Episodes in Consciousness viewed as Brain State Space Transitions

Gerhard Werner
Department of Biomedical Engineering
University of Texas, Austin TX.
gwer1@mail.utexas.edu

Abstract:
In the framework of Dynamical Systems Theory and based on evidence for processes of consciousness being discontinuous, it is here proposed to consider episodes of conscious awareness in the light of non-equilibrium brain state space transitions. The relevant theoretical background is reviewed as are the known landmarks of brain network configurations and their role for conveying oscillatory activity of certain frequencies. This raises the questions: what kind of state space transformations can such systems undergo, and could the relation between sensory-motor cognition and its conscious registration be of the same category of events as are exemplified by state transitions in non-equilibrium physical systems: the formation of a novel physical reality at the macroscopic level, characterized by a different qualitative phenomenology? Approaches for empirical validation of this view by suitably designed brain imaging studies, and for computational simulations of the proposed principle are discussed; the latter guided by new directions in the Physics of complex systems, in place of the extant forms of reductionist explanations.

Introduction:
Two current trends in the discourse in Cognitive Neuroscience and its participating disciplines are, first, the realization that Cognitive Systems are intrinsically dynamical systems (Giunti, 1997, Port & van Gelder, 1995; van Gelder, 1998) which function in interaction with the environment and as embodied actors (Varela et al, 1991; Chiel & Beer, 1997; Thompson, 2007); and, second, the formulation of the Global Workspace Model by Baars (1988, 1997, 2002): both have become reference points for theory formation on brain-mind relations. Baars’ "cognitive theory of consciousness" postulates a multitude of special purpose processors, almost always unconscious. Coalitions of such processors can gain access to a limited capacity global workspace for integrating competing and cooperating input and for recruiting additional processors for dealing with novelty and problem solving. The global workspace is considered the gateway to consciousness. Its principles are reflected in the brain style models of Dehaene & Changeux (2004), are consistent with predictions based on the Dynamic Core Hypothesis of Edelman (2003) and motivated mathematical models by Wallace (2005a,b). For a recent overview, see Werner (2007a).

In the following, I shall draw on the conceptual resources of Dynamical System Theory and the methodological tools of Complex System Science for proposing that considerations of certain aspects of statistical Physics, and of a new outlook on the character of explanation in Physics, will broaden the conceptual base of Cognitive Science and establish a framework for formulating the relation between neural and mental events on a common basis. Its antecedents are the investigations of state transitions at the mesoscopic level of nervous systems by Freeman (e.g.: Freeman, 2005a), and the studies of pattern transitions in brain electrical activity and behavior, conducted in the spirit of Coordination Dynamics (Kelso et al, 1992; Wallenstein et al 1995; Jirsa, 2004; Kelso & Engstrom, 2006).
distinction, I will focus on the creation of qualitative novelty in state space transitions, with its own Ontology and autonomous laws, with possible relevance to events in consciousness.

The following Sections are intended to establish the context of observational and conceptual benchmarks from the Neurosciences and Physics, limited only to such aspects in recent literature and current debates as are required for clearing the way to characterizing the theoretical framework which I propose in the Discussion.

**Oscillatory Neural Activity**

Historical landmarks are Freeman’s (1975) notions of “mass action” in the Nervous System, the discovery of stimulus related oscillatory neuronal activity by Gray et al (1989) and Eckhorn et al. (1988), the suggestion by Damasio (1990) of the functional role of synchronicity of neural oscillations, and its elaboration by Varela (1995) and Varela et al., (2001) in the form of Resonant Cell Assemblies and the Brain Web. Oscillatory neural activity is now widely accepted as subserving communication among cell assemblies across a wide range of spatial and temporal scales (Basar et al., 2004; Buzsaki, 2006): their role is attributed to phase synchrony and thought to be twofold: dynamically constituting functionally interacting neuron assemblies, and acting as communication links among disjoint neural centers. Long-distance synchronization of brain activity in human subjects engaged in perceptual-cognitive tasks was reported, among others, by Rodriguez et al (1999). These and the numerous related observations (for review, see Singer (1993)) can now be supplemented with the more recent findings of Koenig et al., 2005): combining electroencephalographic records with functional magnetic resonance imaging identified transient “binding” (see below) between different brain regions through synchronized oscillations. Synchronization of neural activity in disparate cortical regions may then be thought of as a mechanism for constituting transient functional neurocognitive networks. At least in one instance, it was possible to demonstrate in the awake monkey that oscillations in the beta frequency range carry Granger causal influences from primary somatic sensory and inferior parietal cortices to motor cortex (Brovelli et al., 2004).

The ‘communication-through-coherence’ hypothesis of Fries (2005) proposes to close the gap between functional efficacy of neuronal oscillatory activity on the one hand, and a detailed understanding of its mechanism, on the other: phase-locking is thought to enable effective communication by acting as windows of frequency-specific interaction, as a kind of gating mechanisms. This could constitute an intricate dynamic communication structure for selective routing of neuronal activity. The hypothesis can draw on the fact that gamma-band synchronization of synaptic inputs enhances their effective synaptic strength (Salinas & Sejnowski, 2001), and is also supported by the recent finding of Womelsdorf et al (2007): in a set of elegant experiments on the cortices of awake cats and monkeys, the precise timing between rhythmic activity of neuronal groups determined the strength of their mutual interaction, presumably due to frequency specific rhythmic modulation of synaptic inputs. Hence, effective connectivity can be regulated upward and downward through synchronization, depending on the phase relation. One can readily visualize how the temporal and spatial pattern of synchronization imposes a distributed pattern of synaptic gains (a grid of traffic lights, if you will) on the infrastructure of anatomic connections, with the effect of dynamically modifying interactions among neuronal groups. The significance of this point of view will become even more apparent in later sections of this essay.
Although the focus of interest is generally on the synchronicity (phase locking) of oscillatory interrelation among neuron assemblies, alternating periods of asynchronous coupling are of equal importance for it is these which make dynamic re-configuration of communicating neuron assemblies possible: an interrupt mechanisms, as it were, for enabling swift adaptive changes by shifting coherence patterns from one communicating cell assembly to another (Breakspear et al., 2004; Friston, 2000). The capacity for alternation between synchrony and asynchrony is also one aspect of the metastable coupling/uncoupling of integration and differentiation in coordination Dynamics (Kelso & Engstrom, 2006). Simulated adaptive network controllers evolve towards increasing “fitness” with shifting the dynamical integration/differentiation balance (Seth & Edelman, 2004), reflecting changes in mutual information of interacting neuron pools which was originally proposed as measure of brain complexity (Tononi et al., 1994) and index of consciousness (Tononi & Edelman, 1998). However, it now appears that neither this nor any other thus far considered measure of the brain’s multidimensional complexity adequately reflects the temporal and spatio-temporal dynamics of neural systems (Seth et al., 2006).

How synchronicity is mediated between disjoint neuron assemblies is the subject of the next section.

**On Brain network connectivity.**

Earlier investigation of the brain’s network structure and connectivity patterns (e.g.: Mesulam, 1998) are now enriched by the insights into the statistical mechanics (Albert & Barabasti, 2002), the development (Dorogovtsev & Mendes, 2001) and structure and function (Newman, 2003) of complex networks. Investigations of Hilgetag et al., (2000) and others generated a wealth of neuroanatomical data which, together with the application of methods of graph theory (Bollobas 1985) and computational analysis of network configurations (Hilgetag et al., 2002) came to constitute the field of Computational Neuroanatomy (Sporns et al., 2000). It is concerned with establishing the relations between classes of network connectivity, network dynamics, and measures of the brain’s complexity (Sporns & Tononi, (2002); Sporns et al., (2004). One of the principal objectives of research in this field is to identify the conditions under which structural-anatomical connection patterns become functionally engaged, and what the properties of the resulting functional networks are.

For the purposes of the following, it suffices to consider two classes of networks: small-world, and scale-free. The characteristic connection topology of the former class of networks was first described by Watts & Stogatz (1998) and Strogatz (2001). It consists in the prevalence of highly clustered, short (near- and next-near-neighbor) connections over long range shortcuts to distant nodes. Their connectivity is intermediate between completely regular and completely random, and their short characteristic path length accounts for what is described by the evocative term of “cliquishness”. Scale-free networks, on the other hand are characterized by a scale-free statistical distribution of numbers of edges per network node, which implies their being poised towards critical state transitions.

In a recent review, Sporns and Zipf (2004) reaffirmed and extended the previously established small-world characteristics of the anatomical inter-area cortical connection patterns, as has Humphries et al (2005) for the structure of the brain stem reticular formation. Sporns (2006) also compared brain organization at different levels of scale, ranging from local neuronal groups to patterns of intra- and inter-area connections up to extended cortical systems, and determined that it displays self-similarity, i.e. a fractal patterns, that is: the pattern of network connectivity is identical at all levels of the central...
nervous system, safe for changes in scale. The small-world properties of anatomical brain networks can be viewed as a developmental compromise between keeping the ‘cost’ of wiring length small, and yet secure also the efficiency of long-distance connections. Within complex networks, recurring patterns of interconnections form ‘motifs’ which appear to function as elementary building blocks (Milo et al., 2002). A small repertoire of such structural motifs can also be identified in brain networks (Sporns & Koetter, 2004).

Network topology and network dynamics are decisively interdependent (Grinstein & Linsker, 2005). Therefore, it is essential to ask: what role does the functional brain network architecture play in brain dynamics?

On theoretical grounds, simulation studies attribute significance to features of small-world networks insofar as it supports fast response and temporally coherent oscillations of the network nodes (Lago-Fernandez et al, 2000; Wang & Chen, 2002). This is thought to be of relevance for ‘feature binding’ (Masuda & Aihara, 2004), efficiency in local and global information exchange (Latora & Marchiori, 2001, Bassett & Bullmore, 2006), and for facilitating global synchrony of coupled phase oscillators (Watts, 1999; Barahona & Pecora, 2001). In general, this collective behavior presupposes the non-locality of at least some connections (Gade & Sinha (2006): a requirement which, of course, small-world networks meet. However, small-world networks (graphs) comprise several classes depending on their vertex connectivity; one of them is identical with ‘scale-free networks’ (Amaral et al., 2000; Newman, 2000). Thus, labeling a network merely as “small world” is not necessarily a complete characterization of its functional repertoire.

Observational data obtained with PET and fMRI imaging methods and their interpretation in terms of network theory have generated important insights into intra-area functional connectivity, and its role for human brain function, both in a so-called ‘resting’ state and states of experimenter-elicited activity. However, ‘Resting State’ needs to be taken with a grain of salt (Gusnard & Raichle, 2001). Consider, for instance the brain activity of ‘wandering minds’, attributed to so-called default network of cortical regions (Mason et al, 2007).

Measuring correlations between “voxel” in MGE’s of brains in subjects carrying out a finger tapping task, Eguiluz et al. (2005) identified functional brain small world networks with scale-free properties. Stam (2004) was guided by the idea that network synchronization at different frequency bands may reflect functional differences; in no-task, eye closed conditions, he found that connectivity patterns below 8 Hz and above 30 Hz display small world properties, but they did not in the alpha – and beta range. Frequency dependent difference in functional network connectivity were also found by Salvador et al.(2005).

A major methodological advance in the form of discrete wavelet decomposition of fMRI time series of the human brain (Bullmore et al. 2004) in combination with graph theoretical methods make it possible to extend the frequency band of analysis over a physiologically relevant range. Achard et al (2006) found small-world topology of sparse connections most salient in the frequency range of 0.03-0.06 Hz. The degree distribution of the highly connected hubs of uni- and heteromodal association cortex reflects an exponentially truncated power law. Extending the scaling regime up to 75 Hz, Bassett et al (2006) identified functional networks with small-world topology and critical dynamics over a frequency range from 1 to 75 Hz. Connectivity between brain regions varied with the frequency range, long range functional connectivity in the resting state being stronger at low frequencies. With
task performance, long range functional connections emerge more strongly at higher frequencies, associated with a significant motor task related increase in characteristic length scale of edges in high frequency motor networks. The network dynamics is thus seen as supporting task-related reconfiguration, potentially also drawing on the reserve of uncommitted connections in the otherwise sparse networks. Adaptive and rapid reconfiguration is possible because the synchronizability of the networks is in all states and at all scales close to the order-disorder transition. Under more complex conditions, tested as ‘task sets’ which require initiation, maintenance and adjustments, a fronto-parietal and a cingulo-opercular network participate, operating independently and on different time scales (Dosenbach et al, 2006, 2007; Fair et al, 2007); their life-span development is associated with decrease of short-range, and increase of long-range connections (Fair et al, 2007).

Computational simulations of relations between network structure and activity patterns add insights into an amazingly intricate and specific dynamics in the temporal and spatial domain. In a network model reflecting segregated regions and interregional pathways of the macaque cortex, based on the database CoCoMac (Koetter, 2004), Honey et al (2007) simulated the distribution pattern of spontaneous (i.e. in absence of external stimulation) cortical dynamics. The neuronal dynamics within each of the cortical regions was emulated as a neural mass model with self-organizing oscillations, serving as a spontaneously active motor, as it were, that drives the entire system. Interregional interactions were measured as Transfer Entropy (Schreiber, 2000) and transient synchrony between pairs of regions was expressed as phase locking value (Lachaux et al., 1999). The cortical ‘resting’ state displayed a rich temporal structure at multiple time scales: at time scales of 10 Hz, the system dynamics consists of multiple metastable states; at lower frequencies, fluctuations in the strength of correlated coupling occur, with regions participating in one or the other cluster in an anticorrelated manner. The group of Zhou and associates (Zhou et al, 2006; Zhou et al., 2007; Zemanova et al., 2006) performed a comparable study with similar methodology, based on a model of the cat brain. Their results underscore the sensitive dependence of the network organization into functional synchronizing communities on the degree of coupling, and on the network topology.

The observations cited in the foregoing section point to the brain’s extraordinary capacity for routing the oscillatory activity of neuronal centers into different channels, and over different ranges to different destinations, depending on the oscillation frequency of the signals they convey: a communication system, dynamically structured by the property of the signals it carries. Equally amazing is the interplay between the brain’s small-world network feature supporting synchronisation, and its being near criticality (the later also associated with scale invariance of node degrees which, in the theory of Albert & Barabasi (2002), may be due to local re-wiring, addition or removal of nodes): while criticality favors state transitions, it also constrains the small world synchronizing propensity of the networks, thus preventing ‘seizures-like’ loss of network frequency specificity (Percha et al. 2005): a proper balance of opposing trends must be maintained.

Albeit held in balance, the scale-free aspect of the functional justifies Chialvo’s (2006) evocative expression of “the brain near the edge”. This is also evident from appropriate analyses of the encephalogram. In numerous publications, Freeman (2004a, 2004b, 2005, 2006) provided solid support for the view that the brain state space dynamics is poised towards a global state of self-organized criticality which affords the possibility for virtually instantaneous reorganization upon changes in external or internal variables. Evidence for scale-free brain dynamics of EEG as signature of self-organized criticality was also obtained by Linkenkaer-Hansen et al, 2001, Gong et al (2003), and Stam & de Bruin, (2004). Transcranial magnetic stimulation was shown to induce switches
between two distinct modes of behavior (Meyer-Lindenberg et al, 2002). The totality of these observations is also consistent with qualitative predictions from Baars’ model which envisions the possibility of abrupt activation of a Global Workspace under appropriate conditions. For relevant details on brain self-organized criticality, see Werner (2007a, b).

**On Cognition and Consciousness.**

Discourse in Cognitive Neuroscience is not consistently sensitive to the distinction between Cognition and Consciousness; in fact, it is frequently blurring or even disregarding this distinction. Yet, the reactive behavior of organism’s sensory-motor interaction with the environment must be clearly differentiated from the qualitative subjective states and processes of sentience and awareness; the latter having their own, distinct Ontology (Searle, 2000). The former are emulated by the impressive showpieces of embodied, situated dynamical systems of robots and other automata (see, for instance, Nolfi & Floreano, 2000) and do not merit the attribute ‘conscious’, despite harnessing the power of evolutionary design in computational neuroethology. But the arguments of Dreyfus & Dreyfus (1988) and of Searle’s (1980) Chinese room for the lack of meaning for the agent apply here as they did to the claims of Artificial Intelligence (Di Paolo, 2004).

One group of investigators attribute the distinction between unconscious and conscious sensory events to recurrent processing: based on studies of the visual system, Lamme and associates (Lamme & Roelfsema, 2000; Super et al., 2001; Lamme, 2006b) claim that reportable conscious visual experiences require that the “feedforward sweep” of neural activation from visual towards motor areas become extended to a “backward sweep” which consists of widespread recurrent activation of frontal, prefrontal and temporal cortex, putting “visual information into the context of the system’s current needs, goals and full history” (Lamme, 2003). (see also: Dehaene et al., 2006; Pascual-Leone & Wals, 2001).

However, what appears to be also required (perhaps in addition to the foregoing) is that the somatically embodied and environmentally reactive behavior be associated with adaptive autonomy and internal bioregulatory processes of emotional embodiment. Based on thoughtful analysis of extended clinical observations, Damasio (1990, 1994, 2001, 2003; Damasio et al., 2000), and more recently Rudrauf & Damasio (2006) view subjectivity and the experiencing “I” as essential aspect of Consciousness. It is a “feeling of knowing” that makes the relation to a perceived object salient to the organism. Feelings are grounded in the multidimensional, moment-to-moment shifting physiological internal body environment, associated with the subcortical regulatory structures of, for instance, Hypothalamus and Brain Stem Tegmentum, and the cortical formations of Insula and Cingulate regions. The act of a subject’s conscious cognition is thought to consists of two mutually interdependent and interacting components: sensory events in transaction with the external world, and a ‘feeling’ which is the concurrent reading of the body’s internal environment; the latter in some sense anchoring the subject’s ownership of the former. This theory is consistent with a wealth of observations in Clinical Neurology, and with neuro-anatomical and neuro-physiological data (Craig 2002, Critchley, 2004).
Processes in consciousness: discontinuities as transitions in brain state space.

In the following discussion, I adopt Searle’s (2000) notion of the state of wakefulness as basal (background) consciousness, a kind of unified field (see also: Revonsuo, 2006). Presumably, this state overlaps with the condition of vigilance in the terminology of Dehaene & Changeux (2003). Specific sensory events would then punctuate, as it were, the steady state of the unified field, as the basis of discrete subjective experiences. This notion of discrete events in Consciousness tallies with observations of Fingelkurts & Fingelkurts (2006) of discontinuities in the EEG which they identify as transient operational brain microstates, reflecting shifting activation of neuronal networks; and of Lehmann and associates (2006, 1998) who describe punctuated abrupt changes in EEG activity: presumably distinct steps in mental information processing. Sergent & Dehaene (2004) take their findings with the attentional blink test to be concordant with the notion of consciousness being discontinuous, possibly the result of nonlinear bifurcations in brain neural activity.

Applying insights from the dynamics of state transitions, I will now indicate their potential heuristic merits for characterizing discrete brain events of the kind adumbrated in the foregoing. The following sketch of the relevant landmarks of the theory of dynamic critical phenomena is intended to set the stage. The familiar observation that many conservative (equilibrium) physical systems can undergo a deep structural modification when an external parameter is changed beyond a “critical” value is the subject of the (classical) Theory of Dynamic Critical Phenomena (Hohenberg & Halperin, 1977; Stanley 1999). The state transition of magnetization of ferromagnets is an illustrative example. One of the central notions of this theory is the concept of universality: it allows to group the great variety of critical phenomena of different materials into a small number of universality classes, each characterized by a few basic features (determined as ‘scaling functions’) that obtain in the vicinity of the critical point (i.e. the singularity of state transition), irrespective of the details of the system’s microscopic structure and composition.

For complex systems far from equilibrium, most of the analytical and numerical methods of the ‘classical’ theory remain useful, except for some restrictions: relaxing the balance condition of conservation and the absence of thermodynamic energy parameters. But lack of analytical solutions requires that in most instances numerical simulations must be used to identify a system’s critical behavior. Universality classes in non-equilibrium state transitions are more diverse than in the equilibrium case, and their full classification is still lacking. However, on the positive side, new possibilities arise in the form of state transitions to absorbing (i.e. irreversible) states, and to novel spatial structures which are influenced by the system’s history (Hinrichsen 2006; Luebeck 2004; Odor 2004). Examples of notable universality classes of far-from-equilibrium systems are: Directed percolation, the so-called voter universality class, and the Domany Kinzel cellular automaton. Coupled map lattice models for spatiotemporal intermittency under an asymmetrical updating regime belong to the universality class of directed percolation (Rolf et al., 1998); universal critical behavior is also reported for two-dimensional coupled map lattices (Kuznetsov ,1992; Marcq et al., 1996; Just & Schmueser , 2005). Under certain conditions, self-organized criticality can be understood in terms of absorbing state phase transitions (Dickman et al., 1998).

In the present context, these considerations are of relevance for dynamic processes on brain networks, such as referred to earlier. Topology and dynamics are inseparably linked in the constitution of networks and reflected by their universality class’s defining scaling relation (Albert & Barabasi, 2002; Dorogovtsev et al., 2000). One of the most intensely studied processes taking place on networks
is Percolation (Stauffer & Aharony, 1991/1994). Basically, a percolation process consists of the evolution of patterns in arrays of elements under a given dynamics. The task consists in determining rules of transformation (i.e.: a dynamics) that would generate certain types of patterns (e.g. a Giant Component, i.e. certain patterns of connectivity within the network). If this is achieved, one speaks of a State Space Transition for which Universality classes and critical parameters can be computationally established. For reasons stated earlier, models of this kind on small-world networks are of particular interest (Newman & Watts, 1999; Moore & Newman, 2000; Newman et al, 2002). With a view of simulating state transitions in models of the cerebral cortex, Kozma et al. (2004, 2005), examined percolation in random cellular automata. Wallace’s (2005b) approach to modeling Baars’ Global Workspace envisions the formation of Giant Components whereby the dynamics of their formation relates to the mutual information among the interacting modules.

The state space approach: on neural and mental spaces.

The rich source of recent observations with Brain Imaging, together with the long standing evidence on oscillatory brain activity from Electroencephalography and Coordination Dynamics contributed in the past two decades to the growing realization that the brain should be viewed as a complex dynamical system of unprecedented complexity (John, 2002). This recommends adopting the state (phase) space approach for characterizing the brain’s state as points (or circumscribed regions), and changing activity unfolding in time as trajectories in a high dimensional space. It is then also possible to conceive of behavior, cognitive functions and phenomenal experience similarly in terms of state space dynamics, and to examine corresponding trajectories in brain and mental spaces. Examples of this are: Wackermann’s (1999) assessment of electroencephalographic field changes as state space trajectories, Hobson’s (2000) view of different stages of wakefulness and sleep in terms of state space dynamics; the demonstration of global brain state transitions occurring simultaneously across multiple forebrain areas (Gervasoni et al, 2004); Fell’s (2004) mappings between brain states and phenomenal experience, and P. Churchland’s (1989, 2006) long standing explorations with comparing human phenomenological with neural-network activation spaces. The merits of Dynamical System Theory for mapping Psychological Theories on to neural structures were extensively reviewed by Lewis (2005). As indicated in the Introduction, application of state space concepts in the form of Coordination Dynamics has significantly enriched the understanding of relations between global brain dynamics and behavior.

Here I will pursue additional aspects of state space strategies which Mahner and Bunge (1997) consider an ontological primitive in Biophilosophy. The set of all potentially accessible states of a system constitutes its state space of which each dimension corresponds to a system variable; the system variables can be viewed as components of the state vector. Dynamic system theory is concerned with the progression in time of state vectors in phase space, describing the trajectory of the system’s evolution. Ideally, a system’s state space representation should be based on the dimensions which are intrinsic to the system’s operation, but in actual practice, observables accessible to the observer ‘on the outside’, so to say, must be used. In such cases, the state space is observer relative. Consider, for instance, that state space coordinates in neural activity space are usually based on measures of neural activity in accord with the investigator’s notion of relevance and by accessibility to direct measurement. Thus, depending on the selected measure of neural activity (e.g. neural spike rate, spike interval, LFP, oscillation phase, etc), the neural activity state space is observer relative, and disregards the fundamental fact that the thus constructed state space prejudgets the variables that are
intrinsic to the system’s operation: an instance of the notorious Endo-Exo Physics problematic (Roessler, 1994).

A mental phase space can be chosen to represent attributes of one or another cognitive function (perception, memory, etc) or attributes of subjective (phenomenal) awareness and levels of consciousness. It is then at least in principle possible to evaluate topological features of the trajectory of the chosen mental function with that in neural activity space as a “Neural Correlate” (Chalmers, 2000), subject to the aforementioned observer relativity of dimensionality (scale) selection. P. Churchland (1989, 1995) applied this approach to a characterization of neural states to color perception, and Fell (2004) sketched the application of this approach to compare the frequency of EEG oscillations with levels of phenomenal awareness. If such a comparison is successful, it raises the question whether the trajectory in neural space stands in a relation of causal explanation to that in mental state space. Within the state space approach, this could only be decided on the basis of the temporal relation between the two trajectories or some other source indicating information or energy transfer between the trajectories (Fell et al, 2004). Any decision on this point entails a philosophic commitment: either one conceives of the trajectory pair as reflecting the same process measured in two scales (the neural and the mental) as an Identity theory would demand, or one is liable to slip into if not overt though, at least, some form of crypto-Dualism. Accordingly, even the state space approach does not resolve the prevailing conceptual muddle of “the neural correlate of Consciousness” (Noe and Thompson, 2004) which, however, is perhaps best put to rest with the telling argument of Humphrey (2006), based on the mismatch in the dimensionalities of first and third person phenomena.

The state space approach advocated in this section invites the consideration of exploring an alternative to the extant approaches to ‘Psycho-Neural Reduction’, which will be the subject of some general and specific considerations in the next Section.

**On Theory Reduction, Emergence and asymptotic explanations in Neuroscience**

Philosophers of Science frequently have in mind a logical relation between theories with focus on deducibility of one theory from another. But the majority of the purported intertheoretic reductions advanced to illustrate this alleged relation (Hempel, 1965; Nagel, 1961) are now recognized in Physics to be too simplistic and failing to meet standards of rigor (Primas, 1998). Bickle’s (1998) proposes an form of Reductionism as the ‘New Wave’ approach to Psychoneural Reduction which seeks to overcome the flaws of earlier reductionist theories; and Craver (2007) turned to causal-mechanical explanations in the systems tradition of Simon, Wimsatt and others. While superior in many instances, I submit that both approaches fall short of producing satisfactory accounts of the entangled interactions among the components of complex of nonlinear systems, and the dynamics of their state transitions.

An alternative framework of Scientific Reasoning is rooted in Contemporary Physics and rests on the notion that reality is composed of a hierarchy of scales: the theory on each scale emerges from the theory on the next finer scale by ignoring some of the lower-level details which are irrelevant for the higher level (Anderson, 1977; Laughlin & Pines, 2000; Sokal & Bricmont, 2004). In this view, each level has its own description, and a theory that accounts for that description. For an intuitive appreciation of this, recall the frequently cited case of the “Coast of England” as an example of fractals: a coastline is shown under different magnification, each rendering a different description, and associated with it a “theory” (i.e. a formal rule) that accounts for it. Putting it more generally: each
scale presents itself with its own pattern and with its own formal rules that characterize it. The so-called higher level description is not an approximation of the fundamental (low level) theory, but represents (qualitatively) new patterns of reality (Primas, 1998). In the case of physical matter, the re-scaling is associated with state transitions. As an empirical fact, near the critical point of state transition, certain parameters of the system assume critical values. The important point is that such critical values define entire classes of system (Universality Classes) which present identical macro levels despite a wide range of differences at the micro level (in terms of composition, physical properties etc). The important point is the multiple realizability of emergent new macroscopic phenomena despite microscopic diversity (for details, see: Batterman, 1995, 2000, 2002). The idea underlying this principle can be illustrated with an example from Probability Theory: when sets of multiple independent population samples are subsumed under some statistical (say Gaussian) Distribution, then the parameters of the distribution characterize a (kind of) Universality Class, i.e. the ensemble of which the individual sets of samples are independent realizations.

At stake is the basic issue of delimiting the dominant from irrelevant parameters that shape the system’s behavior at the macro level, for the purposes of categorizing universality classes: i.e. classes of systems which approach asymptotically in the limit identical macroscopically observable behavior despite diversity at the microscopic level. Most of the successful applications of this framework come from Thermodynamics and Statistical Mechanics near critical state transitions. Universality and multiple realization designate that 1) some details of the system which would figure in a detailed causal-mechanistic explanation of the system’s behavior, are in the limit largely irrelevant for characterizing the macroscopic phenomenology of interest; and 2) different systems with vastly different “micro” details can exhibit identical behavior at the macroscopic level. The methods to elucidate both aspects are “coarse graining” (e.g.: Schulman & Gaveau, 2001) and the strategies of Renormalization Group Theory (Wilson, 1979): the latter essentially, re-scaling by successive coarse graining while maintaining self-similarity. Renormalization theory is a “general framework for extracting phenomenological relations among macroscopic observables from microscopic models that may not even be precisely definable” (Goldenfeld et al., 1989). Thus, if what one is interested in is the macroscopic phenomenology, then the adoption of renormalization-like procedures and of asymptotic arguments is the path to follow. Ross & Spruett (2004) touched briefly on related considerations in reference to cognitive and behavioral Sciences.

**Discussion and Conclusions**

The aim of this essay is to assemble evidence suggestive of viewing episodes in consciousness in the framework of critical brain state transitions: reactive sensory-motor cognition and subjective awareness would be associated with different organizations of brain activity at the micro- and mesoscopic level on the one hand, and correspondingly different macroscopic phenomenology.

I pursued this objective in the foregoing in four stages: 1) reviewing the evidence that the dynamics of the brain is poised towards critical state transitions; 2) suggesting that critical state transitions may entail the principle of a “backward sweep” such as envisaged by Lamme and associated for the visual system, and/or involve the inclusion of interoceptive functionality in the active brain state. A quote from a paper by Baars et al (2003) relates to this point: “Conscious perception, like the sight of a coffee cup, seems to involve the brain identifying the stimulus. But conscious input activates more brain regions than are needed to identify coffee cups and faces”. 3) that at certain ‘tipping points’ (Gladwell, 2000), a brain state transition would be associated with a
qualitatively novel macro-level configuration with the phenomenological attributes of discrete events in consciousness; and, finally, 4) that the relation between micro- and macro level is one of descriptions at different scales, with the macro level in the asymptotic limit being a coarse grained approximation of the micro level’s universality class: a level at which microscopic details become largely irrelevant, and where multiple realizability and different laws obtain. In this view, brains of adequate size belong to a Universality Class which can on phase transition exhibit at the macro level the phenomenal attributes associated with consciousness.

The state space framework of micro- and macro levels implies inevitably an intersection of Complexity Science with the perennial philosophical problem of Emergence (Bedau, 2002). Kim (1999) identified five main tenets of “the doctrine of Emergentism”, singling out as defining features the coming-together of lower-level entities in new structural configurations; the origin of “higher level” properties, their unpredictability and irreducibility and, finally, the causal efficacy of emergent properties of their own. Bedau (2002) distinguishes strong and weak emergence: the former having irreducible causal powers—a dramatic form of ontological novelty—, the latter being consistent with and functioning via micro level laws. However, in light of the foregoing discussion of the singular nature of the limiting asymptotic relation between the finer and the coarser theory, it appears that the singular limits are in fact the interface where new objects, properties and laws originate. Taking the view of reality as a hierarchy of scales as basis, objects and properties in reality appear at each level in the hierarchy with their own organizational laws and structure, and with the propensity to undergo sharp state transitions. (see e.g.: Laughlin, 2005). States (phases) are cases of emergence, based on Nature having walls of scales: microscopic rules can operate at one level and, yet, be relatively irrelevant to macroscopic phenomena.

What, if anything, can be learned from applying this line of thought to the questions of Cognition and Consciousness? The intent all along was to adduce plausibility to viewing episodes in consciousness as global brain state space transitions: the emergence, under certain circumstances, of a new organizational structure with novel qualitative phenomenology at the macro level. At the micro level, the state transition would consist of including additional neuronal centers in the zone of active neural processing (see e.g.: Baars et al, 2003), or redirecting the processing flow (Lamme, 2006 b). Could such ‘backward flow’ be comparable to the activation of V1 in the imagination of visual scenes (Kosslyn et al, 2001) ; and could such configuration changes be considered as cases of apparent ‘downward causation’?

The general outline of this proposal entails practical consequences for the conduct of research: in the first place, to seek evidence in the temporal evolution of magnetoencephalographic signals from human brain that could reflect the accretion of Giant Components, especially those which would also encompass brain regions which signal interoceptive physiological conditions of the body. An illustrative example of the type of study design that may be informative is provided by the work of Braeutigam et al (2001, 2004) on identifying stages in real life decision processes. Dehaene et al. (2006) list several methodological refinements which would also yield illuminating insights.

The second consequence follows, in principle, the recommendation of Atkinson et al (2000) to investigate (loosely stated) ‘computational correlates of consciousness’. Numerous specific questions arise the answers to which are beyond one’s intuitive grasp: take for instance the observations of Basset, Achard, Salvador and other investigators cited earlier, indicating the presence of multiple functional small-world networks, each carrying oscillatory activity of different frequencies. What is
the space of potential dynamic evolutions such a highly differentiated system can undergo? What are their “tipping points” (Gladwell, 2000) for state space transitions? What is the space of potential reconfigurations that such a systems can undergo under perturbation? How could reversal of activity flow in networks come about? Under what conditions sustain such systems stability? What forms of interactions may obtain among networks conveying different oscillatory frequencies? What if different active networks share some nodes? Being poised towards instability, how would state transition in one network affect others? Or are the frequency-specific networks functionally segregated to the extent that one may view them as chords in a string instrument? Surely, to come to appreciate the space of possibilities in brain dynamics, these kinds of questions need exploring in their own right. In addition, one may also expect that simulations probing these kinds of questions will provide useful pointers for interpreting imaging data and, perhaps, also raise probing questions for empirical analysis.

Whatever the specific question under study, it amounts in the interpretive framework of this essay to exploring the space of brain-like non-equilibrium dynamical systems on lattices, identify universality classes, and establish the conditions for state space transitions. This approach is guided by the ground rules of Renormalization and asymptotic arguments in the study of the phenomenology of macrostates of complex dynamic systems, where microscopic details become largely irrelevant.

Bear in mind that the main thrust of this essay is the notion of emergence of new patterns of reality in the evolution of complex dynamic systems that cannot be deduced from their microscopic configurations, but can, at best, be approximated by equivalence classes of microscopic models. It is then a pragmatic issue to select from among candidate models those with best predictive value for macrosystem performance, and in closest accord with features and constraints imposed by the system’s known micro- and mesoscopic organization.

Some final questions: can it be shown that brains belong to a universality class with phenomenal attributes of consciousness (as here circumscribed in the foregoing sections), or are they unique and in a class by themselves? If so, on account of what property? And conversely: do there exist micro level states of other matter which, on state transition, constitute a brain-like universality Class?

References:


Chialvo DR (2006). The brain near the edge. 9th *Granada seminar on Computational
Physics, Granada Spain.


Time scales measured with EEG. *Phil.Trans.R.Soc.B* 360:1015-1023.


observed in monkey primary visual cortex (V1). Nature Neurosci. 4: 304-310.
Thompson E, Varela FJ (2001). Radical embodiment: neural dynamics and
USA 91:5033-5037
Synchronization and large-scale integration. Nature Reviews.Neuroscience
2:229-239.
Von der Malsburg C (1981/82, 1994). The correlation theory of brain function . In:
Models of the nervous systems, Vol. 2, E. Domany edit..
Von der Malsburg C (1999). The what and the why of binding: the modeler’s
Wackermann J. (1999). Towards a quantitative characterization of functional states
Of the brain: from non-linear methodology to the global linear description.
Workspace model. New York: Springer.
Consciousness model. //cogprints.org/4578/01/modular72.pdf
Nature 393:440-442.
Werner G (2007a). Metastability, criticality and phase transitions in brain and its
Werner G. (2007b). Brain Dynamics across levels of organization. J. Physiol. Paris,
In print.
Womelsdorf T, Schoffelen J-M, Oostenveld R, Singer W, Desimone R, Engel AK,