Nonlinear Dynamics, Psychology, and Life Sciences, Vol. 5, No. 1, 2001

Complexity Versus Complex Systems: A New Approach to Scientific Discovery

F. T. Arecchi[1]

*Extraction of quantitative features from observations via suitable measuring devices **M** means that the words of science are coded as numbers, and the syntaxis is a set of mathematical rules. Once general premises are available all consequences can be worked in a purely deductive way. This characteristic of science displays two orders of drawbacks, namely, undecidability of deductive procedures, and intractability of computer modelings of complex situations. The way out of such a crisis consists in an adaptive strategy, that is, in a frequent readjustment of **M** suggested by the observed events. As a consequence, **M** provides different data streams (words) for the same observed events, as it is tuned to different resolutions. The adaptive strategy here introduced should by no means be confused with the adaptivity of a learning machine, which—inputted by a data stream—readjusts itself over a class of theoretical explanations in order to select the optimal one, thus providing knowledge conditional on the assigned input. On the contrary, physics aims at extracting regular patterns out of things, by a "trial and error" procedure which includes not only modifications of the explanations for fixed data sets, but also exploring different data sets via modified **M**'s. This **M**-adjustment is a pre-linguistic endeavour, not expressible by a formal language. Such an essential characteristic of the physical program means that physics can not be performed by a machine.*

KEY WORDS: nonlinear dynamics; complexity; epistemology; cognitive sciences.

INTRODUCTION

It is important to distinguish between "Complexity" and "Complex Systems." The first term has been differently defined in formal languages (Hopcroft & Ullman, 1979), computer science (Chaitin, 1966, Kolmorogov,

[1]Department of Physics, University of Firenze, 50125 Italy and Istituto Nazionale di Ottica—Firenze.

1090-0578/01/0100-0021$19.50/0

1965) and nonlinear signal analysis (Abarbanel, Brown, Sidorowich & Tsimring, 1993; Grassberger, Schreiber & Schaffrath, 1991), starting in the early 80's with the intrinsic non–predictability associated with chaotic time series.

"Complexity" is associated with 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? This contextual analysis uncovers the grammatical rules which allow some symbol sequences (words) and forbid some other ones; it also attributes a likelihood to each word and hence it provides predictions about the future of the time series.

This "Complexity" approach has received different formulations, with different solutions leading to automatic procedures on complexity assignment (Badii & Politi, 1997; Casdagli & Eubank, 1992; Crutchfield & Young, 1989; D'Alessandro & Politi, 1990; Grassberger, 1986; Pines, 1987; Zurech, 1990). On the other hand, natural scientists have rather focussed on the ontological fact that reality uncovers lots of complex structures. Two different tasks are thus to be considered, 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) When confronted with a real event, 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 analized as in A)?

Case B is "prior" to A, and it is addressed by adaptive strategies, which we can later formalize as linguistic procedures, but which arise at a prelinguistic level and even determine the same choice of the most appropriate language (Simon, 1982). This more fundamental problem is that of "Truth," defined as "*adaequatio 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 (Crutchfield & Young, 1989). 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.

Our main conclusion is that while there may be a complexity machine for Case A, it is in principle impossible to introduce a science mächine for Case B. Hence Case B remains a human endeavour that is not reducible to automatic procedures. The paper is organized as follows.

The second section reviews current definitions of complexity showing the virtues and intrinsic limitations of a contextual analysis of a data stream. The third section is a dynamic approach to a complex situation. While in the second section no apriori rules are requested but a data stream is preliminary, in the third section 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 the fourth section we present an adaptive strategy that was recently introduced to recognize (Arecchi, Basti, Boccaletti & Perrone, 1994) and control (Boccaletti & Arecchi, 1995) a chaotic dynamics (Arecchi & Boccaletti, 1997). While an adaptive procedure was already incorporated in the complexity machine, in order to fit the theory to the input sequence, we change the same structure of M in order to provide different sets of measurement sequences to be later analyzed.

We conclude in the fifth section of the article with some epistemological insight, recalling that a knowledge program based on assigned input data will 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.

COMPLEXITY OF SYMBOLIC SEQUENCES

In computer science, we define the complexity of a word (symbol sequence) as some indicator of the cost implied in generating that sequence. There is a "space" cost (length of the instruction 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 called AIC (Algorithmic Information Complexity) (Chaitin, 1966; Kolmogorov, 1965) 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 called "logical depth" (Bennett, 1987) is defined as the CPU time required to generate the sequence out of the *minimal* instruction. It is minimal for 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.

While AIC refers to the process of building a single item, logical depth corresponds to finding the properties of all possible outputs from a known source. 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: (a) 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); (b) 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 (Gell-Mann, 1994; Grassberger, 1986) 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 Case A and Case B introduced in the beginning of this article, let us sketch the essential elements of knowledge building in natural sciences. First of all, we realize 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: (a) 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 later on in this article. (b) 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 modeler O_1 is inputted by some sequence s, and it sends an explanation which should enable O_2 to reconstruct an output $s' = s$.

Notice that **M** is a whole class of possible instruments, and different **M**'s will give rise to different data sequences (different words). The explanation (Crutchfield, 1992) 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 $\|x\| = \|m\| + \|e\|$ shorter than the sequence length $\|s\|$.

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

$$Pr(m/s) \qquad (1)$$

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

$$Pr(m|s) = \frac{Pr(s|m)P(m)}{\sum_m P(s|m)} \tag{2}$$

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: (a) x must explain s, that is, O_2 should resynthesize the original data $s' = s$; and (b) the bit length

$$\|x\| = \|m\| + \|e\|$$

must be minimized.

The efficiency of an explanation is given by the compression ratio

$$C(m, s) = \frac{\|x\|}{\|s\|} = \frac{\|m\| + \|e\|}{\|s\|}, \tag{3}$$

where C is a cost function of the "complexity" of the explanation. The optimal model minimizes this cost. 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: $\|e\| = 0$. In this section 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 **M.**

A DYNAMICAL APPROACH TO COMPLEXITY

Modern science started from Galileo's rule: "do not try to grasp the "essence" (i.e. the nature) of a thing, but stick only to its quantitative affections" (Galilei, 1612). Quantities are extracted by a measuring procedure that we denoted by **M.** We call the sequence of **M** outputs as the "words" of science. The most crucial aspect of this procedure is the apparent arbitrariness of **M,** that is, the large number of different scientific words to be associated with an event usually denoted by a single word in any ordinary language. This proliferation of **M**'s has avoided for centuries the question of complexity. Suppose for instance that we look at the cell dynamics by $\mathbf{M}_1$. 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 $\mathbf{M}_2$ which provides different words and hence a different science, namely, "physiology". Similarly if we look at the biomolecules we build a "biochemistry."

We change language depending on the **M** apparatus. A more economic criterion would consist in a reductionist approach. 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 occurs that was, already hinted by Anderson (1972). If our words were a global description of the object in *any* situation (as the philosophical "essences" in Galileo) 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 of organization display features not predictable from the lower ones, is what colloquially we call *complexity.*

When considered in this fashion, 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. Reductionism from large to small as sketched above was accompanied by a logical reduction of the scientific explanation to a deductive task out of a set of axioms. In this spirit, a scientific theory is considered as a set of primitive concepts (defined by suitable measuring apparatuses **M**) plus the mutual relations. Concepts and relations are the axioms of the theory. 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. 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 1 sketches the bifurcation tree of a complex nonlinear dynamics, as one changes a suitable control parameter α. During the dynamical evolution of an open system, the control parameters $\{\alpha\}$ may assume different values, hence the cascade of bifurcations provides a large number of final states starting from a unique initial condition.

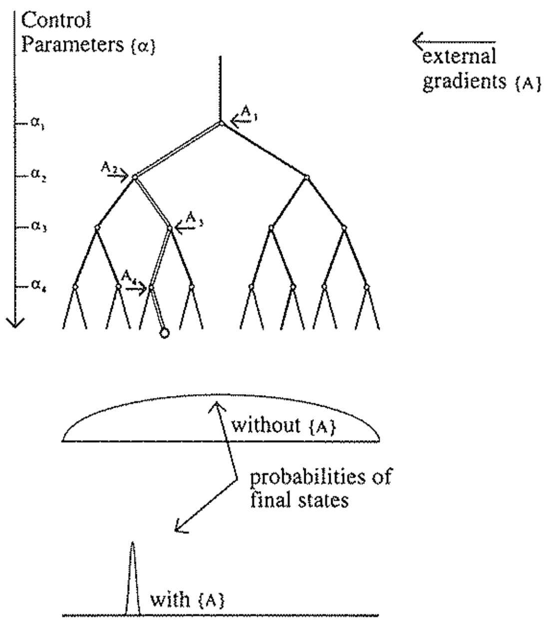

Fig. 1. Bifurcation tree in nonlinear dynamics. As a control parameter α is tuned through different values, novel steady states appear. By tuning α from α_0 to α_N, the system goes from 1 to 2^N different states. We call such dynamical complexity an ambiguity. It is responsible for the failure of reductionism. 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 α_i an external agent A_i has broken the bifurcation symmetry (see next figure).

Thus the reductionistic tentative of explaining reality out of its constituents yields an exponentially high number of possible outcomes, when only one is in fact that observed. This means that, while the theory, that is the syntaxis, 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 nonlinear dynamics for different settings of a control parameter. Notice that dynamical bifurcations in a system of interactive identical particles display specific symmetries (Fig. 2). Only external gradients break this symmetry. Hence, there has been *organization,* that is a unique final state, this means that at each bifurcation vertex of

Fig. 1 the symmetry was broken by an agent external to the system under investigation. We can thus stipulate the following things:

1. A set of control parameters

$$\alpha_1, \alpha_2, \ldots \alpha_N = \{\alpha\}$$

is responsible for successive bifurcations leading to an exponentially high number (in the example, of the order of 20^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.

2. For a system of finite size embedded in an environment, there is a set of external forces

$$A_1, A_2, \ldots A_N = \{A\}$$

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

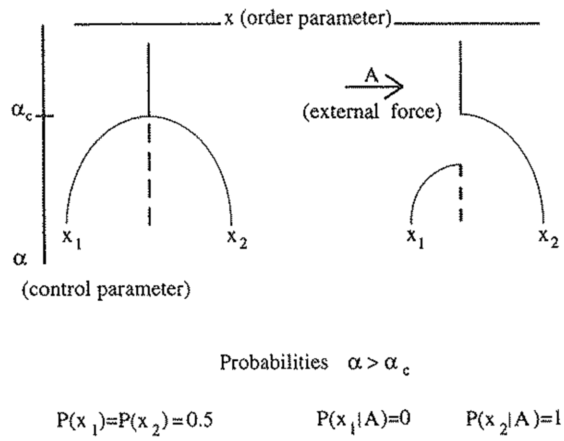

Fig. 2. Examples of bifurcation diagrams. The dynamical variable x (order parameter) varies horizontally, the control parameter α varies vertically. Solid (dashed) lines represent stable (unstable) steady states as the control parameter is changed. Left: symmetric bifurcation with equal probabilities 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 α_c, then the right (left) branch has probability 1 (0).

We are in presence of a conflict between Stipulation 1 syntaxis represented by the set of rules (axioms) $\{\alpha\}$ and Stipulation 2 semantics represented by the intervening external agents $\{A\}$. The syntaxis 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 $\{B\}$ of external agents. Thus, it seems that truth, $\{A\}$ or $\{B\}$, is language dependent! Furthermore, the "emergence" of organization means that we can even build a set of axioms $\{\varepsilon\}$ which succeeds in predicting the correct final state without external perturbations, that is, $\{E\} = \emptyset$ (Fig. 3). This is indeed the pretension of the so called "autopoiesis," or "self-organization" (Krohn, Küppers & Nowotny, 1990).

From a cognitive point of view, the theory $\{\varepsilon\}$ can be reputed to be a "petitio principi," a tricky formulation tailored for a specific purpose and not applicable to slightly different situations. Rather than explicitely detailing the elements of environment which break the symmetry, as e.g. $\{A\}$ for $\{\alpha\}$, the user of language $\{\varepsilon\}$ has already exploited at a pre-formalized level a detailed knowledge of the process in planning the axiom $\{\varepsilon\}$.

An "ad-hoc" model may be appropriate for a specific situation, but in general it lacks sufficient breadth to be considered as a general theory

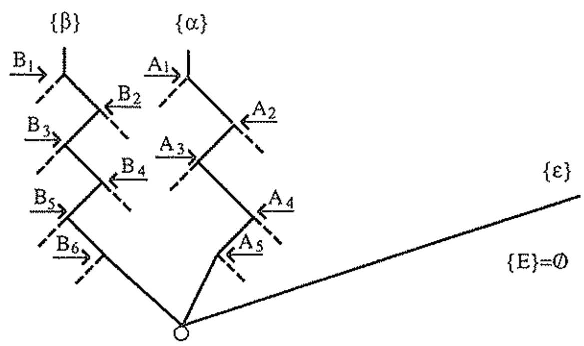

Fig. 3. Different theoretical models may explain the same final state. The backward reconstruction of the dynamical history will then retrieve different classes of external agents.

(think e.g. of the Ptolemaic model of the solar system). However, in describing the adaptive strategy of a living species, or a community etc., "self-organization" may be the most successful description. In other words, once the environmental influences have been known, better to incorporate that knowledge in the model, thus assuring the fast convergence to a given goal.

These pieces of knowledge which preceed axiomatization have received different names, as "abduction" (Hanson, 1958; Peirce, 1931–35) or "tacit dimension" (Polnayii, 1958). 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" (Von Franz, 1988).

AN ADAPTIVE MEASUREMENT PROCEDURE

The measuring apparatus **M** can be modified by changing the number of probes, space or time resolution, or fuzziness (i.e. sharpness in localizing the data within a space-time box). 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. $\{\alpha, A\}$, $\{\beta, B\}$, $\{\gamma, C\}$, 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.

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 4 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,

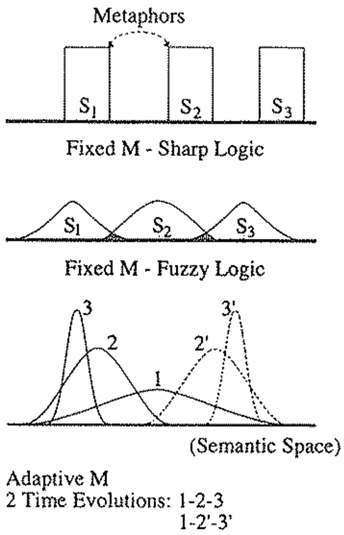

Fig. 4. Fixed measuring apparatuses giving rise to different sciences. Inter communication is based on metaphors (upmost figure) unless by fuzzy logic one accepts ill-defined terms with overlaps (center figure). When the measurement is no longer locked to a fixed position of semantic space, nor to a 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.

not filtered into the technical word. Should we say that two scientists of different areas always communicate by metaphors? A tentative way out, according to fuzzy logic (Zadeh, 1987) is to avoid sharp definitions, so that different disciplinary terms have regions of overlap. However, the most natural approach seems to start with a 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 (Arecchi & Boccaletti, 1997). It consists in developing a measuring procedure whereby we observe a system only at "almost equal" geometric separations in state space. As a consequence, observations are 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: (a) recognizing a chaotic dynamics (Lyapunov exponents, different unstable periodic orbits (UPO's); (b) discriminating determinism from stocastic noise; and (c) controlling chaos, i.e. stabilizing one of the UPO's contained within a chaotic attractor. When applied to the control of chaos (Boccaletti & Arecchi, 1995), this adaptive algorithm is effective for values of the stroboscopic times much larger than the elementary integration time step 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.

A NEW APPROACH TO THE SCIENTIFIC DISCOVERY: THE CARTESIAN CUT VS. THE REALISTIC APPROACH

In the scientific investigation, we select a quantitative feature by application of an apparatus **M.** The emerging description of reality represents an observation "from one point of view" (Agazzi, 1974). 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.

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

In front of the truth problem we can take three attitudes: (a) 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. (b) Consider the truth problem as a metatheoretical problem. At this metalevel, the set of all sets of truth values $\{A\}$, $\{B\}$, . . . has to be considered as the truth, but with the stipulation that any individual set makes sense only if associated with the corresponding theory. (c) A more fundamental approach recovers the polysemicity 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. **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, Russel, Gödel etc., transferred into the heart of the scientific language. 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*.

In the epistemology of scientific discovery, one usually relies on Bayes theorem to corroborate an hypothesis along the following procedure. We start hypothesizing a certain model x out of a possible set $\{x\}$ of models. Let us call $p\{x\}$ the apriori probability of x. As we collect some data d, the conditional probability $p(x/d)$ built upon a Bayes rule is a better approximation to $p(x)$. Thus we expect a convergent feedback loop as the data set $\{d\}$ increases. The Bayes rule was already given in Eq. (2); here we just replace m (models) by x (hypotheses) and the input sequence s with the data d. The rule presupposes the availability of the hypothesis set $\{x\}$ with a pre-assigned measure $p(x)$.

In fact, this approach overlooks the "theory ladenness" of the measuring procedure. Indeed, the apparatus **M** collecting the data $\{d\}$ is built upon the selected theoretical model. Thus, changing the probabilities $p(x/d)$ by conditioning them on the observations $\{d\}$ means also changing the structure of **M**, and hence perturbing the same set $\{d\}$. Hence, there are two feedback loops: (a) The standard one from data $\{d\}$ to models $\{x\}$, which assumes that the set $\{x\}$ is already given and one has just to choose within it; and (b) the new loop from $\{d\}$ to **M**, leading to $\{d'\} \neq \{d\}$.

The second loop cannot be in general formalized, since it works on different classes of **M**, and therefore it requires different operation concepts, or words. Changing point of view means changing language, and only a meta-linguistic approach can cope with variety of languages.

An intermediate compromise is what we called "adaptive procedure." Let us fix some finite bounds in the space of models $\{x\}$ as well in the space of apparatuses $\{M\}$. In such a bounded case, the procedure can be automatized as an algorithm. In the cases discussed by Arecchi and Boccaletti (1997), $\{M\}$ is the class of clocks set on different units of time, and we readjust the observation time upon the dynamical features.

Imposing apriori a bounded variation to **M** (i.e. limiting oneself to a finite number of "points of view") works fine only if you have tested its validity by a pre-formalized session. Indeed, before any scientific enterprise, we have already some non-scientific acquaintance with a problem, motivated by curiosity, practical interest, etc. In a generic scientific endeavour, imposing apriori bounds, would be against wisdom, insofar as it cuts out ranges of phenomena which might be otherwise relevant.

Going back to the title, the truth versus certitude issue can be summarized by the following scheme (Fig. 5). **R** stands for reality (whatever this means), **S** for a symbol interpreter (an intelligent being or even a Turing machine!), and **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 gnoseological 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 Sec. 3, it has to face a grammatical problem inplementable 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

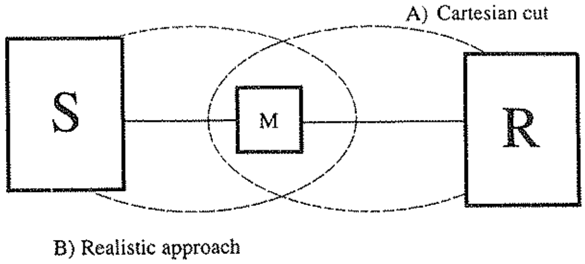

Fig. 5. Knowledge interpretation: **R** = reality, **M** = symbol generator (measuring apparatus), **S** = symbol interpreter (model builder). Dashed line A (Cartesian cut): **M** + **R** provide representations as symbol sequences, which are interpreted by **S**. **S** can be replaced by Turing machine. 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**.

what Atmanspacher called "The Cartesian cut" (Atmanspacher, 1994). 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 operations, **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

Abarbanel, H. D. I., Brown, R., Sidorowich, J. J., & Tsimring, L. S. (1993). The analysis of observed chaotic data in physical systems. *Review of Modern Physics, 65,* 1331–1450.

Agazzi, E. (1974). *Temi e problemi di filosofia della fisica [Topics and problems in the philosophy of physics],* Sec. 50. Rome: Ed. Abete.

Anderson, P. W. (1972). More is different. *Science, 177,* 393–400.

Arecchi, F. T., Basti, G. F., Boccaletti, S., & Perrone, A. (1994). Adaptive recognition of a chaotic dynamics. *Europhysics Letters, 26,* 327–330.

Arecchi, F. T., & Boccaletti, S. (1997). Adaptive strategies for recognition, noise filtering, control, synchronization, and targeting of chaos. *Chaos, 7,* 621–634.

Atmanspacher, H. (1994). Complexity and meaning as a bridge across the Cartesian cut. *Journal of Consciousness Studies, 1,* 168–183.

Badii, R., & Politi, A. (1997). *Complexity.* Cambridge, UK: Cambridge University Press.

Bennett, C. H. (1987). Dissipation, information, computational complexity and the definition of organization. In D. Pines, (Ed.). *Emerging syntheses in science* (pp. 215–233). Redwood City, CA: Addison-Wesley.

Boccaletti, S., & Arecchi, F. T. (1995). Adaptive control of chaos. *Europhysics Letters, 31,* 127–130.
Casdagli, M., & Eubank, S. (Eds.) (1992). *Nonlinear modeling and forecasting.* Redwood City, CA: Addison-Wesley.
Chaitin, G. J. (1966). On the length of programs for computing binary sequences. *Journal of the Association for Computational Mathematics, 13,* 547–562.
Crutchfield, J. P. (1992). Semantics and thermodynamics. In M. Casdagli & S. Eubank (Eds.) (1992). *Nonlinear modeling and forecasting* (pp. 317–360). Redwood City, CA: Addison-Wesley.
Crutchfield, J. P., & Young, K. (1989). Inferring statistical complexity. *Physics Review Letters, 63,* 105–108.
D'Alessandro, G., & Politi, A. (1990). Hierarchical approach to complexity with applications to dynamic systems. *Physics Review letters, 64,* 1609-1612.
Galilei, G. (1612). Letter to M. Welser "On the sunspots." Italian original in A. Favaro (Ed.) (1932), *Le Opere di G. Galilei* (Vol. V, pp. 186–239). Firenze: G. Barbera.
Gell-Mann, M. (1994). *The quark and the jaguar.* New York: W. H. Freeman.
Grassberger, P. (1986). Toward a quantitative theory of self-generated complexity. *International Journal of Theoretical Physics, 25,* 907–925.
Grassberger, P., Schreiber, T., & Schaffrath, C. (1991). Nonlinear time series sequences analysis. *International Journal of Bifurcation and Chaos, 1,* 521–547.
Hanson, W. R. (1958). *Patterns of discovery: An inquiry into the conceptual foundations of science.* Cambridge, UK: Cambridge University Press.
Hopcroft, J. E., & Ullman, J. D. (1979). *Introduction to automata theory, languages and computation.* Reading, MA: Addison-Wesley.
Kolmogorov, A. N. (1965). Three approaches to the quantitative definition of information. *Problems of Information Transmission, 1,* 4–16.
Krohn, W., Küppers, G., & Nowotny, H. (1990). *Self-organization: Portrait of a scientific revolution.* Dordrecht, Netherlands: Kluwer Academic.
Peirce, P. W. (1931-1935). *Collected papers, Vol. 1-VI.* C. Harshorne & P. Weiss (Eds.). Cambridge, MA: Harvard University Press.
Pines, D. (Ed.) (1987). *Emerging syntheses in science.* Redwood City, CA: Addison-Wesley.
Polanyii, M. (1958). *Personal knowledge: Towards a post-critical philosophy.* London: Routledge & Kegan Paul.
Simon, H. (1982). The architecture of complexity. *Proceedings of the American Philosophical Society, 106,* 467–505.
Shaw, R. (1981). Strange attractors, chaotic behavior, and information flow. *Zeitschrift Naturforschung, 38A,* 80–112.
Von Franz, M. L. (1988). *Psyche and materiel.* Einsiedeln, Switzerland: Eranos.
Zadeh, L. A. (1987). *Fuzzy sets and applications: Selected papers.* New York: Wiley.
Zurek, W. H. (Ed.). (1990). *Complexity, entropy and the physics of information.* Redwood City, CA: Addison-Wesley.