Version 9
Global broadcast and attention can be highly punctuated in fuzzy cognitive systems

Rodrick Wallace
Division of Epidemiology
The New York State Psychiatric Institute
Box 47, 1051 Riverside Dr., New York, NY, 10032
Wallace@nyspi.columbia.edu

Previous theoretical work on consciousness and other punctuated global broadcasts associated with attention states has focused on the evolutionary exaptation of the inevitable signal crosstalk between related sets of unconscious cognitive modules (UCM). This has invoked a groupoid treatment of the equivalence class structure arising from information sources 'dual', in a formal sense, to the UCM, via a standard spontaneous symmetry breaking/lifting methodology abducted from statistical physics, and through an index theorem approach based on an Onsager-like stochastic differential equations model. Surprisingly, similar arguments can be applied to the formally 'fuzzy' generalizations that are likely to better fit actual biological complexities.

I. INTRODUCTION

A vast spectrum of cognitive physiological and other systems engage in dynamic recruitment of lower-level 'unconscious' cognitive modules to form attention-directed, temporary working coalitions that address patterns of threat and opportunity confronting organisms, and do so across a variety of modalities, levels of organization, scales, and time domains, in a highly punctuated manner [1, 2].

This dynamic represents the evolutionary exaptation [3] of the crosstalk inevitable to information transmission. Crosstalk – correlation between interacting channels – is known from electrical engineering practice to require extraordinary efforts for mitigation, and is present across a great variety of biological processes (e.g., [4]).

As McNeill and Woodgett [5] put it, in the context of development, signal transduction pathways interact at various levels to define tissue morphology, size and differentiation during development. Understanding the mechanisms by which these pathways collude has been greatly enhanced by recent insights into how shared components are independently regulated and how the activity of one system is contextualized by others. Traditionally, it was assumed that components of signaling pathways show pathway fidelity with high autonomy. However, there is increasing evidence that components are often shared between multiple biological pathways, and other components talk to each other through multiple mechanisms.

Crosstalk and emergence, however, are not confined to development. Attisano and Wrana [6] argue that complete sequences of animal genomes reveal a remarkably small and conserved toolbox of signaling pathways accounting for all biological diversity. This raises the question as to how such a limited set of cues elaborates so many diverse cell fates and behaviors. It is now clear, in their opinion, that components of signaling pathways are physically assembled into higher order networks that ultimately dictate the biological output of pathway activity.

Previous formal studies of such cognitive global broadcasts have been based on necessary conditions models arising from the asymptotic limit theorems of information theory, following the footsteps of Dretske [7, 8]. These have assumed well-defined, tiling-like, quasi-symmetries inherent to cognitive dynamics, and invoked spontaneous symmetry-breaking arguments akin to those now standard in physical theory to derive the highly punctuated accession to 'consciousness' and overt attention characterizing such phenomena [1, 2, 9, 10]. Here, we extend these arguments to inherently 'fuzzy' systems, where underlying form is smeared out, in a certain sense, and may hence be in better consonance with biological realities.

Fuzzy sets, algorithms, and control systems, were first studied by Zadeh [11, 12], and have received increasing attention (e.g., [13]). Biological global broadcasts that range from the immune system, wound healing, the HPA axis, emotional response, consciousness, and sociocultural distributed cognition [1], are basically control processes, and the application of Zadeh's perspective appears straightforward. In essence, the 'fuzzification' of algebraic structures and relations is based on an extension of the idea of the characteristic function, mapping an arbitrary set G onto the set of integers $\{0, 1\}$, so that $f : G \to 0, 1$. Then, if $x \in G, f(x) = 1$. Otherwise, $f(x) = 0$. The fundamental idea involves letting f map onto the real interval $[0, 1]$ rather than onto a set of integers. Rosenfeld [14] first applied the method to defining fuzzy groups and groupoids, and the construction of group/groupoid representations is relatively direct, although modified by some complexities [15, 16].

II. COGNITION

Here we review, briefly, the development of [1, 2, 9, 10]. Atlan and Cohen [17] argue that cognitive function involves comparison of a perceived signal with an internal, learned or inherited picture of the world, and then

PeerJ PrePrints | https://peerj.com/preprints/34v2/ | v2 received: 2 Sep 2013, published: 2 Sep 2013, doi: 10.7287/peerj.preprints.34v2

choice of one response from a much larger repertoire of possible responses. Thus, instantiating the mechanism of Attisano and Wrana [6], 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, in this model, mixed in an unspecified but systematic manner with a pattern of 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 [9, 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 $B_0 \equiv \{b_0, ..., b_k\}$, and $B_1 \equiv \{b_{k+1}, ..., b_m\}$.

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, let $N(n)$ be 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 $h(x) \in B_1$.

Note that identification of the 'alphabet' of the states a_j, B_k may depend on the proper system 'coarse graining' in the sense of symbolic dynamics [18].

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

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

both exists and is independent of the path x. Recall that $N(n)$ 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 [9, pp. 31-32].

Invoking the spirit of the Shannon-McMillan Theorem, it is 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 [19]

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

This information source is defined as dual to the underlying ergodic cognitive process.

'Adiabatic' means that, when the information source is parameterized according to some appropriate scheme, 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' (roughly) that cross-sectional means converge to long-time averages. Between 'pieces' it is possible to invoke various kinds of phase change formalism [9].

Recall that the 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 [19-21] for the standard details.

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 a high probability meaningful path connecting them to the same origin point. Disjoint partition by equivalence class, analogous to orbit equivalence classes for dynamical systems, defines the vertices of a network of cognitive dual languages. Each vertex then represents a different information source dual to a cognitive process. This is not a representation of a neural network as such, or of some circuit in silicon. It is, rather, an abstract set of 'languages' dual to the set of cognitive biological processes.

Such a set of equivalence classes generates a groupoid, whose algebraic properties – an important extension of the idea of both a symmetry group and an equivalence class – are summarized in [1]. An essential point is that products need not be defined globally [22-24].

We now allow generalization of these ideas to fuzzy groupoids, in Rosenfeld's sense [14].

III. TUNING

Given a set of cognitive biological modules that become linked to solve a problem, the no free lunch' theorem of Wolpert and Macready [25, 26] provides a context. They established that there exists no generally superior computational function optimizer. There is no 'free lunch' in the

PeerJ PrePrints | https://peerj.com/preprints/34v2/ | v2 received: 2 Sep 2013, published: 2 Sep 2013, doi: 10.7287/peerj.preprints.34v2

sense that an optimizer 'pays' for superior performance on some functions with inferior performance on others. Thus gains and losses balance precisely, and all optimizers have identical average performance: superiority on one subset of functions necessarily implies inferiority on some complementary subset.

From the no free lunch argument, it is clear that different challenges facing an organism must be met by different arrangements of cooperating 'low level' cognitive modules, and that changes in those arrangements must be fairly rapid. It is possible to make a very abstract picture of this phenomenon, not based on anatomy, but rather on the network of linkages between the information sources dual to the physiological and learned 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 physiological UCM, with the network graph edges measured by the rate of information crosstalk between sets of nodes representing the dual information sources. A more complete treatment of coupling can be given in terms of network information theory [19], incorporating the effects of embedding contexts – signals from the environment.

The possible expansion of a closely linked set of information sources dual to the UCM into a global broadcast depends on the underlying network topology of the dual information sources and on the strength of the couplings between the individual components of that network.

For random networks the results are well known [27-29]. Assume there are n network nodes connected to each other with uniform probability p. For the simple random case, parameterize as $p = c/n$. The essential finding is that the behavior of the random network has three sections:

1. If $c < 1$, all the linked subnetworks are very small, and no global broadcast can take place.

2. If $c = 1$, there is a single large interlinked component of a size $\approx n^{2/3}$.

3. If $c > 1$, then there is a single large component of size yn – a global broadcast – where y is the positive solution to the equation $\exp(-cy) = 1 - y$, which can be calculated explicitly in terms of the Lambert W function.

For a highly nonrandom network, a star-of-stars-of-stars in which every node is directly or indirectly connected with every other one, there is no threshold, only a single giant component, showing that the emergence of a giant component in a network of information sources dual to the UCM is dependent on a network topology that may itself be tunable. A generalization of this result is possible.

The random network argument above is predicated, however, on there being a variable average number of fixed-strength linkages between components. Clearly, the mutual information measure of cross-talk is not inherently fixed, but can continuously change magnitude. This suggests a parameterized renormalization. In essence, the modular network structure linked by mutual information interactions 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. A value of 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, will be locked out and not be 'consciously' perceived. 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 global broadcast. That topology is the tunable syntactic filter across the underlying modular structure, and variation in Ω is only one aspect of a much more general topological shift. Further analysis can be given in terms of a topological rate distortion manifold [30, 31].

IV. GLOBAL BROADCAST AS PHASE TRANSITION

It is possible to apply a statistical mechanics analog, using Landau's spontaneous symmetry breaking/lifting approach via a Morse Theory argument [1, 32]. See [32, 33] for a summary of standard material on Morse Theory. In general, very many Morse functions can be constructed under a given circumstance, and what is perhaps the simplest can be assembled using representations of the appropriate fuzzy groupoids. Again, representations of groupoids and fuzzy groupoids are similar to those of groups, although there are necessary modifications [15, 16].

Taking an appropriate fuzzy groupoid representation in a particular matrix, function, or other algebra, now, following the example of [34], construct a 'pseudo probability' $\mathcal{P}$ for fuzzy groupoid element ν as

$$\mathcal{P}[\nu] = \frac{\exp[-|\chi_\nu|/\kappa\Omega]}{\sum_r \exp[-|\chi_r|/\kappa\Omega]} \,. \tag{3}$$

χ_ϕ is the character of the fuzzy groupoid element ϕ in that representation, i.e., the trace of the matrix or function assigned to ϕ, and $|...|$ is the norm of the character, a nonnegative real number. For systems that include compact groupoids, the sum may be a generalized integral.

The central idea is that F in the construct

$$\exp[-F/\kappa\Omega] = \sum_\nu \exp[-|\chi_\nu|/\kappa\Omega] \tag{4}$$

is a Morse Function in the crosstalk free energy rate temperature-analog Ω to which Landau's spontaneous

PeerJ PrePrints | https://peerj.com/preprints/34v2/ | v2 received: 2 Sep 2013, published: 2 Sep 2013, doi: 10.7287/peerj.preprints.34v2

symmetry breaking arguments apply [1, 32, 35]. This leads to the expectation of empirically observable, highly punctuated, structure and reaction dynamics in the index Ω that are the analog to phase transitions in 'simple' physical systems. Recall Landau's central insight: for many physical phenomena, raising the temperature makes accessible higher energy – and more symmetric – states of the system Hamiltonian, the quantum mechanical energy operator, and the inherent symmetry changes are necessarily be punctuated. Here the focus is directly on a Morse Function constructed from a representation of underlying fuzzy groupoids.

Consider, now, an inverse order parameter defined in terms of some system attention index, a nonnegative real number R. Thus R would be a measure of the attention given by the fuzzy cognitive system to the signal defining the tuning parameter temperature-analog Ω. According to the Landau argument, R disappears when $\Omega \leq \Omega_C$, for some critical value. That is, when $\Omega < \Omega_C$, there is spontaneous symmetry breaking: only above that value can a global broadcast take place entraining numerous 'unconscious' cognitive submodules, allowing $R > 0$. Below Ω_C, the system breaks up into a number of separate modules, and attention is fragmented, so that $R = 0$. A classic Landau order parameter might simply be $Q = 2/(1 + \exp[aR])$, or $Q = 1/[1 + (aR)^n]$, where $a, n \gg 1$.

V. A FUZZY INDEX THEOREM MODEL

Topology, as discussed above, has long been known to profoundly affect phase transition, but another route to this result is possible, using the rich structure of stochastic differential equations.

Define a 'symmetry entropy' based on the Morse Function F of equation (4) over a set of structural parameters $\mathbf{Q} = [Q_1, ..., Q_n]$ (that may include Ω) as the Legendre transform

$$S = F(\mathbf{Q}) - \sum_i Q_i \partial F(\mathbf{Q})/\partial Q_i \, . \quad (5)$$

The dynamics of such a system will be driven, at least in first approximation, by Onsager-like nonequilibrium thermodynamics relations having the standard form [36]:

$$dQ_i/dt = \sum_j \mathcal{K}_{i,j} \partial S/\partial Q_j \, , \quad (6)$$

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 fuzzy, an even more fitting approach might be through a set of stochastic differential equations having the form

$$dQ_t^i = \mathcal{K}_i(t, \mathbf{Q})dt + \sum_j \sigma_{i,j}(t, \mathbf{Q})dB^j \, , \quad (7)$$

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 [37], as will be explored below. Thus we have added even more stochastic fuzz to the fuzzy symmetries of the previous section.

Setting the expectation of this system to zero and solving for stationary points gives attractor states, since noise precludes unstable equilibria, although the solution may, in fact, be a highly dynamic strange attractor set. Some of these states may not, however, be stable in higher order moments [38].

But setting the expectation of equation (7) to zero also generates an index theorem [39] in the sense of Atiyah and Singer [40], i.e., an expression that relates analytic results – the solutions of the equations – to an underlying set of topological structures. These are the eigenmodes of a complicated geometric operator whose group/groupoid spectrum – fuzzy or not – represents the symmetries of the possible changes that must take place for information to be transmitted at a large scale, i.e., for a global workspace to be activated.

VI. DISCUSSION AND CONCLUSIONS

Cognition, it has long been asserted, pervades, and indeed may define, the living state [41-43]. Many cognitive phenomena can be associated with dual information sources constrained by the asymptotic limit theorems of information theory [7, 8]. These sources, in turn, build equivalence classes incorporating a groupoid structure whose dynamics can be modeled using spontaneous symmetry breaking arguments, via the crosstalk inevitable between closely associated information channels. Evolutionary exaptation of this phenomenon appears ubiquitous across biology. But biology is not physics, and biological structure and process are inherently messy. Simple spontaneous symmetry breaking arguments can, however, be significantly expanded using Zadeh's fuzzyness methodology, in which a generalization of the characteristic function imposes itself on algebraic structures and resulting dynamics, via groupoids and their representations.

The resulting 'fuzzy' symmetry breaking/lifting phase transition models, while sufficiently messy to perhaps better approximate biological phenomena, nonetheless show the same highly punctuated accession to global broadcast and its associated attention measure as does the simpler treatment. An even fuzzier index theorem approach based on stationary states of a system of stochastic differential equations leads to much the same conclusion.

PeerJ PrePrints | https://peerj.com/preprints/34v2/ | v2 received: 2 Sep 2013, published: 2 Sep 2013, doi: 10.7287/peerj.preprints.34v2

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

2. Wallace, R., 2013, Cognition and biology: perspectives from information theory. In press, Cognitive Processing.

3. Gould, S., 2002, The Structure of Evolutionary Theory, Harvard University Press, Cambridge, MA.

4. McCarthy, N., 2010, Cell signalling: regulation and crosstalk, Nature Reviews Molecular Biology, 11:390.

5. McNeill, H., J. Woodgett, 2010, When pathways collide: collaboration and connivance among signalling proteins in development, Nature Reviews Molecular Cell Biology, 11:404-413.

6. Attisano, L., J. Wrana, 2013, Signal integration in TGF-β, WNT, and Hippo pathways, F1000Prime Reports, 5:17.

7. Adams, F., 2003, The informational turn in philosophy, Minds and Machines, 13:471-501.

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

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

10. Wallace, R., 2007, Culture and inattentional blindness, Journal of Theoretical Biology, 245:378-390.

11. Zadeh, L., 1965, Fuzzy sets, Information and Control, 3:333-353.

12. Zadeh, L., 1968, Fuzzy algorithms, Information and Control, 12:94-102.

13. Jantzen, J., 2007, Foundations of Fuzzy Control, Wiley, New York.

14. Rosenfeld, A., 1971, Fuzzy groups, Journal of Mathematical Analysis and Applications, 35:512-517.

15. Bos, R., 2007, Continuous representations of groupoids, arXiv:math/0612639v3.

16. Houghton, C., 1975, Wreath products of groupoids, Journal of the London Mathematical Society, 10:179-188.

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

18. Beck, C., F. Schlogl, 1993, Thermodynamics of Chaotic Systems, Cambridge University Press, New York.

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

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

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

22. Brown, R., 1987, From groups to groupoids: a brief survey, Bulletin of the London Mathematical Society, 19:113-1134.

23. Cannas da Silva, A., A. Weinstein, 1999, Geometric Models for Noncommutative Algebras, American Mathematical Society, Providence, RI.

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

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

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

27. Bocaletti, S., V. Latora, Y, Moreno, M. Chavez, D. Hwang, 2006, Complex networks: structure and dynamics, Physics Reports, 424:175-208.

28. Erdos, P., A. Renyi, 1960, On the evolution of random graphs. Publications of the Mathematical Institute of the Hungarian Academy of Sciences, 17-61.

29. Spenser, J., 2010, The giant component: the golden anniversary, Notices of the American Mathematical Society, 57:720-724.

30. Glazebrook, J.F., R. Wallace, 2009a, Rate distortion manifolds as model spaces for cognitive information, Informatica, 33:309-346.

31. Glazebrook, J.F., R. Wallace, 2009b, Small worlds and red queens in the global workspace: an information-theoretic approach, Cognitive Systems Research, 10:333-365.

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

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

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

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

36. de Groot, S., P. Mazur, 1984, Non-Equilibrium Thermodynamics, Dover, New York.

37. Protter, P., 1990, Stochastic Integration and Differential Equations: A New Approach, Springer, New York.

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

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

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

41. Maturana, H., 1970, The biology of cognition, BCL Research Report 9.0, University of Illinois.

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

43. Maturana, H., F. Varela, 1992, The Tree of Knowledge, Shambala Publications, Boston, MA.

PeerJ PrePrints | https://peerj.com/preprints/34v2/ | v2 received: 2 Sep 2013, published: 2 Sep 2013, doi: 10.7287/peerj.preprints.34v2