Journal of Physics: Conference Series **174** (2009) 012010 doi:10.1088/1742-6596/174/1/012010

The physics of mental acts: coherence and creativity

F. Tito Arecchi

University of Firenze and INOA (Istituto Nazionale di Ottica Applicata)-Firenze, Italy

E-mail: tito.arecchi@inoa.it

Abstract. Coherence is a long range order absent at thermal equilibrium, where a system is the superposition of many uncorrelated components. To build non-trivial correlations, the system must enter a nonlinear dynamical regime.The nonlinearity leads to a multiplicity of equilibrium states, the number of which increases exponentially with the number of partners; we call complexity such a situation. Complete exploration of complexity would require a very large amount of time. On the contrary, in cognitive tasks, one reaches a decision within a few hundred milliseconds. Neuron synchronization lasting around 301 msec is the indicator of a conscious perception (Gestalt); however, the loss of information in the chaotic spike train of a single neuron takes a few msec, thus a conscious perception implies a control of chaos, whereby the information stored in a brain area survives for a time sufficient to elicit an action. Control of chaos is achieved by the interaction of a bottom-up stimulus with a top-down control (induced by the semantic memory). We call creativity this optimal control of neuronal chaos; it goes beyond the Bayesian inference, which is the way a computer operates, thus it represent a non-algorithmic step.

1. Introduction

The terms introduced in the Abstract are defined as follows:

Coherence = long range order (in space [vision] or time [music]);
Complexity = display of different coherences;
Creativity = jump from one coherence regime to another.

What are the relations among the three concepts? A detailed analysis has been given elsewhere [Arecchi 2007] and here we recall the main points. Coherence is associated with long range correlations, in space or time; complexity arises whenever an array of coupled dynamical systems displays multiple paths of coherence. Creativity corresponds to a selection of a coherence path within a complex nest. As we will see soon, it seems dynamically related to control of chaos.

Exploration of a complex situation would require a very large amount of time, in order to classify all possible coherences, i.e. long range correlations. In cognitive tasks facing a complex scenario, our strategy consists in attaining a decision within a finite short time. Any conscious perception (we define conscious as that eliciting a decision) requires a few hundred milliseconds, whereas the loss of information in the chaotic spike train of a single neuron takes a few msec.

The interaction of a bottom-up signal (external stimulus) with a top-down modification of the control parameters (induced by the semantic memory) leads to a collective synchronization lasting a sizable fraction of a second: this is the indicator of a conscious perception. The operation is a control of chaos, and it has an optimality; if it lasts less than 200msec, no decisions emerge, if it lasts much longer, there is no room for sequential cognitive tasks. We call creativity this optimal control of neuronal chaos. It amounts to selecting one among a large number of possible coherences all present in a complex situation. *The selected coherence is the meaning of the object under study.*

In Sec. 2, we survey the main steps of a perceptual process, with reference to vision. In Sec. 3, we discuss the cognitive process, starting from the most fundamental inference based on Bayes theorem; we show that it is limited to a single model (or likelihood) whereas complex situations require changes of model. This is the creative step peculiar of human decisions and impossible to a Turing machine. In Sec. 4, we review critically a quantum approach to brain processes. In Sec. 5 (conclusions), we compare two pairs of conflicting categories relevant for decision making, namely, partial truth versus global truth and hermeneutic circle versus hermeneutic spiral.

2. How we see: processing visual events

Fig. 1 shows a brain section with the path of the visual information. A given light signal impinging on the eyes retina stimulates the production of electrical pulses (spikes) high about 100 millivolt and lasting about 1 millisecond. The time sequence of these standard pulses represents a neural code, this means that different inputs are coded in different spike sequences. Spikes travel on the neuron axons , like on transmission lines At the axon end (synapse) the electrical signal releases some amount of neurotransmitters, that is, specialized chemicals which diffuse in the inter-neuron space and arrive to a next neuron, where they again stimulate electrical spikes and so on. The brain has a huge number of neurons tangled by a web of mutual couplings. A detailed account of all couplings is out of reach at the present state of neuroscientific investigation. It would require resolving each neuron separately; this can be done in laboratory animals by inserting a sparse number of microelectrodes each one probing a single neuron . Coarse grained information is captured by EEG (electro encephalo-gram) whereby we measure the electrical activity probed by some electrodes glued on the scalp. At variance with the microelectrodes probing the single neuron, the technique is non invasive; however an external electrode sums up the activity of many neurons, with a poor space resolution. The fMRI (functional magnetic resonance imaging) is sensitive to the magnetic properties of oxygenated blood; thus it records the blood intake to a brain area which is in operation and hence requires energy; again, the space resolution is poor.

Let us consider the visual system; the role of elementary feature detectors has been extensively studied [Hubel]. Thus, we know that some neurons are specialized in detecting exclusively vertical or horizontal or tilted bars, or a specific luminance contrast, etc.

However, the problem arises: how elementary detectors contribute to a holistic (Gestalt) perception? A hint is provided by [Singer]. Suppose we are exposed to a visual field containing two separate objects. Both objects are made of the same visual elements, horizontal and vertical contour bars, different degrees of luminance, etc. What are then the neural correlates of the identification of the two objects? We have one million fibers connecting the retina to the visual cortex. Each fiber results from the merging of approximately 100 retinal detectors (rods and cones) and as a result it has its own receptive field. Each receptive field isolates a specific detail of an object. We thus split an image into a mosaic of adjacent receptive fields. Now the feature binding hypothesis consists of assuming that all the cortical neurons whose receptive fields are pointing to a specific object synchronize the corresponding spikes, and as a consequence the visual cortex organizes into separate neuron groups oscillating on two distinct spike trains for the two objects. Direct experimental evidence of this synchronization is obtained by insertion of microelectrodes in the cortical tissue of animals just sensing the single neuron (Fig. 2).

The interaction of a bottom-up signal (external stimulus) with a top-down change of the control parameters (induced by the semantic memory) leads to a collective synchronization lasting 200 msec: this is the indicator of a conscious perception. The operation is a control

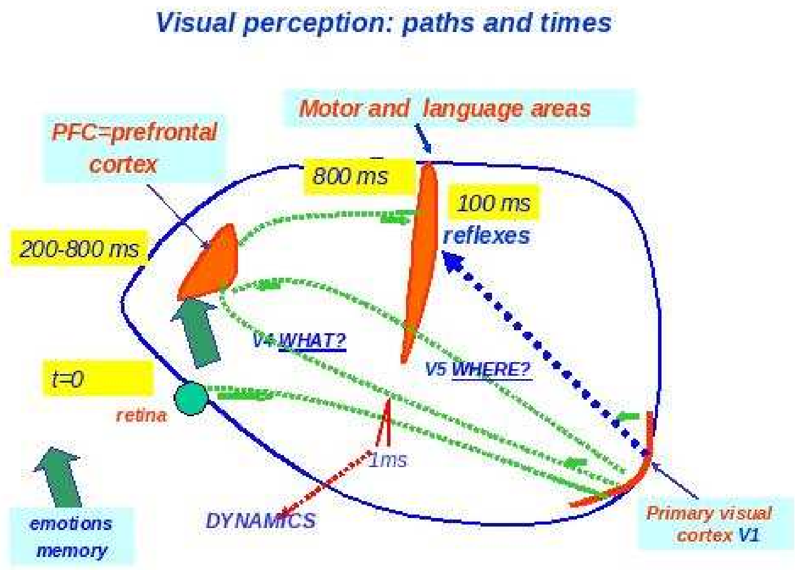


Figure 1. Time sequence of brain events following a visual stimulus on the retina at time t=0. After 200 msec, the signal coded as a train of electrical pulses of 1 msec duration(spikes) traveling on neurons axons, arrives at the primary visual cortex V1; the dynamics of spikes (interspike interval) depends on the sensory stimuli and on the receptive cell (Fig. 2). The signal from V1 is then elaborated in two distinct areas, V4 or WHAT area which recognizes shapes and color, and V5 or WHERE area, which recognizes motion and space relations. The separate information is combined in the PFC (prefrontal cortex) together with top-down signals coming from the inner brain (emotions, memory).This mixing takes about half a second starting from 200 msec; at 800 msec after the sensitization of the retina, a decision emerges activating the motor and language areas, The elaboration in PFC takes so long because the input is compared with previous memories and hence re-coded. Instead vital responses arrive directly to the motor area in even less than 100 msec (reflexes: heavy dotted line).

of chaos, and it has an optimality; if it lasts less than 200msec, no decisions emerge, on the contrary, if it lasts much longer, there is no room for sequential cognitive tasks. The addition of extra degrees of freedom implies a change of code, thus it can be seen as a new level of description of the same physical system.

Based on the neurodynamical facts reported above, we can understand how this occurs [Grossberg]. The higher cortical stages from where decision emerge have two inputs. One (bottom-up) comes from the sensory detectors via the early stages which classify elementary features. This single input is insufficient, because it would provide the same signal for e.g. horizontal bars belonging indifferently to either one of the two objects. However, as we said already, each neuron is a nonlinear system passing close to a saddle point, and the application of a suitable perturbation can stretch or shrink the interval of time spent around the saddle, and thus lengthen or shorten the interspike interval. The perturbation consists of top-down signals corresponding to conjectures made by the semantic memory (Fig. 3). ART (adaptive resonance theory) refers to a re-coding of bottom up stimuli based on previous learning. For simple stimuli which require a fast reaction, there is a direct access from V1 to the motor areas as shown in Fig. 1.

On the other hand, exploration of a complex situation would require a very large amount of

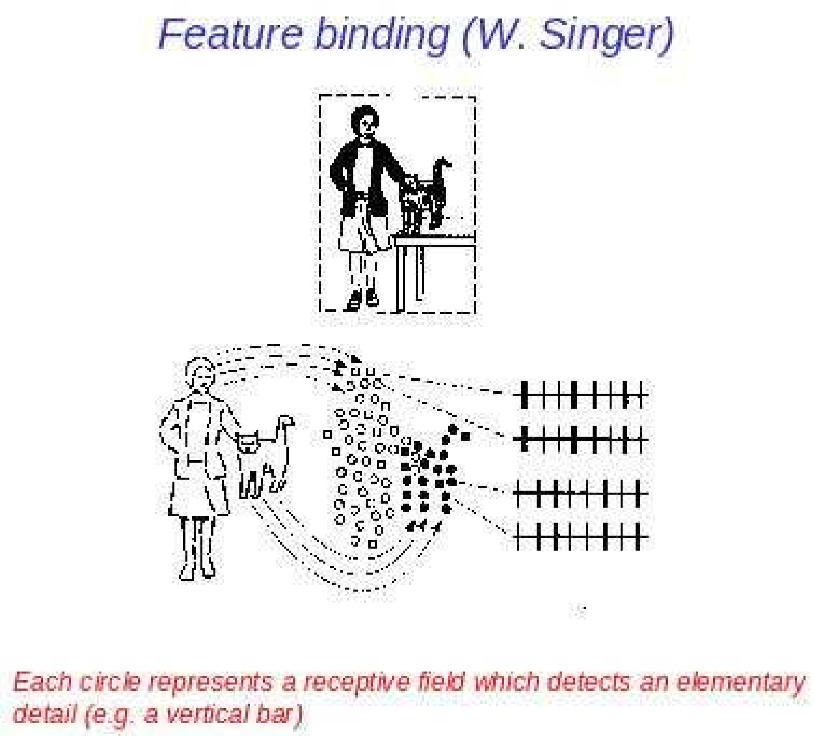


Figure 2. Feature binding: the lady and the cat are respectively represented by the mosaic of empty and filled circles, each one representing the receptive field of a neuron group in the visual cortex. Within each circle the processing refers to a specific detail (e.g. contour orientation). The relations between details are coded by the temporal correlation among neurons, as shown by the same sequences of electrical pulses for two filled circles or two empty circles. Neurons referring to the same individual (e.g. the cat) have synchronous discharges, whereas their spikes are uncorrelated with those referring to another individual (the lady) [from Singer].

time. In cognitive tasks facing a complex scenario, the cognitive strategy consists in converging to a decision within a finite time . Various experiments [Libet, Rodriguez et al.] prove that a decision is taken after a few hundred milliseconds of exposure to a sensory stimulus Thus, any conscious perception (we define conscious as that eliciting a decision) requires a few hundred msec, whereas the loss of information in a chaotic train of neural spikes takes a few msec. There must be a mechanism of chaos control which holds a given information for the time necessary to decide about.

Synchronization of a chain of chaotic lasers provides a promising model for a physics of cognition. Indeed, a peculiar type of chaos, consisting of sequences of identical spikes erratically occurring in time, has been discovered in lasers and called HC [Arecchi and Meucci 2008]. The chaotic train can be regularized by synchronization actions. An array of weakly coupled HC systems represents the simplest model for a physical realization of feature binding. The array can achieve a collective synchronized state lasting for a finite time (corresponding to the physiological 200 ms!) if there is a sparse (non global) coupling, if the input (bottom-up) is applied to just a few neurons and if the inter-neuron coupling is suitably adjusted(top-down control of chaos) [Ciszak, Montina and Arecchi 2009 a,b]. This approach is discussed in the next section.

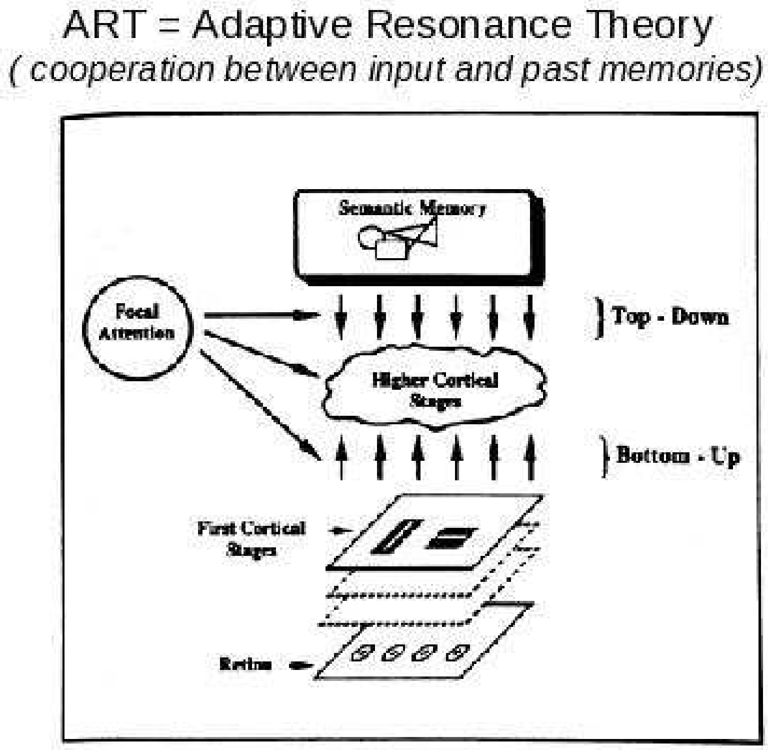


Figure 3. ART = Adaptive Resonance Theory. Role of bottom-up stimuli from the early visual stages an top-down signals due to expectations formulated by the semantic memory. The focal attention assures the matching (resonance) between the two streams.

3. The scientific insight

As Galileo was saying in a letter to Marc Welser, in 1610, a sound scientific approach means avoiding philosophical considerations on the nature of objects, and limiting oneself to measurable data coded in numbers, then connecting these numbers by a mathematical formalism The scientific method of Galileo is based on two pillars:

Sensate esperienze= experiments performed through our senses: if the case helped by equipment, e.g. the telescope and the clock;

Necessarie dimostrazioni = compelling proofs: in a theorem, if one agrees on the premises and understands the proof, than he must necessarily accept the conclusions.

Thus, in science meaning appears irrelevant and only measurable details and repeatable phenomena make sense. The Galileian innovation was to replace words with numbers. Some time later, Newton discovered that assignment of position and velocity of a particle at a given time (the initial condition) determines univocally its future trajectory. All objects in nature are made of elementary particles endowed of this Newtonian property. Thus the future of the whole object can be predicted once we know its initial condition. This was the basis of Laplace determinism (1812) disproved a few decades later by Poincaré.

3.1. Deterministic chaos and its control

For a Newtonian particle, once we know the forces ,the trajectory emerging from a given initial condition ,as the star in Fig. 4, is unique. The coordinates of the initial point are in general

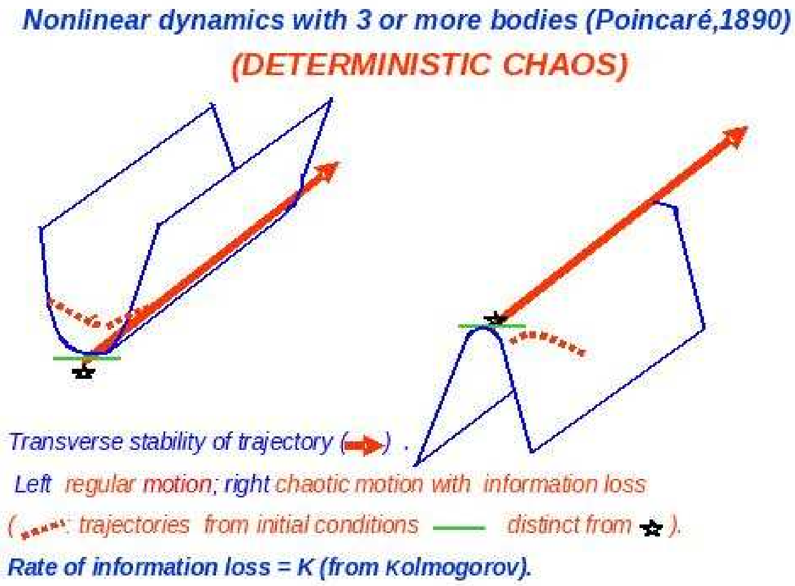


Figure 4. Deterministic chaos-the trajectory emerging from a precise initial condition, as the star, is unique. However, in general the initial condition is spread over a small patch. Points of the patch converge toward the calculated trajectory or diverge from it, depending on whether the transverse stability landscape is a valley (left) or a hill (right). From 3 coupled dynamical degrees of freedom on, the right situation is generic. We call it deterministic chaos. The information loss rate K is called Kolmogorov entropy.

real numbers truncated to a finite number of digits, thus the initial condition is spread over a small patch. Points of the patch converge toward the calculated trajectory or diverge from it, depending on whether the transverse stability yields a landscape like a valley (left) or the ridge of a hill (right). Poincar proved that from 3 coupled dynamical degrees of freedom on, the right situation is generic. Nowadays we call it deterministic chaos. It is a sensitive dependence on initial conditions and implies a loss of information of the initial preparation. The loss rate K is called Kolmogorov entropy. We can adjust its value by adding extra variables which change the slope of the downhill fall without practically perturbing the longitudinal trajectory (control of chaos), as shown in Fig. 5. Control of chaos occurs everywhere in interacting systems and it is essential to establish coherent features.

Suppose we have a N-dimensional chaotic system: adding a few, say p, extra variables we embed the dynamical problem in a $(N+p)$-dimensional space. By careful choice of the control, we can keep the longitudinal trajectory of the original dynamics, modifying however its transverse stability. In the example of Fig. 6, we consider a perceptual feature encoded by a chaotic dynamics which loses information over 2 msec. This time is insufficient to elicit a decision; thus a mechanism of chaos control must stabilize chaos for about 200 msec. The chaos control was initially devised for complete stabilization (as shown on the left of Fig. 6). This strategy however is unfit for perceptual purposes, since the agent should be able to react to a sequence of stimuli, thus it has to fix a given input just for the time necessary to decide about, and then be ready for the next input.

We thus hypothesize that transient control of chaos is the strategy whereby a cognitive agent exploits previously learned resources in order to re-code the input arriving from the early sensory stages. Re-coding over a transient interval can be done in multiple ways. The selected way is the

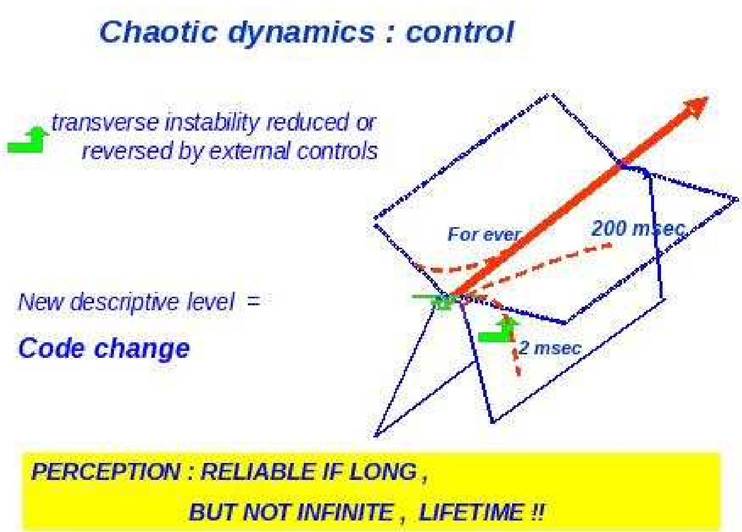


Figure 5. Chaos is controlled by adding extra-dynamic variables, which change the transverse instability without affecting the longitudinal trajectory. In the perceptual case, the most suitable top-down signals are those which provide a synchronized neuron array with an information lifetime sufficient to activate successive decisional areas (e.g. 200 ms), whereas the single HC neuron has a chaotic lifetime of 2 ms. If our attentional-emotional system is excessively cautious, it provides a top-down correction which may stabilize the transverse instability for ever, but then the perceptual area is blocked to further perceptions.

one that best fits the interaction of the agent with its environment. This hypothesis elaborated in the following sub-sections.

3.2. Stabilizing a perception (sensorial input)

From a dynamical point of view, the single neuron is a chaotic system. Even though the electrical spikes are all equal to each other, their time separation is chaotic. As a general consideration, a living being is a semiotic agent; by this we mean that the agent, embedded in an environment, changes the descriptive code in order to stabilize the cognitive dynamics and compensate for unavoidable information loss associated with the chaotic dynamics of its neurons. We have already stated that any conscious perception (we define conscious as that eliciting a decision) requires about 200 msec, whereas the loss of information in a chaotic train of neural spikes takes a few msec.

We have already seen Singer experiment on feature binding. An array of weakly coupled HC systems represents the simplest model for a physical realization of feature binding. The array can achieve a collective synchronized state lasting for a finite time (corresponding to the physiological 200 ms!) if there is a sparse (non global) coupling, if the input (bottom-up) is applied to just a few neurons and if the inter-neuron coupling is suitably adjusted(top-down control of chaos) [Arecchi 2004; Ciszak, Montina and Arecchi 2009 a,b]. The operation is a control of chaos, and it has an optimality ; if it lasts less than 200msec, no decisions emerge, on the contrary, if it lasts much longer, there is no room for sequential cognitive tasks. The addition of extra degrees of freedom implies a change of code, thus it can be seen as a new level of description of the same input.

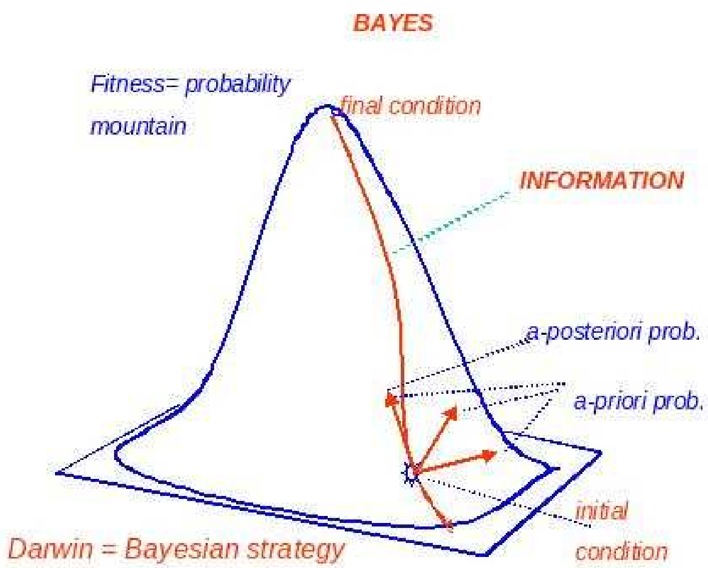


Figure 6. Successive applications of the Bayes theorem to the experiments. The procedure is an ascent of the Probability Mountain through a steepest gradient line. Each point of the line carries an information related to the local probability by Shannon formula. Notice that Darwinian evolution by mutation and successive selection of the best fit mutant is a sequential implementation of Bayes theorem.

3.3. From perception to cognition – Creativity

Considerations analogous to those developed for perception (elaboration of a sensorial input) can be made for cognition (conceptual input). We distinguish two types of cognitive task. In type I, we work within a prefixed framework and readjust the hypotheses at each new cognitive session, by a Bayes strategy. Bayes theorem [Bayes] consists of the relation:

$$P(h|data) = P(h)[P(data|h)/P(data)] \tag{1}$$

That is: the probability $P(h|data)$ of an hypothesis h, conditioned by the observed data (this is the meaning of the bar $|$), called a-posteriori probability of h, is the product of the a-priori probability $P(h)$ of that hypothesis (we assume to have guessed a package of convenient hypotheses with different probabilities), times the ratio of the probability $P(data|h)$ that data is generated by an hypothesis h, (this is the model) to the probability $P(data)$ of the effectively occurred data. When this ratio is largest, the a-posteriori probability is maximum. As shown in Fig. 6, starting from an initial observation and formulating a large number of different hypotheses, the one supported by the experiment suggests the most appropriate dynamical explanation. Going a step forward and repeating the Bayes procedure amounts to climbing a probability mountain along a steepest gradient line. Recent neurological studies [Beck 2008] explain the readiness of reflexes in terms of a fast coding done by a Bayesian strategy which operates very fast, in about 100 msec. Thus fast reactions are NOT the result of an interplay with previous semantic memories at the PCF level (Fig. 1) as modeled in ART. Fast indeed,but:how accurate? What confidence is provided by a Bayesian procedure? To answer the question, we must consider a new feature, namely **complexity**.

A complex problem is characterized by a probability landscape with many peaks, insofar as it can not be described by a single model (Fig. 7). Jumping from a probability hill to another

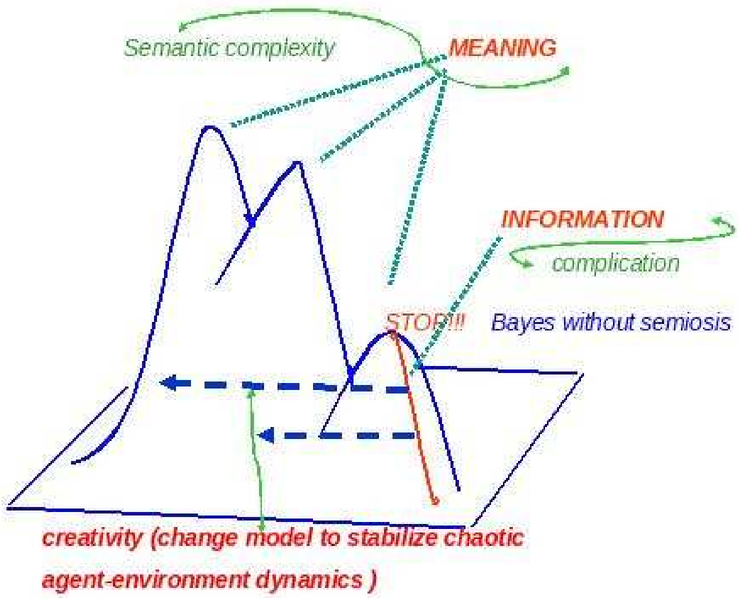


Figure 7. Semantic complexity A complex system is one with a many-peak probability landscape. The ascent to a single peak can be automatized by a steepest gradient program. On the contrary, to jump to other peaks, and thus continue the Bayes strategy elsewhere, is a creativity act, implying a holistic comprehension of the surrounding world (semiosis). We call meaning the multi-peak landscape and semantic complexity the number of peaks. It has been guessed that semiosis is the property that discriminates living beings from Turing machines [Sebeok]; here we show that a non-algorithmic procedure, that is, a non-Bayesian jump from one model to another is what we have called creativity. Semiosis is then equivalent to creativity .The difference between Bayesian strategy and creative jump is the same as the difference between normal science and paradigm shift [Kuhn]

is non- Bayesian; let us call it type II cognition.

In human cognition, type II is driven by hints suggested by the context (semiosis) yet not included in the model. Type II task is a creativity act because it implies a change of code, at variance with type I, which operates within a fixed code. The ascent to a single peak can be automatized in a steepest gradient program; once the peak has been reached, the program stops, any further step would be a downfall.

Question: can the changes of hill in Fig. 7 be handled by a computer? Answer: No. Indeed, as the computer behavior is perturbed by some added noise (as done in Monte Carlo simulations or in genetic algorithms) it explores the immediate surrounding of a Bayes uphill path. A finite step away from it would require a guideline, otherwise it would be whimsical. Thus, a non-deterministic computer cannot perform the jumps of type II, since it intrinsically lacks semiotic abilities. In order to do that, the computer must be assisted by a human operator. We call meaning the multi-peak landscape and semantic complexity the number of peaks.

This intrinsic limitation of a formalism was stated in two famous theorems, namely, the first incompleteness theorem for a consistent theory (Gödel, 1931) and its computer version (Turing, halting problem, 1936) (Fig. 8). Gödel's theorem, stating that there are propositions which appear true to the intuition (semiotic appreciation) of a mathematician endowed with the body of axioms, yet they can not be deduced via the syntax of the formalism, signed a sharp end to the formalistic dream of David Hilbert (1900), to explain all mathematics in terms of logic.

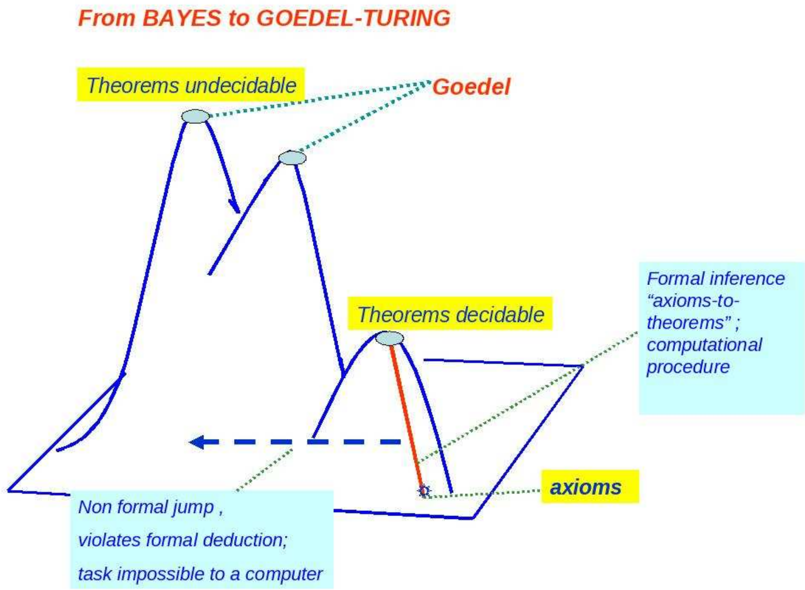


Figure 8. Gödel's first incompleteness theorem (1931): For any consistent formal, computably enumerable theory that proves basic arithmetical truths, an arithmetical statement that is true, but not provable in the theory, can be constructed. "Provable in the theory" means "derivable from the axioms and primitive notions of the theory, using standard first-order logic". Computer equivalent (Turing1936=halting problem): a universal computer, for a generic input, cannot decide to stop.

Since a universal computer is a syntactical machine which grinds the input data via its inner instructions, a similar limitation applies to a computational procedure (Turing). In the space in which we represent Bayesian procedures, the Goedel-Turing theorems are visualized as true propositions (peaks of left-ward hills), accessible from the axioms by a creative jump but not by via the formalism.

Let us discuss in detail the difference between a type I cognitive task, which implies changing hypothesis h within a model, that is, climbing a single mountain, and a type II cognitive task, which implies changing model, that is, jumping over to another mountain. We formalize a model as a set of dynamical variables x_i $(i = 1, 2, N)$, N being the number of degrees of freedom, with the equations of motion

$$\dot{x}_i = F_i(x_1..., x_N; \mu_1..., \mu_M) \tag{2}$$

Where F_i are the force laws and the M numbers represent the control parameters; The set $\{\mathbf{F}, \mathbf{x}, \mu\}$ is the model.

Changing hypotheses within a model means varying the control parameters $\{\mu\}$, as we do when exploring the transition from regular to chaotic motion within some model dynamics.

Instead, *changing code, or model*, means selecting different sets of degrees of freedom $\mathbf{y}$, control parameters μ and equations of motion $\mathbf{G}$ as follows:

$$\dot{y}_i = G_i(x_1..., x_R; \mu_1..., \mu_L) \tag{3}$$

Where R and L are different respectively from N and M. The set $\{\mathbf{G}, \mathbf{y}, \nu\}$ is the new model.

While changing hypotheses within a model is an a-semiotic procedure that can be automatized in a computerized *expert system*, changing model implies catching the meaning of the observed world, and this requires what is called *embodied cognition* [Varela]. Embodied cognition has been developed over thousands of generations of evolutionary adaptation, and we are unable so far to formalize it as an algorithm.

This no-go statement seems to be violated by a class of complex systems, which has been dealt with successfully by recursive algorithms. Let us consider a space lattice of spins, with couplings that can be either ferromagnetic or anti-ferromagnetic in a disordered, yet frozen way (spin glass at zero temperature, with quenched disorder). It will be impossible to find a unique ground state. For instance having three spins A, B, and C in a triangular lattice, if all of them have ferromagnetic interaction, then the ground state will consist of parallel spins, but if instead one (and only one) of the mutual coupling is anti-ferromagnetic, then there will be no satisfactory spin orientation compatible with the coupling (try with: A-up, B-up, C-up; it does not work; then try to reverse a single spin, it does not work either).

This model has a cognitive flavor, since a brain region can be modeled as a lattice of coupled neurons with coupling either excitatory or inhibitory, thus resembling a spin glass, [Hopfield; Amit; Toulouse et al]. We have a large number of possible ground states, all including some frustration. Trying to classify all possible configurations is a task whose computational difficulty (either, program length or execution time) diverges exponentially with the size of the system. Sequentially related changes of code have been successfully introduced to arrive at finite-time solutions. [Mezard et al; Solomon].

Can we say that the mentioned solutions realize the reductionistic dream of finding a suitable computer program that not only climbs the single probability hill, but also is able to chose the best hill? If so, the optimization problem would correspond to understanding the meaning of the object under scrutiny.

We should realize however that spin glasses are frozen objects, given once for ever. A clever search of symmetries has produced a spin glass theory [Mezard et al] that, like the Renormalization Group (RG) for critical phenomena [Wilson] discovers a recursive procedure for changing codes in an optimized way. Even though the problem has a large number of potential minima, and hence of probability peaks, a suitable insight in the topology of the abstract space embedding the dynamical system has led to an optimized trajectory across the peaks. In other words, the correlated clusters can be ordered in a hierarchical way and a formalism analogous to RG applied.

It must be stressed that this has been possible because the system under scrutiny has a structure assigned once for ever. In everyday tasks, we face a system embedded in an environment, which induces unpredictable changes in course of time. This rules out the nice symmetries of hierarchical approaches, and rather requires an adaptive approach. Furthermore, a real life context sensitive system has to be understood within a reasonably short time, in order to take vital decisions about it.

We find again a role of control of chaos in cognitive strategies, whenever we go beyond the limit of a Bayes strategy. We call creativity this optimal control of neuronal chaos. Fig. 9 sketches the reduction of complexity and chaos which results from a creative scientific step. Some examples of scientific creativity are listed in Fig. 10.

The difference here outlined between a Bayesian strategy, that can be implemented in an expert system, and a creative non-algorithmic jump is alike Kuhns distinction between normal science and paradigm shift [Kuhn].

The difference occurs also in translating a literary piece from a language to another. The translator does not trivially replace each word by the corresponding one of the new language. Rather he/she lives a given situation in the original language and then narrates it in the new language.

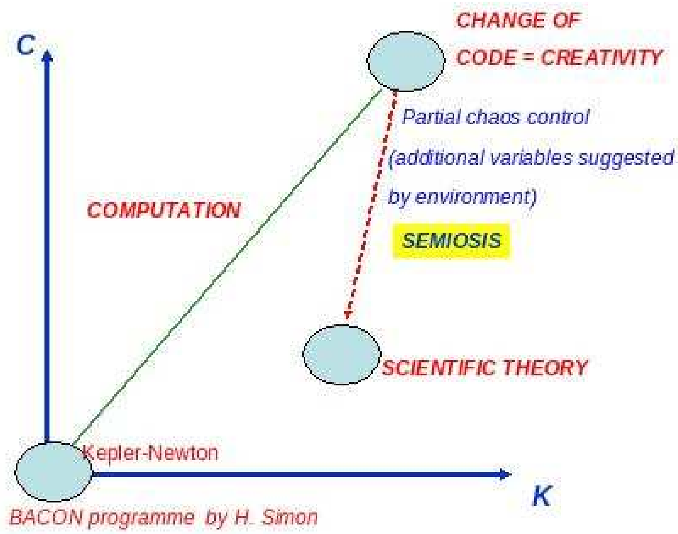


Figure 9. C-K diagram (C=computational complexity; K=information loss rate in chaotic motion): Comparison between the procedure of a computer and a semiotic cognitive agent (say: a scientist). The computer operates within a single code and C increases linearly with K. A scientist explores how adding different degrees of freedom one can reduce the high K of the single-code description. This is equivalent to the control operation of Fig. 6; it corresponds to a new model with reduced C and K. The BACON program [Simon] could retrieve automatically Keplers laws from astronomical data just because the solar system, approximated by Newton two-body interactions, is chaos-free.

1 - electricity; magnetism; optics	*Maxwell equations*
2 - Mendeleev table	*Quantum atom (Bohr,Pauli)*
3 - zoo of 100 elementary particles	*SU(3) - quarks (M Gell Mann)*
4 – scaling laws in phase transitions	*Renormalization group (K.Wilson)*

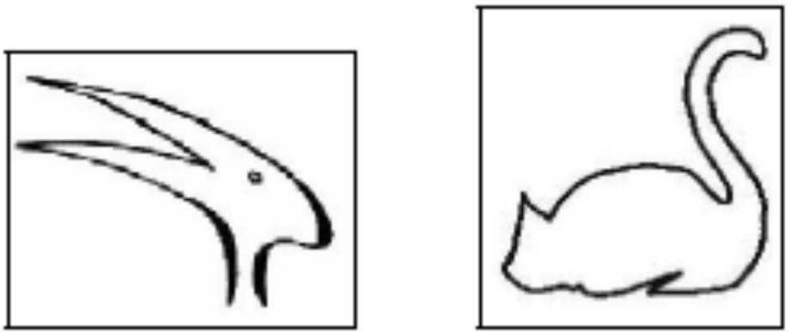
Figure 10. Reduction of complexity by code change.

4. Adaptive creative jumps vs. quantum approaches to mind

Quantum approaches to mind processes have a long story [Stapp]. I do not plan to survey the many different trials. I myself was puzzled by an uncertainty relation between the information

Journal of Physics: Conference Series **174** (2009) 012010 doi:10.1088/1742-6596/174/1/012010

handled in a brain task and the time duration of that task [Arecchi 2005]; but later I realized that the uncertainty was like a thermodynamical limitation that constrains the size of two correlated variables, without any quantum flavour. Indeed, the distinctive mark of a quantum dynamics is to be ruled by a wave-like behaviour which is responsible for interference.



From E Conte et al. arxiv:0710.5092

Figure 11. Two ambiguous figures: left, either rabbit or stork; right, either swan or cat.

A recent series of experiments has reported the alleged evidence of interference in brain processes [Conte et al], based on the following facts. Let us take two ambiguous figures (Fig. 11). The first one can be read either as a rabbit with long ears or as a stork with along beak; the second one either as a swan with a long neck or as a cat with a long tail. In standard experiments on bistable perceptions, a subject is exposed to just one of the two pictures and he/she reports the switching time between the two perceptions. For a standard subject, the average switching time is about 2 or 3 seconds, with strong statistical fluctuations around the average. Of course, over a long run there is equal likelihood for either one of the two perceptions, unless the experimenter introduces a bias which forces , say, the rabbit over the stork. This means that $p(x_1) = p(x_2) = 0.5$, where x_1 and x_2 are the two outcomes, and the probabilities are defined as Eq.(4), where the apex a(b) denote the left or the right figure, respectively.

$$p^a(x) = \frac{\text{the number of results a} = x}{\text{the total number of observations}},\ x \in X. \tag{4}$$

Let us now alternate the presentation of a with that of b. If we evaluate $p(x_1)$ conditional on perceiving y_1 or y_2 in the b presentation we would expect that summing over all y occurrences on the channel b we recover the previous unbiased value, as Eq.(5).

$$p^a(x) = p^b(y_1)p^{a|b}(x|y_1) + p^b(y_2)p^{a|b}(x|y_2),\ x \in X. \tag{5}$$

Instead ,the perturbation due to intercalating the other picture in the sequence of observations is expressed by an extra-term, which occurs with opposite signs for x_1 and x_2, thus unbalancing the two outcomes associated with the first figure. Eq.(5) must be replaced by Eq.(6)

$$p^a(x) = p^b(y_1)p^{a|b}(x|y_1) + p^b(y_2)p^{a|b}(x|y_2) + 2\lambda(a = x|b, C)\sqrt{p^b(y_1)p^{a|b}(x|y_1)p^b(y_2)p^{a|b}(x|y_2)} \quad (6)$$

The extra-term is like interference in QM; being limited between 0 and 1, it can be written as the cosine of an angle, hence justifying a Hilbert space picture. There is however a logical error. In fact, in physics two states of Hilbert space can be measured simultaneously by different observers; here instead, the two events are NOT synchronous.

In ambiguous patterns (bistable percepts) the two patterns do NOT form a superposition in the quantum sense, since they are NOT different states of the same Hilbert space; in fact, they appear successively in course of time, and one cannot speak of quantum interference. The extra-term in Eq.(6)implies a bias of the memory mechanisms, modifying the balance.

More generally, all authors who resort to quantum explanations are violating the Occam razor. Indeed, our previous explanation of non-algorithmic jumps as transient chaos control operations does not require any quantum assumption.

In a similar fashion, quantum decision approaches [Yukalov and Sornette] seem elegant formalisms implying extra assumptions with respect to those necessary to understand the decision process; thus they conflict with Occams economy principle.

5. Conclusions

From what we have discussed, the search for meaning, i.e. a creative decision, is NON-Bayesian.

The world is complex, that is , not grasped by a unique model .How to choose among alternative models? There are two ways:
i) arbitrarily (kind of oracle), this amounts to **relativism**;
or
ii) along some guidelines (fertility and latitude of the new explanation; useful outcomes); since any cognitive decision modifies (through our actions) the environment, we perform several trials with the aim of stabilizing the chaotic agent-environment dynamics. Thus, a **semiotic** approach to cognition discloses hard facts, it is a hint of an **ontology**.

The same re-coding freedom inherent in the transient stabilization of chaos holds for ethical decisions. But behold, our freedom is CONDITIONAL on the environmental features in which we are embedded. Let me quote two philosophers:
- Aristotle: thought and desire play a coordinate role (Nicomachean Ethics)
- Vico: poetic wisdom is the result of perception, memory and imagination
- "The New Science": our freedom is by NO means ABSOLUTE, as instead stated by
- Kant: freedom = independency from cause-effect chains = not-caused cause (Critique of Practical Reason).

The creative jump is not arbitrary, but guided by the situation in which we are embedded. We re-adjust the code until we find a satisfactory reading of the world around us.

Let us return to Fig. 7 but with a large number of peaks, to depict a complex situation (Fig. 12). The ascent to a single Bayes peak yields certainty, i.e. fidelity to a protocol, NOT truth but PAC (probably approximately correct) knowledge [Vidyasagar]. As we look for meaning, we jump non-algorithmically across several peaks, exploring a sizable region of a complex space; we speak of partial truth, from a certain point of view [Agazzi]. This is the only truth accessible to a scientific program; it is a modern formulation of the old definition of truth as adaequatio intellectus et rei [Thomas Aquinas]. Global truth is not accessible to us, as a complex universe is beyond our investigation power Would this mean a relativistic attitude? Not al all, since any starting point X has a pre-scientific motivation. It corresponds to a different choice of an investigation area, thus it corresponds to an ethical, better to say, ethological motivation, depending on personal motivations. For instance, in selecting an area of physics, I have devoted

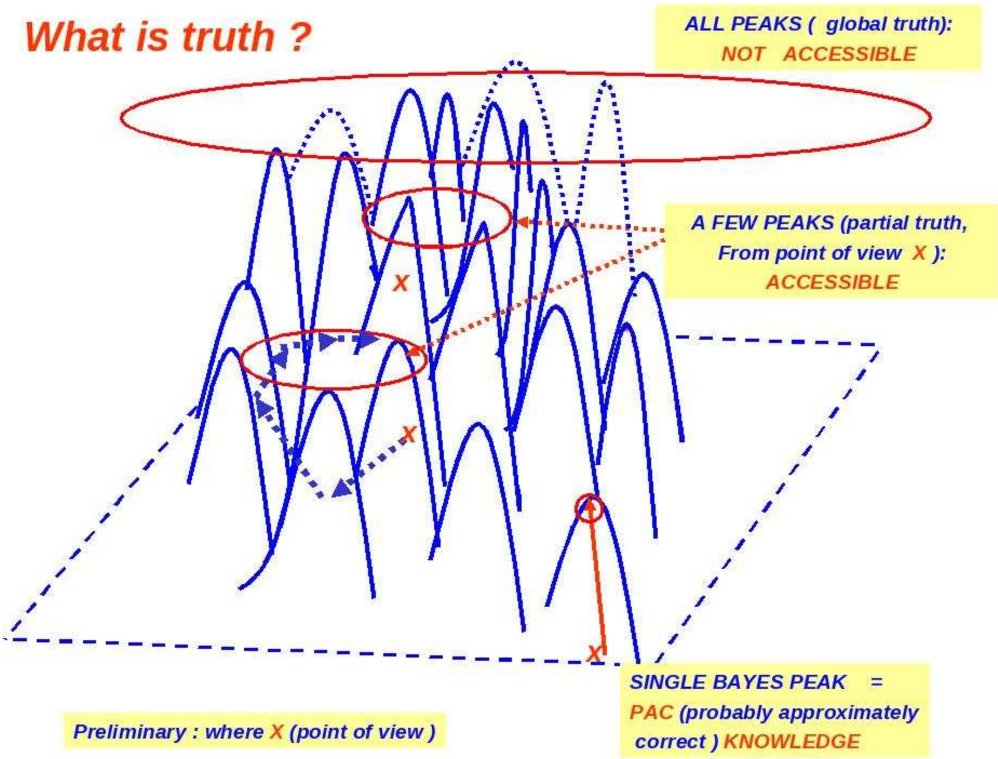


Figure 12. Comparisson of three procedures i) Climbing the single Bayes peak, leading to PAC Knowledge (what a computer does); ii) Different creative endeavors, starting from different X and exploring separate regions of a complex space = partial truth (what a creative mind does); iii) Exhausting all peaks of a complex problem, thus reaching a global truth (what nobody can do).

my efforts to quantum optics and chaotic phenomena, staying away from elementary particles and cosmology. Once made clear my X, I can compare my results (my partial truth) with that of other colleagues.

As a final remark, we compare the hermeneutic circle, peculiar of cognitive agents (animals or robots) with a limited repertoire, with what I call the hermeneutic spiral (Fig. 13) , which corresponds to the human language ability to make an infinite use (in course of time) of finite resources. The words polysemy implies that we in general attribute several meanings or connotations to the same word. If these connotations are frozen once for ever, as in a historical dictionary, then we are in the upper case. Suppose a grammatical operator takes from a sub-meaning of the word A to a sub-meaning of B; applying the inverse operator we recover the initial connotation of A, either in one or several steps. A neurobiological example is offered by the olfactory system of the locust [Laurent et al]. Exposed to a cherry odour, it responds with a specific sequence of neuronal spikes ; re-exposing the locust to the same odour at a later time, the odour is encoded by the same sequence.

If instead we do the experiment with a more sophisticated animal as a rabbit [Freeman], any successive exposure to the same odour is encoded in different spike patterns. The rabbit feels that some time has passed by and that it is experiencing the same perception but not for the first time. This spiral along time, whereby we never repeat exactly the same connotation, is peculiar of human creativity, in language (we make new poetry with the same lexicon) or music

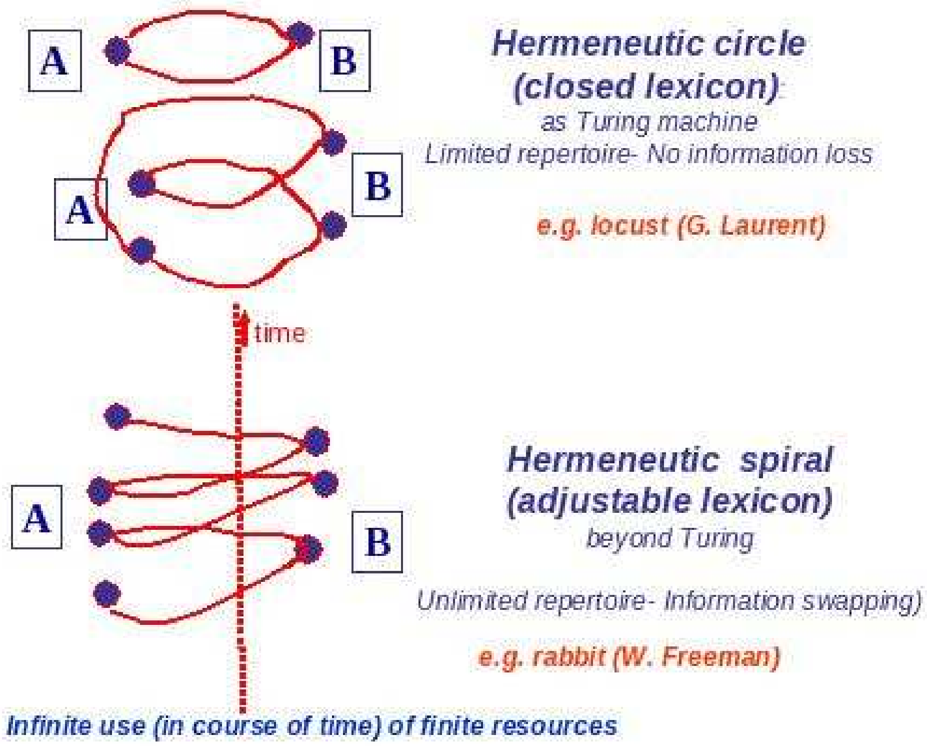


Figure 13. Difference between hermeneutic circle and hermeneutic spiral. Let us consider an algorithm which takes from an item A to an item B. It the items are words of an ordinary language, they are polysemic, that is both A and B do not have a unique connotation. Selecting a specific connotation of A the algorithm arrives at a specific connotation of B. By reversing the algorithmic procedure, one returns at the original connotation of A either in one loop (uppermost figure) or in a finite number of loops (central figure).These cases correspond to a cognitive agent with a limited repertoire and no information loss. If, as decribed in non-Bayesian cases, there information swapping, i.e. re-coding, the number of connotations of the words A and B is for all purposes unlimited. As time flows ,there is a semiotic re-adjustment between cognitive agent and environment which yields new connotations for the same lexical terms, as it occurs in poetry and music.

(we make new pieces with the same notes). Indeed ,it is our common experience that revisiting after a while a piece dear to us, e.g. a Bach Cantata, we discover new meanings, since our inner universe has grown in the meantime.

Acknowledgments
The Author acknowledges the financial support of Ente Cassa di Risparmio di Firenze under the project "dinamiche cerebrali caotiche".

References
Since the topics of this review track my research line along the years, quite a few considerations can be found scattered in my research papers. My publications can be downloaded from my homepage: www.inoa.it/home/arecchi, see: List of publications - Research papers in Physics
Furthermore, these topics have been dealt with extensively in the book:

- Arecchi F T 2007 *Coerenza, Complessità, Creatività* (Roma: S Di Renzo)
All other references are listed here below in alphabetic order:
- Agazzi E 1974 *Temi e problemi di Filosofia della fisica* (Roma: Edizioni Abete)
- Amit D J, Gutfreund H and Sompolinski H 1985 *Phys. Rev* A **32** 1007
- Arecchi F T 2004 *Physica* A **338** 218-237
- Arecchi F T 2005 *Braz. J. Phys.* **35** 253
- Arecchi F T and Meucci R 2008 *Scholarpedia* **3**(9) 7066
- Bayes T 1763 *Phil. Trans. Royal Soc. (London)* **53** 370-418
- Beck J M *et al* 2008 *Neuron* **60** 1142-1152
- Chaitin G J 1987 *Algorithmic information theory* (Cambridge, UK: Cambridge U. Press)
- Ciszak M, Montina A and Arecchi F T 2009 a *Cognitive Processing* **10**(1) 33-39
- Ciszak M, Montina A and Arecchi F T 2009 b Control of transient synchronization with external stimuli *Chaos* **19** to appear
- Conte E *et al* 2008 Mind states follow quantum mechanics during perception and cognition of ambiguous figures *Preprint* arXiv:0802.1835
- Freeman W J 1991 *Scientific American* **264**(2) 78-85
- Grossberg S 1995 *The American Scientist* **83** 439
- Hopfield J J 1982 *Proc. Nat. Acad. Sci. (USA)* **79** 2554
- Hubel D H 1995 *Eye, Brain and Vision* Scientific American Library, No. **22** (New York: W. H. Freeman)
- Kuhn T S 1962 *The Structure of Scientific Revolutions* 1st ed (Chicago: U. of Chicago Press)
- Laurent G, Wehr M and Hananel Davidowitz H 1996 *Journ. Neurosci.* **16** 3837
- Libet B, Wright E W, Feinstein F and Pearl D K 1979 *Brain* **102** 193
- Mezard M, Parisi G and Virasoro M 1987 *Spin glass theory and beyond* (Singapore: World Scientific)
- Rodriguez E, George N, Lachaux J P, Martinerie J, Renault B and Varela F 1999 *Nature* **397** 340343
- Sebeok T A 2001 *Semiotica* **1341**/4 61-78
- Shilnikov A, Shilnikov L and Turaev D 2004 *Int. J. Bif. And Chaos* **14** 2143
- Simon H A 1980 *Cognitive Science* **4** 33-46
- Singer W and Gray E C 1995 *Annu. Rev. Neurosci.* **18** 555
- Solly D R *et al* 2007 *Nature* **450** 1054
- Solomon S 1995 *Ann. Rev. of Comp. Physics II* (Singapore: World Scientific) 243-294
- Stapp H 1993 *Mind, matter and quantum mechanics* (Berlin: Springer)
- Thomas Aquinas 1992 *Quaestiones de Veritate* (Bologna: Ed. Studio Domenicano)
- Toulouse G, Dehaene S and Changeux J P 1986 *Proc. Nat. Acad. Sci. (USA)* **83** 1695
- Varela F, Thompson E and Rosch E 1991 *The Embodied Mind* (Cambridge, MA: MIT Press)
- Vidyasagar M 1997 *A Theory of Learning and Generalization, with Applications to Neural Networks and Control Systems* (Berlin: Springer)
- Wilson K 1983 *Rev. Mod. Phys.* **55** 583
- Yukalov V I and Sornette D 2008 *J. Math. Phys.* to appear (*Preprint* arXiv:0808.0112)