Version 4.2

Metabolic free energy and deterministic-but-for-error biological codes: a 'Data Rate Theorem' aging model

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

March 17, 2014

Abstract

The living state is cognitive at every scale and level of organization. Since it is possible to associate a broad class of cognitive processes with 'dual' information sources, many pathologies can be addressed using statistical models based on the Shannon Coding, the Shannon-McMillan Source Coding, the Rate Distortion, and the Data Rate Theorems, as these provide powerful necessary condition constraints on all information generation and exchange, and on system control. Deterministic-but-for-error codes, although they may be fruitfully studied using information theoretic methods, do not fall so easily within these models. Such codes do not invoke cognition, although they may become essential subcomponents within larger cognitive processes. A formal argument, however, places code stability within a recognizably similar framework, with metabolic free energy serving as a 'control signal' stabilizing efficient operation of complex biochemical coding machinery. Demand beyond available energy supply then expresses itself in punctuated destabilization of the coding channel, affecting gene expression, protein folding, or the operation of the glycan/lectin cell interface. Aging, normal or prematurely driven by psychosocial or environmental stressors, would be expected to eventually interfere with routine code operation, triggering onset of many chronic diseases usually associated with senescence that involve failures of these mechanisms. Amyloid fibril formation is reviewed from this perspective.

Key Words: amyloid; chronic disease; gene expression; glycan code; information theory; protein misfolding; senescence

1 Introduction

Tlusty's (2007) information theoretic topological analysis of the genetic code relies on minimizing certain characteristic error measures. Wallace (2012a) examined the role of the availability of metabolic free energy in the evolution of such codes, using similar methods. Here we first generalize the argument, taking a somewhat more sophisticated approach based on a Black-Scholes 'cost' analysis. We then explore a model of punctuated code failure under free energy constraint that is roughly analogous to Data Rate Theorem (DRT) limitations in control theory (e.g., Nair et al. 2007). This will suggest a deeper understanding of the onset of the chronic diseases of aging, and of those driven by psychosocial or environmental stresses that cause premature aging.

The essential point of the DRT is the unification of control and information theories, finding that certain kinds of unstable systems cannot be stabilized if the rate of control information is below a critical limit, defined as the 'topological information' generated by the unstable system. Metabolic free energy plays a surprisingly similar role in stabilizing deterministic-but-for-error biological codes.

Tlusty's (2007) central idea is that

> To discuss the topology of errors we portray the codon space as a graph whose verticies are the codons... Two codons... are linked by an edge if they are likely to be confused by misreading... We assume that two codons are most likely to be confused if all their letters except for one agree and therefore draw an edge between them. The resulting graph is natural for considering the impact of translation errors on mutations because such errors almost always involve a single letter difference, that is, a movement along an edge of the graph to a neighboring vertex.
>
> The topology of a graph is characterized by its genus γ, the minimal number of holes required for a surface to embed the graph such that no two edges cross. The more connected that a graph is the more holes are required for its minimal embedding... [T]he highly interconnected 64-codon graph is embedded in a holey, $\gamma = 41$ surface. The genus is somewhat reduced to $\gamma = 25$ if we consider only 48 effective codons...
>
> The maximum [of an information-theoretic functional] determines a single contiguous domain where a certain amino acid is encoded... Thus every mode corresponds to an amino acid and the number of

*Wallace@nyspi.columbia.edu, rodrick.wallace@gmail.com

modes is the number of amino acids. This compact organization is advantageous because misreading of one codon as another codon within the same domain has no deleterious impact. For example, if the code has two amino acids, it is evident that the error-load of an arrangement where there are two large contiguous regions, each coding for a different amino acid, is much smaller than a 'checkerboard' arrangement of the amino acids.

This is analogous to the well-known topological coloring problem. However, in the coding problem one desires maximal similarity in the colors of neighboring 'countries', while in the coloring problem one must color neighboring countries by different colors. After some development (Tlusty 2008), the number of possible amino acids in this scheme is determined by Heawood's formula (Ringel and Young 1968).

More explicitly, Tlusty (2007) models the emergence of the genetic code as a transition in a noisy information channel, using an approach based on the Rate Distortion Theorem, with the optimal code is described by the minimum of a 'free energy'-like functional, allowing description of the code's emergence as a transition akin to a phase transition in statistical physics. The basis for this is the observation that a supercritical phase transition is known to take place in noisy information channels. The noisy channel is controlled by a temperature-like parameter that determines the balance between the information rate and the distortion in the same way that physical temperature controls the balance between energy and entropy in a physical system. Following Tlusty's equation (2), the free energy functional has the form $D - TS$ where D is the average error load', equivalent to average distortion in a rate distortion problem, S is the 'entropy due to random drift', and T measures the strength of random drift relative to the selection force that pushes towards fitness maximization. This is essentially a Morse function (Pettini 2007; Matsumoto 2002). According to Tlusty's analysis, at high T the channel is totally random and it conveys zero information. At a certain critical temperature T_c the information rate starts to increase continuously.

The average distortion D measures the average difference between the genetic 'message' sent by a complicated codon 'statement' and what is actually expressed by the genetic (and epigenetic) translation machinery in terms of an amino acid sequence. See figure 1.

Here we envision a multi-step process in which the rate distortion function $R(D)$ between codon sequence and amino acid sequence plays the central role. In the first step, $R(D)$, a nominally extensive quantity, but one physically limited by the channel construction of figure 1, serves as a temperature-analog in a one-parameter distribution of information source uncertainties representing different coding strategies, from which a free energy functional is constructed. While $R(D)$ is not 'temperature-like' – e.g., under a given circumstance it can be increased as much as one likes by establishing parallel channels – the physical structure of translation constrains that approach, ensuring the 'locally intensive' nature of the rate distortion function. Pettini's (2007) 'topological hypothesis'

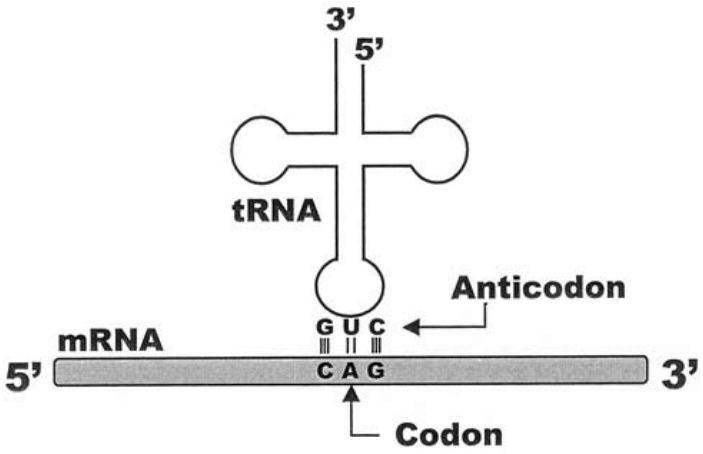

Figure 1: Adapted from fig. 1.8 of Shmulevich and Dougherty (2007). DNA meets RNA in modern protein synthesis. The anticodon at one end of a tRNA molecule binds to its complementary codon in mRNA derived directly from the genome. The average distortion D is a measure of the difference between what is supposed to be coded by a genome sequence and what is actually expressed as an amino acid sequence. Sequence-to-sequence translation is not highly parallel, in this model, and the process can be well characterized by the rate distortion function $R(D)$ representing the minimum channel capacity needed to produce average distortion less than D.

implies that topological shifts in code structure accompany phase transitions in a free energy functional constructed from the distribution of information source uncertainties arising from possible code topologies.

The second stage of the argument revolves around the relation between intensive indices of metabolic free energy availability – e.g., underlying energy per molecular transaction, and/or efficiency of its use – and $R(D)$, leading to a second free energy-like functional that undergoes another set of punctuated phase changes.

While the genetic code has received much attention, Hecht et al. (2004) note that protein α-helices have the 'code' 101100100110... where 1 indicates a polar and 0 a non-polar amino acid. Protein β-sheets have the simpler coding 10101010... Wallace (2010), in fact, extends Tlusty's topological analysis via Heawoods's graph genus formula to the more complicated protein folding classifications of Chou and Maggiora (1998). Wallace (2012b) argues, in addition, that a similar argument must apply to the basic monosaccharides associated with glycan 'kelp frond' production at the surface of the cell. Again, here we shall be interested in calculating metabolic costs necessarily associated with limiting error across such biological codes, and will model both punctuated code evolution and a form of instability triggered by metabolic energy restriction, or by the growth of energy demand beyond available resources.

2 Rate Distortion theory

The existence of a code implies the existence of an information source using that code, and the behavior of such sources is constrained by the asymptotic limit theorems of information theory. That is, the interaction between biological subsystems associated with a code can be formally restated in communication theory terms. Essentially, observation of a code directly implies existence of an information source using it.

Think of the machinery listing a sequence of codons as communicating with machinery that produces amino acids, and compare what is actually produced with what should have been produced, perhaps by a simple survival of the fittest selection mechanism, perhaps via some more sophisticated error-correcting systems.

Suppose a sequence of signals is generated by a biological information source Y having output $y^n = y_1, y_2, ...$ – codons. This is 'digitized' in terms of the observed behavior of the system with which it communicates, for example a sequence of 'observed behaviors' $b^n = b_1, b_2, ...$ – amino acids. Assume each b^n is then deterministically retranslated back into a reproduction of the original biological signal, $b^n \rightarrow \hat{y}^n = \hat{y}_1, \hat{y}_2, ...$

Define a distortion measure $d(y, \hat{y})$ which compares the original to the retranslated path. Many distortion measures are possible. The Hamming distortion is defined simply as

$$d(y, \hat{y}) = 1, y \neq \hat{y}$$

$$d(y, \hat{y}) = 0, y = \hat{y}.$$

For continuous variates the squared error distortion is just $d(y, \hat{y}) = (y - \hat{y})^2$.

There are many possible distortion measures. The distortion between paths y^n and $\hat{y}^n$ is defined as $d(y^n, \hat{y}^n) \equiv \frac{1}{n} \sum_{j=1}^{n} d(y_j, \hat{y}_j)$.

A remarkable characteristic of the Rate Distortion Theorem is that the basic result is independent of the exact distortion measure chosen (Cover and Thomas 2006).

Suppose that with each path y^n and b^n-path retranslation into the y-language, denoted $\hat{y}^n$, there are associated individual, joint, and conditional probability distributions $p(y^n), p(\hat{y}^n), p(y^n, \hat{y}^n), p(y^n|\hat{y}^n)$.

The average distortion is defined as

$$D \equiv \sum_{y^n} p(y^n) d(y^n, \hat{y}^n) \tag{1}$$

This is essentially the 'error load' of Tlusty's (2007) equation (1).

It is possible to define the information transmitted from the Y to the $\hat{Y}$ process using the Shannon source uncertainty of the strings:

$$I(Y, \hat{Y}) \equiv H(Y) - H(Y|\hat{Y}) = H(Y) + H(\hat{Y}) - H(Y, \hat{Y}),$$

where $H(..., ...)$ is the standard joint, and $H(...|...)$ the conditional, Shannon uncertainties (Cover and Thomas 2006).

If there is no uncertainty in Y given the retranslation $\hat{Y}$, then no information is lost, and the systems are in perfect synchrony.

In general, of course, this will not be true.

The rate distortion function $R(D)$ for a source Y with a distortion measure $d(y, \hat{y})$ is defined as

$$R(D) = \min_{p(y,\hat{y}); \sum_{(y,\hat{y})} p(y)p(y|\hat{y})d(y,\hat{y}) \leq D} I(Y, \hat{Y}) \tag{2}$$

The minimization is over all conditional distributions $p(y|\hat{y})$ for which the joint distribution $p(y, \hat{y}) = p(y)p(y|\hat{y})$ satisfies the average distortion constraint (i.e., average distortion $\leq D$).

The Rate Distortion Theorem states that $R(D)$ is the minimum necessary rate of information transmission which ensures the communication between the biological vesicles does not exceed average distortion D. Thus $R(D)$ defines a minimum necessary channel capacity. Cover and Thomas (2006) or Dembo and Zeitouni (1998) provide details. The rate distortion function has been calculated for a number of systems.

Cover and Thomas (2006, Lemma 13.4.1) show that $R(D)$ is necessarily a decreasing convex function of D for any reasonable definition of distortion.

That is, $R(D)$ is always a reverse J-shaped curve. This will prove crucial for the overall argument. Indeed, convexity is an exceedingly powerful mathematical condition, and permits deep inference (e.g., Rockefellar 1970). Ellis (1985, Ch. VI) applies convexity theory to conventional statistical mechanics.

For a Gaussian channel having noise with zero mean and variance σ^2,

$$R(D) = 1/2 \log[\sigma^2/D], 0 \leq D \leq \sigma^2$$
$$R(D) = 0, D > \sigma^2 \tag{3}$$

For the 'natural' channel that seems to describe compression of real images (ref.)

$$R(D) = \sigma^2/D^\alpha \tag{4}$$

with $\alpha \approx 1$.

Recall the relation between information source uncertainty and channel capacity (Cover and Thomas 2006):

$$H[X] \leq C \tag{5}$$

where H is the uncertainty of the source X and C the channel capacity. Remember also that

$$C \equiv \max_{P(X)} I(X|Y) \tag{6}$$

where $P(X)$ is chosen so as to maximize the rate of information transmission along a channel Y.

Note that for a parallel set of noninteracting channels, the overall channel capacity is the sum of the individual capacities, providing a 'consensus average' that does not apply in the case of modern molecular coding.

Finally, recall the analogous definition of the rate distortion function above, again an extremum over a probability distribution.

Recall also the homology between information source uncertainty and free energy density. More formally, if $N(n)$ is the number of high probability 'meaningful' – that is, grammatical and syntactical – sequences of length n emitted by an information source X, then, according to the Shannon-McMillan Theorem, the zero-error limit of the Rate Distortion Theorem,

$$\begin{aligned} H[X] &= \lim_{n\to\infty} \frac{\log[N(n)]}{n} \\ &= \lim_{n\to\infty} H(X_n|X_0,...,X_{n-1}) \\ &= \lim_{n\to\infty} \frac{H(X_0,...,X_n)}{n+1} \end{aligned} \tag{7}$$

where$H(...|...)$ is the conditional and $H(...,...)$ is the joint Shannon uncertainty.

In the limit of large n, $H[X]$ becomes homologous to the free energy density of a physical system at the thermodynamic limit of infinite volume. More explicitly, the free energy density of a physical system having volume V and partition function $Z(\beta)$ derived from the system's Hamiltonian – the energy function – at inverse temperature β is (e.g., Pettini, 2007)

$$F[K] = \lim_{V\to\infty} -\frac{1}{\beta}\frac{\log[Z(\beta,V)]}{V} \equiv \lim_{V\to\infty} \frac{\log[\hat{Z}(\beta,V)]}{V}$$

with $\hat{Z} = Z^{-1/\beta}$. The latter expression is formally similar to the first part of equation (7), a matter having deep implications: Feynman (2000) describes in great detail how information and free energy have an inherent duality. Feynman, in fact, defines information precisely as the free energy needed to erase a message. The argument is surprisingly direct, and for very simple systems it is easy to design a small (idealized) machine that turns the information within a message directly into usable work – free energy. Information is a form of free energy and the construction and transmission of information within living things consumes metabolic free energy, with nearly inevitable losses via the second law of thermodynamics. If there are limits on available metabolic free energy there will necessarily be limits on the ability of living things to process information.

3 Groupoid symmetry shifting

Here we follow, in part, the argument of Wallace (2012a). The direct model finds codons generated by a black box information source whose source uncertainty is constrained by the richness of the coding scheme of Tlusty's analysis. More complex codes will be associated with higher information source uncertainties, i.e., the ability to 'say' more in less time, using a more complicated coding scheme. Suppose there are n possible coding schemes. The simplest approach is to assume that, for a given rate distortion function and distortion measure, $R(D)$, under the constraints of figure 1, serves much as an external temperature bath for the possible distribution of information sources, the set $\{H_1,...,H_n\}$. That is, low distortion, represented by a high rate of transmission of information between codon machine and amino acid machine, permits more complicated coding schemes according to a classic Gibbs relation

$$Pr[H_j] = \frac{\exp[-H_j/\lambda R(D)]}{\sum_{i=1}^{n}\exp[-H_i/\lambda R(D)]} \tag{8}$$

where $Pr[H_j]$ is the probability of coding scheme j having information source uncertainty H_j.

We assume that $Pr[H_j]$ is a one parameter distribution in the 'extensive' quantity $R(D)$ (monotonic convex, however, in D) rather than a simple 'intensive' temperature-analog. This is permitted under the 'structurally intensive' circumstance of figure 1.

The free energy-like Morse Function F_R associated with this probability is defined as

$$\exp[-F_R/\lambda R(D)] = \sum_{i=1}^{n}\exp[-H_i/\lambda R(D)] \tag{9}$$

Applying Landau's spontaneous symmetry lifting argument to F_R (Pettini 2007) generates topological transitions in codon graph structure as the 'temperature' $R(D)$ increases, i.e., as the average distortion D declines, via the inherent convexity of the Rate Distortion Function. That is, as the channel capacity connecting codon machines with amino acid machines increases, more complex coding schemes become possible:

1. The genus of the embedding surface for a topological code can be expressed in terms of the Euler characteristic of the manifold, $\gamma = 1 - \frac{1}{2}\chi$.

2. χ can be expressed in terms of the cohomology structure of the manifold (Lee 2000, Theorem 13.38).

3. By the Poincare Duality Theorem, the homology groups of a manifold are related to the cohomology groups in the complementary dimension (Bredon 1993, p. 348).

4. The (co)homology groupoid can be taken as the disjoint union of the (co)homology groups of the embedding manifold.

One can then invert Landau's Spontaneous Symmetry Breaking arguments and apply them to the (co)homology groupoid in terms of the rising 'temperature' $R(D)$, to obtain a punctuated shift to increasingly complex genetic codes with increasing channel capacity. See Wallace (2012a) for a summary of standard material on groupoids and on Landau's phenomenological theory.

What, then, drives $R(D)$, as this, in turn, drives punctuated changes in the genetic code? Here we will significantly diverge from the arguments in Wallace (2012a), invoking a Black-Scholes formalism for 'cost' in terms of demand for metabolic free energy. Later, we will use a similar argument to examine failures in the dynamics of evolutionarily fixed codes under free energy restraints.

4 Metabolic energy costs

Suppose that metabolic free energy is available at a rate $\mathcal{H}$ determined by environmental structure and previous evolutionary trajectory, which may be prior to the emergence of efficient photosynthesis, predation, mutualism, parasitism, and the like. We iterate the treatment and consider $\mathcal{H}$ as the

temperature analog in a Landau model on the Rate Distortion Function itself. That is, let $R(D)$ be the Rate Distortion Function describing the relation between system intent and system impact. This is essentially a channel capacity, and information transmission rate between the coding machine and the structure or structures that biological code is attempting to affect.

The distortion represents the dynamics of the disjunction between the intent of a code and the actual productions of the system. Let R_t be the RDF of the channel connecting them at time t. The relation can, under conditions of both white noise and volatility, be expressed as

$$dR_t = f(t, R_t)dt + bR_t dW_t \tag{10}$$

Let $\mathcal{H}(R_t, t)$ represent the rate of incoming metabolic free energy that is needed to achieve R_t at time t, and expand using the Ito chain rule:

$$d\mathcal{H}_t = [\partial\mathcal{H}/\partial t + f(R_t, t)\partial\mathcal{H}/\partial R + \frac{1}{2}b^2R_t^2\partial^2\mathcal{H}/\partial R^2]dt + [bR_t\partial\mathcal{H}/\partial R]dW_t \tag{11}$$

Define $\mathcal{L}$ as the Legendre transform of the free energy rate $\mathcal{H}$, a kind of entropy, having the form

$$\mathcal{L} = -\mathcal{H} + R\partial\mathcal{H}/\partial R \tag{12}$$

Using the heuristic of replacing dX with ΔX in these expressions, and applying the results of equation (11), produces:

$$\Delta\mathcal{L} = (-\partial\mathcal{H}/\partial t - \frac{1}{2}b^2R^2\partial^2\mathcal{H}/\partial R^2)\Delta t \tag{13}$$

Analogous to the Black-Scholes calculation, the terms in f and dW_t cancel out, so that the effects of noise are subsumed in the Ito correction involving b. Clearly, however, this also invokes powerful regularity assumptions that may often be violated. Matters then revolve about model robustness in the face of such violation.

$\mathcal{L}$, as the Legendre transform of the free energy rate measure $\mathcal{H}$, is a kind of entropy that can be expected to rapidly reach an extremum at nonequilibrium steady state. There, $\Delta\mathcal{L}/\Delta t = \partial\mathcal{H}/\partial t = 0$, so that

$$\frac{1}{2}b^2R^2\partial^2\mathcal{H}/\partial R^2 = 0 \tag{14}$$

having the solution

$$\mathcal{H}_{NSS} = \kappa_1 R + \kappa_2 \tag{15}$$

a perhaps not unexpected result. This outcome, however, permits a Landau-analog phase transition analysis in which the metabolic free energy available from the embedding ecosystem serves to raise or lower the possible richness of a system's possible biological codes. As Wallace (2012a) argues, if $\mathcal{H}$ is relatively large then there are many possible complex codes. If, however, sufficient metabolic free energy is not available, then the system can only entertain a few simplified codings.

While the aerobic transition apparently enabled endosymbiotic processes producing eukaryotic organisms, it may also have enabled evolution of the extraordinarily rich glycan/lectin cell surface codings essential to all multicellular organisms. Wallace (2012b) argues, however, that full coding, having 5,000-7,500 'glycan determinant' amino acid analogs made up of the appropriate basic set of monosaccharides, would require a coding manifold with genus in the millions, suggesting the need for an intermediate layer of cognitive mechanism at the cell surface.

5 Code/translator stability

Van den Broeck et al. (1994, 1997), Horsthemeke and Lefever (2006), and others, have noted that analogous results relating phase transition to driving parameters in physical systems can be obtained by using the rich stability criteria of stochastic differential equations.

The motivation for this approach is the observation that a Gaussian channel with noise variance σ^2 and zero mean has a Rate Distortion Function $R(D) = 1/2\log[\sigma^2/D]$ using the squared distortion measure for the average distortion D. Defining a 'Rate Distortion entropy' as the Legendre transform

$$S_R = R(D) - DdR(D)/dD = 1/2\log[\sigma^2/D] + 1/2 \tag{16}$$

the simplest possible nonequilibrium Onsager equation (de Groot and Mazur 1984) is just

$$dD/dt = -\mu dS_R/dD = \mu/2D \tag{17}$$

where t is the time and μ is a diffusion coefficient. By inspection, $D(t) = \sqrt{\mu t}$, the classic solution to the diffusion equation. Such 'correspondence reduction' serves as a base to argue upward in both scale and complexity.

But deterministic coding does not involve diffusive drift of average distortion. Let $\mathcal{H}$ again be the rate of available metabolic free energy. Then a plausible model, in the presence of an internal system noise β^2 in addition to the environmental channel noise defined by σ^2, is the stochastic differential equation

$$dD_t = (\frac{\mu}{2D_t} - M(\mathcal{H}))dt + \frac{\beta^2}{2}D_t dW_t \tag{18}$$

where dW_t represents unstructured white noise and $M(\mathcal{H}) \geq 0$ is monotonically increasing.

This has the nonequilibrium steady state expectation

$$D_{NSS} = \frac{\mu}{2M(\mathcal{H})} \tag{19}$$

Using the Ito chain rule on equation (18) (Protter 1990; Khasminskii 2012), one can calculate the variance in the distortion as $E(D_t^2) - (E(D_t))^2$.

Letting $Y_t = D_t^2$ and applying the Ito relation,

$$dY_t = [2\sqrt{Y_t}(\frac{\mu}{2\sqrt{Y_t}} - M(\mathcal{H})) + \frac{\beta^4}{4}Y_t]dt + \beta^2 Y_t dW_t \tag{20}$$

where $(\beta^4/4)Y_t$ is the Ito correction to the time term of the SDE.

A little algebra shows that no real number solution for the expectation of $Y_t = D_t^2$ can exist unless the discriminant of the resulting quadratic equation is ≥ 0, producing a minimum necessary rate of available metabolic free energy for stability defined by

$$M(\mathcal{H}) \geq \frac{\beta^2}{2}\sqrt{\mu} \tag{21}$$

Values of $M(\mathcal{H})$ below this limit will trigger a phase transition into a less integrated – or at least behaviorally different – system in a highly punctuated manner, much as in the Landau example.

From equations (15) and (19),

$$M(\mathcal{H}) = \frac{\mu}{2\sigma^2}\exp[2(\mathcal{H}-\kappa_2)/\kappa_1] \geq \frac{\beta^2}{2}\sqrt{\mu} \tag{22}$$

Solving for $\mathcal{H}$ gives the necessary condition

$$\mathcal{H} \geq \frac{\kappa_1}{2}\log[\frac{\beta^2\sigma^2}{\sqrt{\mu}}] + \kappa_2 \tag{23}$$

for there to be a real second moment in D, under the subsidiary condition that $\mathcal{H} \geq 0$.

Given the context of this analysis, failure to provide adequate rates of metabolic free energy $\mathcal{H}$ would represent the onset of a regulatory stability catastrophe. The corollary, of course, is that environmental influences increasing β, σ, the κ_i, or reducing μ, would be expected to overwhelm internal controls, triggering similar instability.

Variations of the model are possible, for example, applying the formalism to the 'natural' channel, having the rate distortion function $R(D) = \sigma^2/D$. The calculation is direct.

Equation (23), in fact, represents a close analog to the data rate theorem (Nair et al. 2007, Theorem 1): the implication is that there is a critical rate of available metabolic free energy below which there does not exist any quantization, coding, or control scheme, able to stabilize an (inherently) unstable biological system.

Normal, or stress-induced, aging would, at some point, be expected to affect the magnitudes of the parameters on the right hand side of the expression in equation (23), while simultaneously decreasing the ability to provide metabolic free energy – decreasing $\mathcal{H}$. This would result in onset of serious dysfunctions across a range of scales and levels of organization, from genetic to protein folding to cell surface signaling.

6 Extending the model

It is possible to reinterpret the results of equation (23) from the perspective of Section 3, producing a more general picture of code failure under metabolic energy limitations. Suppose we agree that equation (15) is only a first approximation, and that we can take the Rate Distortion Function R as a monotonic increasing function of available metabolic free energy rate $\mathcal{H}$ that we begin to interpret as an effective system 'temperature'. Suppose also there are very many more possible 'modes' of code behavior, in addition to the simple stability/instability break point implied by equation (23). That is, we now expect complex 'phase transitions' in code function with either changing demand for, or ability to provide, rates of metabolic free energy to the coding/translating machine(s).

Given a large enough set of possible modes of code/translation behavior, we write another pseudoprobability like equation (8),

$$Pr[H_j] = \frac{\exp[-H_j/\omega\mathcal{H}]}{\sum_{i=1}^{n}\exp[-H_i/\omega\mathcal{H})]} \tag{24}$$

where H_j is the source uncertainty to be associated with each functional mode j.

This leads to another 'free energy' Morse Function, $\mathcal{F}$, defined now in terms of the rate of available metabolic free energy

$$\exp[-\mathcal{F}/\omega\mathcal{H}] = \sum_{i=1}^{n}\exp[-H_i/\omega\mathcal{H}] \tag{25}$$

Certain details of information phase transitions for this system can be calculated using 'biological' renormalization methods (Wallace, 2005, Section 4.2) analogous to, but much different from, those used in the determination of physical phase transition universality classes (Wilson, 1971).

Given $\mathcal{F}$ as a free energy analog, what are the transitions between functional realms? Suppose, in classic manner, it is possible to define a characteristic 'length', say l, on the system. It is then possible to define renormalization symmetries in terms of the 'clumping' transformation, so that, for clumps of size L, in an external 'field' of strength J (that can be set to 0 in the limit), one can write, in the usual manner (e.g., Wilson, 1971)

$$\begin{aligned}\mathcal{F}[Q(L), J(L)] &= f(L)\mathcal{F}[Q(1), J(1)] \\ \chi(Q(L), J(L)) &= \frac{\chi(Q(1), J(1))}{L}\end{aligned} \tag{26}$$

where χ is a characteristic correlation length and Q is an 'inverse temperature measure', i.e., $\propto 1/\omega\mathcal{H}$.

As described in Wallace (2005), very many 'biological' renormalizations, $f(L)$, are possible that lead to a number of quite different universality classes for phase transition. Indeed, a 'universality class tuning' can be used as a tool for large-scale regulation of the system. While Wilson (1971) necessarily uses $f(L) \propto L^3$ for simple physical systems, following Wallace (2005), it is possible to argue that, since $\mathcal{F}$ is so closely related to information measures, it is likely to 'top out' at different rates with increasing system size, so other forms of $f(L)$ must be explored. Indeed, standard renormalization calculations for $f(L) \propto L^\delta, m\log(L)+1$, and $\exp[m(L-1)/L]$ all carry through.

This line of argument leads to complex forms of highly punctuated phase transition in code/translator function with changes in demand for, or supply of, the metabolic free energy needed to run the machine.

7 Amyloid fibril formation

Another possible inference from the considerations of Sections 3 and 6 is that, under metabolic free energy inadequacy, grossly simplified 'de-facto' codes may sometimes begin to operate in place of the full code. The most direct example, perhaps, is the collapse of the 'protein folding code' from the relatively complicated symmetries described in Wallace (2010) to β-sheet amyloid plaques and fibrils in many protein folding disorders.

More specifically, globular proteins, following the observations of Chou and Maggiora (1998), have four major, and perhaps as many as another six minor, classifications. This suggests a Tlusty code error network that is, essentially, a large 'sphere', having one minor, and possibly as many as three more subminor attachment handles, according to Heawood's formula. These basic structures build a highly complicated 'protein world' that cannot be simply characterized.

The prebiotic 'amyloid world' of Maury (2009), in contrast, is built on an 'error code' having a single β-sheet structure, and shows, in its full extent, a remarkably simple eight-fold 'steric zipper' (Sawaya et al., 2007).

As Goldschmidt et al. (2010) put the matter,

> We found that [protein segments with high fibrillation propensity] tend to be buried or twisted into unfavorable conformations for forming beta sheets... For some proteins a delicate balance between protein folding and misfolding exists that can be tipped by changes in environment, destabilizing mutations, or even protein concentration...
>
> In addition to the self-chaperoning effects described above, proteins are also protected from fibrillation during the process of folding by molecular chaperones...
>
> Our genome-wide analysis revealed that self-complementary segments are found in almost all proteins, yet not all proteins are amyloids. The implication is that chaperoning effects have evolved to constrain self-complementary segments from interaction with each other.

Clearly, effective chaperoning requires considerable metabolic energy, and failure to provide levels adequate for both maintaining and operating such biochemical translation machinery would trigger a canonical 'code collapse'.

8 Discussion and conclusions

As has long been maintained (e.g., Maturana and Varela 1980; Atlan and Cohen 1998; Wallace 2012c, 2014a), the living state is cognitive at every scale and level of organization. Since it is possible to associate a broad class of cognitive processes with 'dual' information sources (e.g., Wallace 2005, 2007, 2012c, 2014a), many phenomena – most particularly, complex patterns of behavioral pathology – can be addressed using statistical models based on the Shannon Coding, the Shannon-McMillan Source Coding, the Rate Distortion, and the Data Rate Theorems (S/SM/RD/DR), as these provide powerful necessary conditions on all information generation and exchange, and on the role of information in system control (e.g., Wallace 2014b, c).

Strictly speaking, deterministic-but-for-error codes, although they may be partly analyzed using information theoretic methods as Tlusty does, do not fall so easily within direct characterization by S/SM/RD/DR models. Such codes do not, in fact, invoke cognition, although, as with cognitive gene expression (e.g., Wallace and Wallace 2010), they may become essential subcomponents within larger cognitive processes.

Nonetheless, the argument leading to the Data Rate Theorem analog of equation (23) – and the generalization of Section 6 – place code stability within a recognizably similar framework, with metabolic free energy serving as a 'control signal' stabilizing efficient operation of complex biochemical coding machinery. Demand beyond available metabolic energy supply then expresses itself in punctuated destabilization, degradation, or pathological simplification, of the code/translation channel, possibly affecting gene expression, protein folding, or the operation of the glycan/lectin cell interface.

Normal aging, or its acceleration driven by psychosocial or environmental stressors, must eventually interfere with routine code operation, triggering onset of many chronic diseases associated with senescence that involve failures of these mechanisms.

References

Atlan, H., I. Cohen, 1998, Immune information, self-organization, and meaning, International Immunology, 10:711-717.

Bredon, G., 1993, Topology and Geometry, Springer, New York.

Chou, K.C., G. Maggiora, 1998, Domain structural class prediction, Protein Engineering, 11:523-528.

Cover, T., Thomas, J., 2006, Elements of Information Theory, Second Edition, Wiley, New York.

de Groot, S., P. Mazur, 1984, Nonequilibrium Thermodynamics, Dover, New York.

Dembo, A., Zeitouni, O., 1998, Large Deviations and Applications, 2nd. ed. Springer, New York.

Ellis, R., 1985, Entropy, Large Deviations, and Statistical Mechanics, Springer, New York.

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

Goldschmidt, L., P. Teng, R. Riek, D. Eisenberg 2010, Identifying the amylome, proteins capable of forming amyloid-like fibrils, Proceedings of the National Academy of Sciences, 107:3487-3492.

Hecht, M., A. Das, A. Go, L. Aradely, Y. Wei, 2004, De novo proteins from designed combinatorialy libraries, Protein Science, 13:1711-1723.

Horsthemeke, W., R. Lefever, 2006, Noise-induced Transitions, Vol. 15, Theory and Applications in Physics, Chemistry, and Biology, Springer, New York.

Khasminskii, R., 2012, Stochastic Stability of Differential Equations, Second Edition, Springer, New York.

Landau, L., Lifshitz, E., 2007, Statistical Physics, Part I, Elsevier, New York.

Lee, J., 2000, Introduction to Topological Manifolds, Graduate Texts in Mathematics Series, Springer, New York.

Levitt, M., C. Chothia, 1976, Structural patterns in globular proteins, Nature, 261:552-557.

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

Maturana, H., F. Varela, 1980, Autopoiesis and Cognition, Reidel, Netherlands.

Maury, C., 2009, Self-propagating β-sheet polypeptide structures as prebiotic informational molecular entities: the amyloid world, Origins of Life and Evolution of Biospheres, 39:141-150.

Nair, G., F. Fagnani, S. Zampieri, R. Evans, 2007, Feedback control under data rate constraints: an overview. Proceedings of the IEEE, 95:108-137.

Pettini, M., 2007, Geometry and Topology in Hamiltonian Dynamics, Springer, New York.

Protter, P., 1990, Stochastic Integration and Differential Equations, Springer, New York.

Ringel, G., J. Young, 1968, Solutions of the Heawood map-coloring problem, Proceedings of the National Academy of Sciences, 60:438-445.

Rockafellar, R., 1970, Convex Analysis, Princeton University Press, Princeton.

Sawaya, M., S. Sambashivan, R. Nelson et al., 2007, Atomic structures of amyloid cross-β splines reveal varied steric zippers, Nature, 447:453-457.

Shmulevich, I., Dougherty, E., 2007, Genomic Signal Processing, Princeton University Press, Princeton.

Tlusty, T., 2007, A model for the emergence of the genetic code as a transition in a noisy information channel, Journal of Theoretical Biology, 249:331-342.

Tlusty, T., 2008, A simple model for the evolution of molecular codes driven by the interplay of accuracy, diversity and cost, Physical Biology, 5:016001; Casting polymer nets to optimize noisy molecular codes, Proceedings of the National Academy of Sciences, 105:8238-8243.

Van den Broeck, C., J. Parrondo, R. Toral, 1994, Noise-induced nonequilibrium phase transition, Physical Review Letters, 73:3395-3398.

Van den Broeck, C., J. Parrondo, J., R. Toral, R. Kawai, 1997, Nonequilibrium phase transitions induced by multiplicative noise, Physical Review E, 55:4084-4094.

Wallace, R., 2005, Consciousness: A Mathematical Treatment of the Global Neuronal Workspace Model, Springer, NY.

Wallace, R., 2007, Culture and inattentional blindness: a global workspace perspective, Journal of Theoretical Biology, 245: 378-390.

Wallace, R., 2010, A scientific open season, Physics of Life Reviews, 7:377-378.

Wallace, R., 2012a, Metabolic constraints on the evolution of genetic codes: did multiple 'preaerobic' ecosystem transitions entrain richer dialects via serial endosymbiosis? Transactions on Computational Systems Biology XIV, LNBI 7625:204-232.

Wallace, R., 2012b, Extending Tlusty's rate distortion index theorem method to the glycome: do even 'low level' biochemical phenomena require sophisticated cognitive paradigms?, BioSystems, 107:145-152.

Wallace, R., 2012c, Consciousness, crosstalk, and the mereological fallacy: an evolutionary perspective, Physics of Life Reviews, 9:426-453.

Wallace, R., 2014a, Cognition and biology: perspectives from information theory, Cognitive Processing, 15:1-12.

Wallace, R., 2014b, Embodied cognition, embodied regulation, and the Data Rate Theorem, Biologically Inspired Cognitive Architectures, doi: 10.1016/j.bica.2014.02.003.

Wallace, R., 2014c, Canonical failure modes of real-time control systems: insights from cognitive theory, International Journal of Systems Science. In press.

Wallace, R., D. Wallace, D, 2010, Gene Expression and its Discontents: The Social Production of Chronic Disease, Springer, New York.

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