SELF-ORGANIZATION OF COMPLEX STRUCTURES
From Individual to Collective Dynamics

Edited by
Frank Schweitzer
Humboldt University, Berlin

Foreword by
Hermann Haken

GORDON AND BREACH SCIENCE PUBLISHERS
Australia • Canada • China • France • Germany • India • Japan • Luxembourg
Malaysia • The Netherlands • Russia • Singapore • Switzerland • Thailand
United Kingdom
1. TRUTH AND CERTITUDE IN THE
SCIENTIFIC LANGUAGE

F.T. ARECCHI

1. INTRODUCTION

A wealth of speculations has recently appeared on (A) "Complexity" and (B) "Complex Systems".

The first term has been differently defined in formal languages [1], computer science [2] and nonlinear signal analysis [3], starting in the early 80's with the intrinsic non predictability associated with chaotic time series [4].

(A) is associated with gnoseological, or epistemic, processes. Once a time series of data, coded in a given alphabet, has been assigned, can one retrieve the meaning of the message just by perusal of that sequence? Is "meaning" just (i) discovering the grammatical rules which allow some symbol sequences (words) and forbid some other ones, or also (ii) attributing a likelihood of occurrence to each word and hence attempting predictions about the future of the time series?

This (A) approach has received different formulations, with different solutions leading to automatic procedures on complexity assignment [5–9].

On the other hand, natural scientists have rather focussed on the ontological fact that reality uncovers lots of complex structures. Even prior to any encoding into some alphabet, and a consequent mathematical elaboration, any holistic Gestalt appreciation shows what we mean by a complex system, independently from any quantitative indicator.

Two different tasks are then faced by investigators of (A) and (B) problems, namely,

(A) given an input, what is the optimal use we can make of it? We call "certitude" the subjective confidence that we have done the best in grasping the inner rules of the input;

(B) facing a piece of world, can we express our knowledge of it in a suitable language, i.e. encode phenomena into symbol sequences from some alphabet (which later will be analyzed as in (A))?

As we see, (B) is "prior" to (A). It appears as the problem that any living being has to solve, and it is usually faced by adaptive strategies, which of course we can later formalize as linguistic procedures, but which arise at a pre-linguistic level and even determine the same choice of the most appropriate language [10]. This more fundamental problem is that of "Truth", defined as "ratioaequatio intellectus et rei" that is "adjustment of our expectations to the changing world".

The two approaches are altogether different. In case (A) a learning machine can be foreseen which automatizes the quest for complexity [7]. On the other hand, case (B) hints at the crucial role of a pre-linguistic stage where we still have to decide how to encode the stream of perceived phenomena into a linguistic sequence, which is then exposed to the inquiry of the complexity machine (A).
We do not want here even to tackle the wealth of philosophical problems involved by this distinction. We just limit to saying that doing physics is (B), that is, making it possible to encode our perceptions into a suitable language, not just building theoretical models to uncover rules and make predictions with regard to given sequences as (A).

Our main conclusion is that while there may be a complexity machine for (A), it is in principle impossible to introduce a science machine for (B). Hence (B) remains a human endeavour not reducible to automatic procedures.

The paper is organized as follows.

Section 2 reviews current definitions of complexity showing the virtues and intrinsic limitations of a contextual analysis of a data stream (by “contextual” we mean that we must rely on correlations and symmetries already built in the symbol sequence without the power to modify it). Section 3 is a historical survey of the birth of modern science seen as the emergence of a formalized language out of pre-scientific observations already expressed in the ordinary language. Section 4 is a dynamic approach to a complex situation. While in Section 2 no apricri rules are requested but a data stream is preliminary, in Section 4 we take for granted the general rules of dynamics and prospect a variety of possible data streams, that is, a variety of different physical situations. This variety appears as a natural implementation of the intuitive concept of complexity, as it is representative of what we call complex systems, from economy and sociology to biology and physics.

In Section 5 we present an adaptive strategy recently introduced to recognize [11] and control [12] a chaotic dynamics. While an adaptive procedure was already incorporated in the complexity machine of Section 2, in order to fit the theory to the input sequence, here, in a more radical way, we change the same structure of the measuring apparatus M in order to provide different sets of measurement sequences to be later analyzed.

We conclude in Section 6 with some epistemological insights, recalling that a knowledge program based on assigned input data is how to make the best use of our mental representations according to a subjectivistic gnoseology started by Descartes and continued by Hume and Kant. On the contrary, a knowledge based on readjustment of our measuring procedures appears as the most natural attitude of living beings. It seems more appropriate to the psychology of cognition as described in classical philosophy (Aristotle, Thomas Aquinas) and as recovered by many naturalistic approaches, under the name of “evolutionary epistemology” [13].

2. COMPLEXITY OF SYMBOLIC SEQUENCES

In computer science, we may define as complexity of a word (symbol sequence) some indicator of the cost implied in generating that sequence. There is a “space” cost (length of the instructions stored in the computer memory) and a “time” cost (the CPU time for generating the final result out of some initial instruction).

A space complexity \( C_1 \) [2] is defined as the length in bit of the minimal instruction which generates the wanted sequence. This indicator is maximum for a random number, since there is no compressed algorithm (that is, shorter than the number itself) to construct a random number.

A time complexity \( C_2 \), called “logical depth” (Bennett in Ref. 5a) is defined as the CPU time required to generate the sequence out of the minimal instruction. \( C_2 \) is minimal for a random: PRINT IT
Of course, problem, b
An exam
While C properties to a process to arrive at a class of objects assigned:
In fact, natural science:
(i) to try a way, good
(ii) to plan forth
For this complex
In view of essential e may just be in format.

We thus
(i) as we; we c; have tryin;
Secti
(ii) As of to O:
of a t
The me:
\( D: n \)
\( e: r: \)
\( r: t: \)
\( \beta = \)
of ea speci
a random number, indeed, once the instruction has stored all the digits, just command:
PRINT IT

Of course, for simple dynamical systems as a pendulum or the Newtonian two-body problem, both complexities are minimal.

An example of a large $C_2$ is offered by Wolfram cellular automaton n.86 [14].

While $C_1$ refers to the process of building a single item, $C_2$ corresponds to finding the properties of all possible outputs from a known source. Following Simon [10], $C_1$ refers to a process description and $C_2$ to a state description. Indeed $C_1$ corresponds to the effort to arrive at a specific object, and $C_2$ corresponds to the mental representation of the whole class of objects, that is, for n.86, to all those site states which are necessary for the central site assignment after $N$ time steps.

In fact, the exact specification of the final outcome is too much for the ambition of the natural scientist, whose goal is more modest. It may be condensed in the two following items:

(i) to transmit some information, coded in a symbol sequence, to a receiver in a compact way, possibly economizing with respect to the actual string length, that is, making good use of the redundancies (this requires a preliminary study of the language style);

(ii) to predict a given span of future, that is, to assign with some likelihood a group of forthcoming symbols;

For this second goal, introduction of a probability measure is crucial [6] in order to design a complexity-machine, able to make the best informational use of a given data set.

In view of the difference between (A) and (B) introduced in Section 1 let us sketch the essential elements of knowledge building in natural sciences. First of all, we realize (it may just be a Gedankenexperiment) a device, or measuring apparatus $M$, whose output is informationally equivalent to what is going on in the observed piece of world.

We thus attribute to knowledge two different meanings:

(i) as we face a phenomenon, $M$ captures (presumably) the relevant aspects of it, so that we can transfer sufficient information, either to another partner or to ourselves if we have to reflect in order to build a possible theory. Knowledge improvement implies trying with different $M$ apparatuses by a suitable program that we will explore in Section 5.

(ii) As observer $O_1$ is exposed to a symbol stream, it has to transmit a compact explanation to $O_2$, so that $O_2$ is able to retrieve the same input sequence. The explanation consists of a tentative theory that we call model $m$.

The measuring apparatus $M$ is characterized by the following elements [15]:

$D$: number of probes (dimensionality of the measurement space),

$\epsilon$: resolution in the projected state space (total number of cells is $\epsilon^{-D}$),

$\tau$: time resolution,

$\beta = 1/T$ ($T$ = noise temperature): fuzziness associated with a non sharp boundary of each resolution box, yielding some ambiguity in the assignment of an event to a specific cell of state space.
At any time slot of width \( r \), we extract \( e^{-D} \) different space data, that we can encode in a suitable one-dimensional string of symbols of an alphabet (e.g., binary).

The modeler \( O_1 \) isinput by some sequence \( s \), and it sends an explanation which should enable \( O_2 \) to reconstruct an output \( s' = s \).

Notice that \( M = M(D, e, \tau, \beta) \) is a whole class of possible instruments, and different individuals will give rise to different data sequences (different words).

The explanation [15] consists of a theoretical guess (model \( m \)) whose validity is tested by simulating an output and comparing it with the actual input data \( s \). The difference yields an error signal \( e \). Observer \( O_2 \) is provided with both \( m \) and \( e \) and it can reconstruct \( s' = s \) upon this information.

The virtue of the explanation \( X \) is to have a bit length \( \| x \| = \| m \| + \| e \| \) much shorter than the sequence length \( \| s \| \): this amounts to extract a relevant semantics out of the redundant features of \( s \).

The explanatory machine is complex in so far as it spans over a whole class of models \( m \). If one had access to a complete probabilistic description of the modeling universe, then the goal would be to maximize the probability of \( m \) conditional to the input \( s \)

\[
Pr(m|s)
\]

This ideal complete description is not available, but an approximation can be obtained by the Bayes' rule

\[
Pr(m|s) = \frac{Pr(s|m)P(m)}{\sum_{m'} P(s|m')}
\]

All these probabilities are conditioned on the choice of the model class. Furthermore, all terms on the right hand side refer to a single data stream \( s \). Here \( Pr(s|m) \) is the probability that a model \( m \) produces the given data. With sufficient effort, \( Pr(s|m) \) can be estimated. Finally the normalization in the denominator depends only on the given data and so can be dropped as a constant.

The most likely explanation corresponds to the shortest code of length \( \| x \| = \log_2 Pr(m|s) \).

There are two criteria for a good explanation:

(i) \( x \) must explain \( s \), that is, \( O_2 \) should resynthesize the original data \( s' = s \)

(ii) the bit length

\[
\| x \| = \| m \| + \| e \|
\]

must be minimized.

The efficiency of an explanation is given by the compression ratio

\[
C(m, s) = \frac{\| x \|}{\| s \|} = \frac{\| m \| + \| e \|}{\| s \|}
\]

\( C \) is a cost function. The optimal model minimizes this cost.

There are the error cl \( \| e \| = 0 \).

The recovery of transitio

A suitable \( G = \{ V, I \} \) of going fr

complexity

Accordi symbolic of the single il

class.

However, sequence, forthcoming every day k

strategy to

3. FROM

The word c

an object p

an infinitel

"semantic e

Usually

different ev

perusal of i

a histogram

histogram i

of available

is more app

of the mom

Thus, we feeding the

finitisic ap

as we reflect

Whence

attributed to

Indo-European

by a "filter

specification.

A discou

a wide river
There are two limit cases. When the model is trivial ($\|m\| = 0$), the entire data are on the error channel: $\|e\| = \|s\|$. On the contrary a tautological model $m = s$ has no error: $\|e\| = 0$.

The reconstruction of a state space out of an assigned time series and the assignment of transition probabilities among states is a task faced in many ways (see e.g. [5], [6]). A suitable topological machine can be foreseen [15] which amounts to a labeled graph $G = (V, E)$ made of vertices $v$ connected by labeled edges which assign a probability of going from vertex $v$ to $v'$ on observing a symbol $s$. Out of this labeled graph, suitable complexity indicators can be extracted.

According to the title of this section, here we have explored the complexity of a given symbolic sequence either from the computational point of view, aimed at reconstructing the single item, or from the probabilistic point of view, aimed at selecting a model within a class.

However preliminary to that, the problem arises of how we have obtained a given sequence, and this implies a critical analysis of the measurement apparatus. In the forthcoming section we show how measuring apparatuses are suitable formalizations of the every day knowledge expressed in the ordinary language, and we will provide an adaptive strategy to optimize the measurement performance in view of some specific goals.

3. FROM THE ORDINARY LANGUAGE TO THE SCIENTIFIC LANGUAGE

The word of the ordinary language is polysemic. In general it does not denote univocally an object per se, but rather the object embedded in different environments, which means an infinitely large variety of different situations, that we call "events", distributed on a "semantic space" (Figure 1a).

Usually the same word can be attributed with different degrees of appropriateness to different events. The different attributions have a different probability, as it emerges from a perusal of the historical dictionary of a given language, where the probabilities are given as a histogram of frequencies of occurrence in the literary texts of that language. In fact, the histogram is a finitistic approximation, a kind of coarse graining, due to the limited number of available texts. If however we consider the everyday use, the continuous probability curve is more appropriate, since the environment includes the observer with his (her) own mood of the moment, which nobody would dare to reduce to a countable number of states.

Thus, while an artificial cognitive agent (a collection of detectors with fixed resolution feeding the input of a universal computer) would extract a histogram, thus justifying a finitistic approach, instead finitism seems excluded from the human everyday experience, as we reflect on the variety of nuances which qualifies a poem, or even a private conversation.

Whence the problem of interpretation, that is, of what is the right meaning to be attributed to a word, within the wide support subtended by its probability distribution. In Indo-European languages, a quasi-univocal, or narrow range space of meaning is obtained by a "filtering" operation, applying to the word a sufficient number of attributions or specifications, as sketched in the figure.

A discourse, seen as a flow of different words connected by grammar rules, appears as a wide river-bed within which everybody can cut out a different interpretation (Figure 1b).
Figure 1  (a) Polysemic word as a probability distribution over a semantic space. Further attributions narrow its semantic range ("oligo-semic" word). (b) A discourse, as the connection of successive words, yields a wider riverbed allowing for different interpretations.

Figure 2  membership between u

As well k sources of term.

Such a philology through 1 measuring to quantify apparentl away all
As well known, there is no unique sense of a given discourse, but one must recur to other sources of information, besides the text itself, in order to narrow the semantic range of each term.

Such ambiguities of the ordinary language were discussed at length by many Renaissance philologists, and were well known to Galileo. He provided [16] a way out of the ambiguities through his suggestion of "naming" an event via the number extracted by a suitable measuring apparatus M applied to the event itself (Figure 2a), that is, limiting himself to quantitative "affections" rather than attempting to grasp the "essence". This procedure apparently provided univocal meanings, since it filtered out a single denotation, clearing away all the context. As an example, a physicist does not speak of a "table" but of the
“weight” or the “length” of the table. In this new language, the syntactical rule connecting two words becomes a mathematical relation connecting the output numbers from two measuring apparatuses related to two “objects”.

This provides a solid framework for any scientific description, in terms of well established existence and uniqueness theorems. Thus, the flow of a scientific discourse consists of sharp, necessary connections among point-like objects of different semantic spaces, corresponding to different measurements, as shown by the solid line in Figure 2b. That solid line seems a great progress as compared to the wide river-bed of Figure 1b. It means that the scientific language is free from interpretational ambiguities.

The most crucial aspect of Galileo’s self-limitation is the apparent arbitrariness in placing $M$ over the semantic space, that is, the large number of different univocal words extractable for a class of events usually denoted by a single word of an ordinary language. This proliferation of $M$‘s has avoided for centuries the question of complexity. Suppose for instance (Figure 3) that by $M_1$ we look at the cell dynamics. Thereby, we build the specific language “cytology”, with its own words. We may now realize that many cells form an organ, but the organ is observed by a different $M_2$ which provides different words and hence a different science: “physiology”. Similarly if we look at the biomolecules we build a “biochemistry”.

We have thus obeyed Occam razor (Entia non sunt multiplicanda praeter necessitatem) in an economic way, that is, changing language whenever it is no longer appropriate.

![Diagram](image)

**Figure 4** of the small absence of

A more we can bu should det [17], occu philosoph: particles v self-limita only from the affecti human bel.

The fact the lower.

This wi relation w self-limita
A more strict obedience would consist in a reductionistic approach. As shown in Figure 4, we can build a hierarchy from large to small and say that the behaviour of smaller objects should determine that of larger ones. But here a perverse thing, already hinted by Anderson [17], occurs. If our word were a global description of the object in any situation (as the philosophical "essences" in Galileo's letter) then, of course, knowledge of elementary particles would be sufficient to make predictions on animals and society. In fact Galileo's self-limitation to some "affections" is sufficient for a limited description of the event, but only from a narrow point of view. Even though we believe that humans are made of atoms, the affections that we measure in atomic physics are insufficient to make predictions on human behaviour.

The fact that higher levels in the ladder of Figure 4 display features not predictable from the lower ones, is what colloquially we call complexity.

This way complexity is not a property of things (like being red or hot) but it is a relation with our status of knowledge, and for modern science it emerges from Galileo's self-limitation.
4. A DYNAMICAL APPROACH TO COMPLEXITY

A theory is successful if it is a "compressed" description of the world, that is, if the length of its initial assumption is much shorter than a detailed description of the events themselves. At the start, a physical theory is just mathematics. It becomes a model, that is, it acquires semantic values, whenever we interpret the objects of the theory as elements of reality [18].

Therefore a scientific theory must be considered a set of primitive concepts (defined by suitable measuring apparatus as M of Figure 2a) related by axioms. The deduction of all possible consequences (theorems) provides predictions which have to be compared with the observations. If the observations falsify the expectations, then one tries with different axioms.

The deductive process is affected by a Gödel undecidability like any formal theory, in the sense that it is possible to build a well formed statement, but the rules of deduction are unable to decide whether that statement is true or false.

Figure 5 Bifurcations tree in nonlinear dynamics. As a control parameter $\alpha$ is tuned through different values, novel steady states appear. By tuning $\alpha$ from $\alpha_0$ to $\alpha_N$, the system goes from 1 to $2^N$ different states. We call dynamical complexity such an ambiguity. It is responsible for the failure of reductionism. In the absence of external gradients, all final outcomes are equally probable. We call "organization" the occurrence of just one event out of $2^N$. This implies that at each $\alpha_i$ an external agent $A_i$ has broken the bifurcation symmetry (see Figure 6).

Figure 6 horizontal: steady state
Left: symm
Right: asym right and left

Besides increase of evolution, has to per type "if-th as one cha
Going b we find ar observed. to all bran final state
It is her branches c ingredient of interacti gradients l
Thus, dt ($\alpha$) are tu change the exponenti
Figure 6 Examples of bifurcation diagrams. The dynamical variable $x$ (order parameter) varies horizontally, the control parameter $\alpha$ varies vertically. Solid (dashed) lines represent stable (unstable) steady states as the control parameter is changed.

Left: symmetric bifurcation with equal probabilities for the two stable branches.
Right: asymmetric bifurcation in presence of an external field $A$. If the gap introduced by $A$ between right and left branch is wider than the range of thermal fluctuations at the transition point $\alpha_e$, then the right (left) branch has probability 1 (0).

Besides that, a second drawback is represented by intractability, that is, by the exponential increase of possible outcomes among which we have to select the final state of a dynamical evolution. Figure 5 sketches a bifurcation tree well familiar to computer scientists, as one has to perform a complex calculation with branch points implying multiple choices of the type "if-then". I rather consider it as the bifurcation tree of a complex nonlinear dynamics, as one changes a suitable control parameter $\alpha$.

Going back to the reductionistic tentative of explaining reality out of its constituents, then we find an exponentially high number of possible outcomes, when only one is in fact that observed. This means that, while the theory, that is the syntax, would give equal probability to all branches of the tree, in reality we observe an organization process, whereby only one final state has a high probability of occurrence.

It is here necessary to recall some descriptive elements on the bifurcation of the stable branches of a dynamics for different settings of a control parameter. These are the necessary ingredients of any complex dynamics. Notice that dynamical bifurcations in a system of interactive identical particles display specific symmetries (Figure 6a). Only external gradients break this symmetry (Figure 6b).

Thus, during the course of a dynamical evolution, either because some control parameters ($\alpha$) are tuned from the outside to assume different values, or because internal feedbacks change the $\{\alpha\}$ set in course of time, starting from some initial conditions we expect an exponential increase of final states.
Whenever there has been organization, this means that at each bifurcation vertex of Figure 5 the symmetry was broken by an agent external to the system under investigation. We can thus stipulate the following things:

(i) A set of control parameters

\[ \alpha_1, \alpha_2, \ldots, \alpha_N = \{a\} \]

is responsible for successive bifurcations leading to an exponentially high number (in the example, of the order of \(2^N\)) of final outcomes. If the system has no boundary effects (considered of infinite size) then all outcomes have comparable probabilities and we call complexity the impossibility of predicting which one is the state we will observe at the end of the chain of bifurcations.

(ii) A set of external forces

\[ A_1, A_2, \ldots, A_N = \{A\} \]

applied at each bifurcation point break the symmetries, biasing toward a specific choice and eventually leading to a unique final state.

We are in presence of a conflict between (i) syntax represented by the set of rules (axioms) \(\{\alpha\}\) and (ii) semantics represented by the intervening external agents \(\{A\}\). The syntax provides \(2^N\) legal outcomes. But if the system is open to the external world, the presence of which is expressed by \(\{A\}\), then it organizes to a unique final outcome. Once the syntax \(\{\alpha\}\) is known, the final result is therefore an ascertainment that the set of external events \(\{A\}\) must have occurred. Therefore we can take \(\{A\}\) as the element of reality in which our system is embedded.

We define "certitude" the correct application of the rules \(\{\alpha\}\), and "truth" the adaptation to the reality which is expressed by \(\{A\}\).

However the same final outcome would be reached by a different set of rules \(\{\beta\}\). In such a case, retracing back the new tree of bifurcations, we would reconstruct a set \(\{\beta\}\) of external agents. Thus, it seems that truth, \(\{A\}\) or \(\{\beta\}\), is language dependent!

Furthermore, the "emergence" of organization means that we can even build a set of axioms \(\{e\}\) which succeeds in predicting the correct final state without external perturbations, that is, \(\{E\} = \emptyset\) (Figure 7). This is indeed the pretension of the so-called "autoposesis", or "self-organization" [19], to which I have opposed the term "hetero-organization" [20].

From a cognitive point of view, the theory \(\{e\}\) can be reputed to be a "petitio principii", a tricky formulation tailored for a specific purpose and not applicable to slightly different situations. Rather than explicitly listing the elements of reality, as e.g. \(\{A\}\) for \(\{\alpha\}\), the user of language \(\{e\}\) has already exploited at a pre-formalized level the elements of reality, and has made good use of them in planning the axioms \(\{e\}\).

These pieces of knowledge which precede axiomatization have received different names, as "abduction" [21] or "tacit dimension" [22]. Some of them have been memorized as universal tools either in our genetic heritage, or during infancy in our learning age. This seems to be the cognitive valence of Jung’s "archetypes" [23].
5. AN ADAPTIVE MEASUREMENT PROCEDURE

In Section 3 we introduced a measuring apparatus \( M(D, \varepsilon, \tau, \beta) \) which can be modified by changing the number of probes \( D \), space \( \varepsilon \) or time \( \tau \) resolution, or fuzziness \( \beta \). Any change in \( M \) leads to a different set of output numbers, and hence to different sequences \( s \) (words). We can specifically tailor the \( M \) characteristics in order to emphasize a given set of dynamical properties and project out some other ones. Since \( M \) acts as a projector into the subspace of the variables measured by \( M \) itself, changing \( M \) means changing the “point of view” under which we observe the world, and hence making a different theoretical model.

In the previous Section we show an indeterminacy in the reconstruction of the elements of reality \( \{A, B, C, \ldots\} \) which modify the dynamical theories \( \{\alpha, \beta, \gamma, \ldots\} \) of a mentally isolated system. As a result, the truth is represented by a combination of a model developed for the subsystem, plus the external gradients, or boundary conditions, once the subsystem is embedded in a suitable environment.

Which one is the most appropriate among the pairs

\[ \{\alpha, A\}, \{\beta, B\}, \{\gamma, C\}, \text{ etc.} \]

The question is equivalent to asking: among all possible measuring apparatuses \( M \) applicable to an event, which one is the most appropriate?

If we prefer not to decide, we independently make use of different \( M \)'s and correspondingly define different sciences (Figure 3).

In real life, we have to face problems overlapping different separate sciences. For instance, a cardiac disease may be due to a global offset of the pace-maker, or to some local cytopatology or even to a drug effect acting at a biomolecular level.
Figure 8 Fixed measuring apparatuses giving rise to different sciences. Inter communication is based on metaphors (uppermost figure) unless by fuzzy logic one accepts ill defined terms with overlaps (center figure). When the measurement is no longer locked to a fixed position of semantic space, nor to a fixed resolution, then the scientific knowledge can cover, with different degrees of detail, different areas of the semantic space, allowing information exchange within a unique language.

Figure 8 shows the difference between fixed and adaptive $M$. In the first case we sharply define three separate sciences. What can be exchanged among different specialists are not technical words, which are specific of each science, but just the residual metaphorical part, not filtered into the technical word. Should we say that two scientists of different areas always communicate by metaphors? A tentative way out (fuzzy logic [24]) is to avoid sharp definitions, so that different disciplinary terms have regions of overlap.

However the most natural approach seems to start with $M$ at very low resolution, covering all the disciplinary areas, and then — depending upon a specific problem — to zoom toward one or the other narrow point of view.

A successful line of adaptive measurement has been worked out in my group [11–12]. It consists in developing a measuring procedure $M$ whereby we observe a system only at
"almost equal" geometric separations in state space. As a consequence, \( M \) is activated only for short time intervals at irregularly distributed times, separated from each other by unequal intervals. The stroboscopic sequence of these intervals has an information content which provides fast reliable answers to the following questions:

1. recognizing a chaotic dynamics (Lyapunov exponents, different unstable periodic orbits (UPO's));
2. discriminating determinism from stochastic noise;
3. controlling chaos, i.e. stabilizing one of the UPO's contained within a chaotic attractor.

When applied to the control of chaos [12], this adaptive algorithm is effective for values of the stroboscopic times much larger than the Runge-Kutta integration steps and smaller than the periods of all UPO's. In other words, the method introduces a natural adaptation time scale which is intermediate between the minimum resolution time of the dynamics and the time scale of the periodic orbits.

6. CONCLUSION

In this conclusion we discuss the truth value of scientific statements on the basis of the considerations of the previous sections. In the scientific investigation, we select a quantitative feature by application of an apparatus \( M \) at a particular point of the semantic space. The emerging description of reality represents an observation "from one point of view" [25].

Due however to the variety of possible \( M \)’s, we must justify at a metascientific level why we have selected that \( M \) rather than another one. This is a general question dealing with the role of those elements of reality which are preliminary to one particular program.

In Section 5 we have seen that adaptation is a slalom among different sets of rules \( \{\alpha\}, \{\beta\}, \ldots \) under the guidance of a preferential set of external elements which has non zero intersections with \( \{A\}, \{B\}, \ldots \) but does not coincide with neither of them.

In front of the truth problem we can take three attitudes, namely,

(i) Assume the adaptive strategy and its associated reality set as a kind of privileged reference frame. Indeed, being the result of an optimization process, it appears more appropriate than any particular theory, \( \{\alpha\} \), or \( \{\beta\} \) etc.

(ii) Consider the truth problem as a metatheoretical problem. At this metalevel, the set of all sets of truth values \( \{A\}, \{B\}, \ldots \) has to be considered as the truth, but with the stipulation that any individual set makes sense only if associated with the corresponding theory.

(iii) A more fundamental approach recovers the polysemy of the ordinary language as a virtue, not a drawback. More than questioning the power of any specific theory we put into question the same set-theoretical approach to physics. Going back to Figure 2, we have seen that \( M \) provides a sharp connotation which allows to classify any observed entity within an appropriate set. Whence the set-theoretical character of all modern sciences, with the consequent antinomies of modern logic after Cantor, Russell, Gödel etc., transferred into the heart of the scientific language.
Figure 9 Knowledge interpretation: \( R \) = reality, \( M \) = symbol generator (measuring apparatus), \( S \) = symbol interpreter (model builder). Dashed line A (Cartesian cut): \( M + R \) provide representations as symbol sequences, which are interpreted by \( S \). \( S \) can be replaced by a Turing machine.

Dashed line B (realistic approach): \( S + M \) globally face \( R \). Before producing outputs, \( M \) is readjusted among a class of possible measuring apparatuses by a prelinguistic procedure not expressible within the formal language which later \( M \) provides to \( S \).

An adaptive \( M \) means that the localization in semantic space is no longer as sharp as whenever it is defined by a precise stipulation as for the sets. This degree of smoothness seems to me as going back to the polysemytic of the ordinary language. Hence “Epimenides Cretan says: all Cretans are liars” is no longer an antinomy, since Epimenides is not bound to be always a liar, but a liar in general, even though sometimes he can even tell the truth!

Going back to the title, the truth versus certitude issue can be summarized by the following scheme (Figure 9). \( R \) stays for reality (whatever this means), \( S \) for a symbol interpreter (an intelligent being or even a Turing machine!), \( M \) is the measuring apparatus.

In modern science, \( M \) is usually not questioned, and the elaboration takes place on the output of \( M \). This was called as the (A) Complexity approach, leading eventually to certitude, when the \( S \) machine replicates exactly the \( M \) output sequence with the minimal amount of information. From a gnosological point of view, if \( M \) are our senses, as suggested by Hume, then \( S \) (called by Descartes “res cogitans”) has to face *not the world, but the representation* already coded by \( M \). As shown in Section 3, it has to face a grammatical problem implementable by a machine. This means that Descartes’ mind is equivalent to a Turing machine, as already suspected by many experts of Artificial Intelligence. This strong association of \( M \) on the side of \( R \) is equivalent to what Atmanspacher called “The Cartesian cut” [26].

On the contrary, the (B) Complex system approach regards sciences as dealing with the world through an adaptive procedure \( M \), for which however a linguistic foundation does not exist, because any linguistic formulation is subsequent to the operation of \( M \). During the scientific operation, \( S + M \) act in an entangled way. A meta-level of investigation (psychology of cognition) is required to disentangle \( S \) from \( M \).

In summary, there is a nonlinguistic residue in the scientific operation which then precludes a Turing machine from acting as a creative scientist.
REFERENCES

5. See the following contributions in the series of Santa Fe Institute publications. Proceedings Volumes:
15. Crutchfield, J.P. “Semantics and thermodynamics” in Ref. 5c.

1. INTRO

The concept of hierarchy is not as simple as it may seem. In most apparent systems, an essential element in the hierarchy is the order of events. In these systems, the hierarchical structure determines the order of events. In a paper published 50 years ago, it was sufficient to introduce the concept of hierarchy. But it was also necessary to introduce the concept of disorder.