Metastability, Criticality and Phase Transitions in brain and its models

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

Abstract

This essay is designed to organize a range of experimental findings and theoretical insights of the past 25 years into a coherent view of the brain’s style of function. The view that emerges places the brain firmly into the conceptual framework of Nonlinear Dynamics, operating at the brink of criticality which is achieved and maintained by self-organization. Application of the twin concepts of scaling and universality of the theory of non-equilibrium phase transitions will then be introduced as an approach to elucidating the nature of underlying neural processes, notably with reference to the role of reentrant activity in neural circuits of cerebral cortex and subcortical structures.

1. Introduction

The relevance of the Dynamical System framework in Neuroscience has been foreseen by Turing (1950) as a way of affording the nervous system the speed and flexibility required for instantaneous reaction to novelty. Ashby (1952/1960) similarly suggested that brain phenomena might best be understood in terms of dynamic theory. More specifically, Katchalsky et al., (1974) stated that “…waves, oscillations, macrostates emerging out of cooperative processes, sudden transitions, patterning, etc. seem made to order to assist in the understanding of integrative processes of the nervous system …”. Haken (1983, 2002) proposed to view the brain as a pattern forming system that operates close to instability points for flexible and rapid switching between coherent states. Basar (1983) was an early contributor to the dynamical system neuroscience framework. Since 1975, Freeman has produced a steady flow of studies of the dynamic principles of wave patterns in brains which have yielded numerous relevant findings, including characterizations of attractors, bifurcations and critical phase transitions. (1975; citing merely more recent publications: 2000,2003,2004; Freeman & Holmes 2005; Freeman & Vitiello, 2006). The investigations which I will discuss in the following have generated several dynamical hypotheses of brain processes which form a kind of “family resemblance”, though also differing in some important aspects and –where applicable- in their respective computational models. Related issues have recently been reviewed by McKenna et al, (1994), Le van Quyen (2003), and Cosmelli et al (2006).

For clarity of exposition, I will trace the various strands of conceptual development and hypotheses separately, notwithstanding the many cross links among them. My main objective is to distill from these various strands a common theme, and to situate it in an intersection of Nonlinear Dynamics and statistical Thermodynamics. This, I suggest, holds a strategic place for gaining more detailed insight into neural-mechanistic events and processes.

2: Background: Data and Models

2.1: Oscillatory activity and Coordination Dynamics

I take the remarkable discovery of stimulus induced oscillatory and synchronous neuronal activity with the predominant frequency in the 40 Hz range (gamma wave activity) as starting point. In sensory systems, the
phase synchronization of oscillation links neurons functionally together to groups that respond to identical stimulus features (Gray and Singer, 1987, 1989; Eckhorn et al. 1988). The functional neuron assemblies thus formed consist of elements which are distributed in space, but are ‘locked’ together by a common signal phase. The possibility of this synchronous activity of neuronal assemblies being a candidate for linking separate stimulus features to objects of conscious perception was intensively investigated (Engel et al. 1999, Lamme et al. 2000; Lutz et al. 2002; Ribary et al. 1991, Singer et al. 1997; Tallon-Baudry and Bertrand, 1999). However, here, I will direct attention to the two unusual properties of this oscillatory activity as of principled significance for brain theory: first, virtual simultaneity and very short onset latency in the oscillating ensemble, implying a process on a faster time scale than conduction and synaptic delays would permit (Singer, 1998; Roelfsema et al. 1997); and second, their pattern of short oscillatory sequences interrupted by epochs of stochastic activity (Murthy and Fetz, 1992).

Simulation studies of neural assemblies replicated the abrupt switching between synchronous and stochastic activity and attributed it to synaptic nonlinearity and fluctuations from peripheral input and/or interaction with other assemblies (Bauer and Pawelzik, 1993). This justifies subsuming the cortical oscillatory activity under the category of Kuramoto’s (1984) “self-synchronization transitions” as a paradigm for synchronization phenomena and a mode of self-organization in populations of interacting elements (Acebron et al. 2005). A revealing aspect of the oscillatory neural activity under discussion is the already mentioned short onset latency. The occurrence of abrupt changes of state, propagated through the system at short latency suggest that the dynamics of the neuronal aggregates under consideration is on the brink of critical transitions. This line of thought will be amplified in Section 3.2.

Activity in a wider frequency band than gamma oscillations reveals an intricate dynamics of surprisingly long lasting activity patterns of appreciable spatial extent, initiated by a sensory event or in task situations. The time course of signal distribution across the multiple cortical areas of the visual system extends over some 150 msec from time of onset of stimulus, with the sequence of activation departing drastically from what one would expect if the cortical regions were activated sequentially in a pure hierarchy: for example, local field potentials (LFP) in Visual area V4 have a much longer latency than, for instance, the temporal areas, although geographically much closer to the cortical visual input are V1 (Schmolensky et al. 1998). A meta-analysis of visual response latencies in the macaque cerebral cortex ascertained a complex pattern of feed-forward and recurrent activity flow: for several hundred msec after a brief stimulus, the temporal sequence of activation departs from the topological proximity of activated structures (Lamme and Roelfsema, 2000). A complicated dynamics of activation and deactivation of cortical regions is also generally associated with the execution of attention-demanding tasks, with some regions being activated while others recede to relative quiescence. This task-related regional dichotomy of activity is also demonstrable in the resting state, suggesting a cortical network of intrinsically anticorrelated components (Fox et al., 2005). In perceptual tasks, synchronous local field potentials (LFP) oscillating at frequencies from 12 to about 80 Hz signal sequential patterns of co-activation of different brain regions, extending up to 500 msec in successful (but not failed) discrimination performance (Bressler, 1995; Ding et al. 2000; Bressler and Kelso, 2001; Rodriguez et al., 1999) with different neuronal assemblies evolving in different frequency ranges (von Stein and Sarntheim, 2000). Braeutigam et al. (2004) demonstrated that the temporal course of decision making in the selection (construction) of preferences, reflected in magnetoencephalographic activity, extends over several hundred msec. Finally, brain imaging during natural viewing conditions reveal the temporal organization (“chronoarchitecture”) of brain activity since functionally different brain regions exhibit individually distinct temporal activity patterns (Bartels and Zeki, 2004).

The conceptual and methodological implications of these and related findings are consonant with the principles of Coordination Dynamics (CD) of complex systems. It serves as a framework for studying the manner in which interdependencies among a system’s components dynamically evolve in time, which will be examined more closely in see Section 3.2. Suffice it to say at this point that it is concerned with the phase relationship of LFP’s as the significant control parameter (collective variable) for ordering the sequence of
activation patterns in time and space (Varela et al., 2001; Bressler and Kelso, 2001). Unlike the msec and mm range of local coordination, times extend up to several hundred msec, and spatial interdependence may encompass diverse cortical and subcortical zones, and the brain stem. High speed Electroencephalography and Magneto-Encephalography supplied data for a refined analysis of relationships between brain activity and perceptual-cognitive behavior. A review by Bressler and Kelso (2001) summarizes the essential results of numerous studies which show that both behavior and brain activity undergo a sudden transition from one stable state to another at some critical stimulus parameter. The following sketch of a typical experimental paradigm will illustrate the principle: the subject’s task is to flex a finger in response to tone stimuli of a certain frequency, initially in a syncopated mode; neuromagnetic field activity, reflecting the dendritic current flow in the brain is recorded from the scalp. The repetition rate of the tone is then stepwise increased. At a particular critical frequency, a sudden transition from syncopated to synchronous responses is accompanied by an equally abrupt transition of the brain’s electromagnetic field (Kelso et al, 1992) from one to another stable pattern occurs, for which the stimulus frequency is a control parameter. This phenomenon is thought to reflect the brain’s universal property of complex systems operating in a metastable dynamics, presumably due to an interplay between integrating and segregating tendencies (see Section 3.2). Signs for the brain’s metastability are also evident in a recent study of Freeman & Holmes (2005), based on the analysis of spatio-temporal patterns in the 12-80 Hz band of the Electroencephalogram. Earlier, Freeman and Barry (1994) had demonstrated the propensity for dynamic instability, manifest as rapid global state transitions in oscillatory cortical activity. Metastability as basis for formation and rapid switches among different operational modules in the EEG is also the essential theme of the extensive work of Fingelkurts and Fingelkurts (2001, 2004), possibly related to the succession of ‘microstates’ in the records of Low Resolution Electromagnetic Tomography [LORETA] (Lehman et al, 1998, 2006). I will return to questions of origin and nature of metastability, and of microstates in the Discussion section.

2.2: The Dynamic Core Hypothesis (DCH)

Plasticity of synapses and neuron connections afford a causal link between the functional organization of neuron assemblies and the world, adaptive to use and disuse. The Theory of Neuronal Group Selection (TNGS), also referred to as the Darwinian Brain, is an application of this principle (Changeux, 1983; Edelman, 1987,1989, 1993; Tononi and Edelman, 2001]. A primary repertoire of anatomical connections established during development responds to experiential exposure to the environment with differential amplification of synaptic populations. The spontaneous formation of neuronal groups can be attributed to the interplay between spike time dependent plasticity (STDP) and conduction delays (Izhikevich et al., 2004). The second central notion is reentrant mapping: this is a dynamic process that is inherently parallel and distributed. It consists of ongoing signaling between separate neuronal groups in a reciprocal and recursive fashion over cortico-cortical, cortico-thalamic and thalamo-cortical radiations. Neuronal group selection and reentrant mapping, together, are considered the prerequisite for establishing new and sustaining existing statistical signal correlations between groups of neurons. Neuronal groups thereby come to reflect spatiotemporal properties of signals arising in the environment, and serviceable for perceptual categorization.

Generalization of this principle to cross-modal perceptual categorization is accomplished by dynamic structures that encompass multiple reentrant local maps (sensory and motor) and interaction with basal ganglia, brain stem and cerebellum; the latter for perception-action coupling. Within such global mappings, long-term changes in synaptic strength favor the formation of neuron groups with correlated activity as basis for memory. Memory in global mappings is procedural, and requires dynamical re-assembly by rehearsal. Note that each re-assembly of a global memory may be constituted by different neuron populations: a consequence of the degeneracy (redundancy) of neuronal groups.

Reentry is instrumental for generating oscillations in the simulated models: Sporns et al. (1991) and Tononi et al. (1992)] established in their respective models the linking of stimulus features by reentrant circuitry, within and between segregated cortical areas. It depends in these studies on the occurrence of rapid
changes in efficacy of reentrant connections, and is an aspect of segregation and integration of elementary features into objects and background through temporal correlation and phase relationships among neuronal groups. In a very large computer model of 65,000 spiking neurons, topographically organized into a primary and secondary visual cortical area and two associated thalamic regions, synchronous oscillations emerged spontaneously, even though the networks was not designed to produce any form of specific dynamics (Lumer et al. 1997).

Conscious experience is in this theory associated with global properties of large but distinct sets of distributed neuronal groups: the Dynamic Core (DC). The theory associates cognitive events with the formation of distributed clusters of neurons that are intensely interacting with each other (i.e.: integrated) and, at the same time, are quite distinct and differentiated from the rest of the system. Functional segregation is epitomized by stimulus feature detectors in cortical receiving areas; functional integration is expressed in temporal correlations and synchrony in the large-scale, reciprocally interconnected cortical network and thalamic regions. The neuron clusters of DC achieve high integration within hundreds of msec through reentrant interactions in the thalamo-cortical system. DC must be viewed as a process, creating transiently the clusters of neurons which reflect rapidly shifting functional connectivity among distributed neuron groups, not constrained by anatomical proximity.

Several measures have been proposed to characterize quantitatively the interplay between the seemingly opposing tendencies of differentiation and integration. Tononi et al (1994) defined a statistical measure of the amount of ‘structure’ within a system’s dynamics, as expression of its “complexity”: its value is high for systems with subsets of large mutual statistical dependencies. In such systems, different parts can engage in separate activity and yet remain interdependent (Tononi and Edelman, 1998). More recently, Tononi and Sporns (2003) and Tononi (2004) proposed a measure of a system’s capacity to integrate information, and Seth (2005) defined the measure of Causal Density for the fraction of causally significant interactions among a system’s elements. It appears, however, that none of these measures by themselves captures the temporal and recursive aspects of the multidimensional complexity of neural systems that could account for conscious experience (Seth et al., 2006).

The central role attributed in the DC theory to processes of integration and differentiation in neuronal groups directed attention to whether (and if so, in what manner) neuronal networks are equipped to support these processes. In a remarkable convergence of results from Computational Neuroanatomy (Sporns and Koetter, 2004), the statistical mechanics of networks (Albert and Barabasi, 2002) and computer simulations (Sporns et al, 2000) ascertained that the particular form of clustered organization encountered in biological (natural) cortical networks would indeed be supportive of synchronous processing and efficient signal exchange, and superior to other network topologies (for review: Sporns, 2004, and references cited therein).

2.3: Global workspace models

2.3.1: Baars’ Global workspace model

Partly in distinction from and partly overlapping with extant psychological theories of conscious experience, Baars (1988) introduced and later extended (Baars,1997) a “Cognitive theory of Consciousness”. The theory postulates a multitude of relatively small special-purpose processes, almost always unconscious. Coalitions of such processes gain access to a limited capacity global workspace which serves to integrate competing and cooperating input networks and is the gateway to consciousness. Messages from this global workspace are broadcast to all the unconscious processors to recruit their participation for dealing with novel situations or solving the current problem. All this occurs under the auspices of certain contexts, each being itself a coalition of processes, jointly constraining conscious events.
In the original book of 1988 and in numerous subsequent publications, Baars illustrated that the three main construct of the theory (namely Global workspace, unconscious special processors and contexts) and their interaction pattern can be construed to be sufficient for most (possibly all) aspects of Cognition and Consciousness. His original approach was informal, descriptive and primarily appealing to psychological intuition. Nonetheless, in a review published in 2002, Baars could cite a large number of investigators implicating one form or another of a ‘global workspace’ in cognition and consciousness, however not necessarily embracing Baars’ tripartite model. The rallying point appears to be the notion that some form of “global workspace” is required for widespread interaction between otherwise independent brain functions: the idea of a ‘global workspace’ had entered the discourse of consciousness studies, albeit in a variety of shades of meaning.

In subsequent publications, functional interpretations in neurophysiological terms were introduced by Baars and associates in a ‘neural global workspace’ (Newman and Baars 1993), primarily in relation to attention in schemas of neural network models (Newman et al. 1997), and similar to those which Taylor and Alavi (1993) had applied earlier. Franklin and Graesser (1999) and Franklin (2003) took a different approach: they implemented the global workspace concept in several successive generations of software agent models in the tradition of Artificial Intelligence. None of these efforts addresses the dynamics of processing in the global workspace and are, therefore, outside the scope of this essay. Although specifically designed this way, it now appears that organization and function of DCH are compatible with and accord satisfactorily with the basic tenets of the Global Workspace Model (Seth and Baars, 2005).

2.3.2: Wallace’s dynamic global workspace hypothesis

In a series of recent publications Wallace (2005a) supplied a dynamics to a Global Workspace. Here I will offer an intuitive account of the formally developed perspective taken in Wallace (2005 b) which applies information theoretic considerations and draws on resources from theories of network structure and dynamics (Erdos and Renyi, 1960; Newman, 2003).

To begin with, a loosely connected network of unconscious modules is considered, corresponding essentially to the unconscious processors of Baars’ theory. Next, the relation among them is characterized in terms of their mutual information. Non-zero mutual information among modules can formally be considered as linkages among them which constitute fleetingly a (functional) Global Network. Network theory now specifies conditions under which sub-networks can become transiently formed. One of them can encompass the majority of the network’s nodes, forming in network jargon a ‘Giant Component’. In terms of network dynamics, this can be considered a phase transition. The formation of the ‘Giant Component’ is considered to reflect a collective cognitive phenomenon. One can readily imagine the complex dynamics of this model: a continuous flux of mutual information among the special purpose processors, at each point giving rise to a fleeting network which, at certain critical points, coalesces momentarily to a Giant Component, only to be restructured by new states of the information processors.

This sketch of the model is merely intended to convey an intuitive “feel” of the model’s dynamics and fails to do justice to its formal sophistication and to the numerous details in its successful application to, for instance, Generalized inattentional blindness and mental disorders Wallace (Wallace, 2005 c; 2006)

2.3.3: The Global Neuronal Workspace hypothesis (GNWH)

The Global Neuronal Workspace hypothesis was described in details by Changeux & Dehaene (1989) and Dehaene & Naccache (2001) and most recently summarized by Dehaene & Changeux (2004). It relates to Baar’s cognitive theory of consciousness, but is distinguished from it by its particular and detailed functional organization, and its implementations in dynamic computational models.
The hypothesis postulates two computational spaces of distinct patterns of connectivity: 1) a collection of subcortical, automatic processors, each specialized for a particular signal input which is provided via encapsulated local and medium-length connections; and 2) a global neuronal workspace with the capacity for wide-spread, long-range connections for reentrant signal flow between it and the specialized processors. The workspace is a dynamic concept: workspace neurons are not sharply delineated anatomically, but distributed among distant association areas (Dehaene and Changeux, 1997, 2005; Dehaene et al. 1998). The decisive event is the activation of GNW (see Section 3.3). Between episodes of activation, the neurons of GNW are in a state of permanent spontaneous activity which is sustained by ascending neuromodulatory input. When of sufficient intensity, the network will display gamma oscillations of thalamo-cortical origin and possibly sudden surges of activation which may be identified with ‘vigilance’ (Llinas et al, 1998): they enhance the activation of GNW by sensory stimuli. Activation of GNW also occurs with intense sensory stimulation in the absence of facilitation. GNWH postulates that global activation of a GNW is associated with reportability of a subjective experience (Dehaene and Changeux, 2004).

The basic design of GNW was implemented in several Neural Network models, with McCulloch Pitts as computational elements, and proved satisfactory for emulating aspects of human performance in a variety of (effortful) psychological-behavioral tasks (Dehaene et al. 1987,1998). To convey the operational flavor of the theory, I will describe here briefly the most recent model by Dehaene et al. (2003) of a network of single compartment model neurons with explicitly specified ionic conductances and synaptic currents for simulating features of the cortical inter- and intra-columnar connectivity and as cortico-cortical projections. The target of the model was a modified attentional blink paradigm for which conditions for reportability of presence or absence of stimuli were determined in human trials: subjects saw serial visual presentation of distractors, interspersed with two targets T1 and T2; the task was to rate T2 visibility and then to report T1 identity. Typically, reportability of T2 drops at for several hundred msec after T1 presentation.

For the simulation of the human task, the model was placed in a regime of spontaneous thalamo-cortical oscillations. The attentional blink test was simulated by stimulating two groups of thalamic neurons, one coding for T1, the other for T2. The index of model performance was the degree and extent of activity across the cortico-thalamic hierarchy. As to be expected, network activation evoked by T1 stimulation set a long-lasting dynamic brain-state in motion. But the activation elicited by T2 stimulation dependent tightly on its timing: T1 elicited activity prevented T2 activation from propagating to higher cortical levels and abolished part of the top-down amplification in reentrant circuitry, with the global network seemingly acting as a bottleneck (Sigman and Dehaene, 2005). Selective lesions of the long distance connections in the model corrupt the model performance.

3: Discussion

To varying degrees of explicitness, the observations and findings of the foregoing section implied principles of nonlinear dynamics, notably the notion of systems evolving in a phase space defined by their independent variables, and capable of undergoing abrupt transitions. The observed brain events cited in reference to Coordination Dynamics as well as the time scale for constituting the DC and The GNW in the respective models occupy durations in the order of several hundred msec. While this may be surprising, recall that it takes some 500 msec in human subjects for a stimulus to the somatic sensory cortex to become experienced, referenced to the short latency component of the Cortical Evoked Potential (Libet, 2004). Pollen (2004, 2006) attributes this delay to the time required for completion of recursive activity in feed-forward and feedback neuronal circuits. With respect to the time course, the microgenetic approach in perception offers a suggestive parallel: it is based on the idea of a ‘gradual maturing’ of the subjective perceptual experience in the sequence of different components of the cortical evoked potentials, ranging from the fast positive component at 20-60 msec to the negative component at 100-150 msec and the positive deflection at 300 msec (Bakhman, 2000). Plausibly, the time periods under discussion reflect the time required for phase space evolution of
neuronal assemblies, in preparation for phase transitions. If there is any merit to this suggestion, it would have significant consequences for the experimenter: rather than seeking a correspondence between subjective report or behavior with a punctate (in space and time) measurement in brain imaging, genuinely informative data on brain processes would be obtained from following trajectories of phase space evolution in brain space and time (Werner, 2004, 2005).

In this and related situations one must ask the question whether a given system is intrinsically dynamical, or whether it is merely the case that it can be interpreted and simulated as such. Since some of the evidence in support of the former (see below) is framed in terms of Self-Organized Criticality (SOC), I will insert here a brief excursus to some of the essential features of the theory.

As is well known, Bak et al (1987/1988) introduced a theory of SOC to designate the property of systems to exhibit non-equilibrium phase transitions on account of their intrinsic dynamics, without requiring tuning of control parameters by external influences. This was thought to be the distinguishing criterion from the conventional phase transitions in equilibrium systems which require external tuning of control parameters to attain critical state. Systems of this former kind evolve spontaneously to a critical state at which their responses to perturbations display a set of characteristic properties: temporal and spatial scale invariance (i.e. absence of a characteristic scale of length and time, associated with fractals and 1/f noise), drastic reduction of the number of degrees of freedom, and divergence of correlation function as signal for lack of characteristic length. ‘Scaling behavior’ refers to determining whether the temporal (or spatial) pattern of an observable remains identical under scale transformation; ‘scale-free’ then signifies the absence of any characteristic scale. Processes based on SOC are characterized by a power law relation between frequency bands and their respective frequency in the record, usually represented as 1/f relation. This is generally taken as a signature of SOC. (Bak, 1996; for an extensive review on scale invariance in Biology, see Gisinger, 2001). Flyvbjerg (1996) offered a minimal definition of the essential features of SOC as “a self-organizing critical system is a driven, dissipative system consisting if of (1) a medium which has (2) disturbances propagating through it, causing (3) a modification of the medium such that eventually (4) the medium is in a critical state, and (5) the medium is modified no more”. This definition reflects the process of propagation of long-range interactions based on local effects in the medium (as a kind of domino effect) until the state of criticality is attained at which any further disturbance triggers an abrupt, critical phase transitions. This sequence of events is sustained by two concurrent processes with different time constants: a faster disturbance of the dynamic stationary state, and a slower relaxation towards its restoration, often referred to as avalanche (a metaphor based on the sand pile of the original model of SOC). This critical state is then maintained until replaced by circumstances that lead to initiation of another process of the same kind.

With the foregoing criteria for self-organized criticality in mind, it is now possible to examine whether measurements of brain activity and structure comply with the stipulations of the theory. Linkenkaer-Hansen (2001) unequivocally established that the amplitude fluctuations in the 10-20 Hz frequency range obey power law scaling behavior in humans. Scale-free neocortical dynamics was also ascertained by Freeman (2005) in the electroencephalogram of rabbits; a computer model also suggested that neodortex is stabilized in a scale free state of self-organized criticality. Quantitative fMRI analysis of functional connectivity (Eguiluz et al, 2005, Chialvo, 2004) and EEG analysis of functional connectivity (Fingelkurts and Fingelkurts, 2006) supply additional evidence. Sporns et al (2004) reviewed recently the numerous literature sources which identify brain neural networks as ‘scale free’. Finally, Beggs & Plenz (2003, 2004) reported critical behavior in slices of cortical tissue.

It is now possible to return to the issue raised earlier whether there is adequate justification to consider the brain as a complex dynamic system, with an affirmative answer. From the evidence listed, it is in fact possible to characterize the brain as a whole as being in a state of self-organized criticality.

3.1: On Coordination dynamics and metastability
The foundational notions of Coordination Dynamics were initially inspired by the principles of Synergetics (Haken, 2002), and subsequently adapted to application in various experimental situations (Kelso, 1959; Bressler & Kelso, 2001; Kelso & Engstrom, 2006). The concepts of self-organizing Pattern Formation and Pattern Dynamics are the essential components of this approach. The former refers to the spontaneous aggregation of multiple (possibly heterogeneous) elements by non-linear interactions to functional units with reduced degrees of freedom, whose behavior can be characterized by merely a few relevant coordination variables (order parameters). At certain critical values, control parameters trigger qualitative changes of behavior of the coordinated unit as a whole. For illustration, recall that in the example cited in Section 2.1, the tone frequency is the control parameter which, at a critical frequency, induced the transitions at the level of behavior and brain activity from one stable state to another. Pattern dynamics designates the process of evolution of the coordination variables (in the example cited above: the motor events underlying the finger tapping frequency and the relevant neural events in the brain) according to their respective dynamical laws. Following the principles of Dynamical System Theory, the relationship between control parameter and values of the coordination variables can be represented as hyperplane in phase space to delineate regions of stability and zones of transition (bifurcation) between them.

At a neural-mechanistic level of analysis, Coordination Dynamics considers the oscillatory phase relations among distributed brain regions as the prerequisite for a dynamic process of self-assembly (binding) to coherent networks. The rapid creation and dissolution of neural assemblies across distributed brain areas is considered to constitute the non-equilibrium phase transitions which are observable (for instance, as neuromagnetic field) and determine the hyperplane of phase space. An intrinsic feature of this view is the already mentioned duality between integration and separation of system components which is also as measure of complexity in the framework of the Neuronal Group Selection Theory (Tononi et al, 1994, Sporns, 2004).

In the earlier section of this essay, metastability descriptively designated shifts and transitions between stable states. It can now be given an operational meaning: rather than viewing integration-separation as polar opposites, a continuous scale of tendencies for neural groups to coordinate and segregate seems more appropriate. At the level of neuronal dynamics, this continuum appears to be supported by the flexibility of coupling coefficients among diverse neuron groups (Friston, 1997, 2000): that is, in the form of a continuum of dynamically shifting, discrete configurations of neuron networks (for review: see Fingelkurts and Fingelkurts, 2006). The expression of this is the meandering of the system’s representation in phase space, dictated by its intrinsic dynamics and external influences. Phase transitions with epochs of stability will then occur along the way, in accord with the shape of the system’s phase space.

### 3.2 Dynamics in the Dynamic Core and the Global Neuronal Workspace Hypotheses

Here I direct attention to the nature of the reentrant activity for the transient formation of the neural functional complex described as “Reentrant Dynamic Core” (DC) on the one hand, and “Global Neural Workspace” (GNW) on the other. Although differing in many respects, both models attribute an essential role to the reentrant neuronal activity in circuits connecting cortical with other cortical, thalamic and peripheral processor regions. Dehaene et al (1998) speak of “distributed neurons with long distance connectivity that provide a ‘global workspace’ that can potentially interconnect multiple distributed and specialized brain areas in a coordinated though variable manner”, and Dehaene & Changeux (2003) refer to “self-amplifying recurrent activity”. Edelman (2003) speaks of “dynamic reentrant interactions across cortical circuits, that allow synchronous linking and binding to take place among widely distributed brain areas”, and considers reentry “a unique feature of higher brains” (Tononi and Edelman, 2000). Both groups of investigators emphasize the importance of this pattern of connectivity for generation of oscillatory activity.
From their respective publications (see Section 2.2 and 2.3.3), it appears that DC and GNW have a somewhat similar temporal pattern of evolving over a few hundred msec, and persisting for several hundred msec, prior to dissolving. Dehaene & Changeux (2004, 2005) who are more explicit about this than the Edelman group describe this temporal course as sudden onset of coherent synchronized neuron activity in multiple distant cortical areas and peripheral processors, which is sustained for several hundred msec by reentrant thalamocortical signal flow. They refer to ‘phase transition in a metastable dynamic’ and use also the apt expression of “ignition” to convey the abruptness of the transition. It is of some historic interest that neuronal modelers of the 1950s and 1960s were interested in the conditions under which their neuronal models would ‘ignite’ between stable states (e.g.; Rapoport, 1952; Griffith, 1963). To assist with gaining an intuitive grasp of the complex dynamics attributed to their Dynamic Core, Tononi & Edelman (2000) offer a helpful model: envision a large cluster of tense springs, variously connected to each other and surrounded by another set springs, loosely coupled to the former cluster; it is then easy to see that even a small perturbation will spread rapidly and effectively throughout a system of this kind.

The suggestion of Dehaene & Changeux to view the natural history of the formation and dissolution of the neural complex that arises transiently in their neuronal models, and to which they attribute a role for cognitive events, invites an exploration of its dynamic origin. Descriptively, the type of activity they observe resembles the avalanches of neural activity noted by Beggs & Plenz (2003, 2004) in slices of cortical tissue, reviewed by Vogels et al (2005). Taking into account that this activity (and the presumed counterpart in the models of the Edelman group) occurs in nonlinear systems far from equilibrium directs attention to the principle of Self-organized Criticality (SOC) of which avalanche formation is one of its signatures.

In the nearly 20 years since introduction of SOC, critical examination of the claims of Bak et al for universality of SOC have introduced some qualifications in the original theory, and circumscribed the range of its validity (Dickman, 2000; Kadanoff et al., 1989; Jensen, 1998; Grinstein, 1995). The conceptual prototype of SOC was originally the ‘sand pile model’ in which stepwise addition of sand grains on the top leads in the critical state to propagation of avalanches across the pile, which exhibit the properties of scale invariance. Numerous modifications of the original paradigm were instrumental to characterize the boundary conditions under which the theory of SOC applies, while, on the other hand, the signatures of SOC were identified in models not originally considered, such as for instance percolation models (Stauffer & Aharony, 1991/1994, Grimmett, 1989). Parenthetically, it is worth noting that one of the extensions of SOC, designed to replicate the scale invariance of earthquakes (Olami, et al., 1992), shares many features with Tononi and Edelman’s (2000) spring model of reentrant activity, referred to earlier.

After the dust stirred by Bak’s original claim settled, it is now firmly established by the work of Bak’s own Group (Paszuski et al, 1996) and many others that SOC does not exclusively refer to spontaneous, parameter free criticality, nor exclusively to self-tuning, but nevertheless remains a useful concept for describing systems far from equilibrium that will manifest a phase transition when driven from the outside, (Grinstein, 1995; Dickman at el, 2000; Frigg, 2003). Like conventional phase transitions, some forms of SOC are amenable to analysis by Renormalization Analysis (Pietronero et al., 1994; Vespignani et al, 1996), that is: the computational technique that enables the explicit computation of the critical exponents for scale invariance and other critical properties (Kadanoff et al., 1967; Wilson, 1979; McComb, 2004), thus blurring what was thought to be the radical dichotomy between SOC and ‘classical’ (tuning-dependent) phase transitions.

The twin concepts of scaling and universality play an important role in description of dynamical systems for elimination of degrees of freedom and scale transformations at points near critical transition.
(Kadanoff et al, 1989; Kadanoff, 1990). The significance of this lies in the possibility of identifying universality classes (Odor, 2004) to which I will return in Section 4. Although still lacking a comprehensive theory of SOC, it is now firmly rooted in Dynamical Systems Theory by characterizing (specifically in some instances and in others, in principle) the critical state as the system’s attractor, and its fractal structure (Blanchard et al, 2000).

What evidence can be adduced that SOC may indeed be a valid and useful notion for neural systems? In the first place, there is abundant evidence that neural network models can converge to critical states. Bienenstock & Lehman (1998) attribute this property to the Hebbian covariance plasticity rule for learning at the synaptic level. Kenridge (1993) reports critical behavior in random neural nets at appropriate levels of connectivity and under the influence of low-intensity driving. Usher et al (1995) generate metastable behavior, pattern formation and critical behavior in networks of Integrate-and-Fire neurons. Bornhold & Rohlf (2003) extend critical self-organization to the evolution of network topology. Critical network models were also studied by Wakeling (2003) and da Silva (1998). Referring to the previously cited studies of Beggs & Plens (2003), Haldeman and Beggs (2005) contribute the additional important observation that branching network models with recurrent connectivity can account for power law relations at critical points, and display metastable states at branching parameters, intermediate between sub- and supercritical values. Regionally localized avalanche behavior occurs in the self-organizing maps of Zhao and Chen (2003). There is, thus considerable evidence that neural networks are apt to exhibit SOC in a variety of forms and circumstances.

In their totality, concepts and observations sketched in this section are intended to give credence to the notion that the transient configurations of neural activity (designated respectively Reentrant Dynamic Core by Edelman et al, and Global Neuronal Workspace by Dehaene & Chanheux) are manifestation of SOC in the neuronal reentry circuits of the respective models. As such, they require several hundred msec for constitution of their long-range connections to full criticality at which point the characteristic properties of scale invariance, reduced dimensionality and long-range correlations come to obtain for critical state’s duration. While still on the way to criticality, a metastable regime is in effect. The next section will address some implications of this view.

4. Discussion and Conclusions

The framework of Operational Architectonics of brain function, described in several publications by Fingelkurts and Fingelkurts (2001,2004,2005,2006, Fingelkurts et al, 2004) directs attention to relations between the activity of functional neuronal populations on a mesoscale, and phasic epochs/segments in the scalp EEG on the millisecond time scale. The latter are associated with rapid jumps in the EEG amplitude (Kaplan et al, 1997; Fingelkurts and Fingelkurts, 2001). The rapid transition processes occurring in the amplitude of continuous EEG activity mark the boundaries between quasi-stationary segments for this activity. It is assumed that each homogenous segment within a particular EEG frequency band corresponds to a temporary stable microstate in the brain’s activity, i.e. an ‘operation’ in the terminology of Fingelkurts and Fingelkurts (2005). The transition from one segment to another is thought to reflect the moment of switching from one neuronal network to another. Moreover, the synchronization of these segments (i.e. EEG structural synchrony) between different EEG channels would then indicate the synchronization of different brain operations: this is the ‘operational synchrony’ phenomenon of Fingelkurst and Fingelkurts, 2001). This process results in transient metastable states of EEG activity in the form of so-called ‘operational modules’. In a different series of studies, some of which were already mentioned in connection with metastability in section 2.1, Lehman and associates (1998, 2006) identified in the scalp EEG the momentary whole brain intervals of quasistability (microstates), separated by sudden transitions in the vector of the maximal EEG potential. Here I suggest that the EEG manifestations in the work of Fingelkurts and Fingelkurts, and those in the work of the Lehman group may be abrupt manifestation of
neuronal events in the models of the Edelman group on the one hand, and Dehaene & Changeux on the other. If this conjecture has some validity, one might then infer that the rapid EEG manifestations referred to are the expression of phase transitions in SOC at the mesoscopic level.

On a different score, I now turn to the notion of Universality classes, mentioned in passing in Section 3.2. Permitting oneself some levity, Universality classes may be viewed as God’s gift to the physicist! Universality refers in this context to the phenomenon whereby dissimilar systems can exhibit the same characteristic exponents that govern scaling and correlation functions (Binney, 1992). These are the numerical indices that reflect the creation of long range correlations from local interactions, and the manner in which disturbances propagate through the system. These indices are independent of physical nature of the system’s components, and are solely determined by the properties of the components’ interactions. It is empirically established that nonlinear dynamics systems, including those operating far from equilibrium (Odor, 2004), can often be categorized by these critical indices into distinct classes. This means that having ascertained one or the other critical property for a system under study, it is then possible to predict all other critical properties of that system merely on the basis of its class membership. Applying this approach to the reentry circuitry in the neuronal models of DC and GNW would enable characterizing the nature of their dynamics, and its relation to the potential role in the Operational Architectonis of Fingelkurts and Fingelkurts (2005). Among the various candidates that come to mind is the type of percolation studied by Kozma et al (2005) as model of the neuropil, or one of its several variants.

I can now return to the introduction of this essay: there I set out to explore evidence and conjectures pertaining to the style of the brain’s function. This led to the suggestive evidence for metastability and phase transitions in self-organized criticality in transiently formed neuronal assemblies constituted by reentrant circuitry. I then pointed to the possible relation of the phase transitions in these mesoscopic neuronal configurations to transient manifestations in the EEG, and adumbrated the status of metastability in the context of SOC. Finally, I suggested that computational models of Dynamic Core and the Global Neuronal Workspace can serve as windows for gaining insight into the dynamics of neuronal assemblies with established functions in their respective models. This, so it is thought, would substitute, at least at this time, for the direct observation of neural assemblies in situ, as it would reveal indirectly the total range of their dynamic properties, at least in part on the basis of universality class membership.

References.


Freeman, W.J., 2005. A field-theoretic approach to understanding scale-free neocortical dynamics Biol.Cybern. 92, 350-359

Freeman, W.J., Holmes, M.D., 2005. Metastability, instability, and state transitions in neocortex. Neural Networks 18, 497-504.


Acknowledgement: Drs. Andrew and Alexander Fingelkurst kindly offered clarifying comments on an earlier version of the manuscript, which are here incorporated with their permission.