Conference Proceedings Paper – Entropy

A Cybernetics Update for Competitive Deep Learning System

Rodolfo A. Fiorini

Politecnico di Milano, Department of Electronics, Information and Bioengineering (DEIB), Milano, Italy; E-Mail: rodolfo.fiorini@polimi.it; Tel.: +039-02-2399-3350; Fax: +039-02-2399-3360.

Abstract: A number of recent reports in the peer-reviewed literature have discussed irreproducibility of results in biomedical research. Some of these articles suggest that the inability of independent research laboratories to replicate published results has a negative impact on the development of, and confidence in, the biomedical research enterprise. To get more resilient data and to achieve higher result reproducibility, we present an adaptive and learning system reference architecture for smart learning system interface. To get deeper inspiration, we focus our attention on mammalian brain neurophysiology. In fact, from a neurophysiological point of view, neuroscientist LeDoux finds two preferential amygdala pathways in the brain of the laboratory mouse to experience reality. Our operative proposal is to map this knowledge into a new flexible and multi-scalable system architecture. Our main idea is to use a new input node able to bind known information to the unknown one coherently. Then, unknown "environmental noise" or/and local "signal input" information can be aggregated to known "system internal control status" information, to provide a landscape of attractor points, which either fast or slow and deeper system response can computed from. In this way, ideal system interaction levels can be matched exactly to practical system modeling interaction styles, with no paradigmatic operational ambiguity and minimal information loss. The present paper is a relevant contribution to fourth generation adaptive learning and machine intelligence systems to update classic cybernetics towards a new post-Bertalanffy General Theory of Systems.

Keywords: deep learning; cybernetics; social safety and security; computer-brain interface; anticipatory system; computational information conservation theory; CICT; post-Bertalanffy cybernetics; combinatorial optimization.

1. Introduction

Investigations of hypotheses with the aid of models lead to specific predictions that have to be verified experimentally. The comparison of model predictions with experimental data can then be used to refine the hypotheses and to develop more accurate models, or even models that can shed light on different phenomena. Nevertheless, a number of recent reports in the peer-reviewed literature [1-4] have discussed irreproducibility of results in biomedical research. Some of these articles suggest that the inability of independent research laboratories to replicate published results has a negative impact on the development of, and confidence in, the biomedical research enterprise. Furthermore, poor reporting of health research is a serious and widespread issue, distorting evidence, limiting its transfer into practice, and providing an unreliable basis for clinical decisions and further research. A series of papers published by the Lancet [5] in January 2014 highlighted the problems of waste in biomedical research and the myriad of issues that can disrupt completion and use of high quality research. The great potential of nano-bio-technology is based also on the ability to deal with complex hierarchically structured systems from the macroscale to the nanoscale; this requires novel theoretical approaches and the competence to create models able to explain system dynamics in multi-scale simulation. In all cases, from a top-down (TD) point-of-view (POV), the length scale ranges from the centimeter scale (*cm*) (macroscale, organs), to the micrometer scale (μ*m*) (cells and local connectivity), to the nanometer scale (*nm*) (cell subcomponents, proteins) till to single-molecule scale (nano-picoscale).

Certainly, statistical and probabilistic theory, applied to all branches of human knowledge under the "continuum hypothesis" assumption, have reached highly sophistication level, and a worldwide audience. It is the core of classic scientific knowledge; it is the traditional instrument of risk-taking. Many "Science 1.0" researchers and scientists up to scientific journals assume it is the ultimate language of science. The basic framework of statistics has been virtually unchanged since Fisher, Neyman and Pearson introduced it. Later, the application of geometry to statistical theory and practice has produced a number of different approaches. As an example, in 1945, by considering the space of probability distributions, Indian-born mathematician and statistician Calyampudi Radhakrishna Rao (1920-) suggested the differential geometric approach to statistical inference. In this way the so called "Information Geometry" (IG) approach was born [6]. So, the modern geometric science of information (GSI) emerged from the study of the geometrical structure of a manifold of probability distributions under the criterion of invariance [7].

Sometimes, IG tensorial formulation may become cumbersome and computational difficulties do occur: in such problems maximum-likelihood estimators (MLEs) are unsuitable or do not exist. Computational information geometry (CIG) use a top-down (TD) point-of-view (POV). All approaches using a TD POV allow for starting from an exact solution panorama of global analytic solution families, which always offers a shallow downscaling local solution computational precision compromise to real specific needs (overall system information from global to local POV is not conserved [8]). Usually it is necessary further analysis and validation to get localized computational solution of any practical value. To develop resilient and antifragile application, we need stronger biological and physical system computational correlates; we need asymptotic exact global solution panoramas combined to deep local solution computational coherent precision for information conservation and vice-versa.

Early biological concepts based on reductionist inference systemic and classic cybernetic approaches have been largely revisited and overshadowed by more recent molecular and pathogenetic findings for healthcare security and safety management, creating a brand new cultural approach. Nevertheless, the change of paradigm caused by QFT has not yet been completely grasped by many contemporary scientific disciplines and current researchers, so that not all the implications of this big change have been realized hitherto, even less their related, vital applications. So, the discreteness approach, developed under the QT (quantum theory) "discreteness hypothesis" assumption, has been considered in peculiar application areas only. It has been further slowly developed by a few specialists and less understood by a wider audience.

Every approach that uses analytical function applies a TD POV implicitly. These functions belong to the domain of Infinitesimal Calculus (IC). Unfortunately, from a computational perspective, all approaches that use a TD POV allow for starting from an exact global solution panorama of analytic solution families, which offers a shallow local solution computational precision to real specific needs (in other words, from global to local POV overall system information is not conserved, as misplaced precision leads to information dissipation [8]). In fact, usually further analysis and validation (by probabilistic and stochastic methods) is necessary to get localized computational solution of any practical value, in real application. A local discrete solution is worked out and computationally approximated as the last step in their line of reasoning, that started from an overall continuous system approach (from continuum to discrete ≡ TD POV). Unfortunately, the IC methods are NOT applicable to discrete variable. To deal with discrete variables, we need the Finite Differences Calculus (FDC). FDC deals especially with discrete functions, but it may be applied to continuous function too. As a matter of fact, it can deal with both discrete and continuous categories conveniently.

In other words, if we want to achieve an overall system information conservation approach, we have to look for a convenient bottom-up (BU) POV (from discrete to continuum view ≡ BU POV) to start from first, and NOT the other way around! Then, a TD POV can be applied, if needed. Deep epistemic limitations reside in some parts of the areas covered in probability and risk analysis and decision making applied to real problems [9]. As a matter of fact, to grasp a more reliable representation of reality, researchers and scientists need two intelligently articulated hands: both stochastic and combinatorial approach synergic ally articulated by natural coupling [9]; let's say we need a fresh "Science 2.0" approach. We just have to remember the Relativity's father inspiration quote: "We cannot solve our problems with the same thinking we used when we created them." The current paper can give a relevant contribute to that perspective to achieve practical operative results quite quickly. Section 2. offers a brief review about most successful recent deep learning systems. Section 3. Focus on major differences between biological neural net vs. artificial neural net. Section 4. Offers new updating on learning from neuroscience. The new computational information conservation (CICT) approach is presented in Section 5. In Section 6. ODR approach is proposed as a cybernetics update. Section 7. gives an full operative example by CICT recurrence sequence representation. In Section 8. a final post-Bertalanffy cybernetics update is presented and discussed. To show the full flexibility of our approach even at multidisciplinary arbitrary scale modeling, in Section 9. a simple example of advanced management architecture for HRO is given. Last Section 10. sums up our findings from a top level point of view.

2. Deep Learning Brief Overview

Deep learning (DL) is part of a broader family of machine learning (ML) methods based on learning representations of data (DLR). Some representations make it easier to learn tasks from examples. Or instance, one of the promises of deep learning is replacing handcrafted features with efficient algorithms for unsupervised or semi-supervised feature learning and hierarchical feature extraction [10]. Recently, DL is making important strides in natural language processing, especially statistical machine translation [11],[12],[13]. Interestingly, one of the key factors that enabled this major progress has been the advent of Graphics Processing Units (GPUs), with speed-ups on the order of 10 to 30-fold, starting with [14], and similar improvements with distributed training [15],[16]. Various DL architectures such as deep neural networks (DNNs), convolutional deep neural networks (CNNs), deep belief networks or dynamic Bayesian networks (DBNs) and recurrent neural networks (RNNs) have been applied to fields like computer vision, automatic speech recognition, natural language processing, audio recognition and bioinformatics where they have been shown to produce state-of-the-art results on various tasks. DNN have substantially pushed the state-of-the-art in a wide range of tasks, especially in speech recognition [17],[18] and computer vision, notably object recognition from images [19],[20]. DNN have achieved state-of-the-art results in a wide range of tasks, with the best results obtained with large training sets and large models. CNN [21],[22] have become the method of choice for object recognition [23]. They have proved to be successful at a variety of benchmark problems including, but not limited to, natural image classification [24],[25] handwritten digit recognition [26], traffic sign recognition [27], house number recognition [28], as well as for speech recognition [29-32]. Furthermore, video imaging tasks like image caption generation [33],[34], video description generation [35] and object localization/detection [36] have been found to be extremely benefitted by image representations from CNNs trained to recognize objects on a large set of more than one million images [20],[37]. Learning generative models of sequences is a long-standing ML challenge and historically the domain of DBNs such as hidden Markov models (HMMs) and Kalman filters. The dominance of DBN-based approaches has been recently overturned by a resurgence of interest in RNN based approaches. An RNN is a special type of neural network that is able to handle both variable-length input and output. By training an RNN to predict the next output in a sequence, given all previous outputs, it can be used to model joint probability distribution over sequences. Unfortunately, for most DL algorithms training is notoriously time consuming. Since most of the computation in training neural networks is typically spent on floating point multiplications, latest research is focused to investigate different approaches to training that eliminates the need for most of these, like in [38]. Indeed, the ability to train larger models on more data has enabled the kind of breakthroughs observed in the last few years. In the past, GPUs enabled these breakthroughs because of their greater computational speed. Today, researchers and developers designing new deep learning algorithms and applications often find themselves limited by computational capability. This along, with the drive to put deep learning systems on low-power devices (unlike GPUs) is greatly increasing the interest in research and development of specialized hardware for deep networks [39-41]. In the future, faster computation at both training and test time is likely to be crucial for further progress and for consumer applications on low-power devices. As a result, there is much interest in research and development of dedicated hardware for DL.

Most of current DLR are stochastically based and some of ML representations for artificial neural networks (ANNs) are inspired by advances in neuroscience. They are loosely based on interpretation of information processing and communication patterns in a nervous system or biological neural networks

(BNNs), such as neural coding with still a reductionist interpretation (Science 1.0). For instance, many attempts to define a relationship between the stimulus and the neuronal responses to arrive to the relationship among the electrical activity of the neurons in the brain, according to neuroscience research and knowledge till the 1980s [42]. Understanding the human neural coding and brain remains a major challenge of the 21st century [43]. Many groups worldwide are tackling the challenge of accelerating our understanding of the human brain through large-scale data management, mining, knowledge management, collaboration, modeling and simulation. We know that stiff and nonlinear dynamical systems are inefficient on a digital computer. A simple example is the IBM Blue Gene project with 4096 CPUs and 1000 Terabytes RAM, which, to simulate the mouse cortex uses 8 x 10^6 neurons, 2 x 10^{10} synapses, 10^9 Hz clock, 40 Kilowatts on digital support. The brain uses 10^{10} neurons, 10^{14} synapses, 10 Hz clock, and 20 watts [44] on biological support. Biological information processing systems operate on completely different principles from those with which most engineers are familiar [45]. For many problems, particularly those in which the input data are ill-conditioned and the computation can be specified in a relative manner, biological solutions are many orders of magnitude more effective than those we have been able to implement using traditional digital methods. This advantage can be attributed principally to the use of elementary physical phenomena as computational primitives, and to the representation of information by the relative values of analog signals, rather than by the absolute values of digital signals. This approach requires adaptive techniques to mitigate the effects of component differences. This kind of adaptation leads naturally to systems that learn about their environment. Shortly, major current DL approaches limitations are:

1- Need for best architecture selection for a specific task;

2- Computational intensive requirement for best result;

3- Stochastic based clustering only;

4- Time Consuming Training.

To minimize or overcome previous major system limitations and to arrive much closer to fourth generation adaptive learning and real machine intelligence systems, in this paper, we propose the adoption of a new input information aggregation system block, called "recursive interactor" (ReInt).

3. Biological Neural Net vs. Artificial Neural Net

There is a myth that the nervous system is slow, is built out of slimy stuff, uses ions instead of electrons, and is therefore ineffective. As a matter of fact, there are about 10^{16} synapses in the human brain. A nerve pulse arrives at each synapse about ten times/s, on average. So in rough numbers, the brain accomplishes 10^{16} complex operations/s. The power dissipation of the brain is a few watts, so each operation costs only 10^{-16} J. The brain is at least a factor of 1 billion more efficient than our present digital technology, and a factor of 10 million more efficient than the best digital technology that we can imagine. From the first integrated circuit in 1959 until today, the cost of computation has improved by a factor about 1 million. We can count on an additional factor of 100 before fundamental limitations are encountered. At that point, a state-of-the-art digital system will still require 10 MW to process information at the rate that it is processed by a single human brain (about 20 W requirement). The

unavoidable conclusion is that we have something fundamental to learn from the brain about a new and much more effective form of computation. The disparity between the efficiency of computation and learning in the nervous system (BNNs) and that in a computer (ANNs) is primarily attributable not to the individual device requirements, but rather to the way the devices are used in the system. The fact that we can build devices that implement the same basic operations as those the nervous system uses leads to the inevitable conclusion that we should be able to build entire systems based on the organizing principles used by the nervous system. We need to integrate our current neural coding reductionist interpretation with new neuroscience insights [46],[47]. We will refer to these new systems generically as neuromorphic anticipatory learning system (ALS) [48]. Then, we can think of a DL ANN general architecture for DL and recognition, based on a network of ReInts that can replace the ubiquitous convolution+pooling layer of the usual deep CNN. ReInt, as basic block for ALS, allows to develop even multiple dynamic network self-rewiring strategies to strengthen emerging computational results with minimal computational burden. ReInt allows to match exactly ideal system interaction levels to practical system modeling interaction styles, with no paradigmatic operational ambiguity and minimal information loss [49]. ReInt concept is a relevant contribute to update classic cybernetics knowledge towards a new post-Bertalanffy General Theory of Systems, where systemic concepts like self-organization, distributed (self-)control, i.e. bottom-up self-regulating system, self-governance, etc. are at work [49]. That is the main reason why this approach can look quite unusual to ML community, with no apparently significant concept by classic DL point of view (we have to learn to use new eyes and to update our DL understanding). There is nothing that is done in the nervous system that we cannot emulate with electronics if we understand the principles of neural information processing right. What kind of computation primitives are implemented by the device physics we have available in nervous tissue or in a silicon integrated circuit? We have to remember, to rediscover or to invent a convenient representation that takes advantage of the inherent capabilities of the medium, such as the abilities to generate exponentials, to do integration with respect to time, and to implement a zero-cost addition using Kirchhoff's law [44],[45]. Therefore, in order to get better results, current computational neuroscience modeling has to face and to overcome two orders of issues at least, immediately:

1- To minimize the traditional limitation of current digital computational resources that are unable to capture and to manage even the full information content of a single Rational Number Q leading to information dissipation and opacity [50],[51].

2- To develop stronger, more effective and reliable neural correlates by the correct arbitrary multi-scale (AMS) modeling approach for complex system [45],[49],[52];

ReInt with ODR approach [8], from computational information conservation (CICT) [53], can give a sound answer to both issues and to achieve greater operating flexibility in multi-scale learning and modeling, with much less computational requirement than current systems. We describe ReInt basic computational properties and give operative examples.

4. Learning from Neuroscience

To get deeper inspiration towards a strong solution to our problem, we focus our attention on mammalian brain neurophysiology. In fact, from a neurophysiological point of view, neuroscientist

Joseph E. LeDoux finds two amygdala pathways in the brain of the laboratory mouse by the use of fear conditioning and lesion study [46],[47]. Although most of the research on the neural basis of conditioned fear has been conducted on animals, fear conditioning procedures can be used in identical ways in humans, according to LeDoux [46]. Information about external stimuli reaches the amygdala by way of direct pathways from the thalamus (the "low road") as well as by way of pathways from the thalamus to the cortex to the amygdala (the "high road"). The direct thalamo-amygdala is a shorter and thus a faster transmission route than the pathway from the thalamus through the cortex to the amygdala. However, because the direct pathway bypasses the cortex, it is unable to benefit from cortical processing. As a result, it can only provide the amygdala with a crude representation of the stimulus. It is thus a quick and dirty processing pathway. The direct pathway allows us to begin to respond to potentially dangerous stimuli before we fully know what the stimulus is. This can be very useful in dangerous situations. However, its utility requires that the cortical pathway be able to override the direct pathway. It is possible that the direct pathway is responsible for the control of emotional responses that we do not understand. The time saved by the amygdala in acting on the thalamic information, rather than waiting for the cortical input, may be the difference between life and death. It is better to have treated a stick as a snake than not to have responded to a possible snake. Most of what we know about these pathways has actually been learned by studies of the auditory as opposed to the visual system, but the same organizational principles seem to apply. The low road (Fig.1, path A) is a pathway which is able to transmit a signal from a stimulus to the thalamus, and then to the amygdala, which then activates a fear-response in the body. This sequence works without a conscious experience of what comprises the stimulus, and it is the fast way to a bodily response (a more primitive mechanism of defence). The high road (Fig.1, path B) is activated simultaneously. This is a slower road which also includes the cortical parts of the brain, thus creating a conscious impression of what the stimulus is (a more sophisticated mechanism of defence). "Amygdala hijack" is a term coined by psychologist D. Goleman [54]. Drawing on the work of Joseph E. LeDoux, Goleman uses the term to describe emotional responses from people which are immediate and overwhelming, and out of measure with the actual stimulus because it has triggered a much more significant emotional threat.
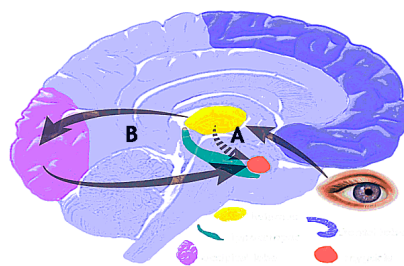

Figure 1. Information about external stimuli reaches the amygdala by way of direct pathways from the thalamus (the "low road", A) as well as by way of pathways from the thalamus to the cortex to the amygdala (the "high road", B) [46],[47],[54].

From the thalamus, a part of the stimulus goes directly to the amygdala (low road) while another part is sent (high road) to the neocortex (the "thinking brain"). If the amygdala perceives a match to the stimulus, i.e., if the record of experiences in the hippocampus tells the amygdala that it is a fight, flight or freeze situation, then the Amygdala triggers the HPA (Hypothalamic-Pituitary-Adrenal) axis and

hijacks the rational brain. This emotional brain activity processes information milliseconds earlier than the rational brain, so in case of a match, the amygdala acts before any possible direction from the neo-cortex can be received. If, however, the amygdala does not find any match to the stimulus received with its recorded threatening situations, then it acts according to the directions received from the neo-cortex. When the amygdala perceives a threat, it can lead that person to react irrationally and destructively. Taking into consideration the neurophysiological findings by LeDoux, differently from the past, it is much better to consider ontological uncertainty [52] as an emergent phenomenon out of a complex system, arriving to the basic schema for Ontological Uncertainty Management (OUM) System [52]. Then, our dynamic ontological perspective can be thought as an emergent, natural trans-disciplinary reality level (TRL) [55] out of, at least, a dichotomy of two fundamental coupled irreducible complementary ideal asymptotic concepts: A) reliable unpredictability and B) reliable predictability [52]. From TD management perspective, the reliable unpredictability concept can be associated to system proactive approach (lead subsystem, Open Logic, to survive and grow, Fig.2) and strategic management techniques. The reliable predictability concept can be referred to traditional system reactive approach (lag subsystem, Closed Logic, to learn and prosper, Fig.2) and operative management techniques. To achieve our final goal, overall system must be provided with smart sensing interface which allow reliable real-time interaction with its environment [56]. To behave realistically, system must guarantee both Logical Aperture (to survive and grow) and Logical Closure (to learn and prosper), both fed by environmental "noise" (better… from what human beings call "noise") [8].
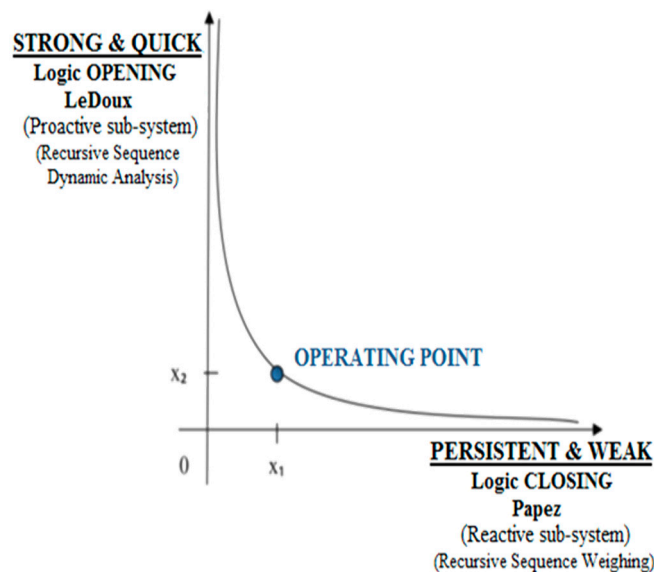

Figure 2. Basic schema for Ontological Uncertainty Management (OUM) System. Operating Point can emerge as a new Trans-disciplinary Reality Level (TRL) [55] from the mutual interaction of Two Complementary Irreducible Management Subsystems with their common environment (see text) [48].

So, according to previous considerations, at brain level, it is possible to refer to LeDoux circuit (Fig.2 path A, "low road", Logical Aperture) for emotional behaviour (i.e. fear, emotional intelligence, etc.) and to Papez circuit (Fig.2 path B, "high road", Logical Closure) for structured behaviour (i.e. rational thinking, knowledge extraction, etc.…) as from Fig.2 [46,47]. Emotional Intelligence (EI) and Emotional

Creativity (EC) [54] coexist at the same time with Rational Thinking in human mind, sharing the same input environment information [56]. Therefore, the mathematical method we are looking for, must possess even anticipatory properties, at design level, needed to the realization of system capable to interact with its environment in real time (leading property) [57]. The notion of anticipation is coming to the foreground as an emerging field of study. The future plays an active way in how we think and act in the present. The traditional understanding, that past events are the primary drivers that influence how we understand the present, is undermined. Both the past and the future are forces that simultaneously and actively influence the present. By interpreting the present as the time where the forces of the past and future meet, our understanding of the present changes from a "thin" (the present as a boundary without any extension between past and future) to a "thick present" (the present as the collection of contemporaneous events). Moreover, by giving the future scientific legitimacy, a novel vision of science arises where a fully scientific (i.e., not allusive, metaphorical or mystical) treatment of "final" causation (= anticipation) is included and not rejected (Science 2.0) as is the case in the traditional scientific paradigm (Science 1.0). An in-depth understanding of the work of theoretical biologist, Robert Rosen (1934–1998), holds important insights for how anticipatory systems can be modeled [57]. Rosen's insight that "science is the art of establishing modelling relations between the natural world and the world of our formalisms" challenges traditional modeling strategies that mainly form simulations of reality, but do not explain causal relations. According to Rosen, the modeling relation (or the main task of "theoretical science") consisted of establishing congruences between "causal relations in the external world, and implicative relations between propositions describing that world." Essentially the mapping relation points to the process we carry out when we "do science" and exposes this process as one in which there can be no biggest model of the world, but only snap-shots thereof. The study of complexity and anticipation can be linked to the modeling relation [57]. The acknowledgement of complexity lays bare the dilemma that there remains a gap between our models and the reality they intend to describe. An irreducible difference exists between the nature of complex reality and our descriptions thereof. By acknowledging that all knowledge of complex, anticipatory systems will always prove to be partial knowledge, one is confronted with the unavoidability of the limitations of human understanding. This recognition opens up a space where the conceptual implications of complexity surpasses epistemological concerns and exposes the normativity that lies in all our modeling strategies. This ethical imperative challenges scholars to engage with the question of re-thinking what it means to be human and calls upon us to proceed differently in this world. Anticipation can be used to proceed differently in the process of "working" with the future when corporate businesses or governments have to come to terms with complexity, risk and uncertainty. Horizon scanning and scenario planning offer the current best futures studies tools for making sense of how one could anticipate the future and make better decisions. Because we cannot have a biggest or best model of the future, it means that futurists cannot predict the future. Instead, their task is to rather help find ways to understand the critical driving forces and uncertainties in the (business) environment and to use this almost BU information to make strategic decisions. By a systemic POV, the logical answer is to design and to use distributed (self-)control, i.e. bottom-up self-regulating systems. Cybernetics (i.e. advanced control theory) and complexity theory tell us that it is actually feasible to create resilient social and economic order by means of self-organization, self-regulation, and self-governance [58],[59]. Scenario planning can be used as a tool for exploring system sustainability transitions. Through anticipatory scenario planning strategies, a more positive vision of

what system interaction evolution could look like, can be developed. Possible change should be scoped out so as to be better prepared to respond to change and surprise and to help influence and drive change along more desirable trajectories, as well as avoid undesirable trajectories.

5. CICT Approach

Recently, CICT (computational information conservation theory) has shown that, by Shannon entropy approach only, even the current, most sophisticated instrumentation system is completely unable to reliably discriminate so called "random noise" (RN) from any combinatorically optimized encoded message, which CICT called "deterministic noise" (DN) [8]. Stochastic vs. Combinatorically Optimized Noise generation ambiguity emphasises the major "information double-bind" (IDB) problem in current most advanced research laboratory and instrumentation system, just at the inner core of human knowledge extraction by experimentation in current science [8]. CICT is a natural framework for arbitrary multi-scale biomedical engineering, computer science and systems biology computational modeling in the current landscape of modern QFT [53]. CICT new awareness of a discrete HG (hyperbolic geometry) subspace (reciprocal space) of coded heterogeneous hyperbolic structures [8], underlying the familiar Q Euclidean (direct space) surface representation can open the way to "holographic information geometry" (HIG) [50],[53]. CICT emerged from the study of the geometrical structure of a discrete manifold of ordered hyperbolic substructures, coded by formal power series, under the criterion of evolutive structural invariance at arbitrary precision [50],[53]. CICT interprets natural rational "OpeRational" (OR, [50] for definition) representation sequences as a language of languages of phased directed number systems quite easily. In fact, we can take the concepts of modular magnitude and direction as basic, and introduce the concept of vector as the basic kind of directed number, with an associated phasing relation. Directed numbers are defined implicitly by specifying rules for adding and multiplying vectors. Furthermore, they can be related uniquely to their remainder sequences to identify "quantum support field" sequences, which subspace inner phased generators can be computed from [60]. Traditional elementary arithmetic long division remainder sequences can be interpreted as OECS (Optimally Encoded Cyclic Sequence) for hyperbolic structures [8], as points on a discrete Riemannian manifold, under HG metric, indistinguishable from traditional random noise sources by classical Shannon entropy, and current most advanced instrumentation approach. CICT defines an arbitrary scaling discrete Riemannian manifold uniquely, under HG metric, that, for arbitrary finite point accuracy level L going to infinity (exact solution theoretically), is isomorphic (even better homeomorphic) to traditional IG Riemannian manifold. In other words, HG can describe a projective relativistic geometry directly hardwired into elementary arithmetic long division remainder sequences, offering many competitive computational advantages over traditional Euclidean approach. It turns out that, while free generator exponentially growing sequences can be divergent or convergent, their closures can be defined in terms of polynomials. Furthermore, combinatorically OECS have strong connection even to classic DFT algorithmic structure for discrete data, Number-Theoretic Transform (NTT), Laplace and Mellin Transforms [8]. In this way, even simple scalar moduli can emerge out from sequences of phased generators. CICT can help to reach a unified vision to many current biophysics and physics problems and to find their optimized solutions quite easily [52]. CICT approach combined to Geometric Algebra and Geometric Calculus [61] unified mathematical language can offer an effective and convenient

"Science 2.0" universal modeling framework, by considering information not only on the statistical manifold of model states but also on the combinatorial manifold of low-level discrete, phased generators and empirical measures of noise sources, related to experimental high-level overall perturbation [52]. Furthermore, due to its intrinsic self-scaling properties, this system approach can be applied at any system scale: from single quantum system application development to full system governance strategic assessment policies and beyond [49],[51]. This approach allows you even to develop more antifragile "anticipatory learning system" (ALS), for more reliable, safe and secure medical application and system (cybersafety) [49]. CICT framework is quite flexible and can be used under a few operational representations. The major CICT representations used are two: formal power series and recurrence sequence.

6. ODR Cybernetics Update

In the past five decades, trend in Systems Theory, in specialized research area, has shifted from classic single domain information channel transfer function approach (Shannon's noisy channel, Fig.3 top diagram) to the more structured ODR Functional Sub-domain Transfer Function Approach (Observation, Description and Representation, Fig.3 middle diagram), according to CICT Infocentric Worldview model (theoretically, virtually noise-free data) [8]. In this way, system resilience and antifragility can be developed quite easily [8].
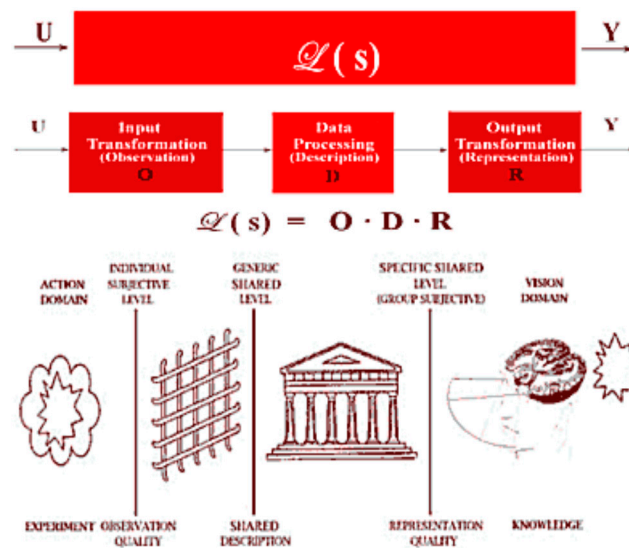

Figure 3. Top Diagram: Traditional Single Domain Channel (SDC) Transfer Function. Middle Diagram: Decomposition of SDC Transfer Function into more structured ODR Functional Sub-domain Transfer Function (Observation, Description and Representation Functional Blocks). Bottom Diagram: ODR Information Channel Co-domain Diagram for System Information Conservation [8].

From an application realization point of view, Canadian ecologist Crawford Stanley (Buzz) Holling (1930-) introduced important ideas in the application of ecology and evolution, including resilience, adaptive management, the adaptive cycle, and panarchy [56],[61]. Shortly, the ODR approach allows for fitting theoretical system design consideration to practical implementation needs much better

(according to information "Input, Processing, Output" paradigm, respectively) than classic single domain channel approach. Thanks to the ODR approach, a deeper awareness about information acquisition and generation limitations by classical experimental observation process has been grown. In fact, CICT showed that classic Shannon entropy is completely unable to reliably discriminate so called computational "random noise" from any combinatorically optimized encoded message by OECS, called "deterministic noise" in [8]. To cope with ontological uncertainty effectively at system level, we must use two asymptotic coupled complementary irreducible information management subsystems, based on the following ideal dichotomy: A) Information Reliable Unpredictability and B) Information Reliable Predictability. In this way, to behave realistically, overall system can guarantee both Logical Aperture (LA) and Logical Closure (LC), both fed by "Environmental Noise." So, a reliable operating point can always emerge as a new trans-disciplinary reality level, out of the interaction of two asymptotic coupled complementary irreducible information management subsystems [8]. As a simple example, the former subsystem has to be an ODR f extension (ODR Co-domain) able to capture as much as possible useful information from the open interaction of two related entities and their environment, and the latter one is the ODR Representation function f.

7. Operative Example

In a continuously changing operational environment, even if operational parameters cannot be closely pre-defined at system design level, we need to be able to design reliable self-organizing, self-regulating and self-adapting system quite easily anyway. In present paper, we like to give and to discuss an operative example based on CICT Recurrence Sequence Representation. We already know about "self-reference" in mathematics as a statement that refers to itself, for example, as a set that contains itself. Traditionally, such statements lead to paradox, a form of inconsistency. In the informal fallacies self-referential statements are considered poor form. That is true in mathematics and arithmetics when you use a continuous support approach and do not take advantage from the finiteness limitations of your real computational resources [50]. According to our Section 4. discussion, to cope with ontological uncertainty effectively at system level, we must use two asymptotic coupled complementary irreducible information management subsystems. So, a reliable operating point can always emerge as a new trans-disciplinary reality level, out of the mutual interaction of two complementary irreducible management subsystems with their common environment. In short, it is necessary to find a mathematical method to aggregate the external information coming from environment with system internal information in a efficient and fast way that is both immune from computational polynomial mirroring and classic noise effect by design [48],[62]. Furthermore, the mathematical method would possess anticipatory properties needed to the realization of system able to interact with its environment in real time (leading property) [57].

7.1. System Input Information Aggregation

Following Holling's approach (Section 6.) [56],[61], our main idea is to introduce an input node able to bind known information to the unknown one in a coherent way. Then, unknown "environmental noise" or/and local "signal input" information can be aggregated to known "system internal control status" information, to provide a self-structuring synthetic attractor point and functional closure. In this way, a

self-organizing landscape of self structuring attractor points can be achieved, and overall system can guarantee both Logical Aperture (LA) and Logical Closure (LC), both fed by "Environmental Noise," to behave realistically (Fig.4).
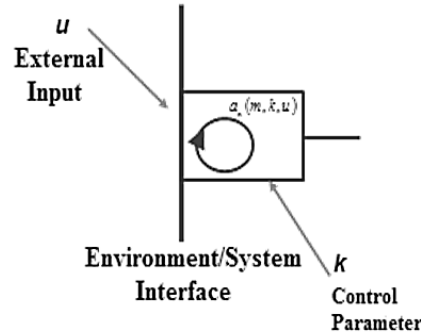

Figure 4. Operative representation of control status k and operative representation of input u are aggregated coherently by recursive subsequence of order m.

A mathematical method that meets these requirements (articulated information aggregation with system information anticipation and no computational polynomial mirroring and noise effect) is a simple recurrence relation, quite well known in mathematics, since long time! In fact, in his book "Liber abaci", for the first time Fibonacci (1170-1250) introduced the concept of recurrence sequence to the Western culture, with the famous sequence:

$$\mathbf{0}, \mathbf{1}, 1, 2, 3, 5, 8, 13, 21, 34, 55.89, 144, \tag{1}$$

in which each term is the sum of the two preceding ones and the numerical sequence composition law can be written as:

$$F_n = F_{n-1} + F_{n-2} \quad \text{with } F_0 = 0 \text{ and } F_1 = 1 \ , \tag{2}$$

and in general as:

$$F_n = k_1{}^* F_{n-1} + k_2{}^* F_{n-2}, \tag{3}$$

where $k_1,\ k_2 = 0,1,2,\ldots\infty$, $k_1,\ k_2 \in N$, for Generalized Fibonacci Sequences. For original Fibonacci sequence $k_1 = k_2 = 1$. Like every sequence defined by a linear recurrence with constant coefficients, the Fibonacci numbers have a closed-form solution. It has become known as Binet's formula, even though it was already known by Abraham de Moivre:

$$F_n = \frac{\varphi^n - \psi^n}{\varphi - \psi} = \frac{\varphi^n - \psi^n}{\sqrt{5}} \tag{4}$$

where

$$\varphi = \frac{1+\sqrt{5}}{2} \approx 1.6180339887 \ \ldots \tag{5}$$

is the golden ratio (τ), discovered by Johannes Kepler (1571–1630) [63], and:

$$\psi = \frac{1-\sqrt{5}}{2} = 1 - \varphi = -\frac{1}{\varphi} \approx -0.6180339887 \ ... \tag{6}$$

In fact, Kepler observed that the ratio of consecutive Fibonacci numbers converges, according to the relation:

$$\lim_{n \to \infty} \frac{F_{(n+1)}}{F_{(n)}} = \varphi \ , \tag{7}$$

providing us with recurrence relation asymptotical functional closure. This convergence holds regardless of the starting values, excluding $F_0 = 0$ and $F_1 = 0$. Another consequence is that the limit of the ratio of two Fibonacci numbers offset by a particular finite deviation in index corresponds to the golden ratio raised by that deviation. Or, in other words:

$$\lim_{n \to \infty} \frac{F_{(n+\alpha)}}{F_{(n)}} = \varphi^{\alpha} \ . \tag{8}$$

The generating function of the Fibonacci sequence is the following power series:

$$s(x) = \sum_{n=0}^{\infty} F_n x^n \ , \tag{9}$$

that converges for $|x| < \frac{1}{\varphi}$, and its sum has the following simple closed-form:

$$s(x) = \frac{x}{1 - x - x^2} \ . \tag{10}$$

If x is the reciprocal of an integer k that is greater than 1, the closed form of the series becomes:

$$s(k) = \sum_{n=0}^{\infty} \frac{F_n}{k_n} = \frac{k}{k^2 - k - 1} \ . \tag{11}$$

Therefore, as a starting point, the relation (1) can be thought as the aggregation of external input information $[F_0, F_1] \equiv [u_1, u_2]$ to internal system status information control vector $\boldsymbol{k} \equiv [k_1, k_2]$ (Fig.4). Please note that recursive sequence information aggregation does not suffer from the computational polynomial mirroring effect (compare eq.(10) to eq.(11)).

7.2. Reliable Unpredictability Information Subsystem

Recurrence sequence asymptotic convergence properties computed by the ratio of successive terms provides us with a recurrence relation asymptotic functional closure η_k, identical for all input vectors $[F_0, F_1] \equiv [u_1, u_2]$ as a function of control vector $\boldsymbol{k}$. For different values of control vector $\boldsymbol{k} \equiv [k_1, k_2]$, we obtain a self-organizing 2-D attractor space as from Fig.5.

k_1 \ k_2	0	1	2	3	4	5	6	7	8	9	$\cdots$
0		$\frac{1+\sqrt{4}}{2}$	$\frac{1+\sqrt{8}}{2}$	$\frac{1+\sqrt{12}}{2}$	$\frac{1+\sqrt{16}}{2}$	$\frac{1+\sqrt{20}}{2}$	$\frac{1+\sqrt{24}}{2}$	$\frac{1+\sqrt{28}}{2}$	$\frac{1+\sqrt{32}}{2}$	$\frac{1+\sqrt{36}}{2}$	$\cdots$
1	$\frac{1+\sqrt{1}}{2}$	$\frac{1+\sqrt{5}}{2}$	$\frac{1+\sqrt{9}}{2}$	$\frac{1+\sqrt{13}}{2}$	$\frac{1+\sqrt{17}}{2}$	$\frac{1+\sqrt{21}}{2}$	$\frac{1+\sqrt{25}}{2}$	$\frac{1+\sqrt{29}}{2}$	$\frac{1+\sqrt{33}}{2}$	$\frac{1+\sqrt{37}}{2}$	$\cdots$
2	$\frac{2+\sqrt{4}}{2}$	$\frac{2+\sqrt{8}}{2}$	$\frac{2+\sqrt{12}}{2}$	$\frac{2+\sqrt{16}}{2}$	$\frac{2+\sqrt{20}}{2}$	$\frac{2+\sqrt{24}}{2}$	$\frac{2+\sqrt{28}}{2}$	$\frac{2+\sqrt{32}}{2}$	$\frac{2+\sqrt{36}}{2}$	$\frac{2+\sqrt{40}}{2}$	$\cdots$
3	$\frac{3+\sqrt{9}}{2}$	$\frac{3+\sqrt{13}}{2}$	$\frac{3+\sqrt{17}}{2}$	$\frac{3+\sqrt{21}}{2}$	$\frac{3+\sqrt{25}}{2}$	$\frac{3+\sqrt{29}}{2}$	$\frac{3+\sqrt{33}}{2}$	$\frac{3+\sqrt{37}}{2}$	$\frac{3+\sqrt{41}}{2}$	$\frac{3+\sqrt{45}}{2}$	$\cdots$
4	$\frac{4+\sqrt{16}}{2}$	$\frac{4+\sqrt{20}}{2}$	$\frac{4+\sqrt{24}}{2}$	$\frac{4+\sqrt{28}}{2}$	$\frac{4+\sqrt{32}}{2}$	$\frac{4+\sqrt{36}}{2}$	$\frac{4+\sqrt{40}}{2}$	$\frac{4+\sqrt{44}}{2}$	$\frac{4+\sqrt{48}}{2}$	$\frac{4+\sqrt{52}}{2}$	$\cdots$
5	$\frac{5+\sqrt{25}}{2}$	$\frac{5+\sqrt{29}}{2}$	$\frac{5+\sqrt{33}}{2}$	$\frac{5+\sqrt{37}}{2}$	$\frac{5+\sqrt{41}}{2}$	$\frac{5+\sqrt{45}}{2}$	$\frac{5+\sqrt{49}}{2}$	$\frac{5+\sqrt{53}}{2}$	$\frac{5+\sqrt{57}}{2}$	$\frac{5+\sqrt{61}}{2}$	$\cdots$
6	$\frac{6+\sqrt{36}}{2}$	$\frac{6+\sqrt{40}}{2}$	$\frac{6+\sqrt{44}}{2}$	$\frac{6+\sqrt{48}}{2}$	$\frac{6+\sqrt{52}}{2}$	$\frac{6+\sqrt{56}}{2}$	$\frac{6+\sqrt{60}}{2}$	$\frac{6+\sqrt{64}}{2}$	$\frac{6+\sqrt{68}}{2}$	$\frac{6+\sqrt{72}}{2}$	$\cdots$
7	$\frac{7+\sqrt{49}}{2}$	$\frac{7+\sqrt{53}}{2}$	$\frac{7+\sqrt{57}}{2}$	$\frac{7+\sqrt{61}}{2}$	$\frac{7+\sqrt{65}}{2}$	$\frac{7+\sqrt{69}}{2}$	$\frac{7+\sqrt{73}}{2}$	$\frac{7+\sqrt{77}}{2}$	$\frac{7+\sqrt{81}}{2}$	$\frac{7+\sqrt{85}}{2}$	$\cdots$
8	$\frac{8+\sqrt{64}}{2}$	$\frac{8+\sqrt{68}}{2}$	$\frac{8+\sqrt{72}}{2}$	$\frac{8+\sqrt{76}}{2}$	$\frac{8+\sqrt{80}}{2}$	$\frac{8+\sqrt{84}}{2}$	$\frac{8+\sqrt{88}}{2}$	$\frac{8+\sqrt{92}}{2}$	$\frac{8+\sqrt{96}}{2}$	$\frac{8+\sqrt{100}}{2}$	$\cdots$
9	$\frac{9+\sqrt{81}}{2}$	$\frac{9+\sqrt{85}}{2}$	$\frac{9+\sqrt{89}}{2}$	$\frac{9+\sqrt{93}}{2}$	$\frac{9+\sqrt{97}}{2}$	$\frac{9+\sqrt{101}}{2}$	$\frac{9+\sqrt{105}}{2}$	$\frac{9+\sqrt{109}}{2}$	$\frac{9+\sqrt{113}}{2}$	$\frac{9+\sqrt{117}}{2}$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Figure 5. An example of the 2-Dimensional Solution Space of Asymptotic Attractor Points (Convergence Rates) System Functional Closure.

As a matter of fact η_k becomes an attractor point for all input sequences $[F_0, F_1] \equiv [u_1, u_2]$ structured by the same 2-D control vector $k \equiv [k_1, k_2]$. In general n-D space is generated by a corresponding n-components control vector. This property allows the creation of a self-structuring system's behavioral space similar, in computational behavior, to the living organism's homeostasis (i.e. the automatic selection of environment's minimum perturbation level that allows optimal interaction between external information from environment (u) and system internal status information (k), as evidenced by Holling [56],[61].
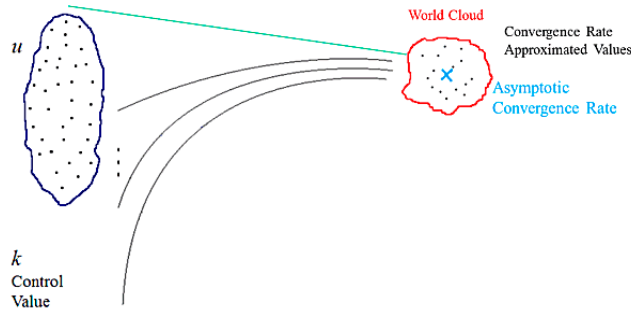

Figure 6. Control vector k allows to arrange external inputs u into symbolic string information with different convergence rate approximation around their convergence rate asymptotic value (World Cloud). For different k values, a landscape of world clouds can start self-organizing as a Baire Space. Control vector k selects recurrence series asymptotic convergence rate (Attractor Point), which different input u approximated rates are self-arranging modularly (World Cloud) around.

System "ecologic interface" allows to reach an operative point in homeostatic equilibrium with its external environment perturbation. In this way, we can allow for articulated information aggregation even in a networking environment and computing their related asymptotic functional closures immediately. Therefore, it is possible to compute recursively further numerical sequences asymptotically converging to irrational limits or completely diverging as function of external information input. In this way system can search automatically for a minimum environmental perturbation level (system internal status) useful to insure sequence asymptotical convergence to get vital information from system

environment (self-regulation and learning as quest for the difference that makes the difference, probing by probing...). Then, homeostatic operating equilibria can emerge out of a self-organizing landscape of self-structuring attractor points with their "World Cloud" (Fig.6). Control vector k allows to arrange external inputs u into symbolic string information with different convergence rate approximation around their convergence rate asymptotic value (World Cloud). For different k values, a landscape of world clouds can start self-organizing as a Baire Space. Control vector k selects recurrence sequence asymptotic convergence rate (Attractor Point), which different input u approximated rates are self-arranging modularly (World Cloud) around (Fig.6).

Irrational numeric limit families, identified by converging recursive numeric sequences allow the structuring of a mathematical Baire's Space. A Baire space consists of countably infinite sequences with a metric defined in terms of the longest common prefix: the longer the common prefix, the closer a pair of sequences. What is of interest to us here is this longest common prefix metric, which we call the "Baire distance", which is an ultrametric distance [64-66]. Baire's Space allows to manage numeric information in a way useful to synthetize quick and raw system primary response "to survive and grow." Furthermore, in this way system can even automatically self-organize and structure numeric families with different numerical closure to conserve overall system information (Generalized Fibonacci Systems, and information conservation by irreducible complementary system) [8]. From the simple 2-term recurrence case, the recurrence aggregation law can be easily extended to 3-term case, by three consecutive terms (starting by trinomial 0,0,1, third order relation, $m=3$), to 4-term case, by four successive terms (starting by quadrinomial 0,0,0,1, fourth order relation , $m=4$), m-term case (by m-nomial formed by j zeroes, plus one 1, $(j+1)^{th}$ order relation, $(m=j+1)$, and in general we can write, in compact form, recurrence relation as function of three parameters, $a_n = a(m, k, u)$, as specified previously, where m is the recurrence relation order. Figure 7. and Figure 8. give two examples of computational system functional closure for two different values of 2-D control vector k with parameter m going from 0 to ∞.
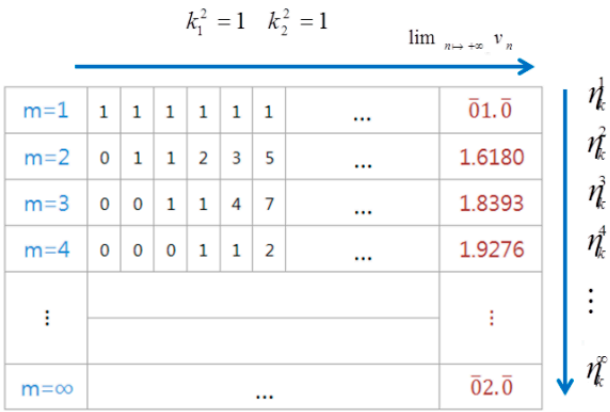

$k_1^2 = 1 \quad k_2^2 = 1$

$\lim_{n \to +\infty} v_n$

m=1	1	1	1	1	1	1	...	$\bar{0}1.\bar{0}$	η_k^1
m=2	0	1	1	2	3	5	...	1.6180	η_k^2
m=3	0	0	1	1	4	7	...	1.8393	η_k^3
m=4	0	0	0	1	1	2	...	1.9276	η_k^4
⋮								⋮	⋮
m=∞				...				$\bar{0}2.\bar{0}$	η_k^∞

Figure 7. System Functional Closures for $k \equiv [1, 1]$ with parameter m going from 0 to ∞. The dynamic and deterministic recognizable numeric behavior, both horizontally and vertically, allows the simple definition of two coordinates for reliable computational bounds to each operating point.
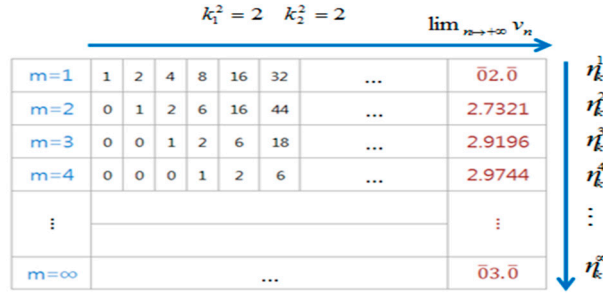
	$k_1^2 = 2$	$k_2^2 = 2$					$\lim_{n \to +\infty} v_n$	
m=1	1	2	4	8	16	32	...	$\bar{0}2.\bar{0}$	n_k^1
m=2	0	1	2	6	16	44	...	2.7321	n_k^2
m=3	0	0	1	2	6	18	...	2.9196	n_k^3
m=4	0	0	0	1	2	6	...	2.9744	n_k^4
⋮								⋮	⋮
m=∞					...			$\bar{0}3.\bar{0}$	n_k^∞

Figure 8. System Functional Closures for $k \equiv [2, 2]$ with parameter m going from 0 to ∞. The dynamic and deterministic recognizable numeric behavior, both horizontally and vertically, allows the simple definition of two coordinates for reliable computational bounds to each operating point.

7.3. System Output Information Anticipation

Rational recurrence sequence represents a convenient mathematical method that holds anticipatory proprieties, because it is possible to implement the anticipatory computation of any recursive sequence's term. Taking arbitrarily any current positional index, it is possible to describe not only the next term but also terms at a certain distance from the current one in an anticipatory way, compared with the current positional index, by implementing its primary relation recursion conveniently. Specifically, starting from the recursive rule that indicates the next term to the current one, it is possible to structure a set of rules that allows to obtain recursive sequence's terms at different distance, defining a set of registers that, working in parallel, are able to provide values with the desired anticipation level immediately. For example, considering second order ($m = 2$) recurrence relations $a_n = a(2, k, u)$, where k and u are 2-D vectors $[k_1, k_2]$ and $[u_1, u_2]$ respectively, and where the *(n + 1)-th* to the current *n-th* term is obtained by the following recurrence relation:

$$a_{n+1} = k_1 * a_n + k_2 * a_{n-1}, \tag{12}$$

where $a_{n-1} = u_1 = 0$ and $a_n = u_2 = 1$, $k_1 = k_2 = 1$, for the Fibonacci sequence. Then, we can specify recursion derived relations to compute appropriate terms at any arbitrary distance from the current position n. As an example, we define the following relations that are valid for computing terms to the distance $n+5$ from the current one, depending, for instance, on control vector $k \equiv [k_1, k_2] = [1, 1]$ and $[k_1, k_2] = [2, 2]$. These recurrence relations can be used in parallel, respectively, to provide the terms of sequence in an anticipatory way simultaneously. So in the case of $[k_1, k_2] = [1, 1]$, recurrence rules are as follows:

$$a_{n+2} = 2 * a_n + 1 * a_{n-1}$$

$$a_{n+3} = 3 * a_n + 2 * a_{n-1}$$

$$a_{n+4} = 5 * a_n + 3 * a_{n-1}$$

$$a_{n+5} = 8 * a_n + 5 * a_{n-1} \tag{13}$$

$$\ldots \qquad\qquad\qquad .$$

In the case of values $[k_1, k_2] = [2, 2]$ we obtain the following recurrence rules:

$$a_{n+2} = \quad 6 * a_n + 4 * a_{n-1}$$

$$a_{n+3} = \quad 16 * a_n + 12 * a_{n-1}$$

$$a_{n+4} = \quad 44 * a_n + 32 * a_{n-1} \tag{14}$$

$$a_{n+5} = 120 * a_n + 88 * a_{n-1}$$

... .

In general for any control vector $k \equiv [k_1, k_2]$ and for any order m, it is always possible to formulate rules associated to their primary recurrence relation: these rules allow the parallel anticipatory computation of recurrence sequence's term at any distance from the current term position n.

7.4. Reliable Predictability Information Subsystem

To synthetize more organized and articulated, but slower, system response "to learn and prosper", it is necessary to structure recursive information into an "ordered polynomial reference", by "polynomial weighing" mapping, to obtain system "coherent perception" [67]. Polynomial weighing is key numeric operation to map recurrence sequence information representation into polynomial format corresponding to combinatorically OECS, folded into rational number operative representation (Fig.9) [8],[53].
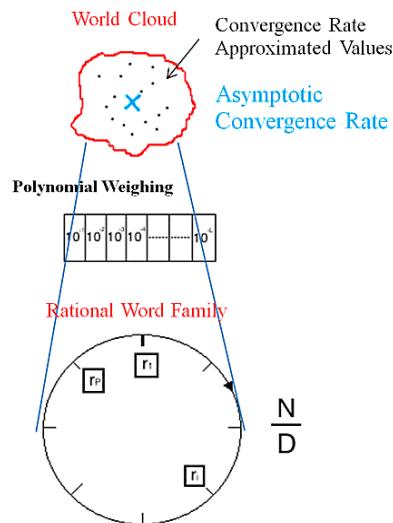
Figure 9. Polynomial Weighing allows the modular mapping of each 2-D approximated algebraic irrational World Cloud into a corresponding Rational Word Family. Denominator D identifies Rational Word Family. Numerator N identifies each algebraic rational approximation of irrational values (attractor points) as component inside that Family.

Polynomial Weighing allows the modular mapping of each 2-D approximated algebraic irrational World Cloud into a corresponding Rational Word Family. Denominator D identifies Rational Word Family. Numerator N identifies each algebraic rational approximation of irrational values (attractor points) as component inside that Family. So, we get a sequence of different structuring operations to get external information more and more coherently formatted to system internal status to arrive to a system

"coherent perception" of external information. In this way, a natural balanced "Operating Point" can emerge, as a new Transdisciplinary Reality Level, from the mutual interaction of two coupled complementary irreducible information management subsystems, immersed within their environment.

8. Post-Bertalanffy Cybernetics Update

As you already guessed, the previous defined overall adaptive and learning system reference architecture is quite flexible. In fact we can use it to develop Anticipatory Learning System (ALS) to get realistic modeling of system natural behavior to be used in High Reliable Organization (HRO) application development. We can even use the same nonlinear logic approach to guess a convenient basic architecture to design anticipatory smart system interface (Interaction Interface System, IIS) as depicted in Fig.10. The simple recursive information aggregation method can be used even for advanced ISS (Inner Safety System) in advanced biomedical and healthcare system development application, from neuroperception to organ level modeling, from bio-transduction functions to anything-you-like, etc. [48],[62].
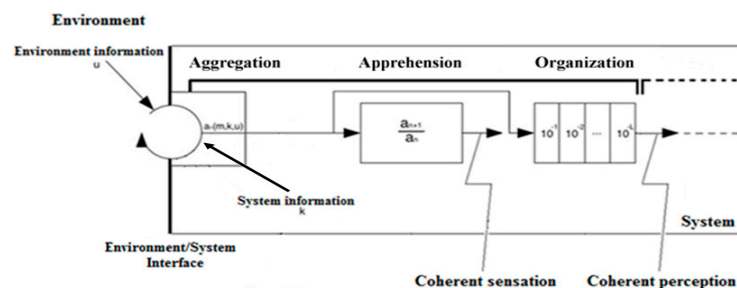
Figure 10. Interaction Interface System (IIS) Reference Architecture [48],[62]: Apprehension (Open Logic Subsystem); Organization (Closed Logic Subsystem).

Due to its intrinsic self-organizing and self-scaling properties, this system approach can be applied at any system scale: from single health application development to full healthcare system governance strategic simulation and assessment application [48]. It can allow both quick and raw system response (Open Logic response, to survive and grow) and slow and accurate information unfolding for future response strategic organization (Closed Logic response, to learn, to adapt and prosper) by coherently formatted operating point [67]. Now, according to previous discussion, it is possible, at systemic level, to envisage a post-Bertalanffy Systemics Framework, with multiscaling properties, able to deal with problems of different complexity in a generalized way when interdisciplinary consists, for instance, of a disciplinary reformulation of problems, like from biological to chemical, from clinical research to healthcare, etc., and transdisciplinary is related to the study of such reformulations and their properties. As shown in Fig.4, our innovative system interaction modality, called "Recursive Interactor," (ReInt for short), corresponds to fourth order biomedical cybernetics in Fig.11. Now, new advanced systemic application can successfully and reliably manage a higher system complexity than current ones, to develop competitive system, with a minimum of design constraints specification and less system final operative environment knowledge at system design level.
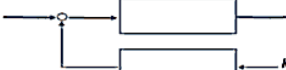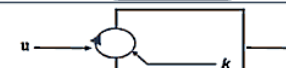
BIOMEDICAL CYBERNETIC ORDER	INTERACTION STYLE	GRAPHIC SYMBOL
Zero	Pure Spectator	
First	Ergodic Observer	
Second	Pulsed Egocentric Interactor	
Third	Iterated Egocentric Interactor	
Fourth	Recursive Interactor	

Figure 11. Our final post-Bertalanffy Systemics Framework [50].

For instance, according to Fig.10, at brain level, it is possible to refer to LeDoux circuit (Logical Aperture) for emotional behavior (i.e. fear, emotional intelligence, etc.) and to Papez circuit (Logical Closure) for structured behavior (i.e. rational thinking, knowledge extraction, etc.…) [46]. Emotional Intelligence (EI) and Emotional Creativity (EC) [47] coexist at the same time with Rational Thinking in human mind, sharing the same input environment information. Then, operating point can emerge as a transdisciplinary reality level from the interaction of two complementary irreducible, asymptotic ideal coupled subsystems. To behave realistically, overall system must guarantee both Logical Aperture (to get EI and EC, to survive and grow) and Logical Closure (to get Rational Thinking, to learn and prosper), both fed by environmental "noise" (better… from what human beings call "noise") [8]. In fact, natural living organism does perturb its environment, but only up to the level it is perturbed in turn by its own environment both to survive and grow, no more [56]. Due to its intrinsic scaling properties, this system approach can be applied at any system scale: from single quantum system application development to full system governance strategic assessment policies and beyond [62]. Expected impacts are multifarious and quite articulated at different system levels: at theoretic level, major one is that, for the first time, Biomedical Engineering ideal system categorization levels can be matched exactly to practical system modeling interaction styles (Fig.11), with no paradigmatic operational ambiguity and information loss, even for living organism application.

9. Multidisciplinary Approach Arbitrary Scale Modeling

ALS, IIS and ISS are three pivotal concepts to develop safer, more adequate, effective and efficient solutions for competitive safety systems and human wellbeing. As a matter of fact their basic operational concepts can be conveniently and successfully extended to many advanced Business and HRO application areas, with no performance or economic penalty, to develop more and more competitive application. For instance, at a higher level of abstraction, environmental noise input information to be aggregated to system internal status information can provide a structured homeostatic synthetic operating point as a reference for further inquiry. Then, System Interaction by internal and external information

aggregation can allow both quick and raw response (Open Logic response, to grow and survive) and slow and accurate information for future response strategic organization (Closed Logic response, to adapt and prosper) by coherently formatted operating point information. As an example of simple advanced management architecture for HRO, for closed logic Reactive Management system, we can choose from different documented operational alternatives offered by literature, like Deming's PDCA Cycle [68], Discovery-Driven Planning [69], etc., while for open logic Proactive Management system, we can choose from Boyd OODA Cycle (1987) [70], Theory-Focused Planning [71], etc.
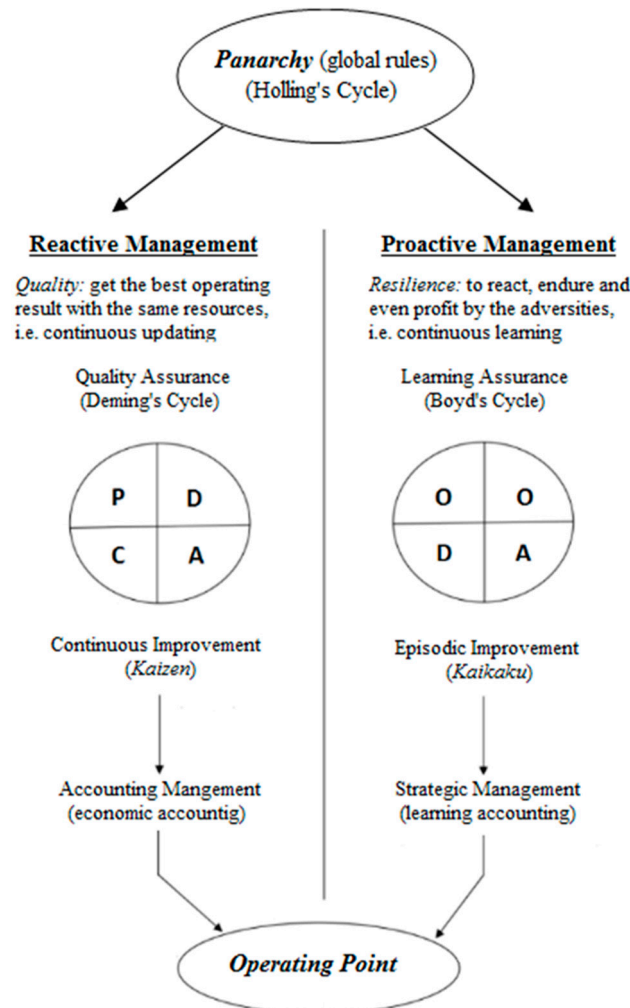

Figure 12. Final Architecture for Advanced Sfety and Security Systemic Governance Framework for HRO [72],[73].

As a simple example, PDCA's cycle (Reactive Management) and OODA's cycle (Proactive Management) can be selected to represent two corresponding complementary irreducible sub-systems for advanced integrated strategic management. Then, our final operative reference architecture, for Safety and Effectiveness Health Systemic Governance, is given as from Fig.12. Specifically, advanced wellbeing applications (AWA), high reliability organization (HRO), mission critical project (MCP) system, very low technological risk (VLTR) and crisis management (CM) system will be highly benefited mostly by CICT newer approach and related techniques.

10. Conclusion

CICT general relationships to compute information coherent functional closure for any computational system from low-level multiplicative noise source generators, related to experimental high-level overall perturbation were presented and discussed. CICT brings classical and quantum information theory together in a single framework, by considering information not only on the statistical manifold of model states but also from empirical measures of low-level multiplicative noise source generators, related to experimental high-level overall perturbation. Traditional elementary arithmetic long division remainder sequences can be interpreted as combinatorically optimized exponential cyclic sequences (OECS) for hyperbolic structures, as points on a discrete Riemannian manifold, under HG metric, indistinguishable from traditional random noise sources by classical Shannon entropy, and current most advanced instrumentation approach [8]. CICT defines an arbitrary-scaling discrete Riemannian manifold uniquely, under HG metric, that, for arbitrary finite point accuracy L going to infinity (exact solution theoretically), is isomorphic (even better homeomorphic) to traditional information geometry Riemannian manifold [53]. In other words, HG can describe a projective relativistic geometry directly hardwired into elementary arithmetic long division remainder sequences, offering many competitive computational advantages over traditional Euclidean approach. It turns out that, while free generator exponentially growing sequences (OECS) can be divergent or convergent, their coherent functional closures can be defined in terms of complete homogeneous polynomial structures. Associated OECS co-domain polynomial information structure can be used to evaluate any computed result, and to compensate for achieving computational information conservation for any computational system in any polynomial base. Due to its intrinsic self-scaling properties, this system approach can be applied at any system scale: from single quantum system application development to full system governance strategic assessment policies and beyond. In this way, we can minimize or overcome previous major system limitations (Section 2.) to arrive much closer to fourth generation adaptive learning and real machine intelligence systems. This approach allows you even to develop more antifragile anticipatory learning system (ALS), for more reliable, safe and secure medical application and system (cybersafety). Specifically, high reliability organization (HRO) [72],[73], mission critical project (MCP) system, very low technological risk (VLTR) and crisis management (CM) system will be highly benefitted mostly by these new techniques. The present paper is a relevant contribution towards a new post-Bertalanffy General Theory of Systems to show how homeostatic equilibria can emerge out of a self-organizing landscape of self-structuring attractor points.

Conflicts of Interest

The author declares no conflict of interest.

References and Notes

[1] Begley, C.G.; Ellis, L.M. Drug development: Raise standards for preclinical cancer research. *Nature*, **2012**, *483*, 531–533.

[2] Ioannidis, J.P.A. Why Most Published Research Findings Are False. *PLoS Med.*, **2005**, *2(8)*, e124.doi:10.1371/journal.pmed.0020124.

[3] Mobley, A.; Linder, S.K.; Braeuer, R.; Ellis, L.M.; Zwelling, L. A Survey on Data Reproducibility in Cancer Research Provides Insights into Our Limited Ability to Translate Findings from the Laboratory to the Clinic. *PLoS ONE*, **2013**, *8(5)*, e63221. doi:10.1371/journal.pone.0063221

[4] Nuzzo, R. Scientific method: Statistical errors. *Nature*, **2014**, *506*, 150-152. Available at URL= <http://www.nature.com/news/scientific-method-statistical-errors-1.14700> (accessed on 25 October 2015).

[5] Chalmers, I.; Bracken, M.B.; Djulbegovic, B.; Garattini, S.; Grant, J.; Gülmezoglu, A.M.; et al. How to increase value and reduce waste when research priorities are set. *Lancet*, **2014**, *383(9912)*, 156-165. PubMed PMID: 24411644

[6] Rao, C.R. Information and the accuracy attainable in the estimation of statistical parameters. *Bull. Calcutta Math. Soc.*, **1945**, *37*, 81–89.

[7] Nielsen, F.; Barbaresco, F., (eds.) *Geometric Science of Information*; Springer: Heidelberg, 2013.

[8] Fiorini, R.A. How Random is Your Tomographic Noise? A Number Theoretic Transform (NTT) Approach. *Fundamenta Infomaticae*, **2014**, *135(1-2)*, 135-170.

[9] Taleb, N.N. Silent Risk, *Lectures on Probability*, 2015, Vol.1, Des Cartes Ed. Available at: URL = <https://drive.google.com/file/d/0B8nhAlfIk3QIR1o1dnk5ZmRaaGs/view?pli=1> (accessed on 25 October 2015).

[10] Song, H.A.; Lee, S.Y. Hierarchical Representation Using NMF. *Neural Information Processing. Lectures Notes in Computer Sciences 8226*; Springer: Berlin, Heidelberg, 2013. pp. 466–473. doi:10.1007/978-3-642-42054-2_58. ISBN 978-3-642-42053-5.

[11] Devlin, J.; Zbib, R.; Huang,Z.; Lamar, T.; Schwartz, R.; Makhoul, J. Fast and robust neural network joint models for statistical machine translation. In *Proc. ACL 2014*, 2014.

[12] Sutskever, I.; Vinyals, O.; Le, Q.V. Sequence to sequence learning with neural networks. In *Proc. NIPS 2014*, 2014.

[13] Bahdanau,D.; Cho, K.; Bengio, Y. Neural machine translation by jointly learning to align and translate. In *Proc. ICLR 2015*, 2015. Available at: URL= <http://arxiv.org/abs/1409.0473> (accessed on 25 October 2015).

[14] Raina, R.; Madhavan, A.; Ng, A.Y. Large-scale deep unsupervised learning using graphics processors. In *Proc. ICML 2009*, 2009.

[15] Bengio, Y.; Ducharme, R.; Vincent, P.; Jauvin, C. A neural probabilistic language model. *Journal of Machine Learning Research*, **2003**, *3*,1137–1155.

[16] Dean, J.; Corrado, G.S.; R. Chen, M.K.; Devin, M.; Le, Q.V.; Mao, M.Z.; Ranzato, M.A.; Senior, A.; Tucker, P.; Yang, K.; Ng, A.Y. Large scale distributed deep networks. In *Proc. NIPS 2012*, 2012.

[17] Hinton, G.; Deng, L.; Dahl, G.E.; Mohamed, A.; Jaitly, N.; Senior, A.; Vanhoucke, V.; Nguyen, P.; Sainath, T.; Kingsbury, B. Deep neural networks for acoustic modeling in speech recognition. *IEEE Signal Processing Magazine*, **Nov. 2012**, *29(6)*, 82–97.

[18] Sainath, T.; Mohamed, A.; Kingsbury, B.; Ramabhadran, B. Deep convolutional neural networks for LVCSR. In *Proc. ICASSP 2013*, 2013.

[19] Krizhevsky, A.; Sutskever, I.; Hinton, G. ImageNet classification with deep convolutional neural networks. In *Proc. NIPS 2012*. 2012.

[20] Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. Technical report, 2014. Available at: URL= <http://arxiv.org/abs/1409.4842> (accessed on 25 October 2015).

[21] Fukushima, K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, **1980**, *36*, 193–202.

[22] Le Cun, Y.; Boser, B.; Denker, J.S.; Henderson, D.; Howard,R.E.; Hubbard, W.; Jackel, L.D. Backpropagation applied to handwritten zip code recognition. *Neural Computation*, **1989**, *1(4)*, 541–551.

[23] Krizhevsky, A.; Sutskever, I.; Hinton, G. ImageNet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems 25*, NIPS 2012, 2012.

[24] Lin, M.; Chen, Q.; Yan, S. Network in network. In *Proceedings of the Second International Conference on Learning Representations*, ICLR 2014, Apr. 2014.

[25] Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. In *Proc. International Conference on Learning Representations*, ICLR 2015, 2015.

[26] Ciresan, D.; Meier, U.; Schmidhuber, J. Multi-column deep neural networks for image classification. Technical report, 2012. Available at: URL= <http://arxiv.org/abs/1202.2745> (accessed on 25 October 2015).

[27] Ciresan, D.; Meier, U.; Masci, J; Schmidhuber, J. Multi-column deep neural network for traffic sign classification. *Neural Networks*, **2012**, *32*, 333–338.

[28] Goodfellow, I.J.; Bulatov, Y.; Ibarz, J.; Arnoud, S.; Shet, V. Multi-digit number recognition from Street View imagery using deep convolutional neural networks. In *International Conference on Learning Representations*, ICLR 2014, 2014.

[29] Abdel-Hamid, O.; Mohamed, A.; Jiang, H.; Penn, G. Applying convolutional neural networks concepts to hybrid nn-hmm model for speech recognition. In *Proc. 2012 IEEE International Conference onAcoustics, Speech and Signal Processing*, ICASSP 2012 , 2012.

[30] Sainath, T.N.; Mohamed, A.-r.; Kingsbury, B.; Ramabhadran, B. Deep convolutional neural networks for lvcsr. In *Proc. 2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, ICASSP 2013, pp.8614–8618.

[31] Sainath, T.N.; Vinyals, O.; Senior, A.; Sak, H. Convolutional, long short-term memory, fully connected deep neural networks. In *Proc. 2015 IEEE International Conference onAcoustics, Speech and Signal Processing*, ICASSP 2015.

[32] Tóth, L. Combining time-and frequency-domain convolution in convolutional neural network-based phone recognition. In *Proc. 2014 IEEE International Conference on Acoustics, Speech and Signal Processing*, ICASSP 2014, pp.190–194.

[33] Vinyals, O.; Toshev, A.; Bengio, S.; Erhan, D. Show and tell: a neural image caption generator, 2014. Available at: URL= <http://arxiv.org/abs/1411.4555> (accessed on 25 October 2015).

[34] Xu, K.; Ba, .L.; Kiros, R.; Cho, K.; Courville, A.; Salakhutdinov, R.; Zemel, R.S.; Bengio, Y. Show, attend and tell: Neural image caption generation with visual attention, 2015. Available at: URL= <http://arxiv.org/abs/1502.03044> (accessed on 25 October 2015).

[35] Yao, L.; Torabi, A.; Cho, K.; Ballas, N.; Pal, C.; Larochelle, H.; Courville A. Video description generation incorporating spatio-temporal features and a soft-attention mechanism, 2015. Available at: URL= <http://arxiv.org/abs/1502.08029> (accessed on 25 October 2015).

[36] Sermanet, P.; Eigen, D.; Zhang, X.; Mathieu, M.; Fergus, R.; LeCun, Y. Overfeat: Integrated recognition, localization and detection using convolutional networks. In *Proc. 2014 International Conference on Learning Representations*, ICLR 2014, 2014.

[37] Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. In *Proc. 2015 International Conference on Learning Representations*, ICLR 2015, 2015.

[38] Lin, Z.; Courbariaux, M.; Memisevic, R.; Bengio, Y. Neural Networks with few Multiplications. Under review as a conference paper at ICLR 2016. Available at: URL= <http://arxiv.org/abs/1510.03009> (accessed on 25 October 2015).

[39] Kim, S.K.; McAfee, .C.; McMahon, P.L.; Olukotun, K. A highly scalable restricted Boltzmann machine FPGA implementation. In *Proc. 2009 IEEE International Conference on Field Programmable Logic and Applications*, FPL 2009, pp.367–372.

[40] Chen, T.; Du, Z.; Sun, N.; Wang, .; Wu, C.; Chen, Y.; Teman, O. Diannao: A small-footprint high-throughput accelerator for ubiquitous machine-learning. In *Proceedings of the 19th ACM international conference on Architectural support for programming languages and operating systems*, ACM, 2014, pp.269–284.

[41] Chen, Y.; Luo, T.; Liu, S.; Zhang, S.; He, L.; Wang, J.; Li, L.; Chen, T.; Xu, Z.; Sun, N.; et al. Dadiannao: A machine-learning supercomputer. In *Proc. 47th Annual IEEE/ACM International Symposium on Microarchitecture*, MICRO 2014, pp.609–622.

[42] Olshausen, B. A. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, **1996**, *381(6583)*, 607–609.

[43] Fiorini, R.A. Entropy, Decoherence and Spacetime Splitting. In *Proceedings of the 2nd Int. Electron. Conf. Entropy Appl.*, 15-30 November 2015; Sciforum Electronic Conference Series, **2015.**

[44] Resconi, G. *Geometry of Knowledge for Intelligent Systems*; Springer International, Vol.407 of Studies on Computational Intelligence, 2012.

[45] Mead, C., *Analog VLSI and Neural Systems* (1st edition); Addison Wesley Publishing Company: New York, N.Y., 1989.

[46] LeDoux, J. *The Emotional Brain, The Mysterious Underpinnings of Emotional Life*; Weidenfeld & Nicolson: Great Britain, 1998.

[47] LeDoux, J. *Synaptic Self, How Our Brains Become Who We Are*; Viking Penguin: New York, NY, 2002.

[48] Fiorini, R.A.; Santacroce, G.F. Economic Competitivity in Healthcare Safety Management by Biomedical Cybernetics ALS. In *2013 Proc. International Symposium, The Economic Crisis: Time For A Paradigm Shift -Towards a Systems Approach*, Universitat de València, January 24-25, **2013**.

[49] Fiorini, R.A. From Systemic Complexity to Systemic Simplicity: A New Networking Node Approach. Chap. 9 in G.F. Minati, M. Abram, E. Pessa, eds., *Towards a Post-Bertalanffy Systemics*, Springer International Publishing: Switzerland, in press.

[50] Fiorini, R.A.; Laguteta, G. Discrete Tomography Data Footprint Reduction by Information Conservation. *Fundamenta Informaticae*, **2013**, *125(3-4)*, 261–272.

[51] Fiorini, R.A. Computerized Tomography Noise Reduction by CICT Optimized Exponential Cyclic Sequences (OECS) Co-domain. *Fundamenta Informaticae,* **2015**, *141*, 115–134, DOI 10.3233/FI-2015-1267.

[52] Fiorini, R.A. More effective biomedical experimentation data by CICT advanced ontological uncertainty management techniques. *Int. J. of Biol and Biomed Eng*, **2015**, *9*, 29-41.

[53] Fiorini, R.A. Computational Information Conservation Theory: An Introduction. In *2014 Proceedings of the 8th International Conference on Applied Mathematics, Simulation, Modelling (ASM '14)*, N.E. Mastorakis, M. Demiralp, N. Mukhopadhyay, F. Mainardi, eds., Mathematics and Computers in Science and Engineering Series, No.34, NAUN Conferences, WSEAS Press, November 22-24, **2014**, Florence, Italy, pp.385-394.

[54] Goleman, D.J. *Emotional Intelligence: Why It Can Matter More Than IQ*; Bantam Books: New York, NY, 1995.

[55] Nicolescu, B. Levels of Complexity and Levels of Reality. In *The Emergence of Complexity in Mathematics, Physics, Chemistry, and Biology*, Bernard Pullman, ed. Proceedings of the Plenary Session of the Pontifical Academy of Sciences, 27-31 October **1992**, Casina Pio IV, Vatican, Ed. Pontificia Academia Scientiarum, Vatican City; Princeton University Press: Princeton, NJ, 1992.

[56] Gunderson, L.H.; Holling, C.S., eds. *Panarchy: understanding transformations in human and natural systems*; Island Press: Washington DC, 2002.

[57] Rosen, R., *Anticipatory Systems*; Pergamon Press: Oxford, 1985.

[58] Ostrom E. *Governing the Commons: The Evolution of Institutions for Collective Action*; Cambridge University Press: Cambridge, UK, 1990.

[59] Ostrom E. Beyond Markets and States: Polycentric Governance of Complex Economic Systems. *American Economic Review*, **2010**, *100(3)*, 1-33. Available at: URL = <http://bnp.binghamton.edu/wp-content/uploads/2011/06/Ostrom-2010-Polycentric-Governance.pdf>(accessed on 25 October 2015).

[60] Fiorini, R.A. GA and CICT for Stronger Arbitrary Multi-Scale Biomedical and Bioengineering Solutions. In *Early Proceedings of the AGACSE 2015 Conference*, July 29-31, **2015**, Barcelona, Spain, pp.153-162.

[61] Gotts, N.M, Resilience, panarchy, and world-systems analysis. *Ecology and Society*, **2007**, *12(1)*, 24. Available at: URL = <http://www.ecologyandsociety.org/vol12/iss1/art24/> (accessed on 25 October 2015).

[62] Fiorini, R.A.; Santacroce, G.F. Systems Science and Biomedical Cybernetics for Healthcare Safety Management. In *2013 Proc. International Symposium, The Economic Crisis: Time For A Paradigm Shift - Towards a Systems Approach*, Universitat de València, January 24-25, **2013**.

[63] Tattersall, J.J., *Elementary number theory in nine chapters*; (2nd ed.). Cambridge University Press: Cambridge, MA, 2005. p.28.

[64] Murtagh, F. On ultrametricity, data coding, and computation. *Journal of Classification,* **2004,** *21*, 167-184.

[65] Murtagh, F. Identifying the ultrametricity of time series. *The European Physical Journal B,* **2005**, *43(4)*, 573–579.

[66] Murtagh, F.; Downs, G.; Contreras, P. Hierarchical clustering of massive, high dimensional data sets by exploiting ultrametric embedding. *SIAM-Journal on Scientific Computing*, **2008**, *30(2)*, 707-730.

[67] Collini, A.V.; Cesario, C. Sistemi a sicurezza intrinseca per la sanità. Master of Science Thesis, Biomedical Engineering Department, Politecnico di Milano, Milano, Italy, 2012.

[68] Ohno, T., *Workplace Management,* Special 100th Birthday Edition, 1 edition; McGraw-Hill Professional: New York, NY, 2012.

[69] McGrath, G.; MacMillan, R.; MacMillan, I.C. Discovery Driven Planning: Turning Conventional Planning on its Head. *Harvard Business Review*, **1995**.

[70] Osinga, F.P.B. *Science, Strategy and War: The Strategic Theory of John Boyd,* 1 edition; Routledge: New York, NY, 2006.

[71] Govindarajan, V.; Trimble, C. Strategic Innovation and the science of learning. *MIT SLOAN Management Review*, **2004**, *45(2)*, 67-75.

[72] Ono, I.; Sagami, F.; Nakamoto K. Citywide Safety and Security Solutions. *Hitachi Review*, **2014**, *63(5)*, 230-235.

[73] Yamaoka, M.; Yoshimura, C.; Hayashi, M.; Okuyama, T.; Aoki, H.; Mizuno, H. New Computing Paradigm for Analyzing Increasingly Complex Social Infrastructure Systems. *Hitachi Review*, **2015**, *64(8)*, 525-531.