Journal of Theoretical Biology 436 (2018) 72–78

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/jtbi

New statistical models of nonergodic cognitive systems and their pathologies

Rodrick Wallace

Division of Epidemiology, The New York State Psychiatric Institute, New York, NY 10032 United States

ARTICLE INFO

Article history:
Received 28 February 2017
Revised 20 September 2017
Accepted 25 September 2017
Available online 29 September 2017

Keywords:
Cognition
Control theory
Data rate theorem
Free energy
Groupoid
Information theory
Phase change
Symmetry breaking
Topology

ABSTRACT

Cognition most singularly involves choice that reduces uncertainty. Reduction of uncertainty implies the existence of an information source 'dual' to the cognitive process under study. However, information source uncertainty for path-dependent nonergodic systems cannot be described as a conventional Shannon 'entropy' since time averages are not ensemble averages. Nonetheless, the essential nature of information as a form of free energy allows study of nonergodic cognitive systems having complex dynamic topologies whose algebraic expression is in terms of directed homotopy groupoids rather than groups. This permits a significant extension of the Data Rate Theorem linking control and information theories via an analog to the spontaneous symmetry breaking arguments fundamental to modern physics. In addition, the identification of information as a form of free energy enables construction of dynamic empirical Onsager models in the gradient of a classic entropy that can be built from the Legendre transform of even path-dependent information source uncertainties. The methodology provides new analytic tools that should prove useful in understanding failure modes and their dynamics across a broad spectrum of cognitive phenomena, ranging from physiological processes at different scales and levels of organization to critical system automata and institutional economics.

1. Introduction

It is not commonly understood that cognitive systems – of any form – can, at least in principle, be described in terms of the 'grammar' and 'syntax' of appropriate information sources: cognition implies choice, choice reduces uncertainty, and the reduction of uncertainty implies the existence of an information source (Atlan and Cohen, 1998; Wallace, 2012; 2015a; 2015b; 2016a; 2016b; 2016c; 2017). Conventional theory focuses, however, on adiabatically piecewise stationary ergodic (APSE) sources, i.e., those that are parameterized in time but remain as close as necessary to ergodic and stationary for the theory to work. 'Stationary' implies that probabilities are not time dependent, and 'ergodic' roughly means that time averages are well represented as ensemble averages. Transitions between 'pieces' can then be described using an adaptation of standard renormalization methods (Wallace, 2005).

See Cover and Thomas (2006), Ash (1990) or Khinchin (1957) for details of information theory. The Wallace references provide details of the 'adiabatic' approximation, much like the Born-Oppenheimer approach to molecular dynamics where nuclear oscillations are taken as very slow in comparison with electron dynamics that equilibrate about the nuclear motions. Here, we examine what would be needed to extend the theory to very highly nonergodic cognitive systems that, as in the case of nonparametric statistics, may encompass more real-world examples than covered by the simpler mathematical models.

Somewhat similar matters have been the focus of increasing attention in economics (Durlauf, 1993). Economic agents are quintessentially cognitive, and the approach can be applied across many scales and levels of biological and other forms of organization. In particular it is possible to describe the dynamics of pathology in such systems using fairly direct methods. For example, Wallace (2015a) applies a 'locally ergodic' formalism to economic problems that is similar to the standard ergodic decomposition methods (Coudene, 2016; Gray, 2011 Lemma 1.5; Gray and Davisson, 1974; Gray and Saadat, 1984; Schonhuth, 2008; VonNumann, 1932) and produces multiple nonequilibrium steady states (nss). These are characterized by assignment of an APSE source to equivalence classes of developmental paths that are represented by groupoid symmetries, leading to groupoid symmetry breaking via an analog of group symmetry breaking in physical systems.

Standard extensions of classic information theory theorems to nonergodic stationary processes, and to asymptotically mean stationary processes, have been in terms of the decomposition of sources into their ergodic components, with averaging across them, a development with a long tradition. Coudene (2016, Section 14.1) puts it

E-mail address: Wallace@nyspi.columbia.edu

https://doi.org/10.1016/j.jtbi.2017.09.024
0022-5193/

> When a system is not ergodic, it is possible to decompose the underlying space into several pieces, so that the transformation is ergodic on each of these pieces. We call this a partition into ergodic components. The number of components may be uncountable, but the resulting partition still satisfies a certain regularity property: it is possible to approximate it with partitions having finitely many pieces.

As Hoyrup (2013) notes, however, while every non-ergodic measure has a unique decomposition into ergodic ones, this decomposition is not always computable. From another perspective, such expansions – in terms of the usual ergodic decomposition or the groupoid/directed homotopy equivalents – both explain everything and explain nothing, in the sense that almost any real function can be written as a Fourier series or integral that retains the essential character of the function itself. Sometimes this helps if there are basic underlying periodicities leading to a meaningful spectrum, otherwise not. The analogy is the contrast between the Ptolemaic expansion of planetary orbits in circular components around a fixed Earth vs. the Newtonian/Keplerian gravitational model in terms of ellipses with the Sun at one focus. While the Ptolemaic expansion converges to any required accuracy, it conceals the essential dynamics.

Here, we use a very general approach adapted from nonequilibrium thermodynamics to study both nonergodic systems and their ergodic components, if such exist, but in terms of inherent groupoid symmetries associated with equivalence classes of directed homotopy developmental pathways. The attack is based on the counterintuitive recognition of information as a form of free energy (Feynman, 2000), rather than an 'entropy' in the physical sense. A central constraint is that, in the extreme case which will be the starting point, only individual developmental paths can be associated with an information-theoretic source function that cannot be represented in terms of a Shannon entropy-like uncertainty value across a probability distribution.

Equivalence classes then must arise via a metric distance measure for which the developmental trajectories of one kind of 'game' are closer together than for a significantly different 'game'. Averaging occurs according to such equivalence classes, and is marked by groupoid symmetries, and by characteristic dynamics of symmetry breaking according to appropriate 'temperature' changes indexing the influence of embedding regulatory mechanisms. We will, however, recover the standard decomposition by noting that larger equivalence classes across which uncertainty measures are constant can be collapsed to single paths on an appropriate quotient manifold.

2. Some formalism

Recall that, for a stationary, ergodic information source **X**, as Khinchin (1957) indicates, it is possible to divide statements of length n – written as $x^n = \{X(0) = x_0, X(1) = x_1, \ldots, X(n) = x_n\}$ – into two sets. The first, and largest, is not consonant with the 'grammar' and 'syntax' of the information source, and consequently has vanishingly small probability in the limit of large n. The second, much smaller set that is consonant and characterized as 'meaningful', has the following essential properties.

If $N(n)$ is the number of meaningful statements of length n, then limits exist satisfying the conditions

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

$H(X_n|X_0, \ldots, X_{n-1})$ and $H(X_0, \ldots, X_n)$ are conditional and joint Shannon uncertainties having the familiar pseudo-entropy form

$$H = -\sum_i P_i \log[P_i]$$
$$0 \le P_i \le 1, \; \sum_i P_i = 1 \quad (2)$$

in the appropriate joint and conditional probabilities (Cover and Thomas, 2006). This limit is called the source uncertainty.

Nonergodic information sources cannot be directly represented in terms of Shannon uncertainties resembling entropies. For such sources, however, a function, $\mathcal{H}(x^n)$, of each path $x^n \to x$, may still be defined, such that $\lim_{n\to\infty} \mathcal{H}(x^n) = \mathcal{H}(x)$ holds Khinchin (1957, p. 72) However, $\mathcal{H}$ will not, in general, be given by the simple cross-sectional laws-of-large numbers analog having the (deceptive) entropy-like form of Eq. (2).

3. Phase transitions: A generalized data rate theorem

Cognitive information sources are characterized by equivalence classes of states and developmental paths in the topological spaces defined by those states (Wallace and Fullilove, 2008; Wallace, 2012; 2015a; 2015b; 2017). Under 'ergodic' conditions, for each of these classes a 'dual' APSE information source can be assigned. Perhaps the simplest example of such an equivalence class would be the set of high probability 'developmental' trajectories from an initial phenotype a_0 to some final phenotype a_∞. Variation in a_0 and a_∞ then produces the set of classes, defining a groupoid (Weinstein, 1996), as opposed to the group symmetries more familiar from standard algebraic topology (e.g., Lee, 2000). Consequently, products may not necessarily be defined between groupoid members (Weinstein, 1996). As discussed elsewhere (e.g., Wallace, 2015a; 2015b; 2017), phase transitions for ergodic cognitive systems are associated with necessary (but not sufficient) changes in underlying groupoid symmetries that are analogous to the spontaneous symmetry breaking of simpler physical systems (e.g., Pettini, 2007).

Consideration of these matters for fully path-dependent nonergodic information sources leads quickly to an analog of the Data Rate Theorem (DRT) that mandates a minimum rate of control information for an inherently unstable system (Nair et al., 2007). A principal tool is directed homotopy, or dihomotopy – the study of topological structure using nonreversible paths rather than complete loops (Fajstrup et al., 2016; Grandis, 2009).

Cognitive systems are embodied: there is no cognition without sensory input, following the basic model of Atlan and Cohen (1998). Sensory information is the tool by which choice-of-action is made, and such choice is the defining characteristic of cognition, reducing uncertainty and implying the existence of a dual information source. For a relatively simple but inherently unstable linear 'plant', clever application of the classic Bode integral theorem implies that the rate of control information must exceed the rate at which that system generates 'topological information', in a particular sense (Nair et al., 2007). Ergodic cognitive processes may be expected to show more complex patterns of behavior, and we will extend the argument to nonergodic cognition.

Again, the central focus is on paths $x_n \to x$ that are consonant with the 'grammar' and 'syntax' of the information source dual to the cognitive process. For these, a fully path-dependent information source function $\mathcal{H}(x)$ can be defined, i.e., it's value, in general, changes from path to path. For an ergodic source, there is only one value possible across an equivalence class of developmental pathways, and it is given by the usual Shannon uncertainty across a probability distribution.