The brain near the edge

Dante R. Chialvo

Department of Physiology, Feinberg Medical School, Northwestern University, 303 East Chicago Ave. Chicago, IL 60611, USA

Abstract. When viewed at a certain coarse grain, the brain seems a relatively small dynamical system composed by a few dozen interacting areas, performing a number of stereotypical behaviors. It is known that, even relatively small dynamical systems can reliably generate robust and flexible behavior if they are located near a second order phase transition, because of the abundance of metastable states at the critical point. The approach pursued here assumes that some of the most fundamental properties of the functioning brain are possible because it is spontaneously located at the border of such instability. In this notes we review the motivation, the arguments and recent results as well as the implications of this view of the functioning brain.

Keywords: Brain, critical phenomena, complex networks

PACS: 87.19.La, 89.75.-k, 89.75.Da

INTRODUCTION

Each year, there are several hundredths of fascinating discoveries of isolated aspects of brain physiology. At the same time, only a handful or reports discusses the reverse process: how the knowledge of isolated pieces can be integrated to explain how the brain works. This is, of course, a well known, hard to tackle challenge which, however, is particularly suitable to a physicist because of its inherent familiarity with ideas of universality and unification. For a newcomer the first concern would be if it is possible to approach the problem of brain function without inventing a new theoretical framework. In other words, is it possible to gain any insight about relevant brain problems by deliberately ignoring-at least for the moment- the soft aspects of brain’s condensed matter?

The approach pursued here assumes that the most fundamental properties of the functioning brain are possible because it is spontaneously located at the border of an instability. Indeed the proposal is that these fascinating properties have no extra cost as they are generic for this state. From this viewpoint, all human behaviors, including thoughts, undirected or goal oriented actions, or simply any state of mind, are the outcome of a dynamical system -the brain- at or near a critical state. The main point is that, as in thermodynamics systems at the critical point, it is only at this state that the largest behavioral repertoire can be attained by the smallest number of degrees of freedom. Behavioral repertoire means the set of actions useful for the survival of the brain and degrees of freedom means the number of (loosely defined) specialized brain areas engaged in generating such actions. By looking at the problem from this angle a number of ideas and results from statistical physics can be used to guide work towards the ultimate goal of understanding how the brain works, without inventing anything new.

This article is dedicated to discuss the basis and the implications of this proposition.
The paper is organized as follows. The second section introduces the main motivation, which is routed in concepts borrowed from complex systems. For completeness we also summarize here some obvious connections with very well known facts from statistical physics. The third section discuss the specific rationale and the next section enumerates evidence that seems to support the usefulness of this approach to brain function. The paper closes with a short discussion of implications.

**FUNDAMENTAL LAWS FOR THE COLLECTIVE**

It is well known that almost all interesting macroscopic phenomena in nature, from gravity to photosynthesis, from superconductivity to muscle contraction are product of an underlying collective phenomena. In this sense, science is the never ending process of explaining macroscopic phenomena observed at one level from fundamental laws uncovered at another level. Neuroscience not being an exception, must explain human behavior, i.e., what we see in terms of the underlying collective which is partially hidden to us. If both phenomena and explanation remains at the same level then nothing is different from the seventeen century understanding of what constituted conscious experience (Fig. 1). The main difficulty, and the concern of this proposal, is that there no fundamental laws yet for the collective of neurons!

However, there are some relevant facts which could be source of inspiration. The brain have, as a collective, some notoriously conflictive demands. On one side it need to be "integrated" while must be able to stay "segregated", as discussed extensively by Tononi and colleagues [27, 28]. This is a non trivial constraint, nevertheless mastered by the brain as it is illustrated with plenty of neurobiological phenomenology. Suffice to think in any conscious experience to immediately realize that always comprises a single undecomposable scene [27], i.e., an integrated state. This integration is such that once a cognitive event is committed, there is a refractory period (of about 150 msec.) in which nothing else can be though of. At the same time the large number of conscious states that can be accessed over a short time interval exemplify very well the segregation property. As an analogy, the integration property we are referring to could be also interpreted as the capacity to act (and react) on an all-or-nothing mode, similar to an action potential or a travelling wave in a excitable system. The segregation property could be then visualized as the capacity to evoke equal or different all-or-nothing events using different elements of the system. This could be more than a metaphor.

While the study of this problem is getting increasing attention, the mechanisms by which this remarkable scenario can exist in the realms of brain physiology is not being discussed as much as it should. Our approach is to look at the integration-segregation dilemma as a generic property of dynamical systems at the critical point of a phase transition. It is our suggestion that at the critical point these and others properties - equally crucial for brain function- appear naturally. If the idea is correct, statistical physics could help to move the current debate from phenomenology to understanding of the lower level brain mechanisms of cognition.
What is special about being critical?

To visualize the potentially useful connections between brain function and thermodynamical systems at a phase transition it is helpful to recall the ferromagnetic-paramagnetic phase transition illustrated in Fig. 2. A material is ferromagnetic if it displays a spontaneous magnetization in absence of any external magnetic field. If we heat up an iron magnet the magnetization gets smaller and finally reaches zero. At low temperature the system is very ordered with only very large domains of equally oriented spins, a state that is practically invariant in time. On the other extreme, at very high temperatures, spins orientation changes constantly, they are correlated at only very short distances and as consequently the mean magnetization vanishes. In between these two homogeneous states, at the critical temperature, the system exhibits very peculiar fluctuations both in time and space. For example, the magnetization temporal fluctuations are known to be scale invariant. Similarly, the spatial distribution of spins clusters show long range (power law) correlations. At the critical point, these large dynamic structures emerge, even though there are only short-range interactions between the systems elements. Thus, at the critical temperature, the system exhibits a greatly correlated (up to the size of the system) state which at the same time is able to wildly fluctuate in time at
FIGURE 2. Ferromagnetic-paramagnetic phase transition. Bottom: Temperature dependence of magnetization $m(T)$ for Fe. Top three panels are snapshots of the spins configuration at one moment in time for three temperatures: subcritical, critical and supercritical from numerical simulations of the Ising model ($d=2$).

all scales. We propose that this dynamical scenario -generic for any second order phase transition- is strikingly similar to the integrated-segregated dilemma discussed above and shown to be relevant for the brain to operate as a conscious device. It is important to note that there is no other conceivable dynamical scenario or robust attractor known to exhibit these two properties simultaneously. Of course, any system could trivially achieve integration and long range correlations in space by increasing link’s strength among faraway sites, but these strong bonds would prevent any segregated state.

By considering the brain embedded in the rest of nature, one adopts the Darwinian view that the brains we see today are the ones that -for whatever means- got an edge and survived. Then we could ask how consistent is our view of the brain near a critical point with these Darwinian constraints. We propose that the brains we see today are critical because the world in which they have to survive is up to some degree critical as well. If the world were sub-critical then everything would be simple and uniform (as in the left panel of Fig. 2) there would be nothing to learn, a brain will be superfluous. In a supercritical world, everything would be changing all the time (as in the right panel of Fig. 2) it would be impossible to learn. Then we have to conclude that the brain is only necessary to navigate in a complex, critical world. In other words we need a brain because the world is critical [2, 3, 4, 9, 18]. Furthermore, a brain not only have
to remember, but also to forget and adapt. In a sub-critical brain memories would be frozen. In a supercritical brain, patterns change all the time so no long term memory would be possible. To be highly susceptible, the brain itself has to be in the in-between critical state.

These ideas are not new at all, indeed almost the same intuition prompted Turing half a century ago to speculate about learning machines using similar terms:

Let us return for a moment to Lady Lovelace’s objection, which stated that the machine can only do what we tell it to do. One could say that a man can “inject” an idea into the machine, and that it will respond to a certain extent and then drop into quiescence, like a piano string struck by a hammer. Another simile would be an atomic pile of less than critical size: an injected idea is to correspond to a neutron entering the pile from without. Each such neutron will cause a certain disturbance which eventually dies away. If, however, the size of the pile is sufficiently increased, the disturbance caused by such an incoming neutron will very likely go on and on increasing until the whole pile is destroyed. Is there a corresponding phenomenon for minds, and is there one for machines? There does seem to be one for the human mind. The majority of them seem to be "subcritical," i.e., to correspond in this analogy to piles of subcritical size. An idea presented to such a mind will on average give rise to less than one idea in reply. A smallish proportion are supercritical. An idea presented to such a mind that may give rise to a whole "theory" consisting of secondary, tertiary and more remote ideas. Animals minds seem to be very definitely subcritical. Adhering to this analogy we ask, "Can a machine be made to be supercritical?"

How things stand today compared with Turing’s days? Very different, because of two important aspects, the first one concerns with all the advances in monitoring brain signals at different resolution and the second concerning the possibility to be guided by the last two decades of results in critical phenomena.

What one should be able to observe?

A number of features, known to be exhibited by thermodynamic systems at the critical point, should be immediately observed in brain experiments, including:

1. At large scale:
   Cortical long range correlations in space and time.
   Large scale anti-correlated cortical states.

2. At smaller scale:
   "Neuronal avalanches", as the normal homeostatic state for most neocortical circuits.
   "Cortical-quakes" continuously shaping the large scale synaptic landscape providing "stability" to the cortex.

3. At behavioral level:
   All adaptive behavior should be "bursty" and apparently unstable, always at the
"edge of failing". Life-long learning should be critical due to the effect of continuously "rising the bar".

In addition one should be able to demonstrate that a brain behaving in a critical world performs optimally at some critical point, thus confirming the intuition that the problem can be better understood considering the environment in which brains evolved.

In the list above, the first item concerns the most elemental facts about critical phenomena: despite the well known short range connectivity of the cortical columns, long range structures appear and disappear continuously. The presence of inhibition as well as excitation together with elementary stability constraints determine that cortical dynamics should exhibits large scale anti-correlated structures as well [15]. The features at smaller scales could have been anticipated from theoretical considerations as well, but avalanches were first observed empirically in cortical cultures and slices by Plenz and colleagues [6]. An important point that is left to understand is how these quakes of activity shape the neuronal synaptic profile during development. At the next level this proposal suggests that human (and animal [7]) behavior itself should show indications of criticality and learning also should be included. For example when teaching any skill one chooses increasing challenge levels which are easy enough to engage the pupils but difficult enough not to bores them. This "rising the bar" effect continues through life, pushing the learner continuously to the edge of failure! It would be interesting to measure some order parameter for sport performance to see if shows some of these features for the most efficient teaching strategies.

**RECENT RESULTS**

**Functional brain networks are complex**

Functional magnetic resonance imaging (fMRI) allows to monitor non invasively spatio-temporal brain activity under various cognitive conditions. Recent work using this imaging technique demonstrated complex functional networks of correlated dynamics responding to the traffic between regions, during behavior or even at rest (see methods in [14]). The data is analyzed in the context of the current understanding of complex networks (for a review see [22]). During any given task the networks are constructed first by calculating linear correlations between the time series of brain activity in each of 36 x 64 x 64 brain sites. After that, links are said to exist between those brain sites whose temporal evolutions are correlated beyond a pre-established value $r_c$.

Fig. 2, show a typical brain functional network extracted with this technique. The top panel illustrates the interconnected network’s nodes and the bottom panel shows the statistics of the number of links (i.e., the degree) per node. There is a few very well connected nodes in one extreme and a great number of nodes with a single connection. The typical degree distribution approaches a power law distribution with an exponent around 2. Other measures revealed that the number of links as a function of -physical- distance between brain sites also decays as a power law, something already confirmed by others [21] using different techniques. Two statistical properties of these networks,
path length and clustering were computed as well. The path length ($L$) between two voxels is the minimum number of links necessary to connect both voxels. Clustering ($C$) is the fraction of connections between the topological neighbors of a voxel with respect to the maximum possible. Measurements of $L$ and $C$ were also made in a randomized version of the brain network. $L$ remained relatively constant in both cases while $C$ in the random case resulted much smaller, implying that brain networks are "small world" nets, a property with several implications in terms of cortical connectivity, as discussed further in [24, 22]. In summary, the work in [14] showed that functional brain networks exhibit highly inhomogeneous scale free functional connectivity with small world properties. Although these results admit a few other interpretations, the long range correlations demonstrated in these experiments are consistent with the picture of the brain operating near a critical point. Of course, further experiments are needed to specifically define and measure some order parameter to clarify the precise nature of these correlations. Furthermore, as more detailed knowledge of the properties of these networks is achieved, the need to integrate this data in a cohesive picture grows as discussed recently by Sporns and colleagues [23].
Cortical networks exhibit neuronal avalanches

Recent experiments from Plenz and colleagues [6] were the first to demonstrate a new type of small scale cortical activity. They showed that under some experimental conditions, the cortex exhibits what they termed neuronal avalanches. This type of population activity seats half way in between two well known patterns: the oscillatory or wave-like highly coherent activity on one side and the asynchronic and uncoherent modality on the other. In each avalanche neuronal activity have a very large probability to engage few neurons and die, and a very low probability to spread and activate the whole system. In very elegant experiments Plenz and colleagues estimated a number of properties indicative of critical behavior including a power law with an exponent $\sim 3/2$ for the density of avalanche sizes (see Fig. 4). This agrees exactly with the theoretical expectation for a critical branching process [31]. Further experiments in other experimental settings, including monkey and rats in vivo recordings, have already confirmed and expanded these initial estimations [19, 25]. An unsolved problem here is to elucidate the precise neuronal mechanisms leading to this behavior. Avalanches of activity such as the one observed by Plenz and colleagues could be the reflection of completely different scenarios. It could reflects a structural (i.e., anatomical) substrate over which travelling waves in the peculiar form of avalanches occur. This will imply that the long range correlations detected are trivially due to long range connections. If
FIGURE 5. McCulloch experiments inducing local seizures by instillation of strychnine. Panel A and B are from chimpanzee experiments [5]. Panel A shows a summary of the sites where the strychnine was applied (filled circle) and the sites of the cortex fired by the topical application. Besides the local ones, long range activations crossing the entire cortex were often observed. Panel B illustrates the adjacency matrix summarizing which areas -on the average- were activated by the strychnine application. Panel C shows similar results obtained by McCulloch and colleagues in Macaca Mulata [17] mapping the entire cortex and basal ganglia. Panel D depicts (note the double logarithmic axis) the edge length density distribution computed from McCulloch’s drawing in Panel A. The dashed line with slope 2 illustrates, for comparison, the average edge-length density found in recent fMRI experiments [14].

that is the case, as was discussed above, this have nothing to do with criticality, and furthermore it will imply that segregation will be impossible. The second possibility is that avalanches occurs over a population of locally connected neurons. Their ongoing collective history will permanently keep them near the border of avalanching and each avalanche will only excite enough neurons to dissipate the excess of activity. Although this is the most likely scenario, which follows the ideas and results put forward by Bak and colleagues [2, 3, 4, 9, 18], there is much theoretical work awaiting to formalize these results.
McCulloch already saw it in 1940

Dusser de Barenne, Warren McCulloch and colleagues [17, 13], more than sixty years ago, experimented inducing local seizures by instilling drops of strychnine in several regions of the monkey cortex while recording cortical electrical activity simultaneously in twenty sites across the entire cortex. This clever technique, mastered by Dusser de Barenne, received the name of strychnine neuronography. In a certain way, these experiments could be considered the earliest attempt to study brain functional connectivity, by inducing some liminal activity in a given area and recording the co-active cortical sites. Typically, they noticed that the initial activity induced by the strychnine remained local, and did not generalized to the entire cortex. However with surprise they noted that, less often, the activity was recorded in very far away locations. Fig. 5 (redrawn from the original sketches in [5]) summarizes these early observations together with our own rough estimations in Panel D. Filled circles in Panel D represent the distribution of edge lengths, computed from the drawing in Panel A as the Euclidean distance (using arbitrary units) between the location of each strychnine instillation and the resulting activation site/s. Note that, despite the scarcity of the data, the results demonstrate long range correlations, the exponent being similar to the estimations using fMRI [14, 21]. For example an application in the frontal cortex induced activity sometimes in the occipital cortex. Nowadays, is not difficult to admit that frontal activation will evoke visual imagery and viceversa, however McCulloch knew that much before us.

Senses are critical

In more than one sense our senses seems to be critical. To move around, to escape from predators, to choose a mate or to find food, the sensory apparatus is critical for any animal survival. But it seems that senses are also critical in the thermodynamic sense of the world. Consider first the fact that the density distribution of the various form of energy around us is clearly inhomogeneous, at any level of biological reality, from the sound loudness any animal have to adapt to the amount of rain a vegetal have to take advantage. From the extreme darkness of a deep cave to the brightest flash of light there are several order of magnitude changes, nevertheless our sensory apparatus is able to inform the brain of such changes. It is well known that isolated neurons are unable to do that because of their limited dynamic range, which spans only a single order of magnitude. This is the oldest unsolved problem in the field of psychophysics, tackled very recently by Kinouchi and Copelli [16] by showing that the dynamics emerging from the interaction of coupled excitatory elements is the key to solve the problem. Their results show that a network of excitatory elements set precisely at the edge of a phase transition - or, at criticality - can be both, extremely sensitive to small perturbations and still able to detect large inputs without saturation. This is generic for any networks regardless of the neurons’ individual sophistication. The key aspect in the model is a local parameter that control the amplification of any initial firing activity. Whenever the average amplification is very small activity dies out; the model is subcritical and not sensitive to small inputs. On the other hand, choosing an amplification very large one sets
up the conditions for a supercritical reaction in which for any - even very small - inputs the entire network fires. It is only in between these two extremes that the networks have the largest dynamic range. Thus, amplification around unity, i.e., at criticality, seems to be the optimum condition for detecting large energy changes as an animal encounters in the real world [11]. It is only in a critical world that energy is dissipated as a fractal in space and time with the characteristic highly inhomogeneous fluctuations. Since the world around us appears to be critical, it seems that we, as evolving organisms embedded in it, have no better choice than to be the same.

OUTLOOK

The preceding section purposely presented only a selection of concrete results inspired in the approach promoted here. They do not probe that the brain is critical, but they demonstrate that there are relevant aspects of brain dynamics which underlying collective is critical in some sense. There are, of course, an increasingly large body of work modelling and explaining further these experimental findings, which we will not enumerate, because this is not an exhaustive review. An excellent survey is in press and we direct the readers to it [20]. Nevertheless we mention, mostly as a guide for further reading, ideas connected with the general framework discussed here. Probably the first to note should be Ashby’s work to understand how the forces of self-organization could shape a brain [1]. The work of Tononi, Edelman and colleagues [26, 27] it is the first to delineate the fundamental problem of integration and segregation and to explore its connection with complexity. The analysis of cortical coordination dynamics discussed by Kelso, Bressler and colleagues [8], are related with this proposal, because it main ingredients, collective variables, and metastable coordination states are all generic of the critical state discussed here. Of note also is Dehaene [12] "workspace" model of conscious experience that resemble the scale free distribution of hubs observed experimentally and discussed above. Most probably a detailed analysis of their specific numerical models would reveal optimum performance near criticality, something worth to pursue. Finally, there is the exhaustive review of Werner [30] advocating to further the study of phase transitions, metastability and criticality in cognitive models and experiments.

The main difference that set apart this proposal from all of the above efforts, is that it does not pretend to be novel or ad hoc. Right or wrong, but deliberately, the proposal is that relevant aspects of brain dynamics can be understood using the same theoretical framework as for any nonequilibrium thermodynamic system at or near the critical point of a second order phase transition.

Arguably, brain theory is still at a stage comparable to physics in "pre-thermodynamic" times. Imagine yourself in days previous to the notion of temperature. Similarities between scalding water and ice will be supported by their similar "burning" (to the touch) properties, when hot or cold were only subjective quantities. Of course, the notion of pressure and temperature together with phases changed everything. Brain theory will eventually undergo such transformation starting with the preliminary definition of order parameters such as Tononi’s Φ [28] and the elaboration of some phase diagram, including degrees of consciousness, modalities of transitions between phases, etc. Until then, pre-thermodynamic debates will surely continue.
ACKNOWLEDGMENTS

Work supported by NIH NINDS of USA (Grants 42660 and 35115). The warm hospitality of the colleagues of Universidad de Granada are also acknowledged. Special thanks to Dr. Dietmar Plenz (NIMH) for stimulating discussions and for providing Fig. 4.

REFERENCES

1. W. Ross Ashby, "Principles of the self-organizing system". In Principles of Self-Organization Transactions of the University of Illinois Symposium, edited by H. Von Foerster and G. W. Zopf, Jr. Pergamon Press, London UK, 1962, pp. 255–278.
2. P. Bak, *How Nature works*. Oxford University Press, Oxford UK 1997, pp. 1–212.
3. P. Bak, C. Tang, and K. Wiesenfeld, *Phys. Rev. Lett.* 59, 381 (1987).
4. P. Bak and D. R. Chialvo, *Phys. Rev. E* 63, 031912 (2001).
5. P. Bailey, G. von Bonin, and W. S. McCulloch, *The isocortex of the chimpanzee*. Univ. of Illinois Press, Urbana USA 1950, pp. 1–440.
6. J. M. Beggs and D. Plenz, *J. Neurosci.* 23, 11167 (2003).
7. D. Boyer, G. Ramos-Fernández, O. Miramontes, J. L. Mateos, G. Cocho, H. Larralde, H. Ramos, F. Rojas. Scale-free foraging by primates emerges from their interaction with a complex environment, *Proceedings of the Royal Society of London 717 Series B: Biological Sciences* In press. Also as http://xxx.lanl.gov/abs/q-bio.PE/0601024.
8. S. L. Bressler and J. A. S. Kelso, *Trends Cogn. Sci.* 5, 26–36 (2001).
9. D. R. Chialvo and P. Bak, *Neuroscience* 90, 1137 (1999).
10. D. R. Chialvo, *Physica A* 340, 756–765 (2004).
11. D. R. Chialvo, *Nature Phys.* 2, 301–302 (2006).
12. S. Dehaene and L. Nagache, *Cognition* 79,1-37 (2001).
13. J.G. Dusser de Barenne, W. Garol, W. S. McCulloch, *J. Neurophysiol.* 4, 324 (1941).
14. V. M. Eguiluz, D. R. Chialvo, G. Cecchi, M. Baliki and V. Apkarian, *Phys. Rev. Lett.* 94, 018102 (2005).
15. M. D. Fox, A.Z. Snyder, J. L. Vincent, M. Corbetta, D. C. van Essen, M. E. Raichle, *Proc. Natl. Acad. Sci. USA* 102, 9673–9678, (2005).
16. O. Kinouchi and M. Copelli *Nature Phys.* 2, 348–352, (2006).
17. W. S. McCulloch, *Physiol. Rev.* 24, 390–407, (1944).
18. M. Paczuski, B. Bak, Self organization of complex systems. In Proceedings of 12th Chris Engelbrecht Summer School. Also as http://www.arxiv.org/abs/cond-mat/9906077.
19. T. Petermann, M. A. Lebedev, M. Nicolelis, D. Plenz, *Soc. Neurosci. Abst.* 531.1 (2006).
20. A. Salvador, J. Suckling, M. R. Coleman, J. D. Pickard, D. Menon, E. Bullmore, *Cereb. Cortex* 15, 1332–1342 (2005).
21. G. Tononi, G. M. Edelman, O. Sporns, *Trends Cogn. Sci.* 2, 474–484 (1998).
22. O. Sporns, D. R. Chialvo, M. Kaiser and C. C. Hilgetag, *Trends Cogn. Sci.* 8, 418–425 (2004).
23. O. Sporns, G. Tononi, R. Kotter, *PLoS Comp. Biol.* 1, 245–251 (2006).
24. O. Sporns, J. D. Zwi, *Neuroinformatics* 2, 145–162 (2004).
25. C. V. Stewart and D. Plenz, *J. Neurosci.* 26, 8148–8159 (2006).
26. G. Tononi, G. M. Edelman, O. Sporns, *Trends Cogn. Sci.* 2, 474–484 (1998).
27. G. Tononi and G. M. Edelman, *Science* 282, 1846–1851 (1998).
28. G. Tononi, *BMC Neurosci.* 5, 42 (2004).
29. A. Turing, *Mind,* 59, 433–460. (1950). I am quoting from E. A. Feigenbaum and J. Feldman (eds.), *Computers and thought*. New York: McGraw-Hill.
30. T. Petermann, M. A. Lebedev, M. Nicolelis, D. Plenz, *Soc. Neurosci. Abst.* 531.1 (2006).
31. S. Zapperi, L. K. Baekgaard, H. E. Stanley, *Phys. Rev. Lett.* 75, 4071–4074 (1995).