Essay

Towards Quantifying a Wider Reality: Shannon Exonerata

Robert E. Ulanowicz [1,2]

1   Department of Biology, University of Florida, Gainesville, FL 32611-8525, USA;
    E-Mail: ulan@cbl.umces.edu
2   University of Maryland Center for Environmental Science, Solomons, MD 20688-0038, USA

Abstract: In 1872 Ludwig von Boltzmann derived a statistical formula to represent the entropy (an apophasis) of a highly simplistic system. In 1948 Claude Shannon independently formulated the same expression to capture the positivist essence of information. Such contradictory thrusts engendered decades of ambiguity concerning exactly what is conveyed by the expression. Resolution of widespread confusion is possible by invoking the third law of thermodynamics, which requires that entropy be treated in a relativistic fashion. Doing so parses the Boltzmann expression into separate terms that segregate apophatic entropy from positivist information. Possibly more importantly, the decomposition itself portrays a dialectic-like agonism between constraint and disorder that may provide a more appropriate description of the behavior of living systems than is possible using conventional dynamics. By quantifying the apophatic side of evolution, the Shannon approach to information achieves what no other treatment of the subject affords: It opens the window on a more encompassing perception of reality.

Keywords: apophasis; constraint; entropy; flexibility; hegelian dialectic; information; meaning; positivism; probability; sustainability

1. A World with Absences

The most important thing about information theory is not information. In today's "Age of Information", as the centennial of the birth of 1960s media guru, Marshall McLuhan [1] is being celebrated, his ideas are enjoying a revival. One of McLuhan's most famous tenets was that truly novel discoveries usually induce numbness in society. In such a benumbed state the usual response is to

interpret that which is new in terms of that which is old and familiar—pouring new wine into old wineskins, so to speak. His favorite example concerned the corporation, International Business Machines (IBM), which early in its existence saw itself as designing and manufacturing machines to facilitate business. The leadership of IBM long remained numb to the actual nature of its activity, and it wasn't until they perceived that their focus was not building machines but processing information that the enterprise grew to mega proportions.

The same numbness can still be seen in conventional evolutionary theory. In 1859 Charles Darwin published his understanding of the *process* of evolution. The notion of process, however, was totally foreign to the existing neoplatonic framework of science, so that those who subsequently interpreted Darwin tried to force process into a Platonic mold. The result has been the grievous minimalism now known as Neo-Darwinism.

The encounter of science with information seems to have elicited the same numbness that McLuhan had suggested. For three centuries now science could be described as almost an entirely positivistic and apodictic venture. No surprise, then, that science should focus entirely on the positivist role of information in how matters transpire. But, in a somewhat ironic reversal of McLuhan's IBM example, some are slowly beginning to realize that a possibly more significant discovery may be the new capability to quantify the absence of information, or "not information".

To assess the importance of the apophatic, or that which is missing, it helps to reframe how Ludwig von Boltzmann [2] treated the subject. Boltzmann described a system of rarefied, non-interacting particles in probabilistic fashion. Probability theory quantifies the degree to which state i is present by a measure, p_i. Conventionally, this value is normalized to fall between zero and one by dividing the number of times that i has occurred by the total number of observations. Under this "frequentist" convention, the probability of i *not* occurring becomes $(1 - p_i)$. Boltzmann's genius, however, was in abjuring this conventional measure of non-occurrence in favor of the negative of the logarithm of p_i. (It should be noted that $-\log(p_i)$ and $(1 - p_i)$ vary in uniform fashion, *i.e.*, a one-to-one relationship between the two functions exists). His choice imposed a strong asymmetry upon matters. Conventionally, calculating the average nonbeing in the system using $(1 - p_i)$ results in the symmetrical parabolic function $(p_i - p_i^2)$. If, however, one calculates average absence using Boltzmann's measure, the result,

$$H = -\sum_i p_i \log(p_i)$$

(1)

becomes skewed towards smaller p_i (or larger $[1 - p_i]$), *i.e.*, towards nonbeing. His H function suited well the phenomenology of the second law of thermodynamics and ex-post-facto accorded with the notion of a universe in which vacuous space is constantly increasing.

2. Confounding Entropy with Information

Claude E. Shannon [3] independently noted that $-\log(p_i)$ also was a suitable measure of the degree of surprise an observer would experience upon an encounter with state i. If $p_i \approx 1$, there is little surprise; however, if p_i is very near zero one experiences major surprise when i occurs. To observe i when p_i is small was said to provide much information. It followed from this reasoning that the average surprisal, which is formally identical to Boltzmann's H function, should provide a convenient

gauge of the total information inherent in the system. Thus it came to pass that the positivist notion of information was confounded with Boltzmann's apophatic measure, H. To make matters worse, John von Neumann suggested (as a joke) to Shannon that he call his function "entropy" following the connection that Boltzmann had drawn with the second law. Sadly, Shannon took the suggestion seriously [4].

Confusion about H stems from the fact that the measure embodies aspects of mutually-exclusive attributes. Ernst von Weizsäcker [5] noted this mutual exclusivity of what he labeled "novelty" and "confirmation" and concluded that "*meaningful* information … does not lend itself as being quantified by one single mathematical expression" [6]. While Weizsäcker's observation may be correct, it does not preclude the possibility that complementarity of the two notions might be apprehended by two separate but related terms. Toward this end, it becomes necessary to segregate the opposing notions within H. A clue to how this might be accomplished comes from noting that the surprise accompanying the observation of i when p_i is small can be assessed only post-facto. In reality, one is comparing the *apriori* probability p_i with the *aposteriori* probability of unity. It is the *change in probability apriori vs. aposteriori* that assigns a magnitude of information to the observation. This relational necessity led Tribus and McIrvine [7] to define information as anything that causes a *change* in probability assignment.

Tribus' definition also identifies *how* information is related to the underlying probability theory. Information deals with the changes in probabilities in the same sense that Newtonian derivatives are related to common algebraic variables. Information is identified not with a probability distribution pre-se, but always *relative* to another distribution. It follows that it is erroneous to identify H with the information inherent in probability distribution p_i. It is possible to speak of information in apodictic fashion only insofar as a given distribution p_i relates to some other distribution, p_i.

It immediately follows that the obverse criticism pertains to Boltzmann's use of H as a general measure of entropy. H is not an appropriate measure of entropy, because the third law of thermodynamics states that entropy can be measured only in relation to some reference state. Although the convention in thermodynamics is to set the reference point as zero degrees Kelvin, more generally the requirement is that some reference state be specified. That Boltzmann may not have been aware of the relativistic nature of entropy is understandable, given as how it was formulated only later by Nernst [8]. Whether von Neumann's joke to Shannon can be as readily forgiven remains, however, subject to debate [4].

3. Parsing What Is from What is Not

It is clear that both information and entropy are relativistic and must always be treated in the context of changing probabilities. Unfortunately, Shannon's "entropy" is identical neither to the common sense of information *nor* to the thermodynamic sense of entropy. The saving grace of Shannon's formulation, however, is that it is built upon solid axiomatic foundations, such as extensivity, additivity and isotropy [3,9]. These mathematical conveniences allow one to parse the Boltzmann/Shannon formula into two independent terms that quantify the relationship of a distributed variable *a* with the distribution of any other variable *b* as,

$$H = A + \Phi \tag{2}$$

where,

$$H = -\sum_{i.j} p(a_i, b_j) \log[p(a_i, b_j)] \tag{2a}$$

$$A = \sum_{i,j} p(a_i, b_j) \log\left[\frac{p(a_i, b_j)}{p(a_i)p(b_j)}\right] \tag{2b}$$

and

$$\Phi = -\sum_{i,j} p(a_i, b_j) \log\left[\frac{p(a_i, b_j)^2}{p(a_i)p(b_j)}\right] \tag{2c}$$

Here the Boltzmann/Shannon measure, H, is applied to the joint probability distribution, p(a_i,b_j) (the distribution of co-occurrences of all possible pairs a_i and b_j).

One notes that if a_i and b_j are completely independent of each other, then p(a_i,b_j) = p(a_i)p(b_j), A = 0 and Φ = H. This is precisely the context (a perfect gas of non-interacting particles) in which Boltzmann developed the measure. If, however, a_i and b_j exert any constraints upon each other whatsoever, then p(a_i,b_j) ≠ p(a_i)p(b_j), and it is possible to prove that 0 < A, Φ < H. In other words, A measures the degree of mutual *constraint* that a_i and b_j exert upon each other (Weizsäcker's "confirmation"). In communication theory, A is assumed to gauge the amount of information that a_i reveals about b_j and *vice-versa*. It is called the *average mutual information* between the distributions a_i and b_j. By contrast, Φ is said to represent the *conditional entropy* between the same distributions.

The particular boundary conditions that Boltzmann chose forced H = Φ. One should note, however, that this equality does not hold for systems of interacting elements [10], so that Φ becomes a more appropriate general measure of entropy. That is, it is more appropriate to call Φ the entropy of *a* in relation to *b*. H, then, is more indicative of the overall *capacity* [11] of a system for either constraint or freedom between *a* and *b*.

4. A Clearer Image of Information

An appreciation for the relativistic nature of information and its measurement resolves several conundrums regarding information and "meaning" [12]. Shannon's colleague, Warren Weaver [13] noted as how "two messages, one heavily loaded with meaning, and the other pure nonsense, can be equivalent as regards information". (Weaver accepted the Shannon formula as a full and complete definition of information.) Along the same lines, if one applies the Shannon formula to the grey scale of pixels on a television screen, the value of H is maximal when there is no signal to the set ("snow") and no correlation between adjacent pixels (A = 0 and Φ = H). It is nonsensical to argue that maximal information inheres in such a display. When an intelligible picture does appear on the screen, the values of adjacent pixels then become correlated (A > 0 and Φ < H).

If one inquires whether the pattern on the screen is *meaningful* to an observer, the answer will depend on how well the image correlates with the perceptual history of the observer. While quantifying such correlation may remain difficult, identifying the possibility of such correspondence is

important, because one of foremost criticisms of the Shannon approach is that it cannot address "meaning" in any realistic way. This criticism is understandable in light of the fact that the original Shannon "entropy" is a conflation of *both* constraint and flexibility. That same criticism, however, need not apply to relativistic formulations of information, as demonstrated by the following numerical examples (that speak cogently to Weaver's "nonsense puzzle"):

The following are three random strings of 200 digits:

Sequence A:

42607951361038986072061245134123237671363847519601557824849686201007746224524209 37159144904694056560480338986072061245134123237671363847519601557824849686201007 7462245242093715914490469405656048033898

Sequence B:

03617746439242093715914490469405656048033898607206124513412323767136384751960155 78248496862010077462245242093715914490469405656048033898607206124513412323767136 38475196015578248496862010077462245242 09

Sequence C:

01475623843789694751743102380318185453848905236473225910906494173735504160210176 85326300670460724247097189694751743102380318185453848905236473225910906494173779 50411021017685326300670460724247097970 8969

The values H for each sequence are 3.298, 3.288 and 3.296 bits, respectively. That no internal order is present in any of the sequences is shown by the average mutual information values of adjacent pairs of digits in each of the three cases (as with the adjacent pixels on a TV screen). These calculate to 10.97%, 10.03% and 9.94% of the respective paired entropies. Each fraction is typical of a random distribution of 200 tokens among 10 types. Relationships between more distant pairs are likewise random.

Next, the correspondences between the three pairs of sequences are examined. Recording how each digit in A pairs with the occupant in its corresponding location in B yields a joint entropy of 5.900 bits, 11.61% of which appears as mutual information (once again, random correspondence). Similar pairings between sequences A and C, however, reveal that fully 91.69% of the joint "entropy" consists of mutual *information* between the sequences. Obviously, the sequences A and C are closely related. In fact, close scrutiny of them shows either to be an arbitrary permutation of the other, along with a handful of "mistakes".

While these comparisons may appear to some as typical exercises in coding/decoding, they actually have deeper implications. Instead of digits, one could have used as categories symbols for codons in a genome (A,C,T,G) or monomers in a protein (Gly, Ala, Leu, Trp, *etc*.) In the latter situation, sequence A might represent the order of proteins on the outer surface of an antibody in the plasma of an organism, while B and C might describe corresponding patterns on the surfaces of microbes present in the same fluid. While B appears to bear no relationship to A, C would match A in almost "hand- in- glove" fashion.

In such a situation the pattern in C would provide ultimate meaning to A. The match would signify the end towards which A was created by the immune system and would initiate a highly directed action on

the part of A (to eliminate the microbe). This significance is clearly apparent in the high value of mutual information between the sequences. Whence, although the primitive Shannon measure does not by itself convey meaning, the relative information indicated by A clearly provides at least a sense of "proto meaning". That such "meaning" for antibodies is but a pale shadow of meaning in the human context only reflects how wanly quantitative models in general prefigure more complicated human situations. In order to get from meaningless physical phenomena to full-blown human semiosis, it is necessary to pass through some inchoate precursor of meaning. Shannon measures, it would appear, are *at least* useful in treating this transitional phase.

5. Quantifying What is Absent

While these two examples highlight more accurately the positive role that information plays in living systems, less attention is usually paid to the residual Φ that represents flexibility. Most would rather ignore Φ in a science that is overwhelmingly positivist and apodictic, because rewards go to those who focus upon identifying the constraints that guide how things happen. The instances where physics addresses anything other than the positivistic are indeed very few—the Pauli Exclusion Principle and Heisenberg uncertainty are the only exceptions within this writer's memory.

Physics, however, deals almost exclusively with the homogeneous, but as soon as one leaves the realm of universals and enters the very heterogeneous world of the living, the *absence* of an object or a trait can loom large [14]. In ecology, for example, the absence of a particular resource or predator often weighs heavily on whether a given population persists or vanishes. For that matter, it is difficult even to talk about patterns without referring to the absences or holes in spatial arrays. Whence, accounting (literally) for absences takes on significant importance in the life sciences.

Such accounting, however, is precisely what Boltzmann initiated (whether consciously or unconsciously). Furthermore, Boltzman weighted non-being so as to skew its importance vis-à-vis that which exists, thereby providing a bias that accords with the second law. Now, it happens that Boltzmann's formula pertains to circumstances far more complex than his rarified, homogeneous and non-interacting example system. Even in highly complex systems, Boltzmann's H can be parsed into separate terms that gauge constraint and flexibility, respectively.

Such parsing requires the comparison of two distributions with one another. There is no prohibition, however, against abstracting the two distributions from the same system. This was done above, for example, when the (non-significant) values of A were calculated on successive pairs of integers within each string of 200 integers. Of possibly greater utility is the comparison of the past (*aposteriori*) of a system with its (*apriori*) future, as can readily be accomplished within networks of interactions [15].

To parse a network in this fashion one considers the interaction strengths, T_{ij}, that join arbitrary components (nodes) i and j. The joint probability that i interacts with j can be estimated in frequentist manner as $p(i,j) \approx T_{ij} / \sum_{m,n} T_{mn}$. The immediate past of j can likewise be estimated as the afferent conditional distribution $p(j) \approx T_{ij} / \sum_{m} T_{mj}$, and the immediate future by its efferent conditional distribution $p(i) \approx T_{ij} / \sum_{n} T_{in}$. One may now substitute these distributions into Equations (2a–c) for H, A and Φ to yield,

$$H = -\sum_{i,j} \left(\frac{T_{ij}}{\sum\limits_{m,n} T_{mn}} \right) \log \left(\frac{T_{ij}}{\sum\limits_{m,n} T_{mn}} \right) \tag{3a}$$

$$A = \sum_{i,j} \left(\frac{T_{ij}}{\sum\limits_{m,n} T_{mn}} \right) \log \left(\frac{T_{ij} \sum\limits_{m,n} T_{mn}}{\sum\limits_{n} T_{in} \sum\limits_{m} T_{mj}} \right) \tag{3b}$$

and

$$\Phi = -\sum_{i,j} \left(\frac{T_{ij}}{\sum\limits_{m,n} T_{mn}} \right) \log \left(\frac{T_{ij}^{2}}{\sum\limits_{n} T_{in} \sum\limits_{m} T_{mj}} \right) \tag{3c}$$

The reader should note that nothing need to be known concerning the particular details of the constraints that guide the constitutive links, nor about the specifics of the degeneracies that contribute to Φ. All one needs to calculate the overall system constraint and flexibility are the phenomenological observations T_{ij}. Such ability to calculate overall properties in abstraction of any micro details is reminiscent of similar calculations in statistical thermodynamics.

Being able to quantify the overall constraint inhering in a system (A) is a major step forward, but it could be argued that the ability to quantify *that which is absent* (Φ) represents an even greater advance. As Terrence Deacon [16] argues, arithmetics had been limited in what it could do until the Ninth Century invention of the cipher 0 to represent absence as a positional number, whereupon a host of arithmetic operations were considerably facilitated. One could argue similarly that the limits that obligate positivism has placed upon the ability of science to address living systems can now be superseded. In the Boltzmann/Shannon approach to information theory one obtains something possibly more important than the quantification of information (constraint) itself—one can now quantify how much is missing from a system.

6. The Necessity of That Which is Absent

In terms of ecological (and likely as well economic, social and immune) systems what is missing can be of critical importance. Parallel redundant pathways, inefficient and incoherent processes all contribute to the magnitude of Φ. While they often hinder the efficient functioning of the system (as gauged by A), it is precisely such "noise" that is required by a system if it is to mount a response to a novel perturbation [17]. Lacking sufficient apophasis, a highly efficient system becomes "brittle" and doomed to collapse at the first new perturbation [18]. It is imperative that living systems retain a degree of *Fehlerfreundlichkeit* ("error-friendliness" [19])

Furthermore, to endure and remain sustainable, it appears that a system must possess even more flexibility (Φ) than constraint (A). Available data on ecosystems indicate that such balance occurs within a narrow range of values of the quotient A/H [20]. Systems that are out-of-balance, such as eutrophic ecosystems, often lack sufficient Φ to persist. Remediation then requires an increase in flexibility (more apophasis, or a decrease in constraint) [21].

The necessity for apophasis bears strongly upon the issue of preserving biodiversity. In recent decades much effort has justifiably been invested at the global level towards the conservation of biodiversity. Society intuitively senses that maintenance of biodiversity is necessary for global ecological health. What is hardly ever mentioned, however, is that solid theoretical justification for preserving biodiversity has been wanting. In retrospect, we see why this is so: Having only positivist tools at one's disposal, one cannot hope to circumscribe the interplay between constraint and looseness that provides sustainability. But the definitions of A and Φ now engender a quantitative methodology with which to follow the dynamics between the apodictic and the apophatic. Furthermore, such analysis often reveals that it is an increase in the latter that becomes necessary for system survival.

7. The Heraclitian Drama

Equation (2) suggests that the relationship between constraint and flexibility is complementary or antagonistic in most systems. Conventionally, attention focuses upon the dynamics inherent in the apodictic variable, A, but it should be clear that not maintaining the complementary Φ leads the system inevitably towards collapse. It is counterproductive, therefore, to regard the dynamics of living systems solely as some mechanical/material juggernaut that grinds inexorably towards some maximal efficiency. Any perspective on ecosystems that ignores apophasis can thereby be labeled "one-eyed-ecology" [22].

Virtually all domains of science remain "one-eyed" in scope, save for the discipline of thermodynamics, where entropy explicitly appears as a manifestation of the apophatic (although it is rarely acknowledged as such). Schroedinger coined the term "negentropy" to refer to the inverse of entropy, and there have been numerous treatments of the entropy-negentropy conversation. Terrance Deacon [23] has spearheaded the need for acknowledging the role of the apophatic in biology (although not in quantitative terms). In economics the role of the apophatic occasionally arises under the rubric of "externalities", but economists are reluctant to divert their attentions from conventional dynamics.

Although the aim of this collection of essays has been a better apodictic notion of information, perhaps a more important goal should be a fuller appreciation for the dialectic between constraint and flexibility. In the end, the metaphor of transaction provides a more appropriate context within which to appraise the dynamics of living systems, because the dynamics of life cannot be minimalized as "matter moving according to universal laws" [24]. It resembles rather a Heraclitian dialectic between the buildup of organization and its decay according to the second law.

To put a finer point on the dialectic, one notes that the opposition between generation and decay is not absolute. In Hegelian fashion, each of the countervailing trends requires the other at some higher level: The development of new adaptive repertoires requires a cache of what formerly appeared as redundant, inefficient, incoherent and dissipative processes. On the other hand, greater constrained performance always generates increased dissipation.

In the dialectical scenario, information as commonly perceived becomes a degenerate subclass of the more general notion of constraint. No longer is it necessary to treat information using the narrow rubrics of communication theory. That Shannon developed his mathematics in that theatre can be

regarded as an historical accident. His ensuing quantifications apply far more broadly, not just to constraint in general, but possibly more importantly, to the lack of constraint as well.

Even though Shannon's H formula rests upon solid axiomatic foundations, it has engendered perhaps as much confusion as it did enlightenment. The selfsame formula was believed to quantify the mutually exclusive attributes of entropy and information. Such contradiction has strained logic and spawned abstruse narration, e.g., Brillouin [25]. The root of the confusion had its origin in von Neuman's most unfortunate suggestion to tie the formula to Boltzmann's entropy. Boltzmann's applied his equation to a hyper-simplistic system wherein H was constrained to represent only the most disorganized state of affairs. More generally, however, an increase in the H function can also pertain to situations in which organization actually increases [22].

The key to resolving confusion about H is to tie the function not only to the second law of thermodynamics, but to connect it to the third law as well. Entropy can never be defined in absolute terms, but acquires *meaning* only in relation to a reference state. Defining H in a relational context obviates any schizoid interpretation by allowing its decomposition into two agonistic (complementary) terms that quantify the degree of system constraint and its residual freedom, respectively. The implication of this decomposition is that the third law applies not only to the concept of entropy, but to the conjugate constraint as well. That is, constraint (and its degenerate subclass, information) has no meaning in abstraction from the third law. Like entropy, information is *always* relative. That is, what is measured as constraint or information of *a* in the context of *b* will generally be different from that between *a* and c [6].

In hindsight it is now clear why the H function alone is a poor surrogate for many of its intended applications (beginning with entropy per se). As demonstrated above, H fails to represent "meaning", whereas its relational component, A, appears capable to the task. Other purported shortcomings of Shannon information should be re-examined as well in the light of equation (2) above.

In conclusion, it is highly premature to dismiss the Shannon/Boltzmann approach for measuring information, because something else as important as information is at stake. Other attempts at improving the apodictic characterization of information fail to encompass the necessary roles for apophasis. The Boltzmann/Shannon mathematics provides in the end a richer and more inclusive vantage on the dynamics of nature—one that allows the scientist to open his/her blind eye towards the broader causality at work in the living world. In that sense, it can truly be said that the most important contribution that information theory makes to science is not information.

Acknowledgments

References

1. McLuhan, M. *Understanding Media*; Mentor: New York, NY, USA, 1964.
2. Boltzmann, L. Weitere studien über das wärmegleichgewicht unter gasmolekülen. *Wien. Ber.* **1872**, *66*, 275-370.

3.   Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379-423, 623-656.

4.   Müller, I. *A History of Thermodynamics—The Doctrine of Energy and Entropy*: Springer: New York, NY, USA, 2007.

5.   von Weizsäcker, E.U. Erstmaligkeit und Bestätigung als Komponenten der Pragmatischen Information. In *Offene Systeme: Band I. Beiträge zur Zeitstruktur von Information, Entropie und Evolution*, von Weizsäcker, E., Ed.; Klett-Cotta: Stuttgart, Germany, 1974; pp. 82-113.

6.   von Weizsäcker, E.U.; von Weizsäcker, C. Information, evolution and „error-friendliness. *Biol. Cybern.* **1998**, *79*, 501-506.

7.   Tribus, M.; McIrvine, E.C. Energy and information. *Sci. Am.* **1971**, *225*, 179-188.

8.   Nernst, W. Über die Beziehung zwischen Wärmeentwicklung und maximaler Arbeit bei kondensierten Systemen. *Ber. Kgl. Pr. Akad. Wiss.* **1906**, *52*, 933-940.

9.   Khinchin, A.I. *Mathematical Foundations of Information Theory*; Dover: New York, NY, USA, 1957.

10. Ulanowicz, R.E. Increasing entropy: Heat death or perpetual harmonies? *Int. J. Des. Nat. Ecodyn.* **2009**, *4*, 83-96.

11. Ulanowicz, R.E. An hypothesis on the development of natural communities. *J. Theor. Biol.* **1980**, *85*, 223-245.

12. Ulanowicz, R.E. *Growth & Development: Ecosystems Phenomenology*; Springer-Verlag: New York, NY, USA, 1986; p. 203.

13. Weaver, W. *Lady Luck: the theory of probability*; Dover:Mineola, NY, 1963.

14. Bateson, G. *Steps to an Ecology of Mind*; Ballantine Books: New York, NY, USA, 1972.

15. Rutledge, R.W.; Basorre, B.L.; Mulholland, R.J. Ecological stability: An information theory viewpoint. *J. theor. Biol.* **1976**, 57, 355-371.

16. Deacon, T.W. *Incomplete Nature: How Mind Emerged from Matter*; W. W. Norton & Co.: New York, NY, USA, 2011.

17. Ulanowicz, R.E. *Ecology, the Ascendent Perspective*; Columbia University Press: New York, NY, USA, 1997; p. 201.

18. Holling, C.S. The Resilience of Terrestrial Ecosystems: Local Surprise and Global Change. In *Sustainable Develoment of the Biosphere*, Clark, W.C., Munn, R.E., Eds.; Cambridge University Press: Cambridge, UK, 1986; pp. 292-317.

19. von Weizsäcker, C.; von Weizsäcker, E.U. Fehlerfreundlichkeit. In *Offenheit-Zeitlichkeit-Komplexitat: Zur Theorie der Offenen Systeme*, Kornwacs, K., Ed.; Klett-Cotta: Stuttgart, Germany, 1984; pp. 167-201.

20. Ulanowicz, R.E. The dual nature of ecosystem dynamics. *Ecol. Model.* **2009b**, *220*, 1886-1892.

21. Ulanowicz, R.E. Quantitative Methods for Ecological Network Analysis and Its Application to Coastal Ecosystems. In *Treatise on Estuarine and Coastal Science*, Wolanski, E., McLusky, D.S., Eds.; Elsevier: Amsterdam, The Netherlands, in press; Section 9.04.

22. Ulanowicz, R.E.; Goerner, S.J.; Lietaer, B.; Gomez, R. Quantifying sustainability: Resilience, efficiency and the return of information theory. *Ecol. Complex.* **2009**, *6*, 27-36.

23. Deacon, T.W. Emergence: The Hole at the Wheel's Hub. In *The Re-Emergence of Emergence: The Emergentist Hypothesis from Science to Religion*, Clayton, P., Davies, P., Eds.; Oxford University Press: New York, NY, USA, 2008; pp. 111-150.

24. Ayala, F.J. Darwin's Revolution. Biological Evolution: Facts and Theories. In *Proceedings of the III STOQ International Conference*, Rome, Italy, 3–9 March 2009**.**

25. Brillouin, L. *Science and Information Theory*; Academic Press: New York, NY, USA, 1956.