example has to do with self-referential propositions. As you may remember, these have always been believed to be the real trouble makers, for instance, the paradoxes of the Epimenides type, one of which I have mentioned before (the barber's difficulty of shaving himself). However, as we shall see in a moment, these situations are not only not irresolvable, as was thought before, but their solutions provide us with insights into other domains.

Consider the following incomplete sentence:

THIS SENTENCE HAS ... LETTERS

and find a number whose name spelled out and inserted into the blank spaces makes this sentence complete and consistent. Clearly, from the infinite reservoir of numbers only a few, if any at all, will fulfill this condition. For instance THIRTY would not do, for the sentence "This sentence has thirty letters" has in fact only 28 letters.

There are two solutions, two "Eigen-values", to this problem which satisfy the conditions of above. One of these is THIRTYONE. Indeed, the sentence

THIS SENTENCE HAS THIRTYONE LETTERS

has 31 letters. Moreover, note that this sentence says what it does!

The other solution I suggest you work out for yourself, because such an exercise drives home forcefully what it means "to make ends meet" (10).

Since in these cases of closure one runs the result of an operation again through this operation, one speaks of "recursive operations" (from re=again, and currere=to run). The theory that provides the formalism for these processes is called "recursive function theory". Today, this mathematical field is a well established and extensive body of knowledge (11), and I shall touch upon it briefly later on.

What are the consequences of all this for management? Let me suggest one which I think has many ramifications:

In a self-organizing managerial system each participant is also a manager of this system.

Such an organizational structure is called a "heterarchy" (heteros = the other, and archein = to rule), for at one time it may be one of your neighbours who is making the decisions, at another you, as the neighbour of others. This organization is, of course, the antipode of a "hierarchy", where the "holy" (hieros) rules, where the boss has all the power, and the line of command is from top down.

The notion of heterarchy was, to my knowledge, first introduced by Warren McCulloch in one of his papers "A Heterarchy of Values Determined by the Topology of Nervous Nets" (12), which to read is an intellectual feast.

As McCulloch told it, he derived the concept of a heterarchy from a principle he very much cherished. It is

The principle of potential command, where information constitutes authority.

As an example of this principle he used to tell the story of the battle of the Midway Islands where the Japanese fleet was out to destroy the American fleet. Indeed, the American flagship went down in the first few minutes, and its fleet was left to organize itself, i.e., to switch from a hierarchy to a heterarchy. What happened now was that the skipper of each vessel, small or big, took the command over the entire fleet whenever he realized that he, because of his position at that moment, knew best what to do. The result, as we all know, was the destruction of the Japanese fleet, and the turning point for the developments of the war in the Pacific.

2. Machines

I am sure you spotted in my presentation the two main themes to which I returned again and again, self-reference and closure, and also sensed my attempt of slipping these two notions into each other. The device I employed in this attempt was "recursion", and I hope that you could taste some of its flavour, because I would like now to demonstrate the power of this concept in the context of our meeting. Since I wish to do that by invoking elementary steps in its formalism, the formalism of recursive computations, I will make first some preliminary remarks on computation in general.

First, let me remind you that the etymological root of "computation" does not in the least confine it to numerical expressions. The word is a merger of "com" = to-

gether, and of "putare" = to contemplate, that is, contemplating things together. Clearly, there is no restriction regarding the "things" contemplated, and I shall use it in this general sense.

As a vehicle for talking about computation I am going to use the idea of a "machine", very much in the sense in which Alan Turing introduced it almost a half century ago, namely, as a conceptual device with well-defined rules of operation. However, I will not describe here a Turing Machine (13), for it would move us too far away from our central topic, but I will give you an account of even more general conceptual computing devices, the so-called "Finite State Machines" (14).

Of those there are two kinds available now, the Trivial and the Non-Trivial Finite State Machine, or the TM and NTM for short. I shall first extoll the charms of the trivial machine (TM), and then develop those of the NTM.

The Trivial Machine

Figure 4 is a schematic representation of a TM, with the labels x , y , f , referring to "input", "output", and "function" of this machine respectively, and the arrows indicating the direction in which the operations are performed. The idea is to have a clear understanding of process. Take, for example, x and y representing the natural numbers 1, 2, 3, 4, ..., and let the function of this machine be the production of an output y that is the square of the input x , i.e., this machine is a "Squaring" TM. Of course, you know what is going on here, and you also know that there is a variety of ways for describing this, some anthropomorphic - or even biomorphic - ways. For instance, if one "feeds" our Squaring machine a 4 ($x=4$), it will "spit out" 16 ($y=16$). Or take another TM, those one sees today at the check-out counters of supermarkets. An item is moved with its code lines over the machine's "sensor" and the printer enters "NOODLES ... \$ 3.50" on the bill (a "Billing" TM). Or kick a ball into the air ($x=kick$) and watch it flying up and falling down (watch y). This is the operation of a "Gravitational-Attraction" TM. Or consider the structure of the deductive syllogism. The classical example is, of course: "All men are mortal" (the major premise); "Socrates is a man" (the minor premise); and how the conclusion: "Socrates is mortal", I call this the "All-Men-Are-Mortal" trivial machine, for whatever you take as an input, as long as it is a man, a (potential) corpse will emerge on the other side; and so on and so forth.



Figure 4: Trivial Machine .

I have chosen this outrageous mix of samples, for I wanted to let the following three points to become utterly, utterly clear.

Number one: In spite of the tremendous variety of context in these examples, the underlying schema of argument, logic, operation, etc., is in all the same: because of the invariable relationship (f) between input (x) and output (y), a y once observed for a given x will be the same for the same x given later. The consequence of this is that all TMs are:

- (i) predictable,
- (ii) history independent.

Number two: Because of the popularity of the inference schema of trivial machines the three the machine determining entities, x , y , and f , depending on the different contexts, appear and re-appear under the most diverse names. Here is an incomplete list:

x	f	y
input	operation	output
independent variable	function	dependent variable
cause	Law of Nature	effect
minor premise	major premise	conclusion
stimulus	C.N.S.	response
motivation	character	deeds
goal	system	action
...
...

Number three: When a TM is synthesized, that is, when the x - y correspondence (i.e., the function f) is established, this machine is then unambiguously defined. One speaks here of a synthetically determined system. A particularly nice feature of these machines is that they are also analytically determinable, for one simply has to record for each given x the corresponding y. This record is then "the machine". Hence, all TMs are

- (iii) synthetically deterministic,
- (iiii) analytically determinable.

I shall summarize this now by inviting you to contemplate a trivial machine that has the following properties: it can distinguish four input states (x): A, U, S, T; and two output states (y): 0, 1. The correspondence between x and y is established through this Table:

f	
x	y
A	0
U	1
S	1
T	0

Hence, from the input sequence of, say, A, U, S, T, the machine will compute the output sequence 0, 1, 1, 0; or from the sequence U, S, A, it will compute 1, 1, 0; and when this sequence is repeated again and again, undisturbed of what may happen in between, we shall obtain again and again, 1, 1, 0, until the Day of Judgement.

Non-Trivial Machines

Obedience is the hallmark of the trivial machine; it seems that disobedience is that of the non-trivial machine. However, as we shall see, the NTM too is obedient, but to a different voice. Perhaps, one could say obedient to its inner voice.

How do NTMs differ from TMs? In fact in a very simple, but profoundly consequential way: a response once observed for a given stimulus may *not* be the same for the same stimulus given later.

The most fruitful way to account for such changes in performance may be through the machine's internal states (z), whose values co-determine its input-output relation (x, y). Moreover, the relationship between the present and subsequent internal states (z, z') is co-determined by the inputs (x). Perhaps the best way to visualize this is by seeing this arrangement as a machine in a machine (see Figure 5). From the outside such a machine looks very much like a trivial machine, with an input x and an output y. However, when the lid is taken off (as in Fig. 5), one can see now the entrails of an NTM. The novel feature here is the place (circle in the centre) that holds the internal state z. This state, together with the input x,

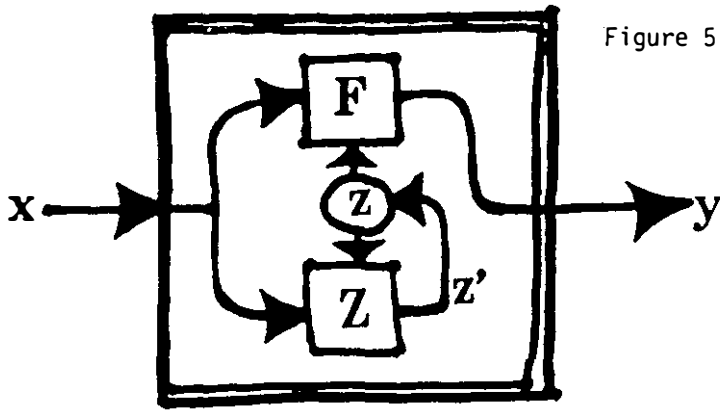


Figure 5: Non-Trivial Machine.

furnish an input - on the one hand - to **F**, a trivial machine computing the NTM's output y , and - on the other hand - to **Z**, another trivial machine computing the subsequent internal state z' . From this it should be clear that the non-trivial machine too is synthetically deterministic.

I will have such a machine running for you in a moment, but would like first to get some terminology out of our way. **F** and **Z** are usually referred to as the Driving Function and the State Function respectively. Algebraically this is expressed by

$$\begin{aligned} y &= \mathbf{F}(x, z), & \text{Driving Function} \\ z' &= \mathbf{Z}(x, z), & \text{State Function.} \end{aligned}$$

Perhaps you noted that the state function **Z** expresses a quantity (z') through itself at an earlier stage (z). This is the essence of recursive computations. I shall talk about these in point number (iii).

Let us construct now a minimal NTM, as closely as possible related to our TM of before. A minimal extension would be to add simply one internal state to that machine so that we have now instead of only one, two internal states. Let them be called I and II, and have the driving and state functions as follows:

When in I			When in II		
x	y	z'	x	y	z'
A	0	I	A	1	I
U	1	I	U	0	II
S	1	II	S	0	I
T	0	II	T	1	II

Now, let us explore the behaviour of this machine. I suggest testing first with the first input symbol A. We present the machine with several A's (A, A, A, ...), and to our satisfaction we get consistently zeros (0, 0, 0, ...). We turn now to a sequence of U's (U, U, U, ...), to which the machine responds with a sequence of ones (1, 1, 1, ...). Confidently we try the input S and obtain 1; but when checking out S again, for one who does not know the inner workings of the machine, something unpleasant is happening: instead of a 1, the machine responds with a 0. We could have predicted that, because the state function switches the machine when in I, given S, into its internal state II, and here the response to stimulus "S" is "0". However, being in II, given S, the machine returns to internal state I, and a new test of S will yield 1, etc., etc., ...

Checking out the patriotic sequence USA, depending upon whether one starts when the machine is in its internal state I or in II, it will respond with either 111,

or else with 000; apparently indicating different political persuasions. Perhaps these examples suffice to justify the qualifier "non-trivial" for these machines.

More important, however, is to see the distinction between the one who knows the driving and state functions of the machine (may be he did the synthesizing), and the other one who has no access to this knowledge and is restricted to observing sequences of input/output pairs as his only base for hypothesizing about the inner workings of this machine.

At first glance, the distinction between the knower and the experimenter may appear to be not too severe. Clearly, the experimenter has the boring task of going through all these sequences to establish the rules that produce them; nevertheless, ultimately he should be able to crack the code of these machines, and their workings will become as transparent for him as for the knower: cumbersome, but possible.

Alas, this is not so.

Let me first turn to "cumbersome". The problem here is to identify among all possible machines with the given number of input and output states the one under investigation. By "identifying" is, of course, meant to infer from the observed sequences of input/output pairs the machine's driving and state functions.

Table T

The Number of Effective Internal States Z , the Number of Possible Driving Functions \mathfrak{R}_D , and the Number of Effective State Functions \mathfrak{R}_S for Machines with One Two-Valued Output and with from One to Four Two-Valued Inputs

n	Z	\mathfrak{R}_D	\mathfrak{R}_S
1	4	256	65536
2	16	$2 \cdot 10^{10}$	$6 \cdot 10^{76}$
3	256	10^{100}	$300 \cdot 10^{10^3}$
4	65536	$300 \cdot 10^{10^3}$	$1600 \cdot 10^{7 \cdot 10^4}$

In table T I have listed the numbers of the possible non-trivial machines with exactly two output states, say 0, 1, as is the case with ours, and with 2, 4, 8, 16, input states ($n = 1, 2, 3, 4$). Our machine has four input states A, U, S, T, ($n = 2$), hence our experimenter must search amongst

$$6 \cdot 10^{76}$$

different machines to find the right one. Cumbersome? No! Transcomputational!

Now to "possible". There exists a large class of machines whose driving and state functions are such that it is *in principle* impossible to infer these functions from the results of a finite number of tests: the general machine identification problem is unsolvable! This also means that there are non-trivial machines that are unknowable.

I shall summarize now the essential features of non-trivial machines, and then conclude with a few comments. In parallel to what I have said earlier about trivial

machines, one can say that all NTM's are:

- (i) synthetically deterministic;
- (ii) history dependent;
- (iii) analytically indeterminable;
- (iiii) analytically unpredictable.

With the principle expressed in (iii) the non-trivial machines join their famous sisters who sing of other limitations:

Gödel: Incompleteness Theorem;
Heisenberg: Uncertainty Principle;
Gill: Indeterminacy Principle.

If one also takes the other unpleasanties of these machines into account, namely, the dependence on their past and their unpredictability, our efforts to remove or suppress all uncertainties in our environment are quite understandable. When we buy a machine we want it to function exactly as intended. When we turn the starter key in our car, it should start; when we dial a telephone number, we want the right connection, etc., etc., we want trivial machines. Hence, we like those guarantees that, in essence, are saying: "... at least for one year this machine will remain a trivial machine." If, in spite of this, it shows non-trivial tendencies (the car won't start, etc.) we call upon a specialist in trivialization who remedies the situation.

This is all very well. However, when we begin to trivialize one another, we shall soon not only be going blind, we shall also become blind to our blindness. Mutual trivialization reduces the number of choices, hence goes counter to the ethical imperative I voiced in the beginning. The task at hand is:

de-trivialization.

3. Recursive Computations

Is the world a trivial or a non-trivial machine? Perhaps Einstein had an answer for this question when he said: "Raffiniert ist der Herrgott, aber boshaft ist er nicht" (Subtle is the Lord, but malicious He is not (15)). And Heisenberg: what would his answer have been after he saw that the interference of an observation leaves the observed in a state of uncertainty. Or should we switch his principle around and say more accurately that the interference of an observation leaves the *observer* in a state of uncertainty?

May be the original question contains an implicit flaw by stipulating a dichotomy between a world observed, and one who makes the observations. Perhaps each of us has first to answer for himself the question: "Am I a part of the Universe, or are we both apart." In other words, should we contemplate an epistemology in which I, the observer, am included in the domain of my observations, or shall we prohibit this re-entry (for ultimately we may see ourselves!).

Since the orthodox position here is to stipulate the separation of the observer from the world observed, a world usually perceived as a trivial machine whose function we are to uncover. Since this perspective is almost all-pervasive, I need not address myself to it.

Instead, I shall expand on the concepts of autology and closure of before, by making full use of the notions of "machines" whose behaviour under closure we are now to explore.

Consider an arbitrary large network of interacting NTMs that are fully connected to each other. By this is meant that each machine's output is an input for some others (or for itself); and each machine's input is an output from some others (or from itself) (see Figure 6a). Since there is no lead to the world outside of this network, this system is closed, it is its own world. Ross Ashby, who was one of the first to study the activity of such nets, referred to them as "systems without input" (16).

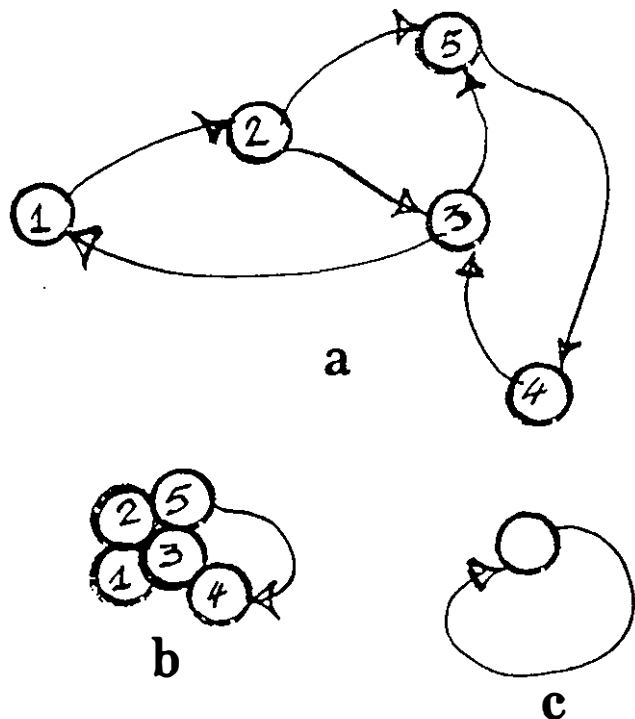


Figure 6: Network of Interacting Machines .

If we were to grab one of the connections between any two machines to observe the signal flow between them, it is irrelevant to how many more they are connected (6b): the whole net acts as a single NTM whose output is its input (6c).

Let us consolidate the operation of the entire net between the chosen points of in- and output into one operator

O_p ,

and let the result of this operation become the beginning of its next operation. In other words, let this become a recursive operation.

At this point I have struggled with myself whether I shall let you go through the paces of an elementary formal approach to recursive function theory, or whether I should make a short cut by just summarizing some results. Since I could not make up my mind, I decided to do both, for you can always skip various steps in the formal arguments, and turn to the summary. Nevertheless, I recommend you come along with me over the four points of this Primer on Recursions, because you will enjoy the consequences of the argument much more after having watched their development.

A Primer on Recursions

Elements of a Formalism.

1 Consider the (independent) variable x_0 (call it the "primary argument", being subscripted with "0" to indicate that this is the variable taken ab ovo).

1.1 As the case may be, this variable may assume numerical values, or it may represent arrangements (e.g., arrays of numbers, vectors, geometrical configurations, etc.); functions (e.g., polynomials, algebraic functions, etc.); behaviours described by mathematical functions (e.g., equations of motion, etc.); behaviours described by propositions (e.g., the McCulloch-Pitts (TRE's temporal propositional expressions) etc.).

2 Consider an operation (transformation, algorithm, functional, etc.):

"Op"

acting upon the variable x_0 ; indicate the action on this operand x_0 by

$$Op(x_0)$$

Call x_1 the values generated by the first application of Op on x_0

$$x_1 = Op(x_0) \tag{1}$$

or graphically

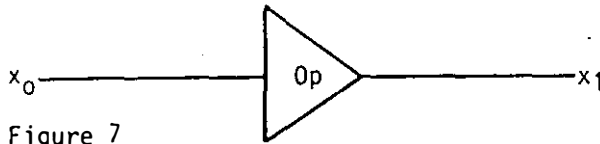


Figure 7

2.1 Apply Op to x_1 , and call x_2 the values so generated:

$$x_2 = Op(x_1), \tag{2}$$

that is, x_2 represents the values generated by having Op applied twice to x_0 . (With Equ. (1) and (2)):

$$x_2 = Op(x_1) = Op(Op(x_0)). \tag{3}$$

2.2 Call $Op^{(n)}$ the n-th application of Op to a variable, then:

$$x_n = Op^{(n)}(x_0), \tag{4}$$

or graphically

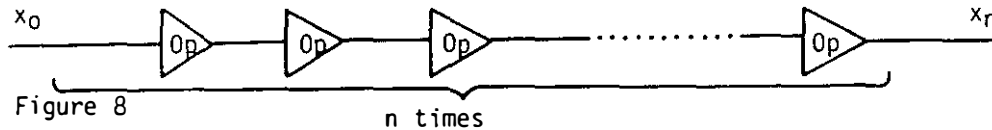


Figure 8

3 Consider the case in which Op is applied indefinitely ($n \rightarrow \infty$) to a variable, say, x_0 :

$$x_\infty = Op^{(\infty)}(x_0) \quad \text{or} \tag{5}$$

$$x_\infty = Op(Op(Op(Op(Op(Op(Op(Op(Op(Op(Op(Op(Op(Op(\dots$$

3.1 Contemplate expression (6) and observe:

3.11 That the independent variable x_0 , the "primary argument" has disappeared;

3.12 That, since x_∞ expresses an indefinite recursion of the operators Op onto operators Op, any indefinite recursion within that expression can be replaced by x_∞ :

$x_0 =$ 1, 2, 3, 4, 5, 6, 7, 8, 9, 0
 $x_1 =$ 1, 0, 6, 4, 5, 6, 7, 8, 9, 0
 $x_2 =$ 1, 0, 3, 4, 5, 6, 7, 8, 9, 0
 $x_3 =$ 1, 0, 2, 4, 5, 6, 1, 8, 9, 0
 $x_4 =$ 1, 0, 2, 4, 4, 6, 1, 0, 9, 0
 $x_5 =$ 1, 0, 2, 0, 4, 6, 4, 0, 9, 0
 (observe the vanishing of odds)
 $x_6 =$ 1, 0, 2, 0, 4, 6, 4, 0, 0, 0
 $x_7 =$ 1, 0, 2, 0, 4, 0, 4, 0, 0, 0
 (observe the emergence of naughts)
 $x_8 =$ 1, 0, 2, 0, 4, 0, 4, 0, 0, 0
:
 $x_{15} =$ 0, 0, 2, 0, 4, 0, 0, 0, 0, 0
:
 $x_{\infty} =$ 0, 0, 0, 0, 0, 0, 0, 0, 0, 0 = E_1

2.) Consider the array

1, 2, 3, 4, 5, 6, 7, 8, 9, 0 .

Apply to it Ashby's "Co-Evolutionary Operator" CE:

CE = "Choose two numbers, α, β , at random; change β to the last digit of

$$\alpha^4 + \beta^4;$$

leave α unchanged.

From the following Table, which lists these last digits for each of the pairs, α, β , one may convince oneself that the Eigen-Arrays either contain 2's and 7's in equal numbers, or else 2's only. (Note that in case the 2's disappear completely, they will be regenerated through the 7's. The converse is not true).

Table

	1	2	3	4	5	6	7	8	9	0
1	2	7	2	7	6	7	2	7	2	1
2		2	7	2	1	2	7	2	7	6
3			2	7	6	7	2	7	2	1
4				7	1	2	7	2	7	6
5					6	1	6	1	6	5
6						2	7	2	7	6
7							2	7	2	1
8								2	7	6
9									2	1

INITIAL Y=3

-0.9899924293	0.6916683255	0.7681274735	0.6948208121
-0.8360218258	0.7701829943	0.6947897149	0.7681687568
0.6704198624	0.6962666018	0.7681883513	0.6948194033
0.6213150305	0.7672419786	0.6948334981	0.7681693438
0.3131136789	0.6941525818	0.7681603226	0.6948198267
0.7264305416	0.7685961014	0.6948133393	0.7681690722
0.7475500224	0.6951266802	0.7681732221	0.6948196347
0.6798440992	0.7679725702	0.6948226158	0.7681692011
0.777670743	0.6946783	0.7681672922	0.6948197269
0.701621614	0.7682593786	0.6948183524	0.7681691408
0.7637965103	0.6948847942	0.7681700157	0.6948196821

I hope that with this brief description of some points in recursive function theory, and with the few examples of its application, you could, at least, get a flavour of this method, and could see in the recursive operation a principle of self-organization that allows certain structures to emerge - to crystallize - from early, arbitrary stages. However, many other interesting results I have not mentioned, results involving multiple eigenvalues, compositions of such states, and many more. Moreover, examples in which the eigenstates are not numerical quantities, but are themselves operators (Eigen-Operators), or of other domains would have been illuminating. This would, however, require a much more elaborate formal apparatus, and for the study of such cases I have to refer to the literature (11) (17) (18).

Nevertheless, I cannot leave this account without a short note on the results of Ashby's studies of the dynamics of large systems without input. In a computer simulation of an arrangement as in Fig. 6a, Ashby connected in one series of experiments 100, in another 1,000 non-trivial machines (essentially computing on their inputs a variety of logical functions), and after setting them at a initial value let them loose. After some transients in the beginning (see also our examples) the systems settled into various eigenbehaviours, i.e. "limit cycles", of different length, in many cases representing large domains of initial values. Polystability was his term for this phenomenon (16). His studies have recently been revived with much faster and larger computers by a French group leading to many new and fascinating results (19).

I shall now conclude my story on recursive computation with a few words on terminology. As I mentioned before, it was David Hilbert who, around the turn of the century, introduced the terms Eigen-value and Eigen-function, terms I find particularly well chosen for representing the logic that is involved here. Somewhat later, another attractive feature of these values, namely the invariance under their corresponding operations, brought them the name "fix points". And recently, some computer buffs discovered these fascinating values for themselves, and since they could not believe their eyes when they saw what they saw, they called these values "strange attractors", a term, I am sorry to say, I find repulsive.

4. Socio-Managerial

Malik and Probst in their article on evolutionary management looked, of course, upon the role of negotiation from the perspective of the firm as an evolving, self-organizing system. I would like to supplement their observations with points that emerge from the strategies I just reported.

I propose to look for the moment at negotiation as an attempt by members of a group to "solve a common problem". The quotation marks here I intended to be acting as

flags, as caution signs, inviting us to re-examine the over-used and over-abused terms so quoted. What is meant by "solve", by "common", by "problem"? Most likely, there is no common problem! each one of the members may have his own; worse, may be he does not *have* a problem, perhaps *he* is the problem; etc.

With this warning in the back of our minds, I propose again to look at negotiations as a problem solving task, where one of the solutions may indeed be the identification of a "common problem".

Small Group Dynamics

I will describe now one of the early experiments in small group dynamics, an experiment which is to my taste much too little known for the many interesting conclusions one can draw from its results. This experiment was designed in the early 'fifties by Alex Bavelas (20), then at M.I.T., who became interested in the evolution of strategies and feelings of people with different expertise, who participate in various tasks whose ends and means are given in terms that span from transparency to opacity. There seems to be a similarity here with the situations in which the Principle of Potential Command may have its application. This, however, is not the case, for here intended actions are controlled in a way, as we shall see in a moment (and for the record monitored).

The task given to the members of a group of five is to find the only common symbol in a deck of cards, of which each member has only one card to look at, but can communicate with others, exclusively through prescribed channels, to get the needed information for the other symbols. Let me first describe the deck of cards, and then the spatial arrangement of this experiment.

Cards: Consider six different symbols, say, a square, a cross, a triangle, etc., which I conveniently will label 1,2,3,4,5,6,. Design six different cards, each with one symbol missing, but showing all other five:

2	1	1	1	1	1
3	3	2	2	2	2
4	4	4	3	3	3
5	5	5	5	4	4
6	6	6	6	6	5

1 2 3 4 5 6

The last line of this schema indicates the missing symbol.

It is clear that by removing from this set one card that lacks, say, symbol 3, the remaining cards of the set will have one, and only one, symbol in common, namely 3.

It is also clear that in this way six sets, or decks, of cards can be generated, each deck distinct from the others by its common symbol.

In a preliminary "get at ease" session each prospective participant is given such a set with the question to identify the common symbol. That takes between 20 and 40 seconds to answer. At that time he is also told that in the actual experimental situation he will see only one card, and has to infer the common symbol through interactions with other members of the group.

Space: Consider two concentric pentagonal cylinders, where the space in between them is subdivided into five identical compartments, each of which can seat com-

fortably one of the five participants. In front of the wall facing the centre is a wide and shallow desk. In the wall above the desk are slots, some open, some closed, through which messages can be sent to, or received from, other participants via tubes that are concealed behind the walls. Communication through these tubes is the only way participants are able to interact, sound proof walls, etc., restricting other means.

Two crucial points in the design of the experiment are (i) the possibility of specifying beforehand (unknown to the participants) the connectivity between compartments, for instance, those of Figures 9; and (ii), the possibility of keeping track of the communication process through numbered and colour coded message pads and pencils.

Experiment: A single session begins with the five participants seated in their compartments and facing one card in the set of five. They may use their note pads for any message, a question, an answer, whatever, being sent to others. As soon as each participant thinks he knows the common symbol, he presses the appropriate key, one of six on his desk. The session ends when all participants have pressed the same key.

Results: Although the experiments yielded a large crop of results, I shall talk only about two kinds of variation within this overall design. One regarding connectivity, the other different symbols. Variation in connectivity, say, from "circle" (Figure 9a) to "star" (Figure 9c), produces changes in performance that are already quite impressive. When varying symbols, the changes were dramatic. In one set of experiments identifiable symbols were used in groups connected differently. In the other set "noise" was introduced into the communication channels - or should I say into the channels of cognition - by using "symbols" that are not only difficult to distinguish, they don't even have names: here sets of differently mottled marbles were used in lieu of symbols.

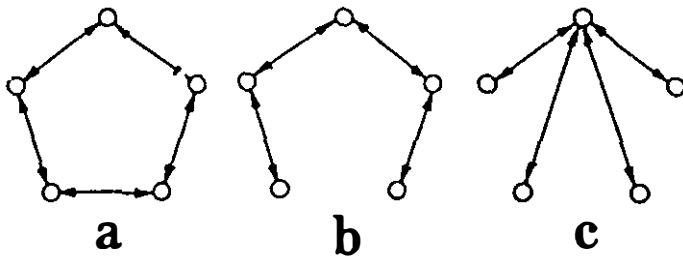


Figure 9: Various Connected Groups

Let me report first about the "noiseless" case, i.e. the one in which the symbols were identified and named. When those who played the "circle" were asked how they felt during the session, how they perceived their performance, etc., they consistently replied that they were feeling fine, that they performed fast and efficient, that they may have done better, etc. When asked whether they could identify a "leader" in the group, the averaging of replies distributed this "leadership" evenly over all five positions.

For the "star" performers the story is almost the opposite of the "circlers". Although their performance is about twice as fast than that of the circle groups, they had the feeling of defeat. They sensed themselves as slow and inefficient. They blame some "idiot" in the team for that. Of the participants 94 % identify the apex of the connectivity to be the leader.

Because of the difference of perceiving the absence or the presence of a leadership, Bavelas and his co-workers nicknamed these two connection schemes "democratic" and "authoritarian".

What happens now when "noise" is introduced? Surprisingly (or perhaps not surprising at all), the democratic group works just as well, however somewhat slower than before. They still feel fine and think they are doing well. The dramatic change is with the authoritarians: depending on the "strangeness" of the symbols, the groups disintegrate sooner or later. Participants walk out in anger, the "idiots" multiply, and blame is passed from one to others. Indeed, when later the communication records are studied, the star performers soon stop talking about symbols, they start calling each other names. There is a fascinating switch from an attention to communicabilia to that of the communicators.

The difference to "democrats" is fundamental. Expanding language is what keeps the people in this configuration going. As the records show, names for the funny looking things are soon invented, some referential, "lion", "cow", etc., or new ones "splops", "bimbim", etc., names that are either kicked around, modified, or kept; and when adopted by the group, the find-the-common-symbol problem is back to finding common *symbols*, by ignoring fuzzy objects.

I went through these exercises at a somewhat greater length, for I felt that these experiments are superbly suited for connecting the four notions, management, self-organization, evolution, and language, very much in the sense of Malik and Probst admonished us in their paper.

There is no doubt that one of the managerial tasks is to generate a climate that fosters communication. One of the outcomes of the Bavelas experiments suggest that interaction structures can be facilitators or inhibitors for communication. It appears that circulatory, recursive interaction patterns are highly stable against perturbations. The important point, however, is that this stability is not through counteracting the perturbing forces, but by utilizing them as a wellspring of creation. And finally, these experiments show again the significance of language in the managerial process (21).

My friend, the composer Herbert Brun, once taught me "...a language learned is a language lost" (22). But here, in one of Bavelas' situations, are instances of language in the make.

Lexically speaking, language is a closed system: ask for the meaning of a word, and you get words. I want to know the meaning of "subsequent". The dictionary (1) tells me "following" (see Figure 10). I want to know what this means, etc. ... Figure 10 tells where this all leads to; one may say to nowhere. Can one get out of this trap?

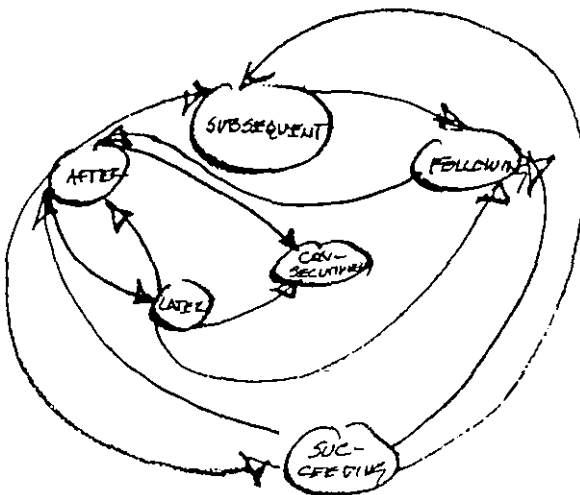


Figure 10: Relational Network of Synonymous Terms.

I suggest one path that was seen by the British philosopher John Langshaw Austin. He observed in our language a peculiar family of utterances that say what they do; or perhaps I should say, they do what they say. Now, how is that?

Imagine yourself in a crowded bus; inadvertently you step on somebody's toes. Politely you say:

"I apologize".

The magic of this utterance is that it is the apology. For obvious reasons, Austin called these utterances "performative" (23). Once one is aware of these utterances in our language, one sees them appearing more and more: "I promise", "I declare" ..., etc. Contemplate for a moment the extraordinary things that are going on when in a marriage ceremony the priest says:

"I declare you husband and wife".

When this formula is uttered, they *are* husband and wife.

The notion of the performative utterance I brought up at the end of my story, for - in proper recursive fashion - it ties me back to the beginning. You may remember the sentence that says of itself how many letters it has. When indeed it says so correctly, we called it an Eigen-value. May be, one should call it an Eigen-utterance, to let the connection with performative utterances become visible. Here, I suggest, is the window through which we can step outside of language. Hence, let me conclude with a reference to your kindness for having invited me, and to your patience in listening to me, in form of a performative utterance:

"Thank you very much!"

References

- (1) The American Heritage Dictionary of the English Language, Houghton Mifflin, Boston (1966).
- (2) Malik, F. and G.J.B. Probst: "Evolutionary Management" Cybernetics and Systems: Int. J., 13, 153-174 (1982).
- (3) von Foerster, H.: "On Constructing a Reality" in Observing Systems, A Collection of Papers by Heinz von Foerster, Intersystems Publications, Seaside (1982).
- (4) Gunther, G.: "Cybernetic Ontology and Transjunctional Operations" in Beiträge zur Grundlegung einer operationsfähigen Dialektik, I, Gotthard Gunthers gesammelte Werke, Felix Meiner Verlag, Hamburg (1976).
- (5) Personal communication from Professor Lars Lofgren, Dept. for Automata Theory and General Systems, Building E, University of Lund, Box 725, S-220 07 LUND, Sweden.
- (6) Varela, F.G.J.: "A Calculus of Self-Reference" Int. J. General Systems, 2, 5-24 (1975).
- (7) Psk. G.: "The Meaning of Cybernetics in the Behavioural Sciences" in Progress in Cybernetics, J. Rose (ed.), Gordon and Breach, New York, 1, 15-44 (1969).

- (8) von Foerster, H.: "Foreword" in Rigor and Imagination, Essays from the Legacy of Gregory Bateson, C. Wilder-Mott and John H. Weakland (eds.) Praeger, New York (1981).
- (9) Wittgenstein, L.: Philosophical Investigations, G.E.M. Anscombe (tr), The Macmillan Company, New York (1953).
- (10) Hofstadter, D.R.: "Metamagical Themas" Sc.Am.Jan., 12-332 (1981), and Jan., 16-28 (1982).
- (11) Davis, M.: Computability and Unsolvability, McGraw-Hill, New York (1958).
- (12) McCulloch, W.S.: "A Heterarchy of Values Determined by the Topology of Nervous Nets" in Embodiments of Mind, MIT Press, Cambridge (1965).
- (13) Turing, A.M.: "On Computable Numbers, with an Application to the Entscheidungsproblem" Proc. London Math. Soc., ser. 2, 42, 230-265 (1936-1937).
- (14) Gill, A.: Introduction to the Theory of Finite State Machines, McGraw-Hill, New York (1962).
- (15) Pais, A.: "'Subtle is the Lord ...' The science and the Life of Albert Einstein", Oxford University Press, New York (1982).
- (16) Walker, C.C./Ashby, W.R.: "On Temporal Characteristics of Behavior in Certain Complex Systems" Kybernetik, 3 (2), 100-108 (1966).
- (17) von Foerster, H.: "Objects: Tokens for (Eigen-)Behaviors" in Observing Systems (see Ref. 3.).
- (18) Hofstadter, D.R.: "Metamagical Themas" Sc.Am. Nov., 22-43 (1981).
- (19) Fogelman-Soulie, F., Goles-Chacc, F. and G. Wissbuch: "Specific Roles of the Different Boolean Mappings in Random Networks" Bull. Math. Biol., 44 (5), 715-730 (1982).
- (20) Bavelas, A.: "Communication Patterns in Problem-Solving Groups" in Cybernetics, Heinz von Foerster (ed.), Josiah Macy Jr. Foundation, New York (1952).
- (21) Request (and indeed request!) literature from: Hermenet, 1750 Union Street, San Francisco, California 94123; Attention Dr. F.C. Flores.
- (22) Brun, H.: "Futility 1964" (side 5, band 3) in Compositions by Herbert Brun, Non Sequitur Records, Box 872, Champaign, IL 61820 (1983).
- (23) Austin, J.L.: "Performative Utterances" in J.L. Austin: Philosophical Papers. J.O. Urmson and G.J. Warnock (eds.), At the Clarendon Press, Oxford (1961).