Version 3.2

Hunter-gatherers in a howling wilderness: Neoliberal capitalism as a language that speaks itself

Rodrick Wallace, PhD
Division of Epidemiology
The New York State Psychiatric Institute *

February 7, 2011

Abstract

The 'self-referential' character of evolutionary process noted by Goldenfeld and Woese (2010) can be restated in the context of a generalized Darwinian theory applied to economic process through a 'language' model: The underlying inherited and learned culture of the firm, the short-time cognitive response of the firm to patterns of threat and opportunity that is sculpted by that culture, and the embedding socioeconomic environment, are represented as interacting information sources constrained by the asymptotic limit theorems of information theory. If unregulated, the larger, compound, source that characterizes high probability evolutionary paths of this composite then becomes, literally, a self-dynamic language that speaks itself. Such a structure is, for those enmeshed in it, more akin to a primitive hunter-gatherer society at the mercy of internal ecological dynamics than to, say, a neolithic agricultural community in which a highly ordered, deliberately adapted, ecosystem is consciously farmed so as to match its productivity to human needs.

Key Words: economics, evolution, information theory, large deviations, Morse Function, punctuated equilibrium, renormalization, universality class tuning

1 Introduction

Haldane and May (2011), taking the 'econophysics' perspective of Caccioli et al. (2009), recently explored risk in banking ecosystems, adopting tools from network theory to study the effects of interaction between individual subcomponents leading to the propagation of shocks within large-scale financial structures. Other approaches to the origin and propagation of such 'shocks' arise more naturally from the generalized Darwinian perspective of Aldrich et al. (2008), based on a necessary-conditions application of the Modern Evolutionary Synthesis to economic phenomena.

Wallace (2010a) has proposed expanding the Modern Synthesis itself by introducing 'The principle of environmental interaction,' i.e., that individuals and groups engage in powerful, often punctuated, dynamic mutual relations with their embedding environments that may include the exchange of heritage material between markedly different organisms. Wallace (2011) applies the expanded model to the generalized Darwinism of Aldrich et al. (2008). Escaping the intellectual straightjacket of mathematical population genetics and analogous forms of replicator dynamics – or at least exchanging it for a slightly larger one – the approach characterizes the heritage system of the firm, the cognitive process by which the firm responds to patterns of threat and opportunity, and embedding socioeconomic environment, as interacting information sources constrained by the asymptotic limit theorems of information theory. This leads to an inherently coevolutionary system described in terms of a formalism quite similar to that of Onsager's nonequilibrium thermodynamics, having quasi-stable 'coevolutionary' states coupled by highly structured large deviations, all much in the sense of Champagnat et al. (2006). The possibility arises that such structured large deviations, rather than merely expressing the self-dynamic processes of a language that speaks itself, can be harnessed by an external 'farmer', that is, regulated to produce a directed socioeconomic ecosystem akin to the primitive neolithic agriculture that enabled the construction of richer social and cultural milieus.

The Wallace work introduces powerful methods from the statistical physics of phase transitions into generalized Darwinian evolutionary theory, much in the spirit of the recent paper by Goldenfeld and Woese (2010), who focus on evolution 'as a problem in nonequilibrium statistical mechanics, where the key dynamical modes are collective'. They provide a central insight:

> ...[T]he genome encodes the information which governs the response of an organism to its physics and biological environment. At the same time, this environment actually shapes genomes through gene transfer processes and phenotype selection. Thus,

*Contact: wallace@pi.cpmc.columbia.edu. Box 47, 1051 Riverside Drive, New York, New York, 10032.

> we encounter a situation where the dynamics must be self-referential: the update rules change during the time evolution of the system, and the way in which they change is a function of the state and thus the history of the system... self-referential dynamics is an inherent and probably defining feature of evolutionary dynamics and thus biological systems.

Here we explore such self-referential dynamics explicitly from the perspectives of Wallace (2010a, 2011), recognizing that the representation of fundamental biological and socioeconomic processes in terms of information sources restrains their inherent nonequilibrium nature. That is, although the operation of information sources is both nonequilibrium and irreversible in the most fundamental sense (e.g., few and short palindromes), the asymptotic limit theorems of information theory beat back, somewhat, the mathematical thicket surrounding such phenomena. The theorems permit something of a formal regularization of inherently nonequilibrium processes under proper circumstances that may lead to the development of new statistical tools for the study of empirical data beyond the narrow confines of network theory.

2 Basic formalism

The evolutionary process of generalized Darwinism, in the sense of Aldrich et al. (2008), as envisioned by Wallace (2011), involves dynamic interplay between (at least) three information sources representing transmission of corporate heritage, the cognitive response of a corporation to patterns of threat and opportunity, and embedding environment, given that both corporation and environment 'remember', producing serial correlations in time. We suppose it possible to coarse-grain observational measures of those three processes, representing the results in terms of some 'alphabet' of possible states. Our interest is in (properly characterized, and possibly very long) temporal paths beginning at some initial state a_0, and having the form

$$x_n \equiv \{a_0, a_1, ..., a_n\},$$

where the a_j are possible elements of the coarsegrained alphabet.

Given a particular tripartite starting point, a_0, evolution, being inherently path dependent, must build on what has gone before. Thus, crudely, subsequent paths can be divided into two classes, a vast set having vanishingly small probability, and a much smaller high probability set that, we suppose, follows something like the regularities of information theory that govern the three component information sources. That is, if $N(n)$ is the number of high probability paths of length n, then there exists a *path independent* limit H such that

$$H = \lim_{n \to \infty} \frac{\log[N(n)]}{n}. \quad (1)$$

Below we will indicate how the restriction of path independence might be lifted, somewhat.

We assume that, associated with each path x_n of length n, there is an information source X_n producing it that is defined in terms of the joint and conditional probabilities

$$P(a_0, a_1, ...a_n)$$

and

$$P(a_n|a_{n-1}, ..., a_1, a_0),$$

such that appropriate Shannon uncertainties may be defined (e.g., Ash, 1990; Khinchin, 1957; Cover and Thomas, 2006), and that the Shannon-McMillan Theorem holds:

$$H = \lim_{n \to \infty} \frac{\log[N(n)]}{n} =$$

$$\lim_{n \to \infty} H(X_n|X_{n-1}, ..., X_0) =$$

$$\lim_{n \to \infty} \frac{H(X_0, X_1, ..., X_n)}{n+1}. \quad (2)$$

We now shift perspective, defining equivalence classes of paths, and an associated symmetry groupoid (simplest example, a disjoint union of groups: see the Mathematical Appendix) that will be needed for the characterization of collective phenomena, much in the sense that a symmetry group is needed for Landau's theory of phase transition.

We call two states a_j and a_k equivalent if there is a high probability path beginning with a_0 that reaches them. The set of high probability paths beginning at a_0 defines the possible evolutionary processes that start at that state, and the set of equivalence classes defines a groupoid in a standard manner that characterizes the information source H associated with them.

We can now index the set of possible evolutionary information sources by the groupoids defining the equivalence classes of high probability paths associated with them.

Next, allow the initial state to vary, that is, allow different starting points, a_0, across the system. This produces an even larger groupoid that will enable our analysis of certain collective phenomena.

3 Punctuated equilibrium: phase transitions in evolution

As Feynman (2000) argues, based on work by Bennett (1988), information is simply another form of free energy, and the information in a message is quite precisely measured by the free energy needed to erase it. Indeed, Feynman (2000) shows how to construct an (idealized) machine that directly converts the information in a message to work.

But there are subtleties. First, information sources are already inherently irreversible dynamic systems. For example, in spoken or written English, the short sequence ' the ' has much higher probability than its time reversed ' eht '. There is no local reversibility, and adaptation of methods from nonequilibrium statistical mechanics or thermodynamics will not be graced with 'Onsager reciprocal relations'.

Another subtlety is that, in spite of the inherently nonequilibrium dynamic nature of an information source, the asymptotic limit theorems defining information source uncertainty appear to permit 'nonequilibrium equilibria' in a certain sense.

We suppose there to be some monotonic increasing measure of available free energy M, $Q(M), Q(0) = 0$. We assume that possible generalized Darwinian trajectories are constrained by the availability of resources, so that the probability of an (inherently irreversible and highly dynamic) information source associated with groupoid element G_j, at a fixed $Q(M)$, is given, in a first approximation, by the standard expression for the Gibbs distribution

$$P[H_{G_j}] = \frac{\exp[-H_{G_j}/Q]}{\sum_i \exp[-H_{G_i}/Q]}.$$

(3)

As Goldenfeld (2010) has pointed out, the Gibbs distribution appears to be not really appropriate for systems evolving in an open manner, and we will generalize the treatment somewhat, using an adiabatic approximation in which the dynamics remain 'close enough' to a form in which the mathematical theory can work, adapting standard phase transition formalism for transitions between adiabatic realms. In particular, rather than using exponential terms, one might well use any functional form $f(H_{Gi}, Q)$ such that the sum over i converges.

In essence, however, by adopting an information source perspective on evolutionary process we implicitly incorporate the possibility of 'nonequilibrium equilibria' in the sense of Eldredge and Gould (1972).

As we shall show, the 'E-property' that Khinchin (1957) identifies – the division of paths into high and low probability sets – the limiting relation

$$\lim_{n\to\infty} \frac{\log[N(n)]}{n} = H$$

and its variants for all high probability paths generated by an ergodic information source, permit imposition of a powerful regularity onto inherently nonequilibrium evolutionary processes.

The partition function-analog of this strange system is, as usual, defined as

$$Z_G(Q) = \sum_i \exp[-H_{G_i}/Q].$$

(4)

We can now define a highly simplified evolutionary 'groupoid free energy', F_G, *constructed over the full set of possible evolutionary trajectories as constrained by available free energy*, as

$$\exp[-F_G/Q] \equiv \sum_i \exp[-H_{G_i}/Q],$$

(5)

so that

$$F_G(Q) = -Q \log[Z_G(Q)].$$

(6)

This is to be taken as a Morse Function, in the sense of the Mathematical Appendix. As we shall show below, other – essentially similar – Morse Functions may perhaps be defined on this system, having a more 'natural' interpretation from information theory.

Argument is now by abduction from statistical physics (Landau and Lifshitz, 2007; Pettini, 2007). The Morse Function F_G is seen as constrained by free energy availability in a manner that allows application of Landau's theory of punctuated phase transition in terms of groupoid, rather than group, symmetries.

Recall, now, Landau's perspective on phase transition (Pettini, 2007). The essence of his insight was that certain physical phase transitions took place in the context of a significant symmetry change, with one phase being more symmetric than the other. A symmetry is lost in the transition, i.e., spontaneous symmetry breaking. The greatest possible set of symmetries being that of the Hamiltonian describing the energy states. Usually, states accessible at lower temperatures will lack the symmetries available at higher temperatures, so that the lower temperature state is less symmetric, and transitions can be highly punctuated.

Here, we have characterized the dependence of evolutionary process on the availability of metabolic free energy in terms of groupoid, rather than group, symmetries, and the argument by abduction is essentially similar: Increasing availability of free energy – rising $Q(M)$ – will allow richer interactions between the three basic economic information sources, and will do so in a highly punctuated manner, as in Eldredge and Gould (1972).

4 Extending the model

4.1 Kadanoff theory

Given F_G as a free energy analog, we are interested in a mathematical treatment of transitions between adiabatic realms and suppose it possible to define a characteristic 'length', say r, on the system, as more fully described below. We can then define renormalization symmetries in terms of the 'clumping' transformation, so that, for clumps of size R, in an external 'field' of strength J (that we can set to 0 in the limit), one can write, in the usual manner (e.g., Wilson, 1971)

$$F_G[Q(R), J(R)] = f(R)F_G[Q(1), J(1)],$$

$$\chi(Q(R), J(R)) = \frac{\chi(Q(1), J(1))}{R}, \qquad (7)$$

where χ is a characteristic correlation length.

As Wallace (2005) shows, following Wilson (1971), very many 'biological' renormalizations, $f(R)$, are possible that lead to a number of quite different universality classes for phase transition. Wallace (2005) and Wallace and Fullilove (2008) describe how 'universality class tuning' can be used as a tool for large-scale regulation of the system. See the Mathematical Appendix for a summary.

In order to define the metric r, we impose a topology on the system, so that, near a particular 'language' A defining some H_G there is (in an appropriate sense) an open set U of closely similar languages $\hat{A}$, such that $A, \hat{A} \subset U$.

Since the information sources are 'similar', for all pairs of languages $A, \hat{A}$ in U, it is possible to:

1. Create an embedding alphabet which includes all symbols allowed to both of them.

2. Define an information-theoretic distortion measure in that extended, joint alphabet between any high probability (grammatical and syntactical) paths in A and $\hat{A}$, which we write as $d(Ax, \hat{A}x)$ (Cover and Thomas, 2006). Note that these languages do not interact, in this approximation.

3. Define a metric on U, for example,

$$r(A, \hat{A}) = |\lim \frac{\int_{A,\hat{A}} d(Ax, \hat{A}x)}{\int_{A,A} d(Ax, A\hat{x})} - 1|, \qquad (8)$$

using an appropriate integration limit argument over the high probability paths. Note that the integration in the denominator is over different paths within A itself, while in the numerator it is between different paths in A and $\hat{A}$. Consideration suggests r is indeed a formal metric.

Clearly, other approaches to metric construction on U seem possible, and other approaches to renormalization than outlined by equation (7).

4.2 Nonergodic information sources

The ergodic nature of an information source is a generalization of the law of large numbers and implies that the long-time averages can be closely approximated by averages across the probability spaces of those sources. For non-ergodic information sources, a function, $\mathcal{J}(x_n)$, of each path $x_n \to x$, may be defined, such that $\lim_{n\to\infty} \mathcal{J}(x_n) = \mathcal{J}(x)$, but $\mathcal{J}$ will not in general be given by the simple cross-sectional laws-of-large numbers analogs above (Khinchin, 1957).

Let $s \equiv d(x, \hat{x})$ for high probability paths x and $\hat{x}$, where d is a distortion measure, as described in Cover and Thomas (2006). For 'nearly' ergodic systems one might use something of the form

$$\mathcal{J}(\hat{x}) \approx \mathcal{J}(x) + sd\mathcal{J}/ds|_{s=0}$$

for s sufficiently small. The idea is to take a distortion measure as a kind of Finsler metric, imposing a resulting 'global' structure over an appropriate class of non-ergodic information sources. One question obviously revolves around what properties are metric-independent, in much the same manner as the Rate Distortion Theorem.

These heuristics can be made more precise:

Take a set of 'high probability' paths $x_n \to x$.

Suppose, for all such x, there is an open set, U, containing x, on which the following conditions hold:

1. For all paths $\hat{x}_n \to \hat{x} \in U$, a distortion measure $s_n \equiv d_U(x_n, \hat{x}_n)$ exists.

2. For each path $x_n \to x$ in U there exists a pathwise invariant function $\mathcal{J}(x_n) \to \mathcal{J}(x)$, in the sense of Khinchin (1957, p.72). While such a function will almost always exist, only in the case of an ergodic information source can it be identified as an 'entropy' in the usual sense.

3. A function $F_U(s_n, n) \equiv f_n \to f$ exists, for example,

$$f_n = s_n, \log[s_n]/n, s_n/n,$$

and so on.

4. The limit

$$\lim_{n\to\infty}\frac{\mathcal{J}(x_n)-\mathcal{J}(\hat{x}_n)}{f_n}\equiv\nabla_F\mathcal{J}|_x$$

exists and is finite.

Under such conditions, standard global atlas/manifold constructions are possible. Again, $\mathcal{J}$ is not simply given by the usual expressions for source uncertainty if the source is not ergodic, and the phase transition development above may be correspondingly more complicated. Restriction to high probability paths simplifies matters considerably, although precisely characterizing them may be difficult, requiring extension of the Shannon-McMillan Theorem and its Rate Distortion generalization.

An essential question is under what circumstances this differential treatment for 'almost' ergodic information sources permits something very much like what Khinchin (1957, p. 54) calls the 'E property' enabling classification of paths into a small set of high probability and a vastly larger set of vanishingly small probability (Khinchin, 1957, p. 74).

4.3 Network information theory: toward more 'natural' Morse Functions

As Goldenfeld (2010) has pointed out, equation (3), the Gibbs distribution, seems, on the surface, not really appropriate for a system evolving in an open manner, although, as we have argued, the regularities imposed by the asymptotic limit theorems of information theory permit study of 'nonequilibrium equilibria' in a standard way via the interpretation of equation (6) as a Morse Function. For example, the Gibbs distribution approach has had considerable success in reframing key results in protein folding dynamics (Wallace, 2010b). Here we extend that treatment, adopting a perspective from network information theory (e.g., Cover and Thomas, 2006; El Gamal and Kim, 2010). The theory is, however, much a work in progress, with many unsolved difficulties. As El Gamal and Kim note, the simplistic model of a network consisting of separate links and naive forwarding nodes does not capture many important aspects of real world networked systems that involve multiple sources with various messaging requirements, redundancies, time and space correlations, and time variations. As they note, the goal in many information systems is not merely to communicate source information, but to make a decision or coordinate an action – in our context, cognitive process. Indeed, the first paper on network information theory was by Claude Shannon himself, who did not solve the question of optimal rates, a matter that remains open (Shannon, 1961), along with many others.

We suppose that a measure of available free energy is itself associated with an information source, Z, representing the intents of an external 'farmer' who provides regulation to the system. This source represents an identifiable subset of the environmental dynamics and provides an embedding context for evolutionary process. It defines jointly typical paths (Cover and Thomas, 2006) for an associated set of economic information sources.

Given three interacting information sources, Y_1, Y_2, Z, the splitting criterion for tripartite jointly typical sequences, taking Z as an external context, is (Cover and Thomas, 2006, p. 524)

$$I(Y_1;Y_2|Z)=H(Z)+H(Y_1|Z)+H(Y_2|Z)-H(Y_1,Y_2,Z),$$

(9)

where $H(...|...)$ and $H(...,...,...)$ represent conditional and joint uncertainties (Ash, 1990; Khinchin, 1957; Cover and Thomas, 2006).

This presumably generalizes to something like

$$I(Y_1;...;Y_n|Z)=H(Z)+\sum_{j=1}^{n}H(Y_j|Z)-H(Y_1,...,Y_n,Z).$$

(10)

More complicated multivariate typical sequences receive much the same treatment (El Gamel and Kim, 2010, p.2-26). Given a basic set of information sources $(X_1,...,X_k)$ that one partitions into two ordered sets $X(\mathcal{K})$ and $X(\mathcal{K}')$, then the splitting criterion becomes $H(X(\mathcal{K})|X(\mathcal{K}'))$. Generalization to more ordered sets is straightforward.

Then the joint splitting criterion – I, H above – however it may be expressed as a composite of the underlying information sources and their interactions, satisfies a relation closely analogous to the first one in equation (2), where $N(n)$ is the number of high probability jointly typical paths of length n. This expression is, then, essentially the same as equations (5) and (6) in that the joint splitting criterion *is given as a functional composition of the underlying information sources and their interactions.*

There are two immediate implications of this insight.

First, I in equation (10) and its generalizations can be considered as Morse Functions in the sense of the Mathematical Appendix that can be parametrized in terms of the monotonic expression involving some appropriate index of available free energy Q. The natural association of equivalence classes of evolutionary states and trajectories with groupoid symmetries then suggests that Landau's spontaneous symmetry breaking arguments, extended to groupoids, will again apply, producing richer and more 'symmetric' socioeconomic processes and structures as Q increases, leading to analogs to serial endosymbiosis and a sequence of 'eukaryotic-like' transitions to more highly structured socioeconomic systems.

Second, since I in equation (10) and its generalizations have the form of a free energy, we can *directly* invoke biological-like renormalization relations like equation (7) (Wallace, 2005), e.g.,

$$I[Q(R), J(R)] = f(R)I[Q(1), J(1)],$$

$$\chi(Q(R), J(R)) = \frac{\chi(Q(1), J(1))}{R},$$

(11)

where we again parametrize by the scalar function Q of available metabolic free energy as above. The splitting criterion I and its generalizations are supposed to be adiabatically piecewise stationary ergodic between phase transitions, so that the asymptotic limit theorems work 'well enough', while the transitions themselves are associated with universality classes according the particular form of $f(R)$. The universality class tuning of Wallace (2005) permits regulation of the phase transitions, and allows another layer of external control.

This reformulation is, then, a more complete answer to the concerns of Goldenfeld (2010) regarding the appropriateness of the Gibbs distribution under these circumstances, although characterization of F_G from equation (6) as a Morse Function might well be a sufficient argument.

In summary, I in equation (10) and the more complicated versions of the splitting criteria for multivariate typical sequences are to be taken as Morse Functions, so that Pettini's (2007) topological hypothesis applies, and Landau's symmetry breaking arguments carry through, albeit in a groupoid context, so that 'symmetry', i.e. evolutionary complexity, can increase with increase in available free energy in an inherently punctuated manner. I and the other splitting criteria analogous to equation (10), however, have, in a sense, a more 'natural' interpretation than F_G.

The inference is that choice of a proper Morse Function may depend strongly on context, with a simple Gibbs distribution sufficient for strongly 'physics-bound' processes such as protein folding (Wallace, 2010b), while more complex splitting criteria are to be associated with more complex biological, social, or economic phenomena.

4.4 Large deviations

Wallace (2010a) has taken a particularly recognizable nonequilibrium statistical mechanics approach to evolutionary dynamics. In that work the interaction of genes, (cognitive) gene expression, and environmental information sources is expressed using the coevolutionary formalism of Chapagnat et al. (2006). The basic idea is to write each information source as a function of those with which it interacts:

$$H_m = H_m(Q_1, ..., Q_s, ...H_j...), j \neq m.$$

where the Q_k represent other relevant parameters. The dynamics of such a system is defined by the usual recursive network of stochastic differential equations, using gradients in a 'disorder' construct as analogs to the more usual gradients in entropy, the thermodynamic forces:

$$S_m \equiv H_m - \sum_j \partial H_m / \partial K_j,$$

(12)

where we have expressed both the H_j and Q_j as driving parameters K_j, again with the proviso that one not express H_m directly as a function of itself.

Then, via the homology between information and free energy, the dynamics become driven by the usual Onsager set of stochastic differential equations,

$$dK_t^j = \sum_i [L_{i,j}(t, ...\partial S_m/\partial K_i...)dt + \sigma_{i,j}(t, ...\partial S_m/\partial K_i...)dB_t^i] =$$

$$L_j(t, K_1, ..., K_n)dt + \sum_i \sigma_{i,j}(t, K_1, ..., K_n)dB_t^i$$

(13)

where we have collected and simplified terms. L_j and the $\sigma_{i,j}$ are functions, and the terms dB_t^j represent different kinds of 'noise' constrained by particular forms of quadratic variation, in the usual manner. Standard texts abound.

Again, since information sources are not locally time-reversible, there are no 'Onsager reciprocal relations'.

Several patterns are obvious.

1. Setting this system of equations to zero and solving for stationary points gives quasi-equilibrium attractor states since the noise terms preclude unstable equilibria. The system then undergoes diffusive drift about the equilibrium configuration.

2. The system may converge to a limit cycle or a pseudo-random strange attractor.

3. What is converged to, however, is not a simple state or set of such states. Rather, this system, via the constraints imposed by the asymptotic limit theorems of information theory, converges to an equivalence class of of highly dynamic information sources coupled by mutual crosstalk, and equivalence classes define groupoids, as above. In effect, via the Shannon-McMillan Theorem that defines the information source uncertainty, we have driven the mathematical thicket one layer back, expressing a dynamical system in terms of a relatively

simple formalism abducted from nonequilibrium statistical mechanics.

As Champagnat et al. (2006) note, however, shifts between the quasi-equilibria of this system of equations can be addressed by the large deviations formalism. They find that the issue of evolutionary dynamics drifting away from trajectories predicted by the canonical equation can be investigated by considering the asymptotic of the probability of 'rare events' for the sample paths of the diffusion.

By 'rare events' they mean diffusion paths drifting far away from the canonical equation. The probability of such rare events is governed by a large deviation principle: when a critical parameter (designated ϵ) goes to zero, the probability that the sample path of the diffusion is close to a given rare path ϕ decreases exponentially to 0 with rate $\mathcal{I}(\phi)$, where the 'rate function' $\mathcal{I}$ can be expressed in terms of the parameters of the diffusion. This result, in their view, can be used to study long-time behavior of the diffusion process when there are multiple attractive evolutionary singularities. Under proper conditions the most likely path followed by the diffusion when exiting a basin of attraction is the one minimizing the rate function $\mathcal{I}$ over all the appropriate trajectories. The time needed to exit the basin is of the order $\exp(V/\epsilon)$ where V is a quasi-potential representing the minimum of the rate function $\mathcal{I}$ over all possible trajectories.

An essential fact of large deviations theory is that the rate function $\mathcal{I}$ which Champagnat et al. invoke can almost always be expressed as a kind of entropy, that is, having the canonical form

$$\mathcal{I} = -\sum_j P_j \log(P_j) \tag{14}$$

for some probability distribution. This result goes under a number of names; Sanov's Theorem, Cramer's Theorem, the Gartner-Ellis Theorem, the Shannon-McMillan Theorem, and so forth (Dembo and Zeitouni, 1998; R. Wallace and R.G. Wallace, 2008).

These considerations lead very much in the direction of equation (13), but now seen as subject to internally-driven large deviations *that are themselves described as information sources*, providing H-parameters that can trigger punctuated shifts between quasi-stable modes, in addition to resilience transitions driven by 'catastrophic' external events or the exchange of heritage information between different classes of 'organisms', in a large sense.

Figure 1 is a schematic that links this perspective to the Morse Theory treatment of section 4.3. I, as a Morse Function, is subject to punctuated transitions in a driving 'metabolic' parameter that we call Q. As Q increases, spontaneous symmetry breaking permits, say, a transition to more complex 'eukaryotic' structures via some analog to serial endosymbiosis: the transition from the lower cluster to the higher. But this is seen to take place via a highly structured large deviation *that is itself constrained as being the output of an information source*. This may be determined by internal self-dynamic forces, or it may be imposed from without by a culturally-specific 'farmer'.

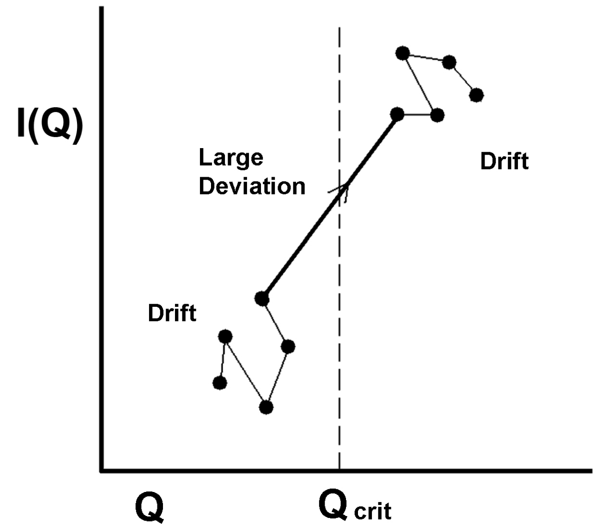

Figure 1: Spontaneous symmetry breaking in I as an approximation to a structured large deviation driven by rise of a monotonic index of available free energy. Unlike a simple physical system, such a transition can occur if Q increases beyond Q_{crit}, but will not do so in the absence of a highly structured large deviation. Rise in Q is therefore a necessary, but not sufficient, condition. In a 'farmed' system the large deviation is directed by the intent of the farmers. In a 'creative destruction' Schumpeterian system, the large deviation is entirely determined by self-referential internal phenomena, independent of the wishes or welfare of those who constitute the individual elements of that system.

The spontaneous symmetry breaking argument is thereby seen as a simplified approximation to the coevolutionary formalism of Champagnat et al. (2006), as adapted by Wallace (2010a). Such transitions can occur, but, unlike simple physical systems, need not occur, in the absence of a large deviation that is itself highly structured. To reiterate, in figure 1, increase of available indices of free energy (or other resources) is a necessary, but not sufficient, condition for punctuated evolutionary change that must be driven by a 'self-dynamic' or 'farmed' large deviation having either its own grammar and syntax or that given it by the farmer.

5 Discussion and conclusions: emerging from the wilderness

Clearly, something analogous to what Goldenfeld and Woese (2010) want to do can, in fact, be done, at least in terms of a generalized Darwinian theory of evolutionary collective phenomena that has roots in physics. But life and socioeconomics are not physics: the self-referential nature of generalized Darwinian evolutionary process is truly something different. While dependent on indices of available free energy and constrained by physical principles, from the perspectives of this analysis, raw evolution is a language that speaks itself. For example, available free energy, written as $Q(M)$ above, can itself be an evolutionary product, as, in biological systems, with the aerobic transition. In socioconomic terms, the acquisition of fire, domestication of farm animals for plowing, development of road systems enabling transfer, hence increased availability of existing energy and resources, development of steam technology, use of fossil fuels, and so on, provide examples. The formal description of such bootstrapping will require more comprehensive methods than are available by abduction from relatively simple physical theory, as Goldenfeld and Woese (2010) have noted.

Again, the example of figure 1 suggests that changes of indices representing available free energy can be a necessary, but not sufficient, condition for eukaryotic-like transitions to greater complexity. Evolution is indeed self referential.

Firms instantiate cognitive processes that take cues from the embedding environment to produce behavioral responses. Modes of such expression having adaptive value can become fixed in the cultural heritage of the firm by learning or selection. But evolution in simple Schumpeterian market economies will remain self-dynamic, self-referential, continually bootstrapping phenomena, in effect, languages that speak themselves, independent of the needs or wishes of those embedded in them.

But a socioeconomic system, unlike possible biological counterparts, is a cultural artifact. There is nothing 'natural' to any particular such construct, although the dynamics are constrained by resource availability in the context of historical trajectory and other cultural factors. Within those riverbanks the socioeconomic stream can flow according to its own dynamics, or it can be subject to rigorous cultural channeling. The metaphors of hunter-gatherer vs. farmer are not inappropriate.

Farmed ecosystems are inherently more productive, from a human perspective, than what can be gathered from raw nature. The transition from literal hunter-gatherer societies to neolithic farming enabled the subsequent construction of rich human ecosystems, including cities, city-states, and more elaborate structures. At present, Western neoliberal ideologies of unregulated Schumpeterian 'free markets' have given unfettered reign to an enormous structure with self-dynamic 'large deviations' that possess a grammar and syntax whose internal logic is unaffected by human needs or concerns. Some billions of us ride a rampant, rapidly evolving, socioeconomic engine that has neither engineer nor conductor. We are, very essentially, a tribe of primitive hunter-gatherers at the mercy of an unstable ecological monstrosity that we do not have the political will to control. Emerging from the present howling wilderness of neoliberal capitalism will require a farmed economic ecosystem, a large-scale agricultural economics that must be culturally tailored to local conditions. As with language, music, art, and all the rest, there can be no one, fixed farmed economy that will fit all needs at all times.

The universality class tuning outlined in the Mathematical Appendix provides some insight into means of regulating otherwise disruptive phase transitions in economic systems.

There is, of course, a cautionary note to what we have done here. Pielou (1977, p. 106) warns that mathematical models in biology and ecology are only useful as subordinate partners in a continuing dialog with data: models can only recommend perspectives for subsequent empirical test that, in turn, can be used to correct the models. Replacing the intellectual straightjacket one set of economic theories with another driven by the asymptotic limit theorems of information theory will not address the essential scientific problems now facing generalized evolutionary theory applied to economic process. These will yield only to data-based empirical study in which mathematical models are only one among many possible tools: the word is not the thing.

6 Mathematical appendix

6.1 Groupoids

Following Weinstein (1996), states a_j, a_k in a set A are related by the groupoid morphism if and only if there exists a high-probability grammatical path connecting them to the same base point, and tuning across the various possible ways in which that can happen parameterizes the set of equivalence relations and creates the groupoid. This assertion requires some development.

Note that not all possible pairs of states (a_j, a_k) can be connected by such a morphism, that is, by a high-probability, grammatical and syntactical path linking them with some given base point. Those that can define the groupoid element, a morphism $g = (a_j, a_k)$ having the natural inverse $g^{-1} = (a_k, a_j)$. Given such a pairing, it is possible to define 'natural' end-point maps $\alpha(g) = a_j, \beta(g) = a_k$ from the set of morphisms G into A, and a formally associative product in the groupoid $g_1 g_2$ provided $\alpha(g_1 g_2) = \alpha(g_1), \beta(g_1 g_2) = \beta(g_2)$, and $\beta(g_1) = \alpha(g_2)$. Then the product is defined, and associative, $(g_1 g_2) g_3 = g_1 (g_2 g_3)$.

In addition, there are natural left and right identity elements λ_g, ρ_g such that $\lambda_g g = g = g \rho_g$ (Weinstein, 1996).

An orbit of the groupoid G over A is an equivalence class for the relation $a_j \sim G a_k$ if and only if there is a groupoid element g with $\alpha(g) = a_j$ and $\beta(g) = a_k$. Following Cannas da Silva and Weinstein (1999), we note that a groupoid is called transitive if it has just one orbit. The transitive groupoids are the building blocks of groupoids in that there is a natural decomposition of the base space of a general groupoid into orbits. Over each orbit there is a transitive groupoid, and the disjoint union of these transitive groupoids is the original groupoid. Conversely, the disjoint union of groupoids is itself a groupoid.

The isotropy group of $a \in X$ consists of those g in G with $\alpha(g) = a = \beta(g)$. These groups prove fundamental to classifying groupoids.

If G is any groupoid over A, the map $(\alpha, \beta) : G \to A \times A$ is a morphism from G to the pair groupoid of A. The image of (α, β) is the orbit equivalence relation $\sim G$, and the functional kernel is the union of the isotropy groups. If $f : X \to Y$ is a function, then the kernel of f, $ker(f) = [(x_1, x_2) \in X \times X : f(x_1) = f(x_2)]$ defines an equivalence relation.

Groupoids may have additional structure. As Weinstein (1996) explains, a groupoid G is a topological groupoid over a base space X if G and X are topological spaces and α, β and multiplication are continuous maps. A criticism sometimes applied to groupoid theory is that their classification up to isomorphism is nothing other than the classification of equivalence relations via the orbit equivalence relation and groups via the isotropy groups. The imposition of a compatible topological structure produces a nontrivial interaction between the two structures. Below we will introduce a metric structure on manifolds of related information sources, producing such interaction.

In essence, a groupoid is a category in which all morphisms have an inverse, here defined in terms of connection to a base point by a meaningful path of an information source dual to a cognitive process.

As Weinstein (1996) points out, the morphism (α, β) suggests another way of looking at groupoids. A groupoid over A identifies not only which elements of A are equivalent to one another (isomorphic), but *it also parameterizes the different ways (isomorphisms) in which two elements can be equivalent*, i.e., all possible information sources dual to some cognitive process. Given the information theoretic characterization of cognition presented above, this produces a full modular cognitive network in a highly natural manner.

Brown (1987) describes the fundamental structure as follows:

> A groupoid should be thought of as a group with many objects, or with many identities... A groupoid with one object is essentially just a group. So the notion of groupoid is an extension of that of groups. It gives an additional convenience, flexibility and range of applications...
>
> EXAMPLE 1. A disjoint union [of groups] $G = \cup_\lambda G_\lambda, \lambda \in \Lambda$, is a groupoid: the product ab is defined if and only if a, b belong to the same G_λ, and ab is then just the product in the group G_λ. There is an identity 1_λ for each $\lambda \in \Lambda$. The maps α, β coincide and map G_λ to λ, $\lambda \in \Lambda$.
>
> EXAMPLE 2. An equivalence relation R on [a set] X becomes a groupoid with $\alpha, \beta : R \to X$ the two projections, and product $(x, y)(y, z) = (x, z)$ whenever $(x, y), (y, z) \in R$. There is an identity, namely (x, x), for each $x \in X$...

Weinstein (1996) makes the following fundamental point:

> Almost every interesting equivalence relation on a space B arises in a natural way as the orbit equivalence relation of some groupoid G over B. Instead of dealing directly with the orbit space B/G as an object in the category S_{map} of sets and mappings, one should consider instead the groupoid G itself as an object in the category G_{htp} of groupoids and homotopy classes of morphisms.

The groupoid approach has become quite popular in the study of networks of coupled dynamical systems which can be defined by differential equation models, (Golubitsky and Stewart, 2006).

6.2 Morse Theory

Morse theory examines relations between analytic behavior of a function – the location and character of its critical points – and the underlying topology of the manifold on which the function is defined. We are interested in a number of such functions, for example a 'free energy' constructed from information source uncertainties on a parameter space and 'second order' iterations involving parameter manifolds determining critical behavior. These can be reformulated from a Morse theory perspective. Here we follow closely the elegant treatments of Pettini (2007).

The essential idea of Morse theory is to examine an n-dimensional manifold M as decomposed into level sets of some function $f : M \to \mathbf{R}$ where $\mathbf{R}$ is the set of real numbers. The a-level set of f is defined as

$$f^{-1}(a) = \{x \in M : f(x) = a\},$$

the set of all points in M with $f(x) = a$. If M is compact, then the whole manifold can be decomposed into such slices in a canonical fashion between two limits, defined by the minimum and maximum of f on M. Let the part of M below a be defined as

$$M_a = f^{-1}(-\infty, a] = \{x \in M : f(x) \leq a\}.$$

These sets describe the whole manifold as a varies between the minimum and maximum of f.

Morse functions are defined as a particular set of smooth functions $f : M \to \mathbf{R}$ as follows. Suppose a function f has a critical point x_c, so that the derivative $df(x_c) = 0$, with critical value $f(x_c)$. Then f is a Morse function if its critical points are nondegenerate in the sense that the Hessian matrix of second derivatives at x_c, whose elements, in terms of local coordinates are

$$H_{i,j} = \partial^2 f / \partial x^i \partial x^j,$$

has rank n, which means that it has only nonzero eigenvalues, so that there are no lines or surfaces of critical points and, ultimately, critical points are isolated.

The index of the critical point is the number of negative eigenvalues of H at x_c.

A level set $f^{-1}(a)$ of f is called a critical level if a is a critical value of f, that is, if there is at least one critical point $x_c \in f^{-1}(a)$.

Again following Pettini (2007), the essential results of Morse theory are:

1. If an interval $[a, b]$ contains no critical values of f, then the topology of $f^{-1}[a, v]$ does not change for any $v \in (a, b]$. Importantly, the result is valid even if f is not a Morse function, but only a smooth function.

2. If the interval $[a, b]$ contains critical values, the topology of $f^{-1}[a, v]$ changes in a manner determined by the properties of the matrix H at the critical points.

3. If $f : M \to \mathbf{R}$ is a Morse function, the set of all the critical points of f is a discrete subset of M, i.e. critical points are isolated. This is Sard's Theorem.

4. If $f : M \to \mathbf{R}$ is a Morse function, with M compact, then on a finite interval $[a, b] \subset \mathbf{R}$, there is only a finite number of critical points p of f such that $f(p) \in [a, b]$. The set of critical values of f is a discrete set of $\mathbf{R}$.

5. For any differentiable manifold M, the set of Morse functions on M is an open dense set in the set of real functions of M of differentiability class r for $0 \leq r \leq \infty$.

6. Some topological invariants of M, that is, quantities that are the same for all the manifolds that have the same topology as M, can be estimated and sometimes computed exactly once all the critical points of f are known: Let the Morse numbers $\mu_i (i = 1, ..., m)$ of a function f on M be the number of critical points of f of index i, (the number of negative eigenvalues of H). The Euler characteristic of the complicated manifold M can be expressed as the alternating sum of the Morse numbers of any Morse function on M,

$$\chi = \sum_{i=0}^{m} (-1)^i \mu_i.$$

The Euler characteristic reduces, in the case of a simple polyhedron, to

$$\chi = V - E + F$$

where V, E, and F are the numbers of vertices, edges, and faces in the polyhedron.

7. Another important theorem states that, if the interval $[a, b]$ contains a critical value of f with a single critical point x_c, then the topology of the set M_b defined above differs from that of M_a in a way which is determined by the index, i, of the critical point. Then M_b is homeomorphic to the manifold obtained from attaching to M_a an i-handle, i.e., the direct product of an i-disk and an $(m - i)$-disk.

Again, Pettini (2007) contains both mathematical details and further references. See, for example, Matusmoto (2002) or the classic by Milnor (1963).

6.3 Universality class tuning

6.3.1 Biological renormalization

Equation (7) states that the information source and the correlation length, the degree of coherence on the underlying network, scale under renormalization clustering in chunks of size R as

$$H[K_R, J_R]/f(R) = H[J, K]$$

$$\chi[K_R, J_R]R = \chi(K, J),$$

with $f(1) = 1, K_1 = K, J_1 = J$, where we have slightly rearranged terms.

Differentiating these two equations with respect to R, so that the right hand sides are zero, and solving for dK_R/dR and dJ_R/dR gives, after some consolidation, expressions of the form

$$dK_R/dR = u_1 d\log(f)/dR + u_2/R$$

$$dJ_R/dR = v_1 J_R d\log(f)/dR + \frac{v_2}{R} J_R.$$

(15)

The $u_i, v_i, i = 1, 2$ are functions of K_R, J_R, but not explicitly of R itself.

We expand these equations about the critical value $K_R = K_C$ and about $J_R = 0$, obtaining

$$dK_R/dR = (K_R - K_C) y d\log(f)/dR + (K_R - K_C) z/R$$

$$dJ_R/dR = w J_R d\log(f)/dR + x J_R/R.$$

(16)

The terms $y = du_1/dK_R|_{K_R=K_C}, z = du_2/dK_R|_{K_R=K_C}, w = v_1(K_C, 0), x = v_2(K_C, 0)$ are constants.

Solving the first of these equations gives

$$K_R = K_C + (K - K_C) R^z f(R)^y,$$

(17)

again remembering that $K_1 = K, J_1 = J, f(1) = 1$.

Wilson's (1971) essential trick is to iterate on this relation, which is supposed to converge rapidly near the critical point, assuming that for K_R near K_C, we have

$$K_C/2 \approx K_C + (K - K_C)R^z f(R)^y. \tag{18}$$

We iterate in two steps, first solving this for $f(R)$ in terms of known values, and then solving for R, finding a value R_C that we then substitute into the first of equations (7) to obtain an expression for $H[K, 0]$ in terms of known functions and parameter values.

The first step gives the general result

$$f(R_C) \approx \frac{[K_C/(K_C - K)]^{1/y}}{2^{1/y} R_C^{z/y}}. \tag{19}$$

Solving this for R_C and substituting into the first expression of equation (7) gives, as a first iteration of a far more general procedure (Shirkov and Kovalev, 2001), the result

$$H[K, 0] \approx \frac{H[K_C/2, 0]}{f(R_C)} = \frac{H_0}{f(R_C)}$$

$$\chi(K, 0) \approx \chi(K_C/2, 0)R_C = \chi_0 R_C, \tag{20}$$

which are the essential relationships.

Note that a power law of the form $f(R) = R^m, m = 3$, which is the direct physical analog, may not be biologically reasonable, since it says that 'language richness' can grow very rapidly as a function of increased network size. Such rapid growth is simply not observed.

Taking the biologically realistic example of non-integral 'fractal' exponential growth,

$$f(R) = R^\delta, \tag{21}$$

where $\delta > 0$ is a real number which may be quite small, equation we can be solve for R_C, obtaining

$$R_C = \frac{[K_C/(K_C - K)]^{[1/(\delta y + z)]}}{2^{1/(\delta y + z)}} \tag{22}$$

for K near K_C. Note that, for a given value of y, one might characterize the relation $\alpha \equiv \delta y + z =$ constant as a 'tunable universality class relation' in the sense of Albert and Barabasi (2002).

Substituting this value for R_C back gives a complex expression for H, having three parameters: δ, y, z.

A more biologically interesting choice for $f(R)$ is a logarithmic curve that 'tops out', for example

$$f(R) = m \log(R) + 1. \tag{23}$$

Again $f(1) = 1$.

Using Mathematica 4.2 or above to solve equation (19) for R_C gives

$$R_C = [\frac{Q}{LambertW[Q \exp(z/my)]}]^{y/z}, \tag{24}$$

where

$$Q \equiv (z/my)2^{-1/y}[K_C/(K_C - K)]^{1/y}.$$

The transcendental function LambertW(x) is defined by the relation

$$LambertW(x) \exp(LambertW(x)) = x.$$

It arises in the theory of random networks and in renormalization strategies for quantum field theories.

An asymptotic relation for $f(R)$ would be of particular biological interest, implying that 'language richness' increases to a limiting value with population growth. Taking

$$f(R) = \exp[m(R-1)/R]$$

(25)

gives a system which begins at 1 when $R = 1$, and approaches the asymptotic limit $\exp(m)$ as $R \to \infty$. Mathematica 4.2 finds

$$R_C = \frac{my/z}{LambertW[A]},$$

(26)

where

$$A \equiv (my/z)\exp(my/z)[2^{1/y}[K_C/(K_C - K)]^{-1/y}]^{y/z}.$$

These developments indicate the possibility of taking the theory significantly beyond arguments by abduction from simple physical models, although the notorious difficulty of implementing information theory existence arguments will undoubtedly persist.

6.3.2 Universality class distribution

Physical systems undergoing phase transition usually have relatively pure renormalization properties, with quite different systems clumped into the same 'universality class,' having fixed exponents at transition (Binney et al., 1986). Biological and social phenomena may be far more complicated:

If the system of interest is a mix of subgroups with different values of some significant renormalization parameter m in the expression for $f(R, m)$, according to a distribution $\rho(m)$, then the first expression in equation (7) should generalize, at least to first order, as

$$H[K_R, J_R] =< f(R, m) > H[K, J]$$

$$\equiv H[K, J] \int f(R, m)\rho(m)dm.$$

(27)

If $f(R) = 1 + m\log(R)$ then, given any distribution for m,

$$< f(R) >= 1+ < m > \log(R)$$

(28)

where $< m >$ is simply the mean of m over that distribution.

Other forms of $f(R)$ having more complicated dependencies on the distributed parameter or parameters, like the power law R^δ, do not produce such a simple result. Taking $\rho(\delta)$ as a normal distribution, for example, gives

$$< R^\delta >= R^{<\delta>}\exp[(1/2)(\log(R^\sigma))^2],$$

(29)

where σ^2 is the distribution variance. The renormalization properties of this function can be determined from equation (19), and the calculation is left to the reader as an exercise, and can be done in Mathematica 4.2 or above.

Thus the information dynamic phase transition properties of mixed systems will not in general be simply related to those of a single subcomponent, a matter of possible empirical importance: If sets of relevant parameters defining renormalization universality classes are indeed distributed, experiments observing pure phase changes may be very difficult. Tuning among different possible renormalization strategies in response to external signals would result in even greater ambiguity in recognizing and classifying information dynamic phase transitions.

Important aspects of mechanism may be reflected in the combination of renormalization properties and the details of their distribution across subsystems.

In sum, real biological, social, or interacting biopsychosocial systems are likely to have very rich patterns of phase transition which may not display the simplistic, indeed, literally elemental, purity familiar to physicists. Overall mechanisms will, however, still remain significantly constrained by the theory, in the general sense of probability limit theorems.

6.3.3 Punctuated universality class tuning

The next step is to iterate the general argument onto the process of phase transition itself, producing a tunable punctuation.

As described above, an essential character of physical systems subject to phase transition is that they belong to particular 'universality classes'. Again, this means that the exponents of power laws describing behavior at phase transition will be the same for large groups of markedly different systems, with 'natural' aggregations representing fundamental class properties (Binney et al., 1986).

It appears that biological or social systems undergoing phase transition analogs need not be constrained to such classes, and that 'universality class tuning', meaning the strategic alteration of parameters characterizing the renormalization properties of punctuation, might well be possible.

Here we focus on the tuning of parameters within a single, given, renormalization relation. Clearly, however, wholesale shifts of renormalization properties must ultimately be considered as well.

Universality class tuning has been observed in models of 'real world' networks. As Albert and Barabasi (2002) put it,

> The inseparability of the topology and dynamics of evolving networks is shown by the fact that [the exponents defining universality class] are related by [a] scaling relation..., underlying the fact that a network's assembly uniquely determines its topology. However, in no case are these exponents unique. They can be tuned continuously...

Suppose that a structured external environment, itself an appropriately regular information source **Y**, 'engages' a modifiable system characterized by an information source. The environment begins to write an image of itself on the system in a distorted manner permitting definition of a mutual information $I[K]$ splitting criterion according to the Rate Distortion or Joint Asymptotic Equipartition Theorems. K is an inverse coupling parameter between system and environment. At punctuation – near some critical point K_C – the systems begin to interact very strongly indeed, and, near K_C, using a simple physical model,

$$I[K] \approx I_0[\frac{K_C - K}{K_C}]^{\alpha}.$$

For a physical system α is fixed, determined by the underlying 'universality class.' Here we will allow α to vary, and, in the section below, to itself respond explicitly to imposed signals.

Normalizing K_C and I_0 to 1,

$$I[K] \approx (1 - K)^{\alpha}.$$

(30)

The horizontal line $I[K] = 1$ corresponds to $\alpha = 0$, while $\alpha = 1$ gives a declining straight line with unit slope which passes through 0 at $K = 1$. Consideration shows there are progressively sharper transitions between the necessary zero value at $K = 1$ and the values defined by this relation for $0 < K, \alpha < 1$. The rapidly rising slope of transition with declining α is of considerable significance:

The *instability* associated with the splitting criterion $I[K]$ is defined by

$$Q[K] \equiv -KdI[K]/dK = \alpha K(1 - K)^{\alpha - 1},$$

(31)

and is singular at $K = K_C = 1$ for $0 < \alpha < 1$. Following earlier work (e.g., Wallace and Fullilove, 2008), we interpret this to mean that values of $0 < \alpha \ll 1$ are highly unlikely for real systems, since $Q[K]$, in this model, represents a kind of barrier for 'social' information systems.

On the other hand, smaller values of α mean that the system is far more efficient at responding to the adaptive demands imposed by the embedding structured environment or regulatory authority, since the mutual information which tracks the matching of internal response to external demands, $I[K]$, rises more and more quickly toward the maximum for smaller and smaller α as the inverse coupling parameter K declines below $K_C = 1$. That is, systems able to attain smaller α are more responsive to external signals than those characterized by larger values, in this model, but smaller values will be harder to reach, probably only at some considerable physiological or opportunity cost. Focused conscious action takes resources, of one form or another.

The more biologically realistic renormalization strategies given above produce sets of several parameters defining the universality class, whose tuning gives behavior much like that of α in this simple example.

Formal iteration of the phase transition argument on this calculation gives a tunable regulation, focusing on paths of universality class parameters:

Suppose the renormalization properties of an information source at some 'time' k are characterized by a set of appropriately coarse-grained parameters $A_k \equiv \alpha_1^k, ..., \alpha_m^k$. Fixed parameter values define a particular universality class for the renormalization. We suppose that, over a sequence of 'times', the universality class properties can be characterized by a path $x_n = A_0, A_1, ..., A_{n-1}$ having significant serial correlations which, in fact, permit definition of *another* adiabatically piecewise stationary ergodic information source associated with the paths x_n. Call that source **X**.

Suppose also, in the now-usual manner, that the set of external (or internal, systemic) signals impinging on the information source of basic interest is also highly structured and forms another information source **Y** that interacts not only with the system of interest globally, but specifically with its universality class properties as characterized by **X**. **Y** is necessarily associated with a set of paths y_n.

Pair the two sets of paths into a joint path, $z_n \equiv (x_n, y_y)$ and invoke an inverse coupling parameter, K, between the information sources and their paths. This leads, by the arguments above, to phase transition punctuation of $I[K]$, the mutual information between **X** and **Y**, under either the Joint Asymptotic Equipartition Theorem or under limitation by a distortion measure, through the Rate Distortion Theorem. The essential point is that $I[K]$ is a splitting criterion under these theorems, and thus partakes of the homology with free energy density which we have invoked above.

Activation of universality class tuning, the mean field model's version of attentional focusing, then becomes itself a punctuated event in response to increasing linkage between the organism and an external structured signal or some particular system of internal events.

This iterated argument exactly parallels the extension of the General Linear Model to the Hierarchical Linear Model in regression theory.

Another path to the fluctuating dynamic threshold might be through a second order iteration similar to that just above, but focused on the parameters defining the universality class distributions given above.

7 Acknowledgments

The author thanks Dr. D.N. Wallace for basic insights from population, community, and ecosystem ecology.

8 References

Albert, R., A Barabasi, (2002) Statistical mechanics of complex networks, Reviews of Modern Physics, 74:47-97.

Aldrich, H., G. Hodgson, D. Hull, T. Knudsen, J. Mokyr, V. Vanberg, (2008), In defense of generalized Darwinism. Journal of Evolutionary Economics, 18:577-596.

Ash, R., (1990), Information Theory, Dover, New York.

Bennett, C., (1988), Logical depth and physical complexity. In The Universal Turing Machine: A Half-Century Survey. R. Herkin (ed.), pp. 227-257, Oxford University Press.

Binney, J., N. Dowrick, A. Fisher, M. Newman, (1986), The Theory of Critical Phenomena, Clarendon Press, Oxford, UK.

Brown, R., (1987), From groups to groupoids: a brief survey. Bulletin of the London Mathematical Society, 19:113-134.

Caccioli, F., M. Marsilli, P. Vivo, Eroding market stability by proliferation of financial instruments, European Journal of Physics B, 71:467-479.

Cannas Da Silva, A., and A. Weinstein, (1999), Geometric Models for Noncommutative Algebras. American Mathematical Society, Providence, RI.

Cover, T., and J. Thomas, (2006), Elements of Information Theory, Second Edition, John Wiley and Sons, New York.

Champagnat, N., R. Ferriere, and S. Meleard, (2006), Unifying evolutionary dynamics: from individual stochastic process to macroscopic models. Theoretical Population Biology, 69:297-321.

Dembo, A., and O. Zeitouni, (1998), Large Deviations and Applications, Second Edition, Springer, New York.

El Gamal, A., and Y. Kim, (2010), Lecture Notes on Network Information Theory, arXiv:1001.3404v4.

Eldredge, N., and S. Gould, (1972), Punctuated equilibrium: an alternative to pyletic gradualism. In T. Schopf (ed.), Models in Paleobiology, Freeman Cooper and Co., San Francisco.

Feynman, R., (2000), Lectures on Computation, Westview Press, New York.

Goldenfeld, N., and C. Woese, (2010), Life is physics: evolution as a collective phenomenon far from equilibrium. ArXiv:1011.4125v1 [q-bio.PE].

Goldenfeld, N., (2010), Personal communication.

Golubitsky, M., and I. Stewart, (2006), Nonlinear dynamics and networks: the groupoid formalism. Bulletin of the American Mathematical Society, 43:305-364.

Haldane, A., R. May, 2011, Systemic risk in banking ecosystems, Nature, 469:351-355.

Kastner, M., (2006), Phase transitions and configuration space topology. ArXiv cond-mat/0703401.

Khinchin, A., (1957), Mathematical Foundations of Information Theory. Dover, New York.

Landau, L., and L. Lifshitz, (2007), Statistical Physics, Third Edition, Part 1. Elsevier, New York.

Matsumoto, Y., (2002), An Introduction to Morse Theory. Translations of Mathematical Monographs, Vol. 208, American Mathematical Society.

Milnor, J., (1963), Morse Theory. Annals of Mathematical Studies, Vol. 51, Princeton University Press.

Odling-Smee, F., K. Laland, and M. Feldman, (2003), Niche Construction: The Neglected Process in Evolution. Princeton University Press, Princeton, NJ.

Pettini, M., (2007), Geometry and Topology in Hamiltonian Dynamics and Statistical Mechanics. Springer, New York.

Pielou, E., (1977), Mathematical Ecology. John Wiley and Sons, New York.

Shannon, C., (1961), Two-way communication channels, in Proceedings of the 4th Berkeley Symposium in Mathematical Statistics and Proability, Vol. 1, University of California Press, pp. 611-644.

Wallace, R., and D. Wallace, (2009), Code, context, and epigenetic catalysis in gene expression. Transactions on Computational Systems Biology XI, LNBI 5750: 283-334.

Wallace, R., (2005), Consciousness: A Mathematical Treatment of the Global Neuronal Workspace Model. Springer, New York.

Wallace, R., (2009), Metabolic constraints on the eukaryotic transition. Origins of Life and Evolution of Biospheres, 38:165-176.

Wallace, R., (2010a), Expanding the modern synthesis. Comptes Rendus Biologies, 333:701-709.

Wallace, R., (2010b), Structure and dynamics of the 'protein folding code' inferred using Tlusty's topological rate distortion approach. In press, BioSystems, doi:10.1016/j.biosystems.2010.09.007.

Wallace, R., (2011), Unnatural selection: A new formal approach to punctuated equilibrium in economic process.

http://precedings.nature.com/documents/5612/version/1

Weinstein, A., (1996), Groupoids: unifying internal and external symmetry. Notices of the American Mathematical Association, 43:744-752.

Wilson, K., (1971), Renormalization group and critical phenomena. I. Renormalization group and the Kadanoff scaling picture. Physical Review B, 4:3174-3183.