Version 8

Body mass control and unresolved sociocultural stress: application of the generalized Data-Rate Theorem

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

May 19, 2014

Abstract

New results from control theory allow construction of necessary conditions statistical models of body mass regulation in the context of interaction with a complex dynamic environment. Focusing on the stress-related induction of central obesity via failure of HPA axis regulation, we explore implications for strategies of prevention and treatment. It rapidly becomes evident that individual-centered biomedical reductionism is an inadequate paradigm. Absent mitigation of HPA axis or related dysfunctions arising from social pathologies of power imbalance, economic insecurity, and so on, it is unlikely that permanent changes in visceral obesity can be maintained without constant therapeutic effort, an expensive – and likely unsustainable – public policy.

Key words: cognition; HPA axis; information theory; visceral obesity; stress

1 Introduction

Body mass control instantiates a sophisticated feedback between interior and exterior, very much a cognitive process in the sense of Atlan and Cohen (1998) and of Maturana and Varela (1980). Debate on internal 'set point' vs. externally-driven 'settling points' resonates across the literature (e.g., Muller et al., 2010; Speakman et al., 2011), and the matter has been the subject of some mathematical modeling (e.g., Tam et al., 2009).

Taking a distinctly different perspective, Bjorntorp (2001) viewed visceral obesity – the most dangerous form – in terms of unresolved 'flight-or-fight' activation of the HPA axis. Our own analyses in this direction can be found in Chapters 1 and 7 of Wallace and Wallace (2010).

Following Bjorntorp, when the input of noxious signals is prolonged, the HPA axis reactivity changes from normal and relatively transient attempts to maintain homeostasis or allostasis with temporary peaks of cortisol secretion first, to a state of sensitization, which reacts with exaggerated cortisol secretion after a given input. This occurs during the most active phase of the HPA axis – the early morning in humans.

When repeated too often and with sufficient strength of the input, the first sign of malfunction is a delayed down-winding, so that cortisol secretion stays elevated for a prolonged period of time. Subsequently, the normal diurnal pattern is disrupted, and morning values tend to be lower. This subsequently develops into a low, steady, rigid diurnal cortisol secretion with little reactivity, a 'burned out' HPA axis. In parallel, the controlling, central glucocorticoid receptors become less efficient, and down-regulated.

Further challenges are followed by atrophy of the entire system, often found after long-term, severe hypercortisolaemia as in Cushing's syndrome, melancholic depression, post traumatic stress disorder (PTSD), and the aftermath of war. Much research shows that lowered sex steroid and growth hormone secretions have the same consequence, because of the insufficient counteraction against cortisol effects, and the combination of these abnormalities powerfully directs a larger than normal fraction of total body fat to visceral deposits.

In sum, increased activity of the HPA axis triggers inhibition of both the pituitary gonadal and growth hormone axes. Stress then may synergistically cause accumulation of visceral fat, via elevated cortisol secretion and decrease of sex steroid and growth hormones. Bjorntorp (2001) concludes in particular that the deposit of central body fat, which is closely correlated with general measures of obesity, can serve as a reasonable approximation to the long-term endocrine abnormalities associated with stress and often-repeated or chronic activation of the HPA axis. That is, stress literally writes an image of itself onto the body as visceral fat accumulation, first having written an image of itself onto the HPA axis. The phenomenon can be interpreted as the transmission of a structured signal between communicating systems, in a large sense, from the embedding psychosocial structure to an individual's HPA axis, as has been mathematically modeled in Chapter 7 of Wallace and Wallace (2010).

An enormous literature supports the relation between stress and obesity (e.g., Dallman et al., 2003; Barrington et al., 2012;

*Box 47, NYSPI, 1051 Riverside Dr., NY, NY, 10032, USA. Wallace@nyspi.columbia.edu, rodrick.wallace@gmail.com

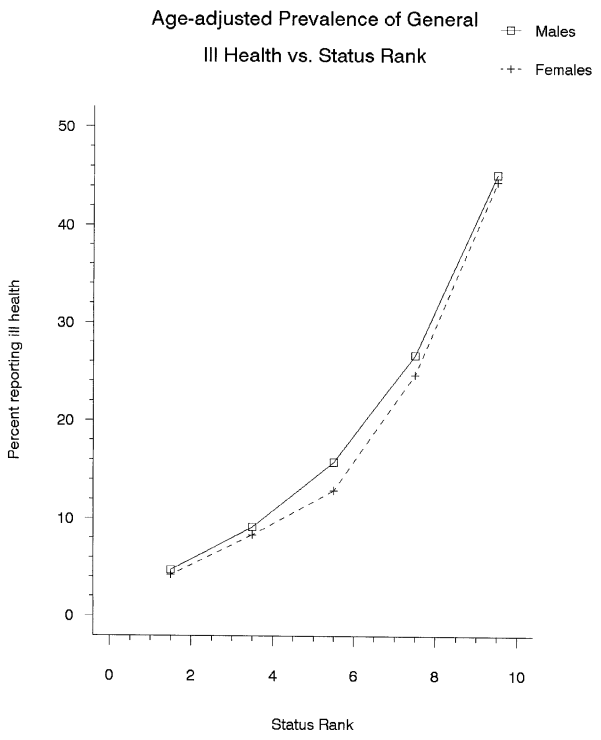

Figure 1: Redisplay of data from Singh-Manoux et al. (2003). Sex-specific dose-response curves of age-adjusted self-reported ill-health vs. self-reported status rank, for the Whitehall II cohort, 1997 and 1999. 1 is high status, 10 low status, indexing a higher level of unresolvable stress. The curve is approaching the LD-50 at which half the population suffers the impact of a poison.

Kouvonen et al., 2005; Tomiyama et al., 2012; etc.).

Singh-Manoux et al. (2003) provide something of a 'best case' example of unresolved psychosocial stress in a Western democracy. Figure 1, adapted from their paper, shows a non-linear dose-response relation between age adjusted prevalence of self-reported ill health versus self-reported status rank for government white collar workers in the UK. 1 is high rank and 10 low rank, likely to involve high levels of unresolved HPA axis activation, presumably via progressive loss of control over work. The low status group approaches the 'LD-50' level at which half the population shows response to dosage.

Animal models provide some support. For example, Tasker and Herman (2011) write that

> Stress activation of the hypothalamic-pituitary-adrenal... axis culminates in increased circulating corticosteroid concentrations. Stress-induced corticosteroids exert diverse actions in multiple target tissues over a broad range of timescales, ranging from rapid actions, which are induced within seconds to minutes and gene transcription independent, to slow actions, which are delayed, long lasting, and transcription dependent...

Such inherently cognitive phenomena, in the sense of Atlan and Cohen (1998) as described below, are necessarily constrained, Dretske (1994) has noted, by certain asymptotic limit theorems of probability:

> Communication theory can be interpreted as telling one something important about the conditions that are needed for the transmission of information as ordinarily understood, about what it takes for the transmission of semantic information. This has tempted people... to exploit [information theory] in semantic and cognitive studies...
>
> ...Unless there is a statistically reliable channel of communication between [a source and a receiver]... no signal can carry semantic information... [thus] the channel over which the [semantic] signal arrives [must satisfy] the appropriate statistical constraints of information theory.

Intersection of that theory with the formalisms of feedback system control may provide insight into HPA axis response to unresolved 'flight-or-fight' signals causing punctuated or progressive increase of body mass.

Something of this approach has become a touchstone in recent debate on the future direction of the nutrition sciences. Schubert et al. (2011) in particular call for placing disciplines like cultural sociology, human geography, social anthropology, political science, and health economics at the center of a new nutrition paradigm. In their view, social science needs to be central to the enterprise of the nutrition sciences as a necessary filter between atomistic, biomedical 'basic' science and useful policy and practice.

Here, we will provide a deep formal engagement of precisely that enterprise.

In sum, then, we extend recent work relating control theory to information theory, and apply the resulting necessary conditions statistical constraints in the construction of a class of conceptual models characterizing body mass dysfunctions in terms of their interpenetration and synergism with complex, dynamic, social and cultural environments.

We begin with a review of recent developments in control theory.

2 The Data-Rate Theorem

The data-rate theorem, a generalization of the classic Bode integral theorem for linear control systems (e.g., Yu and Mehta, 2010; Kitano, 2007; Csete and Doyle, 2002), describes the stability of linear feedback control under data rate constraints (e.g., Mitter, 2001; Tatikonda and Mitter, 2004; Sahai, 2004; Sahai and Mitter, 2006; Minero et al., 2009; Nair et al., 2007; You and Xie, 2013). Given a noise-free data link between a discrete linear plant and its controller, unstable modes can be stabilized only if the feedback data rate $\mathcal{H}$ is greater than the rate of 'topological information' generated by the unstable system. For the simplest incarnation, if the linear matrix

equation of the plant is of the form $x_{t+1} = \mathbf{A}x_t + ...$, where x_t is the n-dimensional state vector at time t, then the necessary condition for stabilizability is

$$\mathcal{H} > \log[|det\mathbf{A}^u|], \quad (1)$$

where *det* is the determinant and $\mathbf{A}^u$ is the decoupled unstable component of $\mathbf{A}$, i.e., the part having eigenvalues ≥ 1.

The essential matter is that there is a critical positive data rate below which there does not exist any quantization and control scheme able to stabilize an unstable (linear) feedback system.

This result, and its variations, are as fundamental as the Shannon Coding and Source Coding Theorems, and the Rate Distortion Theorem (Cover and Thomas, 2006; Ash, 1990; Khinchin, 1957).

Standing the approach on its head, in a sense, we will entertain and extend these considerations, using methods from cognitive theory to explore 'failure modes' of body mass control under data-rate overload.

Failure of cognition, it should be recognized, need not be 'graceful degradation under pressure'. Punctuated collapse may often be the norm, for deep reasons related to the inevitability of phase transitions in information systems, a matter long-known and much studied from the viewpoint of computational complexity (e.g., Cheeseman et al., 1991; Hogg et al., 1996; Monasson et al., 1999) that can be extended to very general forms of feedback control.

The essential analytic approach will be something much like Pettini's (2007) 'topological hypothesis' – a version of Landau's spontaneous symmetry breaking insight (Landau and Lifshitz, 2007) – which infers that such punctuated events involve a change in the topology of an underlying configuration space, and the observed singularities in the measures of interest can be interpreted as a 'shadow' of major topological change happening at a more basic level.

The preferred tool for the study of such topological changes is Morse Theory (Pettini, 2007; Matsumoto, 2002), summarized in the Mathematical Appendix, and we shall construct a relevant Morse Function using the Shannon uncertainty of information sources 'dual', in a certain sense, to cognitive processes of interest. Other such Morse Functions could be defined, for example, via the groupoid representation method of Wallace (2012b).

We begin with recapitulation of an approach to cognition using the asymptotic limit theorems of information theory (Wallace 2000, 2005, 2007, 2012). This will later be expanded using an extension of the Data Rate Theorem.

3 Cognition as an information source

Atlan and Cohen (1998) argue that the essence of cognition – biological or otherwise – involves comparison of a perceived signal with an internal, learned or inherited picture of the world, and then choice of one response from a much larger repertoire of possible responses. That is, cognitive pattern recognition-and-response proceeds by an algorithmic combination of an incoming external sensory signal with an internal ongoing activity – incorporating the internalized picture of the world – and triggering an appropriate action based on a decision that the pattern of sensory activity requires a response.

Incoming sensory input is thus mixed in an unspecified but systematic manner with internal ongoing activity to create a path of combined signals $x = (a_0, a_1, ..., a_n, ...)$. Each a_k thus represents some functional composition of the internal and the external. An application of this perspective to a standard neural network is given in Wallace (2005, p.34).

This path is fed into a highly nonlinear, but otherwise similarly unspecified, decision function, h, generating an output $h(x)$ that is an element of one of two disjoint sets B_0 and B_1 of possible system responses. Let

$$B_0 \equiv \{b_0, ..., b_k\},$$

$$B_1 \equiv \{b_{k+1}, ..., b_m\}.$$

Assume a graded response, supposing that if

$$h(x) \in B_0,$$

the pattern is not recognized, and if

$$h(x) \in B_1,$$

the pattern is recognized, and some action $b_j, k+1 \leq j \leq m$ takes place.

Interest focuses on paths x triggering pattern recognition-and-response: given a fixed initial state a_0, examine all possible subsequent paths x beginning with a_0 and leading to the event $h(x) \in B_1$. Thus $h(a_0, ..., a_j) \in B_0$ for all $0 \leq j < m$, but $h(a_0, ..., a_m) \in B_1$.

For each positive integer n, take $N(n)$ as the number of high probability paths of length n that begin with some particular a_0 and lead to the condition $h(x) \in B_1$. Call such paths 'meaningful', assuming that $N(n)$ will be considerably less than the number of all possible paths of length n leading from a_0 to the condition $h(x) \in B_1$.

Identification of the 'alphabet' of the states a_j, B_k may depend on the proper system coarsegraining in the sense of symbolic dynamics (e.g., Beck and Schlogl, 1993).

Combining algorithm, the form of the function h, and the details of grammar and syntax, are all unspecified in this model. The assumption permitting inference on necessary conditions constrained by the asymptotic limit theorems of information theory is that the finite limit

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

both exists and is independent of the path x. Again, $N(n)$ is the number of high probability paths of length n.

Call such a pattern recognition-and-response cognitive process ergodic. Not all cognitive processes are likely to be ergodic, implying that H, if it indeed exists at all, is path dependent, although extension to nearly ergodic processes, in a certain sense, seems possible (e.g., Wallace, 2005, pp. 31-32).

Invoking the Shannon-McMillan Theorem (Cover and Thomas, 2006; Khinchin, 1957), we take it possible to define an adiabatically, piecewise stationary, ergodic information source $\mathbf{X}$ associated with stochastic variates X_j having joint and conditional probabilities $P(a_0, ..., a_n)$ and $P(a_n|a_0, ..., a_{n-1})$ such that appropriate joint and conditional Shannon uncertainties satisfy the classic relations

$$H[\mathbf{X}] = \lim_{n\to\infty} \frac{\log[N(n)]}{n} = \lim_{n\to\infty} H(X_n|X_0, ..., X_{n-1}) = \lim_{n\to\infty} \frac{H(X_0, ..., X_n)}{n}. \quad (2)$$

This information source is defined as *dual* to the underlying ergodic cognitive process (Wallace, 2005).

'Adiabatic' means that, when the information source is properly parameterized, within continuous 'pieces', changes in parameter values take place slowly enough so that the information source remains as close to stationary and ergodic as needed to make the fundamental limit theorems work. 'Stationary' means that probabilities do not change in time, and 'ergodic' that cross-sectional means converge to long-time averages. Between pieces it is necessary to invoke phase change formalism, a 'biological' renormalization that generalizes Wilson's (1971) approach to physical phase transition (Wallace, 2005).

Shannon uncertainties $H(...)$ are cross-sectional law-of-large-numbers sums of the form $-\sum_k P_k \log[P_k]$, where the P_k constitute a probability distribution. See Cover and Thomas (2006), Ash (1990), or Khinchin (1957) for the standard details.

More general biological applications of this perspective can be found in Wallace (2012, 2013b) and Wallace and Wallace, (2013).

Note that we are not limited to the Atlan-Cohen model, involving an internal 'representation' of the world, learned or inherited. A very larger class of cognitive processes, including those without representation, can be mapped onto a dual information source via a category theory functor: cognition inevitably involve choice, choice reduces uncertainty, and this characterizes existence of an information source.

4 Network topology, symmetries, and dynamics

An equivalence class algebra can be constructed by choosing different origin points a_0, and defining the equivalence of two states a_m, a_n by the existence of high probability meaningful paths connecting them to the same origin point. Disjoint partition by equivalence class, analogous to orbit equivalence classes for a dynamical system, defines the vertices of a network of cognitive dual languages that interact to actually constitute the system of interest. Each vertex then represents a different information source dual to a cognitive process. This is not a representation of a network of interacting physical systems as such, in the sense of network systems biology (e.g., Arrell and Terzic, 2010). It is an abstract set of languages dual to the set of cognitive processes of interest, that may become linked into higher order structures.

Topology, in the 20th century, became an object of algebraic study, so-called algebraic topology, via the fundamental underlying symmetries of geometric spaces. Rotations, mirror transformations, simple ('affine') displacements, and the like, uniquely characterize topological spaces, and the networks inherent to cognitive phenomena having dual information sources also have complex underlying symmetries: characterization via equivalence classes defines a groupoid, an extension of the idea of a symmetry group, as summarized by Brown (1987) and Weinstein (1996). Linkages across this set of 'languages' occur via the groupoid generalization of Landau's spontaneous symmetry breaking arguments that will be used below (Landau and Lifshitz, 2007; Pettini, 2007). See the Mathematical Appendix for a brief summary of basic material on groupoids.

Given a set of cognitive modules that are linked to solve a problem, the 'no free lunch' theorem (English, 1996; Wolpert and Macready, 1995, 1997) illustrates how a 'cognitive' treatment extends a network theory-based theory (e.g., Arrell and Terzic, 2010). Wolpert and Macready show there exists no generally superior computational function optimizer. That is, there is no 'free lunch' in the sense that an optimizer pays for superior performance on some functions with inferior performance on others gains and losses balance precisely, and all optimizers have identical average performance. In sum, an optimizer that has superiority on one subset of functions must have inferiority on the complementary subset.

This result is well-known using another description. Shannon (1959) recognized a powerful duality between the properties of an information source with a distortion measure and those of a channel. This duality is enhanced if we consider channels in which there is a cost associated with the different letters. Solving this problem corresponds to finding a source that is right for the channel and the desired cost. Evaluating the rate distortion function for a source corresponds to finding a channel that is just right for the source and allowed distortion level.

Another approach is the through the 'tuning theorem' (Wallace, 2005, Sec. 2.2), which inverts the Shannon Coding Theorem by noting that, formally, one can view the channel as 'transmitted' by the signal. Then a dual channel capacity can be defined in terms of the channel probability distribution that maximizes information transmission assuming a fixed message probability distribution.

From the no free lunch argument, Shannon's insight, or the 'tuning theorem', it becomes clear that different challenges facing any cognitive system – or interacting set of them – must be met by different arrangements of cooperating low level cognitive modules. It is possible to make a very abstract picture of this phenomenon based on the network of linkages between the information sources dual to the individual 'unconscious' cognitive modules (UCM). That is, the remapped network of lower level cognitive modules is reexpressed in terms of the information sources dual to the UCM.

Given two distinct problems classes, there must be two different wirings of the information sources dual to the available UCM, with the network graph edges measured by the amount of information crosstalk between sets of nodes representing the dual information sources.

The mutual information measure of cross-talk is not inherently fixed, but can continuously vary in magnitude. This suggests a parameterized renormalization: the modular network structure linked by mutual information interactions and crosstalk has a topology depending on the degree of interaction of interest.

Define an interaction parameter ω, a real positive number, and look at geometric structures defined in terms of linkages set to zero if mutual information is less than, and 'renormalized' to unity if greater than, ω. Any given ω will define a regime of giant components of network elements linked by mutual information greater than or equal to it.

Now invert the argument: a given topology for the giant component will, in turn, define some critical value, ω_C, so that network elements interacting by mutual information less than that value will be unable to participate, i.e., will be locked out and not be perceived. See Wallace (2005, 2012) for details. Thus ω is a tunable, syntactically-dependent, detection limit that depends critically on the instantaneous topology of the giant component of linked cognitive modules defining the larger regulatory structure. That topology is the basic tunable syntactic filter across the underlying modular structure, and variation in ω is only one aspect of a set of more general topological properties that can be described in terms of index theorems, where far more general analytic constraints can become closely linked to the topological structure and dynamics of underlying networks, and, in fact, can stand in place of them (Atyah and Singer, 1963; Hazewinkel, 2002).

The central point, however, is that, in the context of this study, some topological conformations of cognitive physiological submodules, under the added influence of externally-imposed stress signals, will represent 'normal' metabolic processes, and others metabolic syndrome or worse.

In that regard, we now examine how environmental signals carry messages.

5 Environment as information source

Multifactorial cognitive systems interact with, affect, and are affected by, embedding environments that 'remember' interaction by various mechanisms. It is possible to reexpress environmental dynamics in terms of a grammar and syntax that represent the output of an information source – another generalized language.

For example, the turn-of-the seasons in a temperate climate, for many ecosystems, looks remarkably the same year after year: the ice melts, the migrating birds return, the trees bud, the grass grows, plants and animals reproduce, high summer arrives, the foliage turns, the birds leave, frost, snow, the rivers freeze, and so on. In a social setting, interacting actors can be expected to behave within fairly well defined cultural and historical constraints, depending on context: birthday party behaviors are not the same as cocktail party behaviors in a particular social set, but both will be characteristic. Gene expression during development is highly patterned by embedding environmental context via 'norms of reaction'.

Suppose it possible to coarse-grain the ecosystem at time t, in the sense of symbolic dynamics (e.g., Beck and Schlogl, 1993) according to some appropriate partition of the phase space in which each division A_j represent a particular range of numbers of each possible fundamental actor in the generalized ecosystem, along with associated larger system parameters. What is of particular interest is the set of longitudinal paths, system statements, in a sense, of the form $x(n) = A_0, A_1, ..., A_n$ defined in terms of some natural time unit of the system. Thus n corresponds to an again appropriate characteristic time unit T, so that $t = T, 2T, ..., nT$.

Again, the central interest is in serial correlations along paths.

Let $N(n)$ be the number of possible paths of length n that are consistent with the underlying grammar and syntax of the appropriately coarsegrained embedding ecosystem, in a large sense. As above, the fundamental assumptions are that – for this chosen coarse-graining – $N(n)$, the number of possible grammatical paths, is much smaller than the total number of paths possible, and that, in the limit of (relatively) large n, $H = \lim_{n\to\infty} \log[N(n)]/n$ both exists and is independent of path.

These conditions represent a parallel with parametric statistics systems for which the assumptions are not true will require specialized approaches.

Nonetheless, not all possible ecosystem coarsegrainings are likely to work, and different such divisions, even when appropriate, might well lead to different descriptive quasi-languages for the ecosystem of interest. Thus, empirical identification of relevant coarsegrainings for which this theory will work may represent a difficult scientific problem.

Given an appropriately chosen coarsegraining, define joint and conditional probabilities for different ecosystem paths, having the form $P(A_0, A_1, ..., A_n)$, $P(A_n|A_0, ..., A_{n-1})$, such that appropriate joint and conditional Shannon uncertainties can be defined on them that satisfy equation (2).

Taking the definitions of Shannon uncertainties as above, and arguing backwards from the latter two parts of equation (2), it is indeed possible to recover the first, and divide the set of all possible ecosystem temporal paths into two subsets, one very small, containing the grammatically correct, and hence highly probable paths, that we will call 'meaningful', and a much larger set of vanishingly low probability.

For humans, of course, the most influential environments are social and cultural, in the sense of Schubert et al. (2011).

6 Thermodynamics of regulation and control

Continuing the formal theory, information sources are often not independent, but are correlated, so that a joint information source can be defined having the properties

$$H(X_1, ..., X_n) \leq \sum_{j=1}^{n} H(X_j)\,, \tag{3}$$

with equality only for isolated, independent information streams.

This is the information chain rule (Cover and Thomas, 2006), and has implications for free energy consumption in regulation and control. Feynman (2000) describes how information and free energy have an inherent duality, defining information precisely as the free energy needed to erase a message. The argument is quite direct, and it is easy to design an 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 – the physical instantiation of information – itself consumes considerable free energy, with inevitable – and massive – losses via the second law of thermodynamics.

Suppose an intensity of available free energy is associated with each defined joint and individual information source $H(X,Y), H(X), H(Y)$, e.g., rates $M_{X,Y}$, M_X,M_Y.

Although information is a form of free energy, there is necessarily great entropic loss in its actual expression, so that the probability distribution of a source uncertainty H might be written in Gibbs form as

$$P[H] = \frac{\exp[-H/\kappa M]}{\int \exp[-H/\kappa M] dH}\,, \tag{4}$$

assuming κ is very small.

To first order, then,

$$\hat{H} = \int H P[H] dH \approx \kappa M\,, \tag{5}$$

and, using equation (3),

$$\begin{aligned} \hat{H}(X,Y) &\leq \hat{H}(X) + \hat{H}(Y) \\ M_{X,Y} &\leq M_X + M_Y\,. \end{aligned} \tag{6}$$

Thus, as a consequence of the information chain rule, allowing crosstalk consumes a lower rate of free energy than isolating information sources. That is, in general, it takes more free energy – higher total cost – to isolate a set of cognitive phenomena and an embedding environment than it does to allow them to engage in crosstalk (Wallace, 2012).

Hence, at the free energy expense of supporting two information sources, X and Y together, it is possible to catalyze a set of joint paths defined by their joint information source. In consequence, given a cognitive module (or set of them) having an associated information source $H(...)$, an external information source Y, the embedding environment, can catalyze the joint paths associated with the joint information source $H(..., Y)$ so that a particular chosen developmental or behavioral pathway – in a large sense – has the lowest free energy.

At the expense of larger global free information expenditure, that is, maintaining two information sources with their often considerable entropic losses instead of one, the system can feed, in a sense, the generalized physiology of a Maxwell's Demon, doing work so that environmental signals can direct system cognitive response, thus *locally* reducing uncertainty at the expense of larger global entropy production.

We next examine some details of how such regulation might operate, focusing on the role of feedback information, in the sense of the Data-Rate Theorem.

7 Phase transition

A fundamental homology between the information source uncertainty dual to a cognitive process and the free energy density of a physical system arises, in part, from the formal similarity between their definitions in the asymptotic limit. Information source uncertainty can be defined as in the first part of equation (2). This is quite analogous to the free energy density of a physical system in terms of the thermodynamic limit of infinite volume (e.g., Wilson, 1971; Wallace, 2005).

Recall again Feynman (2000), who provides a series of physical examples, based on Bennett's (1988) work, where this homology is, in fact, an identity, at least for very simple systems. Bennett argues, in terms of idealized irreducibly elementary computing machines, that the information contained in a message can be viewed as the work saved by not needing to recompute what has been transmitted.

We can model a cognitive system – here, body mass control – interacting with an embedding environment using a simple extension of the language-of-cognition approach above. Recall that cognitive processes can be formally associated with information sources. Again, a formal equivalence class algebra can be constructed for a complicated cognitive system by choosing different origin points in a particular abstract 'space' and defining the equivalence of two states by the existence of a high probability meaningful path connecting each of them to some defined origin point within that space.

To reiterate, disjoint partition by equivalence class is analogous to orbit equivalence relations for dynamical systems, and defines the vertices of a network of cognitive dual languages available to the system: each vertex represents a different information source dual to a cognitive process. The structure creates a large groupoid, with each orbit corresponding to a transitive groupoid whose disjoint union is the full groupoid, and each subgroupoid associated with its own dual information source. Larger groupoids will, in general, have 'richer' dual information sources than smaller.

We can now begin to examine the relation between system cognition and the feedback of information from the embedding environment, $\mathcal{H}$, in the sense of equation (1).

With each subgroupoid G_i of the (large) cognitive groupoid associated with body mass control we can associate a dual information source X_{G_i} having source uncertainty H_{G_i}.

Dynamic responses of a cognitive system can be represented by high probability paths connecting 'initial' multivariate states to 'final' configurations, across a great variety of beginning and end points. This creates a similar variety of groupoid classifications and associated dual cognitive processes in which the equivalence of two states is defined by linkages to the same beginning and end states. Thus, it becomes possible to construct a 'groupoid free energy' driven by the quality of information coming from the embedding ecosystem, represented by the information rate $\mathcal{H}$, taken as a temperature analog.

In the particular context of this work, $\mathcal{H}$ will be an index of unresolved flight-or-fight psychosocial stress.

$\mathcal{H}$ is, then, an embedding context for the underlying cognitive processes of interest, here the shifting, tunable control of body mass and the geography of fat deposition. The argument-by-abduction from physical theory is, then, that $\mathcal{H}$ constitutes a kind of thermal bath for the processes of cognition. Thus we can, in analogy with the standard approach from physics (Pettini, 2007; Landau and Lifshitz, 2007) construct a Morse Function by writing a pseudo-probability for the dual cognitive information source H_{G_i} as

$$P[H_{G_i}] = \frac{\exp[-H_{G_i}/\kappa\mathcal{H})]}{\sum_j \exp[-H_{G_j}/\kappa\mathcal{H}]}\,, \tag{7}$$

where κ is an appropriate dimensionless constant characteristic of the particular system. The sum is over all possible subgroupiods of the largest available symmetry groupoid. Again, compound sources, formed by the (tunable, shifting) union of underlying transitive groupoids, being more complex, will have higher free-energy-density equivalents than those of the base transitive groupoids.

A simple Morse Function for invocation of Pettini's topological hypothesis or Landau's spontaneous symmetry breaking is then a 'groupoid free energy' F defined by

$$\exp[-F/\kappa\mathcal{H}] \equiv \sum_j \exp[-H_{G_j}/\kappa\mathcal{H}]\,. \tag{8}$$

We can, using F, apply an analog to Landau's spontaneous symmetry breaking arguments, and Pettini's topological hypothesis, to the groupoid associated with the set of dual information sources.

Recall that Landau's and Pettini's insights regarding phase transitions in physical systems were that certain critical phenomena take place in the context of a significant alteration in symmetry, with one phase being differently symmetric than the other (Landau and Lifshitz, 2007; Pettini, 2007). If a symmetry is lost in the transition, the change is termed spontaneous symmetry breaking. The greatest possible set of symmetries in a physical system is that of the Hamiltonian describing its energy states. Usually states accessible at lower temperatures will lack the symmetries available at higher temperatures, so that the lower temperature phase is less symmetric: The randomization of higher temperatures ensures that higher symmetry/energy states will then be accessible to the system. The shift between symmetries is highly punctuated in the temperature index. Here, however, the larger index represents a different symmetry – constrained external behaviors – resulting in a qualitatively different cognitive groupoid leading to different patterns of body mass and fat deposition.

The essential point is that change in the stress signal $\mathcal{H}$, or in the ability of that signal to influence response, as indexed by κ, can lead to punctuated change in the complex cognitive processes of body mass and fat deposition regulation within the individual, at a variety of scales and levels of organization.

This permits a Landau-analog phase transition analysis in which incoming information from the embedding ecosystem – unresolvable psychosocial stress – serves to alter the body mass regulatory system's cognitive settings. If $\kappa\mathcal{H}$ is relatively large, as perceived by the system – then there are very few active cognitive responses possible: neither flight nor fight are available.

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

These results represent a significant generalization of the Data-Rate Theorem, as expressed in equation (1).

Consider the number of possible overt behavioral responses as an order parameter, a number $\mathcal{R} \geq 0$. Thus $\mathcal{R}$ would measure the external response given the 'control' signal $\mathcal{H}$. According to the Landau argument, $\mathcal{R}$ declines sharply when $\mathcal{H} \geq \mathcal{H}_C$, for some critical value. That is, when $\mathcal{H} \geq \mathcal{H}_C$, there is an large-scale topological change: above that value, a kind of pathological 'global broadcast' takes place, collapsing possible overt behaviors and entraining a characteristic set of unconscious internal cognitive submodules into dysfunctional fat accumulation and deposition.

8 Another approach

Here we use the rich vocabulary associated with the stability of stochastic differential equations to model, from another perspective, phase transitions in the composite system of individual body and environment (e.g., Horsthemeke and Lefever, 2006; Van den Broeck et al., 1994, 1997).

Define a 'symmetry entropy' based on the Morse Function F of equation (8) over a set of structural parameters $\mathbf{Q} = [Q_1, ..., Q_n]$ (that may include $\mathcal{H}$) as the Legendre transform

$$S = F(\mathbf{Q}) - \sum_i Q_i \partial F(\mathbf{Q})/\partial Q_i\,. \tag{9}$$

The dynamics of such a system will be driven, at least in first approximation, by Onsager-like nonequilibrium thermodynamics relations having the standard form (de Groot and Mazur, 1984):

$$dQ_i/dt = \sum_j \mathcal{K}_{i,j} \partial S/\partial Q_j\,, \tag{10}$$

where the $\mathcal{K}_{i,j}$ are appropriate empirical parameters and t is the time. A biological system involving the transmission of information may, or may not, have local time reversibility: in English, for example, the string ' eht ' has a much lower probability than ' the '. Without microreversibility, $\mathcal{K}_{i,j} \neq \mathcal{K}_{j,i}$.

Since, however, biological systems are quintessentially noisy, a more fitting approach is through a set of stochastic differential equations having the form

$$dQ_t^i = \mathcal{K}_i(t, \mathbf{Q})dt + \sum_j \sigma_{i,j}(t, \mathbf{Q})dB^j \,, \tag{11}$$

where the $\mathcal{K}_i$ and $\sigma_{i,j}$ are appropriate functions, and different kinds of 'noise' dB^j will have particular kinds of quadratic variation affecting dynamics (Protter, 1990).

Several dynamics become immediately evident:

1. Setting the expectation of equations (11) equal to zero and solving for stationary points ultimately must give attractor states since the noise terms preclude unstable equilibria. This result, however, requires some further development.

2. This system may converge to limit cycle or pseudo-random 'strange attractor' behaviors similar to thrashing in which the system seems to chase its tail endlessly within a limited venue – a kind of 'Red Queen' pathology.

3. What is converged to in both cases is not a simple state or limit cycle of states. Rather it is an equivalence class, or set of them, of highly dynamic cognitive modes coupled by mutual interaction through crosstalk and other interactions. Thus 'stability' in this structure represents particular patterns of ongoing dynamics rather than some identifiable static configuration or 'answer'. These are 'nonequilibrium steady states', to use the physicists' terminology.

4. Applying Ito's chain rule for stochastic differential equations to the $(Q_t^j)^2$ and taking expectations allows calculation of variances. These may depend very powerfully on a system's defining structural constants, leading to significant instabilities depending on the magnitudes of the Q_i, as in the Data Rate Theorem (Khasminskii, 2012). That is, the stability of states found by setting the expectation of equation (11) to zero may be strongly parameterized.

5. Following the arguments of Champagnat et al. (2006), this is very much a coevolutionary composite structure, where fundamental dynamics are determined by the feedback between internal and internal and between internal and external.

In particular, setting the expectation of equation (11) to zero generates an index theorem (Hazewinkel, 2002) in the sense of Atiah and Singer (1963) that relates analytic results, the solutions of the equations, to underlying topological structure, the eigenmodes of a complicated geometric operator whose groupoid spectrum represents symmetries of the possible changes that must take place for a tunable workspace of cooperating cognitive modules to become assembled and activated. The outcome of such activation need not be beneficial to the organism on any given timescale.

9 A simple example

Consider a body mass measure, R, assuming some R_1 is an 'optimum' level. Suppose, once triggered, the reverberation of cognitive attention to an incoming unresolved stress signal signal is self-dynamic but that the response rate is determined by the magnitude of the signal $\kappa\mathcal{H}$, and affected by noise, so that, as a first approximation,

$$dR_t = -\kappa\mathcal{H}|R_t - R_1||R_t - R_2|dt + \beta R_t dW_t \,, \tag{12}$$

where dW_t represents white noise, and all constants are positive. At equilibrium, the expectation of equation (12) – the mean body mass – is either the 'normal' value R_1 or the canonical excitation level $R_2 > R_1$. Note that, for the purposes of calculation, one may replace the absolute value by the positive square root of the square, an expression more friendly to some computer algebra programs.

But Wilson (1971) invokes fluctuation at all scales as the essential characteristic of physical phase transition, with invariance under renormalization defining universality classes. Criticality in biological or other cognitive systems is not likely to be as easily classified, e.g., Wallace (2005, Section 4.2), but certainly failure to have a second moment seems a good analog to Wilson's instability criterion. As discussed above, analogous results relating phase transitions to noise in stochastic differential equation models are widely described in the physics literature.

To calculate the second moment in R, now invoke the Ito chain rule, letting $Y_t = R_t^2$. Then

$$dY_t = (-2\kappa\mathcal{H}|R_t - R_1||R_t - R_2|R_t + \beta^2 R_t^2)dt + 2\beta R_t^2 dW_t \,, \tag{13}$$

where $\beta^2 R_t^2$ in the dt term is the Ito correction due to noise. Again taking the expectation at equilibrium, no second moment can exist unless the expectation of R_t^2 is greater than or equal to zero, giving the condition

$$\kappa\mathcal{H} > \frac{\beta^2}{2(R_2 - R_1)} \,. \tag{14}$$

Suppose, now, that the new 'set point' R_2 is itself determined by the magnitude of $\kappa\mathcal{H}$ as

$$(R_2 - R_1) = \alpha\kappa\mathcal{H} \,. \tag{15}$$

Then the new condition becomes

$$\kappa\mathcal{H} > \frac{\beta}{\sqrt{2\alpha}} \,. \tag{16}$$

Thus, in consonance with the direct phase transition arguments in $\mathcal{H}$, there is a minimum stress signal level necessary to support a self-dynamic shift to higher body mass, in this model. For a given level of 'noise', the larger α, the smaller the needed unresolved stress signal strength to trigger punctuated body mass increase.

More complex algebraic relations between $\kappa\mathcal{H}$ and R than equation (15) are, of course, quite likely. Indeed, hints of such a mechanism are in the literature, as indicated in figure 1.

10 Therapeutic intervention

As Champagnat et al. (2006) describe, shifts between the quasi-equilibria of a coevolutionary system like that of equation (11) can be addressed by the large deviations formalism. The dynamics of drift 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.

'Rare events' are the diffusion paths drifting far away from the direct solutions of the canonical equation. The probability of such rare events is governed by a large deviation principle, driven by a 'rate function' $\mathcal{I}$ that can be expressed in terms of the parameters of the diffusion.

This result can be used to study long-time behavior of the diffusion process when there are multiple attractive singularities, here, multiple set/settling points. 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.

An essential fact of large deviations theory is that the rate function $\mathcal{I}$ almost always has the canonical form

$$\mathcal{I} = -\sum_j P_j \log(P_j) \tag{17}$$

for some probability distribution in the P_k (Dembo and Zeitouni, 1998).

The argument relates to equation (11), now seen as subject to large deviations that can themselves be described as the output of an information source L_D having source uncertainty $\mathcal{I}$, driving Q^j-parameters that can trigger punctuated shifts between quasi-stable topological modes of the system of interacting cognitive submodules determining body mass.

That is, we can now write a joint source uncertainty for bodymass regulation in the context of therapeutic intervention as

$$H(X_{G_i}, L_D)\,, \tag{18}$$

and carry through the arguments leading to equations (7), (8), and (11).

It should be clear that both other internal and feedback signals, beyond $\mathcal{H}$, and independent, externally-imposed perturbations, can cause such transitions in a highly punctuated manner. Some of these may, in such a coevolutionary system, be highly pathological over a developmental trajectory, necessitating therapeutic counterinterventions – an imposed $\mathcal{I}$ – over a subsequent trajectory.

Similar ideas are now common across much of systems biology (e.g., Kitano 2004).

11 Discussion and conclusions

Two factors determine possible dynamics, in the simplest version of the model: the magnitude of of the environmental feedback signal – the unresolved HPA axis stress $\kappa\mathcal{H}$, and the inherent structure of the groupoid defining the cognitive free energy analog of the body mass regulatory system, F. Higher levels of stress will impose markedly different possible behaviors and phenotypes, triggering body mass pathologies. This has profound implications for understanding disorders of body mass as gestalt processes, involving not just an atomized individual, but the individual-in-context. Such context, for humans, inevitably includes both cultural and social expectations, interactions, and constraints, as in figure 1, and as emphasized by Schubert et al. (2011).

Equations (11-18) expand the argument to model in more detail the effects of pathological developmental perturbations and therapeutic interventions designed to counter them, finding that body mass can be driven by unresolved sociocultural stressors, according to various possible dose-response relations.

Nunney (1999) examines cancer occurrence in terms of increasingly elaborate tissue-specific control mechanisms that must have evolved along with increase in the size of organisms. Wallace (2005b) takes Nunney's viewpoint regarding mechanisms that must have evolved to stabilize animal consciousness, and proposes a 'cancer model' for mental disorders in which failure of stabilization lies at the heart of the dysfunction phenotypes. Following Bjorntorp (2001), we argue here that collapse of 'flight-or-fight' mechanisms under imposed psychosocial stress can lie at the heart of regulatory failure in body mass dysfunction, and model the outcome using new formal developments relating control theory to information theory.

We have, in some measure, extended the criticisms that Bennett and Hacker (2003) made of contemporary neuroscience. They explored the 'mereological fallacy' of a decontextualization that attributes to 'the brain' what is really the province of the whole individual. Here, we argue that, in terms of body mass regulation and the geography of fat deposition, for humans, the 'whole individual' involves essential interactions with embedding social and cultural milieu, power relations between groups and individuals, and the lasting influence of path-dependent historical trajectory determining those relations. Understanding, prevention, and treatment of body mass disorders cannot be disentangled from the role of feedback from the embedding sociocultural and other environments that necessarily incorporate power relations, as Schubert et al. (2011) argue. Currently popular views on these matters often restrict consideration of the social environment to 'fast food', over-large portions, too much fructose, seductive advertising, lack of exercise, and so on (e.g., Hill et al. 2003). People do not self-medicate with food, alcohol, tobacco, prescription and illegal drugs, television, and the like, simply because these may be readily available. People do so in response to persistent unresolved stress at various scales and levels of organization.

For the study of body mass dysfunctions, individual-centered biomedical reductionism is an inadequate paradigm. Indeed, absent mitigation of unresolved socially-induced HPA axis or related stresses arising from pathologies of power, socioeconomic deprivation, and so on, it seems unlikely that permanent changes in visceral obesity can be maintained without constant therapeutic effort – a persistent perturbing informa-

tion source L_D. This would constitute an exceedingly expensive public policy, unlikely to be sustainable in the long term.

12 Mathematical appendix

12.1 Morse Theory

Morse Theory explores 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 information source uncertainty on a parameter space and possible iterations involving parameter manifolds determining critical behavior. An example might be the sudden onset of a giant component. These can be reformulated from a Morse Theory perspective (Pettini, 2007).

The basic 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

$$\mathcal{H}_{i,j} = \partial^2 f / \partial x^i \partial x^j,$$

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 $\mathcal{H}$ 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 $\mu_i (i = 0, ..., m)$ 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,

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

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.

Pettini (2007) and Matsumoto (2002) contain details and further references.

12.2 Groupoids

A groupoid, G, is defined by a base set A upon which some mapping – a morphism – can be defined. Note that not all possible pairs of states (a_j, a_k) in the base set A can be connected by such a morphism. 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 $\lambda_g g = g = g \rho_g$.

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 $\beta(g) = a_k$. 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. For example, 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.

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.

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., in our context, 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.

13 Acknowledgments

The author thanks J.F. Glazebrook, B. Sobral, and D.N. Wallace for useful discussions.

References

Arrell, D., A. Terzic, 2010, Network systems biology for drug discovery, Clinical Pharmacology and Therapeutics, 88:120-125.

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

Atiyah, M., I. Singer, 1963, The index of elliptical operators on compact manifolds, Bulletin of the American Mathematical Society, 69:322-433.

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

Baillieu, J., 2001, Feedback disigns in information based control. In Stochastic Theory and Control: Proceedings of a Workshop Held in Lawrence, Kansas, B. Pasik-Duncan (ed.), Springer, New York, pp.35057.

Barrington, W., R. Ceballos, S. Bishop, B. McGregor, S. Beresford, 2012, Perceived stress, behavior, and body mass index among adults participating in a worksite obesity prevention program, Seattle, 2005-2007, Preventing Chronic Disease, 9:120001.

Beck, C., F. Schlogl, 1995, Thermodynamics of Chaotic Systems, Cambridge University Press, 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, New York.

Bennett, M., P. Hacker, 2003, Philosiphical Foundations of Neuroscience, Blackwell Publishing, London.

Bjorntorp, P., 2001, Do stress reactions cause abdominal obesity and comorbidities? Obesity Reviews, 2:73-86.

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

Champagnat, N., R. Ferrier, S. Meleard, 2006, Univying evolutionary dynamics: from individual stochastic process to macroscopic models, Theoretical Population Biology, 69:297-321.

Chandra, F., G. Buzi, J. Doyle, 2011, Glycolytic oscillations and limits on robust efficiency, Science, 333:187-192.

Cheeseman, P., R. Kanefsky, W. Taylor, 1991, Where the really hard problems are, Proceedings of the 13th International Joint Conference on Artificial Intelligence, J. Mylopolous, R. Reiter (eds.), Morgan Kaufmann, San Mateo, CA, pp. 331-337.

Cover, T., J. Thomas, 2006, Elements of Information Theory, 2nd Edition, Wiley, New York.

Csete, M., J. Doyle, 2002, Reverse engineering of biological complexity, Science, 295:1664-1669.

Dallman, M., N. Pecoraro, S. Akana, S. La Fleur, F. Gomez, H. Houshyar, M. Bell, S. Bhatnagar, D. Laugero, S. Manalo, 2003, Chronic stress and obesity: a new view of 'comfort food', Proceedings of the National Academy of Sciences 100:11696-11701.

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

Dretske, F., 1994, The explanatory role of information, Philosophical Transactions of the Royal Society A, 349:59-70.

English, T., 1996, Evaluation of evolutionary and genetic optimizers: no free lunch. In Evolutionary Programming V: Proceedings of the Fifth Annual Conference on Evolutionary Programming, Fogel, L., P. Angeline, T. Back (eds.), pp. 163-169, MIT Press, Cambridge, MA.

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

Hazewinkel, M., 2002, Encyclopedia of Mathematics, 'Index Formulas', Springer, New York.

Hill, J., H. Wyatt, G. Reed, J. Peters, 2003, Obesity and the environment: where do we go from here? Science, 266:853-858.

Hogg, T., B. Huberman, C. Williams, 1996, Phase transitions and the search problem, Artificial Intelligence, 81:1-15.

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, Springer, New York.

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

Kitano, H., 2004, Biological robustness, Nature Genetics, 5:826-837.

Kouvonen, A., M. Kivimaki, S. Cox, T. Cox, J. Vahtera, 2005, Relationship between work stress and body mass index among 45,810 female and male employees, Psychosomatic Medicine, 67:577-583.

Landau, L., E. Lifshitz, 2007, Statistical Physics, 3rd Edition, Part I, Elsevier, New York.

Matsumoto, Y., 2002, An Introduction to Morse Theory, American Mathematical Society, Providence, RI.

Maturana, H., F. Varela, 1980, Autopoiesis and Cognition, Reidel Publishing Company, Dortrecht.

Minero, P., M. Franceschetti, S. Dey, G. Nair, 2009, Data Rate Theorem for stabilization over time-varying feedback channels, IEEE Transactions on Automatic Control, 54:243-255.

Mitter, S., 2001, Control with limited information, European Journal of Control, 7:122-131.

Monasson, R., R. Zecchina, S. Kirkpatrick, B. Selman, L. Troyansky, 1999, Determining computational complexity from characteristic 'phase transitions', Nature, 400:133-137.

Muller, M., A. Bosy-Westphal, S. Heymsfield, 2010, Is there evidence for a set point that regulates human body weight? F1000 Medicine Reports, 2:59.

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

Nunney, L., 1999, Lineage selection and the evolution of multistage carcinogenesis, Proceedings of the Royal Society B, 266:493-498.

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

Sahai, A., 2004, The necessity and sufficiency of anytime capacity for control over a noisy communication link, Decision and Control, 43rd IEEE Conference on CDC, Vol. 2, 1896-1901.

Sahai, A., S. Mitter, 2006, The necessity and sufficiency of anytime capacity for control over a noisy communication link Part II: vector systems, http://arxiv.org/abs/cs/0610146.

Schubert, L., D. Gallegos, W. Foley, C. Harrison, 2011, Reimagining the 'social' in the nutrition sciences, Public Health Nutrition, 15:352-359.

Seaton, T., J. Miller, T. Clarke, 2013, Semantic bias in program coevolution, K. Krawiec et al. (eds.), EuroGP 2013, LNCS 7831:193-204, Springer-Verlag, Berlin.

Shannon, C., 1959, Coding theorems for a discrete source with a fidelity criterion, Institute of Radio Engineers International Convention Record Vol. 7, 142-163.

Singh-Manoux, A., N. Adler, M. Marmot, 2003, Subjective social status: its determinants and its association with measures of ill-health in the Whitehall II study, Social Science and Medicine, 56:1321-1333.

Speakman, J., D. Levitsky, D. Allison, etc., 2011, Set points, settling points and some alternative models: theoretical options to understand how genes and environments combine to regulate body adiposity, Disease Models and Mechanisms, 4:733-745.

Tam, J., D. Fukumura, R. Jain, 2009, A mathematical model of murine metabolic regulation by leptin: energy balance and defense of a stable body weight, Cell Metabolism, 9:52-63.

Tasker, J., J. Herman, 2011, Mechanisms of rapid glutocorticoid feedback inhibition of the hypothalamic-pituitary-adrenal axis, Stress, 14:398-406.

Tatikonda, S., S. Mitter, 2004, Control over noisy channels, IEEE Transactions on Automatic Control, 49:1196-1201.

Tomiyama, A., E. Puterman, E. Epel, D. Rehkopf, B Laraia, 2012, Chronic psychological stress and racial disparities in body mass index change between Black and White girls aged 10-19, Annals of Behavioral Medicine, DOI 10.1007/s12160-012-9398-x.

Touchette, H., S. Lloyd, 2004, Information-theoretic approach to the study of control systems, Physca A, 331:140-172.

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, R. Toral, R. Kawai, 1997, Nonequilibrium phase transitions induced by multiplicative noise, Physical Review E, 55:4084-4094.

Wallace, R., 2000, Language and coherent neural amplification in hierarchical systems: renormalization and the and the dual information source of a generalized spatiotemporal stochastic resonance, International Journal of Bifurcation and Chaos, 10:493-502.

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

Wallace, R., 2005b, A global workspace perspective on mental disorders, Theoretical Biology and Medical Modelling, 2/1/49.

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

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

Wallace, R., 2012b, Spontaneous symmetry breaking in a nonrigid molecule approach to intrinsically disordered proteins, Molecular BioSystems, 8:374-377.

Wallace, R., 2013, Canonical failure modes of real-time control systems: cognitive theory generalizes the data-rate theorem. Submitted.

Wallace, R., 2013b Cognition and biology: perspectives from information theory, Cognitive Processing, DOI 10.1007/s10339-013-0573-1.

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

Wallace, R., D. Wallace, 2013, A Mathematical Approach to Multilevel, Multiscale Health Interventions: Pharmaceutical Industry Decline and Policy Response, Imperial College Press, London.

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.

Wolpert, D., W. MacReady, 1995, No free lunch theorems for search, Santa Fe Institute, SFI-TR-02-010.

Wolpert, D., W. MacReady, 1997, No free lunch theorems for optimization, IEEE Transactions on Evolutionary Computation, 1:67-82.

Wong, W., R. Brockett, 1999, Systems with finite communication bandwidth constraints II: stabilization with limited information feedback, IEEE Transactions on Automation and Control, 44:1049-1053.

You, K., L. Xie, 2013, Survey of recent progress in networked control systems, Acta Automatica Sinica, 39:101-117.

Yu, S., P. Mehta, 2010, Bode-like fundamental performance limitations in control of nonlinear systems, IEEE Transactions on Automatic Control, 55:1390-1405.