A New Hypothesis for Sleep: Tuning for Criticality

Barak A. Pearlmutter
Hamilton Institute, NUI Maynooth, Co. Kildare, Ireland
barak@cs.nuim.ie

Conor J. Houghton
School of Mathematics, Trinity College Dublin, Ireland
houghton@maths.tcd.ie

September 22, 2008

Abstract

We propose that the critical function of sleep is to prevent uncontrolled neuronal feedback while allowing rapid responses and prolonged retention of short-term memories. Through learning the brain is tuned to react optimally to environmental challenges. Optimal behavior will often require rapid responses and the prolonged retention of short-term memories. At a neuronal level, these correspond to recurrent activity in local networks. Unfortunately, when a network exhibits recurrent activity, small changes in the parameters or conditions can lead to runaway oscillations. Thus, the very changes that improve the processing performance of the network can put it at risk of runaway oscillation. To prevent this, stimulus-dependent network changes should only be permitted when there is a margin of safety around the current network parameters. We propose that the essential role of sleep is to establish this margin by exposing the network to a variety of inputs, monitoring for erratic behavior, and adjusting the parameters. When sleep is not possible, an emergency mechanism must come into play, preventing runaway behavior at the expense of processing efficiency. This is tiredness.

1 Introduction

As bedtime approaches, we begin to experience tiredness. The symptoms of tiredness include yawning, a tendency to fall asleep and a degradation of some cognitive abilities, particularly those related to short term memory and attention. These symptoms tend to become more acute if we are unable to sleep; there is a progressive reduction in coordination, judgment, and reaction time. After a prolonged period without sleep, alertness is often interrupted by micro-sleeps: brief intervals of sleep which we experience as lapses in attention and the intrusion of dream-like thoughts.

One striking property of tiredness is that its symptoms can be temporarily reversed: if we are engaged in a stimulating or crucial task our feeling of tiredness diminishes and our cognitive abilities return to normal (Hartmann, 1973). Furthermore, it is possible to stay awake for days at a time: soldiers in combat sometimes go long periods without sleep, and there are documented cases of extreme sleep deprivation lasting a week or longer. During extreme sleep deprivation people suffer hallucinations and are described as lacking affect but recover after sleep (Ross, 1965).

The normal consequence of tiredness is sleep. There are two main sleep types, rapid eye movement sleep (REM) and slow wave sleep (SWS) during which the brain steps through a variety of slow wave oscillation states. Beginning with SWS, the two sleep types alternate with a ninety minute period. About a fifth of each cycle is REM, with that fraction increasing through the sleep period. The total amount of SWS increases with illness and physical exertion (Hartmann and Brewer, 1976) while the fraction of sleep which is REM along with the total time spent sleeping tends to decrease through life: newborn babies often sleep 18 hours each day and spend three-fifths of that time in REM.

Electroencephalography shows that during SWS neuronal activity is synchronized at low frequency with individual neurons alternating between periods of activity and inactivity (Destexhe et al., 2007). In contrast, during REM, neurons are active and unsynchronised. Heart rate and breathing become irregular in a way that also mimics waking, the eyes move rapidly behind closed eyelids and the penis is erect or the clitoris engorged. There is a loss of muscle tone and the body is paralyzed.

It is believed that we spend most of REM dreaming: dreams are illusionary experiences and are one of the most intriguing features of sleep. While they are particularly associated with REM, dreams have been reported upon waking from SWS. Dreams often involve bizarre sequences of events; while dreaming, however, dreamers are not usually disturbed by the weirdness of the dream and are swept along by the dream narrative. Dreams can produce a strong emotional response. Typically, dreams do not directly revisit recent experiences, rather they often have elements which refer to the recent or distant past. Although some dream narratives are common: being naked in public, running but getting nowhere, flying and falling, dream content shows great variation from person to person and dream to dream (Hartmann, 1973). We do not usually remember our dreams and, if we do, the memory tends to fade rapidly.

Sleep is common to much of the animal kingdom and the human pattern of sleep, with SWS and REM alternating, is common to all terrestrial mammals and to birds (Roth et al., 2006). There is, however, a huge variation in how much different mammals sleep and what fraction of their sleep is REM (Siegel, 2008). Although ranking learning ability or intelligence across species is a somewhat subjective enterprise, it is apparent that the amount of REM sleep is not correlated with intelligence.

1.1 Why do we sleep?

While it is clear that we rest while we are asleep, it is not at all clear why we need to sleep in order to rest. It is only on reflection that we realize what a peculiar phenomenon sleep is. Most obviously, it is dangerous in that there is a substantial threshold of arousal and a period of befuddlement on awakening. Unlike hibernation, which has an obvious metabolic explanation, sleep cannot be explained from a purely ecological point of view and must be the only practical way of performing some particular essential function. Furthermore, while it may at first seem vacuous to point out that the consequence of sleeplessness is tiredness, it should not be forgotten that a theory of sleep should explain the symptoms of tiredness.

The currently established candidate for the essential function of sleep is memory consolidation. According to this hypothesis, sleep is an essential part of the process by which long-term memories are formed. This is a popular and widely tested theory (Fishbein and Gutwein, 1977; Gutwein and Fishbein, 1980; Sejnowski, 1995; Sejnowski and Destexhe, 2000), but the experimental evidence is complex and inconclusive (Allen et al., 1972; Horne and McGrath, 1984; Vertes and Siegel, 2005; Foster and Wilson, 2006; Maquet, 2001) and the original hypothesis has been weakened to the point where some consider it vacuous (Siegel, 2005). There appears to be little evidence that sleep is vital to the consolidation of declarative memories (Smith, 2001; Walker, 2005). In the case of procedural memory for learned motor responses, where evidence for consolidation is strongest, sleep appears important in enhancing response speed, but not quality (Fischer et al., 2002; Walker et al., 2002; Finelli and Sejnowski, 2005). It would seem strange if so extreme and odd an adaptation as sleep was explained by such specific aspects of memory and learning.

Coenen (2005) argues that the classical interference theory of sleep can account for the relationship between sleep and memory. According to the interference theory, the main requirement for memory consolidation is time, and sleep is merely a convenient time period for consolidation to occur. During sleep there is little interference from either sensory input or conscious rumination. The analogy of digestion is given: much of our food is digested during sleep, however, while rest aids digestion, sleep is not vital. Since digestion is not the primary purpose of sleep, the phenomenology of sleep and digestion is fragmentary and complex, as it is for memory consolidation (Coenen, 2005).

This is an extreme view; it is probable that aspects of sleep have evolved to facilitate memory consolidation and that significant amounts of memory consolidation occurs during sleep. However, we contend that memory consolidation is an example of a "piggy-backed" function. Sleep seems to have evolved to subserve a number of functions in addition to its core essential function. For example, it seems likely that the penile erections which occur during REM sleep serve a function which could also be performed by penile erections during wakefulness, and that their occurrence during sleep is a matter of convenience rather than necessity. In this paper, we propose a new hypothesis concerning the core function of sleep.

We propose that the neuronal plasticity rules that subserve learning would inevitably result in harmful runaway oscillations unless there were intervals of sleep during which different protective rules operated.

2 Criticality

In physics, criticality describes the behavior of a statistical system at the boundary between two phases: for example, water at its boiling point. As the temperature is increased, the properties of water change gradually, but, at the temperature critical point there is a sudden, discontinuous qualitative change as the water turns to steam. We are not in a position here to define and study criticality for the brain in the same precise way that it can be defined and studied for physical systems such as the Ising model. However, a neuronal network has dynamics determined by the neuronal parameters, such as synaptic efficiencies and membrane time constants. Here, we will use criticality to refer to a border in the neuronal parameter space, separating networks with different qualitative behavior. Specifically, we use criticality to describe a border which separates networks with persistent but controlled activity from networks with pathological behavior.

It is clear that changes to the synaptic efficacies in a neuronal network can lead to more persistent activity and to a more rapid response of neurons to afferent spiking. We argue that, in aggregate, the synaptic changes that occur during normal awake behavior increase the persistence of activity and the rapidity of responses: increasing the strength of the connections between neurons decreases the effective time constants for the local neuronal circuits. Persistence and rapidity are clearly positive performance attributes for the brain, useful for the integration of separate inputs and for timely reactions to sudden events. It is also clear that the type of synaptic changes that increase persistence of activity and rapidity of response are likely to bring the network closer to criticality. As such, they risk creating a situation where normal sensory data can cause uncontrolled feedback. In short, it seems likely that there is a performance benefit to synaptic changes which approach the critical boundary separating persistent or rapid activity from uncontrolled feedback. Although the rate of plastic changes in the brain is greatest during development, they are by no means exclusive to it. The brain is continually changing, these changes are presumably to the benefit of performance and, therefore, we would argue, likely to bring the brain closer to the critical boundary.

It would be extremely damaging for the brain to suffer uncontrolled feedback. If the changes in the synapses that occur while awake tend to bring the neuronal networks in the brain closer to the critical boundary, then there must be some mechanism to protect the brain from straying over that boundary. We hypothesize that this is the primary function of sleep and that different plasticity regimes at work in the sleep-wake cycle conspire to tune the brain for criticality. Neuronal circuits move closer to criticality while awake. During sleep, the activity of the sleeping brain serves to discover network dynamics whose behavior shows too close a proximity to the critical boundary and to remove the consequent risk of potentially harmful self-reinforcing loops though protective plastic changes. This is illustrated in Fig. 1.

Regarded as a dynamical system, the brain is, of course, extremely complex. At the cellular level, there is a huge range of different dynamical parameters including membrane constants, parameters describing the density and the time course of ion gates along with synaptic efficiencies and other synaptic parameters. Neurons participate in local networks and these, in turn, can give rise to a huge variety of different local circuits with different, and subtle, dynamical behaviors. We do not wish to pin our hypothesis to any particular proposed local neuronal circuit; instead, we hope to advance our hypothesis as a general principle with validity across a large range of circuit types. Optimal behavior in different circuits will be associated with different sorts of tuning of the effective dynamical parameters, such as the time constants or the relative degree of excitation and inhibition. The persistence of short term memories would, for example, require dynamical circuit with long time constants. In contrast, rapid responses to stimuli would require short time constants. What we propose is that across many local circuits, the on-going plastic changes which adjust parameters to optimize behavior will bring the system closer to super-criticality, super-critically which might in some case mean an uncontrolled growth in activity, in others, over-persistent or over-synchronized oscillations.

As discussed in the introduction, it is probably that tuning for criticality is not the only function of sleep. It is quite possible that memory consolidation also occurs during sleep, it is even possible that the plastic changes that occur during sleep are chosen to simultaneously promote the consolidate of memories and protect the brain from super-criticality. However, our primary intention here is to argue that tuning-for-criticality is the main reason for sleep.

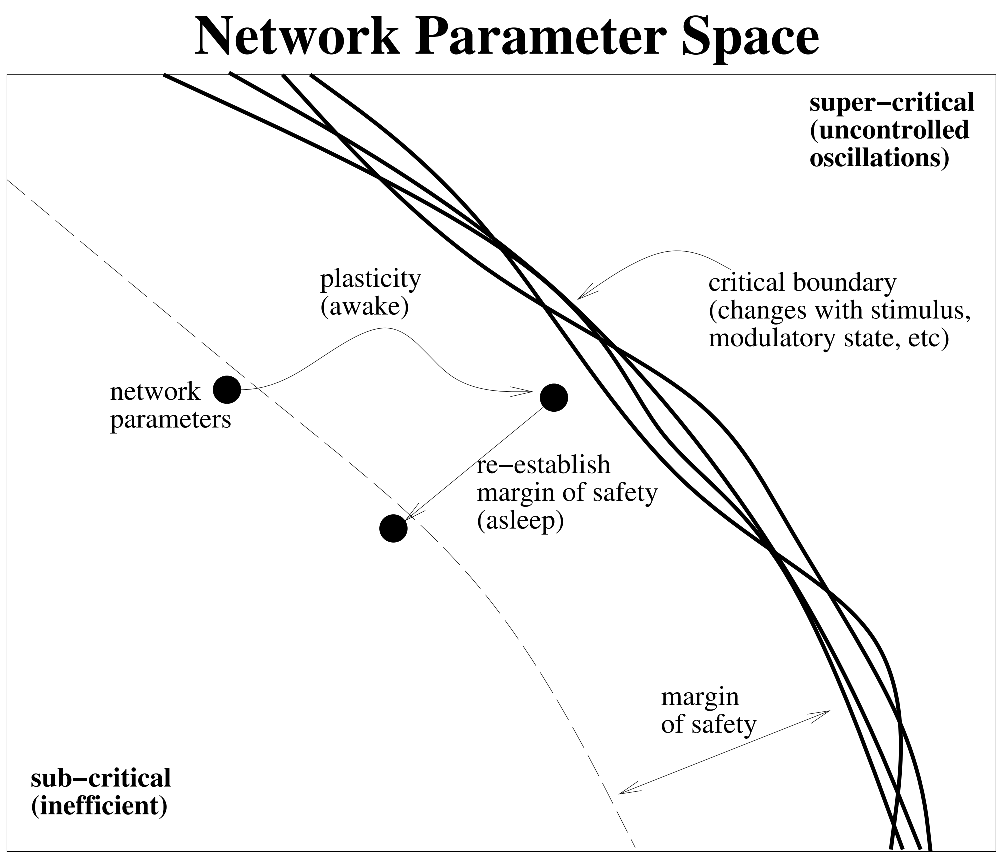

Figure 1: A cartoon representation of the network parameter space and the wake-sleep cycle. Each point in this space represents a different network with different network parameters. In reality, of course, this space would have a huge number of dimensions. The boundary between the sub-critical and super-critical region is represented by a cluster of interwoven lines, this indicates that the network dynamics depend on stimulus and on modulatory behavior. Some points in parameter space might have apparently safe, sub-critical, behavior for most stimuli but produce a super-critical response to some unusual events. During learning the network moves towards this critical boundary, sleep moves it back. This does not reverse the effect of learning, but it does change the network to ensure that there is a margin of safety between it and the critical boundary.

2.1 Tiredness

We propose that tiredness is an emergency response to a lack of sleep. Tiredness is not just a craving: it degrades our performance. Hunger, for example, does not have physical ill-effects until the lack of food begins to cause a metabolic deficit. The symptoms of tiredness include reduced mental ability. This seems to indicate that tiredness occurs when some important resource has been exhausted. Here, we are suggesting that this resource is the margin of safety separating the brain from the boundary of criticality, and that tiredness occurs when it would be dangerous for the brain to continue to operate at its full ability. If we remain awake for too long, synaptic plasticity is reduced to avoid increasing the risk of super-critical behavior and, crucially, neuronal parameters are reversibly down-regulated in a way that reduces the persistence of activity across the brain and therefore the risk of uncontrolled feedback. Such a protective response is tiredness. This is illustrated in Fig. 2. This hypothesis is consistent with our ability, when required, to go without sleep and to briefly throw off the negative effects of tiredness.

2.2 Epilepsy

Dinner and Lüders (2001) note that there ‘is a very close relationship between epilepsy and sleep that

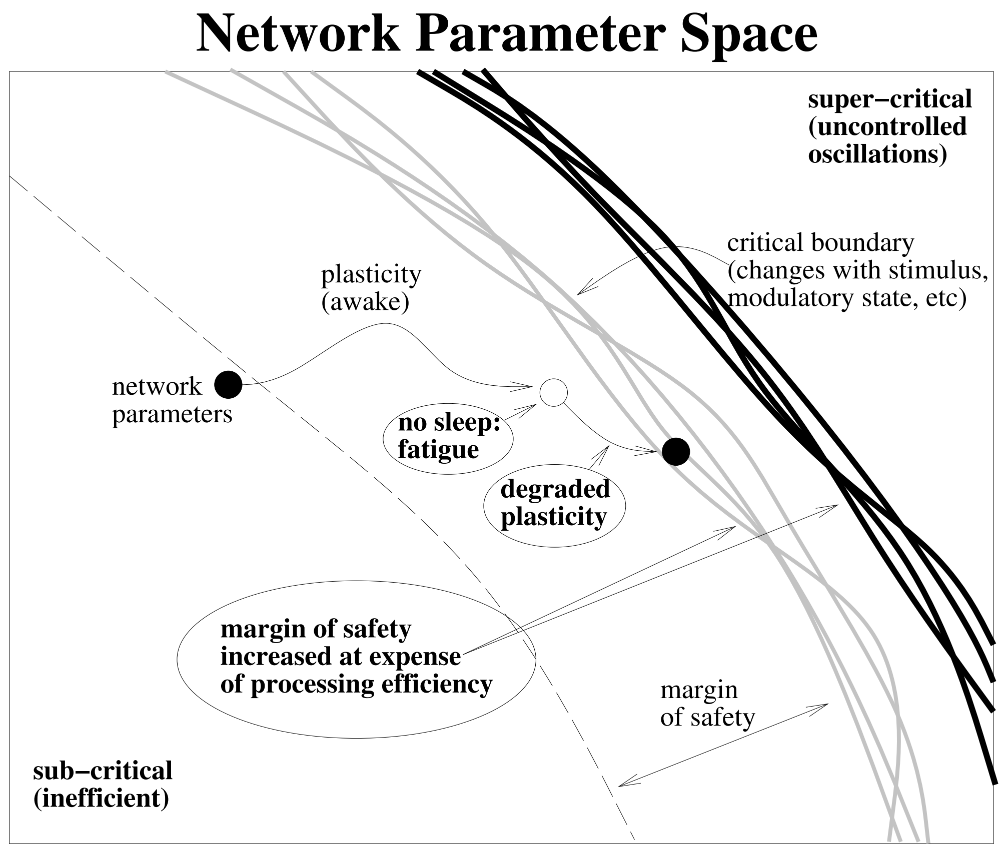

Figure 2: A cartoon representation of the effect of tiredness. In this illustration the network has approached the critical boundary and is at the point where sleep normally restores the margin of safety. However, in the absence of sleep, an emergency response is required, the network is modulated so that performance is degraded and the critical boundary moves further away. Synaptic plasticity is also down regulated, preventing the network from straying too much further into the danger region.

has been recognized since antiquity.' In epileptics, tiredness and sleep deprivation are often triggers of seizures and 'in a few patients, recognition and treatment of a coexisting sleep disorder can make the difference between complete seizure control and refractory epilepsy' (Bazil, 2002). There is evidence that some epileptic seizures are accompanied by an increase in the spatial and temporal correlation of neuronal activity (Lehnertz and Elger, 1998) and we hypothesize that these seizures are the result of super-critical behavior. Furthermore, seizures are often followed by a postictal period: a period of tiredness and listlessness lasting up to 48 hours during which the epileptic often sleeps (Kolb and Litt, 2002); perhaps this period serves to reduce the risk of further super-critical behavior by an extended search for near critical loops. Tiredness is also a significant trigger for migraine and migraine is also often followed by a postictal period (Mulder et al., 1999) making it possible to speculate that migraine is also related to the risk of super-criticality.

2.3 Depression

Clinical depression is exacerbated by too much sleep, and treatments include sleep deprivation (Wirz-Justice and den Hoofdakker, 1999) and modafinil, a drug which suppresses tiredness (Menza et al., 2000). Perhaps one contributing cause of depression is a tuning-for-criticality mechanism that is too strong. This results in a system that is overly sub-critical whose sub-critical dynamical constants cause an inability to sustain concentration, a reduced short-term memory capacity, and lethargy. Furthermore, the fact that the depressed brain is sub-critical could reduce the need for sleep below what is regulated for by the circadian rhythm, resulting in insomnia.

2.4 Dreams

Dreaming fits naturally into this tuning-for-criticality theory of sleep. In the context of our hypothesis, dreams stimulate the brain in a manner that mimics awake cognition as part of a search for self-reinforcing loops. Dreams are a guess, based on recent and distant memories, of possible future sensory stimuli. They need not have a coherent and logical narrative structure to effectively test whether the neuronal circuits in the brain can be stimulated to produce feedback behavior. However, it would be advantageous for dreams to include a wide variety of pointed and provocative content. This is consistent with observation.

By claiming that dreaming is a mechanism for locating the most proximate parts of the critical boundary, we appear to be restricting our hypothesis to REM. Of course, it is possible that SWS does have an unrelated critical function and that REM and SWS have different functional origins. It is also possible that the oscillations that occur in SWS broaden or refine the search for harmful super-criticality in the brain. Another possibility is that the search for synaptic changes that are required to protect the brain from the risk of super-critical behavior takes place while awake and during REM, but the actual changes occur during SWS. Perhaps while awake and REM, local neuronal behavior typical of cells at risk of super-critical behavior is tagged biochemically and, during SWS, these markers result in protective synaptic changes. This is consistent with the reduction in REM fraction seen in polyphasic sleep (Vogel, 1978). However, this is just speculation; of main proposal is that tuning for criticality is the main purpose of sleep and that sleep structure is, in some way, optimized to that purpose.

2.5 Cetaceans and unihemispherical sleep

Cetaceans cannot sleep in the normal way: they must remain swimming in order to surface and breath and an unconscious cetacean would drown. In fact, many cetaceans sleep with half their brains at a time: one half of the brain remains awake, and one eye remains open, enabling the whale or dolphin to continue to surface and sound (Mukhametov, 1985). This example alone allows us to discount ecological theories of sleep: if sleep where simply a mechanism to conserve energy and encourage stillness during periods when activity is not advantageous, dolphins would not sleep. One disadvantage of uni-hemispherical sleep is that it appears to make REM impossible; at least, it has not been observed in cetaceans (Mukhametov, 1985; Lyamin et al., 2000; Mukhametov, 1984; Shurley et al., 1969; Mukhametov, 1988). According to the tuning-for-criticality theory of sleep, this requires that the dolphin brain be less subtly tuned for criticality. This, in turn, implies that the dolphin brain is less efficient than the brains of terrestrial mammals. In fact, cetaceans in general, and in particular cetacean, such as the killer whale, narwhal and dolphin, that hunt and that live in challenging environments, are unusually highly encephalized.

One common measure of encephalization, thought to give a useful measure of the available processing power of an animals brain, is the encephalization quotient (EQ) defined (Jerison, 1973) by

$$\mathrm{EQ} = \frac{\text{brain mass}}{0.12(\text{body mass})^{2/3}} \tag{1}$$

In Table 1 we exhibit the EQ for a range of animals. This appears to roughly list the animals in order of what we would intuitively estimate as their cognitive ability. However, the dolphin is surprisingly high, leading to the suggestion (Kesarev, 1971) that the massive development of the neocortex in dolphins may be related to more active participation of the central nervous system in the regulation of visceral functions and their correlation with somatic functions. It is hard to see why this would be the case, and we claim that dolphins have a high EQ because they are unable to benefit from the same, presumably optimal, sleep structure found in other mammals.

Unihemispherical sleep is not unique to Cetaceans, it is also observed in eared seals, manatees and many species of bird, however, other birds and mammals which exhibit unihemispherical sleep also sleep bihemispherically and experience REM sleep. This presumably reflects the undesirability of exclusively unihemispherical sleep as an adaptation; only the particular challenge presented to mammals by a wholly aquatic exist results in unihemispherical sleep.

Of course, it is possible that the dolphin's high EQ is an accurate reflections of its cognitive and sensory ability or that the specific EQ scaling factors are chosen specifically to order terrestrial mammals, and

Table 1: Encephalization Quotient (EQ) of Selected Mammals (MacPhail, 1982)

Species	EQ
Man	7.44
Dolphin	5.31
Chimpanzee	2.49
Rhesus Monkey	2.09
Elephant	1.87
Whale	1.76
Dog	1.17
Cat	1.00
Horse	0.86
Sheep	0.81
Rat	0.40
Rabbit	0.40

that marine animals, with different propulsive mechanisms and surface to volume ratios, would suit a different scaling factor. However, what is certain is that sleep time and profile varies hugely from mammal to mammal and a convincing theory of REM sleep function should explain these great differences.

2.6 Learning and memory in Zebra finch

Adult males zebra finches sing, and have a single identifying song which they learn under the tutelage of an adult male. At 30 days after hatching juvenile finches begin producing unstructured sounds; after tutoring the sounds become structured through frequent rehearsal and by 90 days the song has stabilized, a process referred to as song crystallization. Derégnaucourt et al. (2005) describes a comprehensive study of sleep, learning and memory in zebra finch, in which juvenile male zebra finches were separated from their colony at day 30 after hatching. Tutoring began at day 43 and, from that time until crystallization, all sounds produced by the finch were recorded and analyzed.

Using automated quantitative analysis of song quality and song similarity (Ho et al., 1998; Tchernichovski et al., 2000), it was observed that song quality degrades during sleep and songs produced after waking are less similar to the model songs than those produced the preceding night; see, for example, Fig. 3. Great care was taken in this study to rule out explanations other than sleep time degradation of performance, for example, the possibility that the passing of time, rather than actually sleep, is responsible for the effect, was studied and excluded, as was the possibility that the vocal apparatus or morning grogginess, rather than song learning, was at fault for poor morning performance. However, no attempt was made to manipulate sleep structure; this may a useful future possibility, zebra finches are known to have a similar sleep structure to mammals (Low et al., 2008).

While it is possible to imagine explanations within the memory consolidation theory, this degradation of performance is hardly a natural prediction of a theory which posits that memory and learning is stabilized and even improved during sleep. It is however a natural prediction of our hypothesis. Song learning is accompanied by substantial plastic change in the brain, necessitating a substantial sleep-time retreat from the critical boundary.

2.7 Avalanches

Striking evidence for critical behavior is provided by the observation of avalanche-like spiking in cortical slices (Beggs and Plenz, 2003, 2004). Extended recordings are made of slices using an electrode array. It is observed that spiking is intermittent, with periods of activity interleaved by periods of quiescence. By dividing time into short bins and counting how many electrodes are active in each bin, it is possible to define the length and size of the periods of activ-

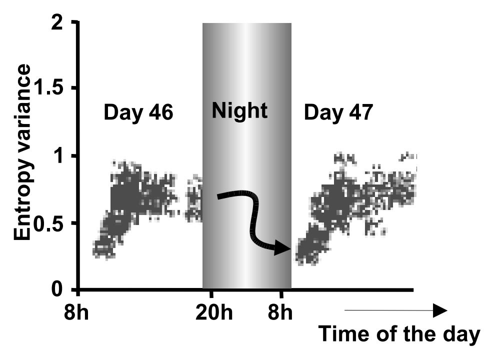

Figure 3: Zebra finch song quality monitored during acquisition. "Entropy variance" was a measure of imitation accuracy. Learning was rapid after awakening, then slowed. Sleep degraded song performance. Birds with the most degradation by sleep ultimately attained the best performance. Adapted by permission from Macmillan Publishers Ltd: Nature (Derégnaucourt et al., 2005) copyright (2005)

ity, or avalanches. Both the number of avalanches of a given duration and the number of a given size decrease according to a power law: $N \propto n^{\gamma} + C$ where C is a constant, N is the number of events and n the duration or size with scaling exponents $\gamma = -2$ for duration and $\gamma = -3/2$ for size. Although power-law scaling behavior does not necessarily mean the system is critical, it is typical of critical systems and this specific power law behavior arises in the critical regime of a simple network model (Beggs and Plenz, 2003, 2004; Zapperi et al., 1995).

3 Discussion

3.1 Related theories of the brain

This tuning-for-criticality hypothesis regarding the role for sleep is novel. However, the idea that the brain is a near-critical network is not. For example, a classic paper by Turing (1950) considers in passing how criticality might apply to the brain. More recently, "Computation at the edge of chaos" has gained considerable attention in the study of complex systems (Chialvo, 2004). It has been proposed that, in order to support computation, the brain exists near criticality (Greenfield and Lecar, 2001). As we have discussed, criticality has engineering advantages; furthermore, some learning algorithms for temporal sequences, such as liquid state machines (Maass et al., 2002) and echo state networks (Jaeger and Haas, 2004), perform well only when their recurrent network modules are near the boundary of criticality. It has also been proposed that the brain self-organizes for criticality, in other words, that near-critical behavior is the consequence of the synaptic plasticity rules involved in learning (Bak, 1997). Self-organization for criticality is an intriguing proposal, but one that differs from the hypothesis that we advance here: we propose that near-critical behavior is the result of tuning during the wake-sleep cycle.

A discussion of neuronal network criticality, which closely anticipates our proposal, occurs in a paper concerning a spiking network capable of detecting spatio-temporal patterns (Hopfield and Brody, 2001). It is noted that if too many patterns are learned, there is a phase transition into a regime in which the network loses its selectivity. This super-critical regime is likened to epilepsy. It is noted that more patterns can be stored if a few neurons are removed: the ones involved in the encoding of the largest number of patterns and therefore most implicated in the "epileptic" behavior. This is an example of an intervention which prevents criticality in order to improve performance. In our proposal, the neurons would not be removed; rather synaptic strengths would be depressed during a sleeping phase.

Earlier theories have also suggested that the function of sleep is related to unlearning. Based on an analogy between the cerebral cortex and artificial neural networks, it has been proposed that sleep removes certain undesirable modes of interaction between cortical cells, namely interactions that may become parasitic (Crick and Mitchison, 1983). The proposed function of REM sleep is thus to tune the network using random neuronal firings, helping to eradicate spurious connections. One could imagine this as a kind of reverse learning mechanism, in which spurious connections are attenuated, rather than strengthened.

The risk of uncontrolled behavior presented by recurrent connections is emphasized in Crick and Koch (1998). The authors note that "neurons are usually kept in a jittery state near the edge of excitation in order to make them respond more promptly" and that makes it likely that strong excitatory loops "would throw the cortex into uncontrolled oscillations, as in epilepsy". They also note that excitatory feedback makes it difficult to design recurrent neural networks and compare recurrent models to the "early radio sets that emitted howls (caused by positive feedback) if the volume control was turned too high." Crick and Koch (1998) proposes that the absence of strong-excitatory loops is a design principle in cortical and thalamic projection; here we are proposing that the plastic changes which occur during sleeping ensure that there are no strong-excitatory loops which are likely to lead to uncontrolled oscillations.

There has been considerable study of learning algorithms for continuous networks with hidden units, time delays, and asymmetric connections (Pearlmutter, 1995). One issue with such systems is that some of the most practical learning algorithms assume the system settles to a fixed point, that is, that it is an attractor network, and ignore the path taken to reach the fixed point (Feynman, 1939; Jordan, 1986; Pineda, 1987; Almeida, 1987). This motivated exploration of how it could be ensured that a network of this sort would indeed settle to a fixed point, and how the time taken to reach the fixed point could be controlled. This is very closely related to super-criticality, as making these trajectories faster improves performance, but the changes required to make them faster, taken too far, lead to oscillation and chaos. Atiya and Abu-Mostafa (1990, Eq. 3) derived sufficient conditions on the weights in such a network to guarantee stability, but these were quite conservative: many stable networks used in practice violate them. Simard et al. (1990, 1991) develops methods to measure and control the spectrum of the Jacobian of the evolution function in such networks. Monte Carlo exploration of the parameter space of such networks revealed complex structure unlikely to be captured by any simple prescription (Renals and Rohwer, 1990). Local exploration is the only known practical way to identify when a complex network is near to criticality.

The length of the impulse response of a recurrent network determines the maximum temporal separation which two stimuli can have in order for a gradient-based method to learn an association between them. This makes for a delicate trade-off: a short impulse response makes the network stable, and allows a more aggressive learning rate, but gives it a brief short-term memory and makes it unable to learn longer temporal associations. A longer impulse response requires a slower learning rate, and also puts the network at risk of infinite impulse responses, i.e., oscillatory or chaotic behavior. These issues were noted and motivated the construction of a specialized recurrent architecture which allows very long impulse responses Hochreiter and Schmidhuber (1997).

It has been shown that even in a simple associative neural network model without hidden units, spurious attractors can be removed by a phase of unlearning of noise input, which was suggested to correspond to REM sleep (Hopfield et al., 1983). Many modern unsupervised machine learning algorithms, such as the Boltzmann Machine (Ackley et al., 1985; Hinton, 1989) or the Helmholtz Machine (Hinton et al., 1995; Dayan et al., 1995; Dayan and Hinton, 1996), require alternating “awake” and “hallucinating” phases. The purpose of these “hallucinating” phases is, however, quite different from the purpose we propose for sleep.

3.2 Partial evidence which may support our theory

There is evidence that synaptic strengths are down-regulated during SWS (Cirelli and Tononi, 2000; Cirelli et al., 2004; Tononi and Cirelli, 2006). This supports our hypothesis. It is further suggested (Tononi and Cirelli, 2006) that synapses are weakened in proportion to their strength and that this reduces energy and space costs and prevents synaptic saturation. In contrast, according to our hypothesis, synapse-specific changes are made to reduce the risk of network super-criticality: we contend that this is more consistent with the phenomenology of sleep and tiredness.

Sleep in adults is typically driven by a circadian rhythm. Apart from the apparent ecological role of a circadian rhythm in some animals, according to this tuning-for-criticality theory we might say that tiredness is a mechanism to avoid super-criticality, which is very dangerous; and that the circadian rhythm is a mechanism to avoid tiredness, which is itself risky. Furthermore, during the evolutionary emergence of sleep, the alteration of night and day presumably entrained the tuning cycle and resolved the competing need to remain close to the critical boundary by sleeping frequently and the need to have unbroken periods of wakefulness during which to perform tasks.

Monoamine oxidase (MAO) inhibitors are a class of anti-depressants which inhibit the breakdown of monoamine neurotransmitters. Although an effective therapy, MAO inhibitors are now rarely prescribed because they interact dangerously with some foods and require the patient to follow a restricted diet. Surprisingly, REM sleep is virtually abolished by MAO inhibitors (Cohen et al., 1982). This is very hard to explain within memory consolidation theories of sleep since these would predict that MAO inhibitors would cause memory impairment when, in fact, there is some evidence that MAO inhibitors produce memory improvement (Georgotas et al., 1983). There is also a famous case of a man who almost completely lacks REM sleep following brain stem lesion from shrapnel fragments, but who does not seem to suffer a detectable cognitive impairment (Lavie et al., 1984).

It must be acknowledged that the elimination of REM also poses a challenge to our hypothesis, and indeed to any theory of sleep that assigns a function to REM. We contend, of course, that REM must serve a function and those who avoid REM must suffer some negative consequence. However, all that we seem to know is that this negative consequence is not related to memory. We have suggested that protective synaptic changes occur during SWS in response to biochemical markers laid down during wake and dreaming. In the context of this version of our hypothesis, it is possible that MAO inhibitors slow the breakdown of those markers, resulting in protective changes which respond only to stimuli experienced during wakefulness. This would imply that those avoiding REM by taking MAO inhibitors, or indeed, by adopting a polyphasic sleep pattern, should have an increased susceptibility to tiredness, and even an increased risk of epilepsy, when confronted by situations poorly foreshadowed by normal awake life.

Yawning remains a mystery. Yawning is a symptom of tiredness, boredom, stress and migraine, and people most frequently yawn just before going to sleep and just after waking up. Yawning is also suggestive: we are prompted to yawn by the sight of someone else yawning, or even by *reading* about yawning. It is not know why we yawn; one intriguing suggestion is that it is a thermo-regulatory mechanism for controlling brain temperature (Gallup and Gallup, 2007). The dynamics of neurons in the brain are temperature sensitive, and increasing the temperature of the brain might be expected to decrease thresholds of activation. If the purpose of yawning is to reduce brain temperature, then the increased rate of yawning that accompanies tiredness would be consistent with the tuning-for-criticality hypothesis, as raising such thresholds would be particularly important when the brain is at risk of super-critical behavior.

3.3 Predictions

The most direct prediction derived from our hypothesis is that different plasticity regimes operate during wakefulness, tiredness and slow-wave sleep. There are obvious technical difficulties to the experimental study of the differences in plasticity during wakefulness versus sleep. However, given a strong motivating framework, such studies may be within the capabilities of current technology.

Some experiments that relate to less direct predictions have already been conducted for other purposes. For instance, we would expect that the impulse response of the cortex after extended wakefulness would be near the boundary of criticality, and that during slow wave sleep the impulse response would be moved further from the boundary. This prediction is consistent with some experimental data (Beggs and Plenz, 2003, 2004; Massimini et al., 2005).

In addition to its success in explaining some of the established phenomenology of sleep and tiredness, the tuning-for-criticality hypothesis makes a number of testable psychophysical predictions. Moreover, introducing a plausible rival theory to the currently dominant memory consolidation theory may help give a framework for future research in this area.

One success of our hypothesis is that it explains tiredness as an emergency protection against supercriticality. If this is true, then it may be risky to suppress tiredness. Combating the symptoms of tiredness through constant stimulation or by using drugs should increase the risk of suffering epilepsy. An increased risk of epilepsy is likely to be a consequence of regular sleep deprivation, particularly among babies and children, whose brains are undergoing more substantial awake synaptic changes. Similarly the use of melatonin to improve sleep patterns should help control epilepsy.

According to our hypothesis, down-regulation of synaptic plasticity is another symptom of tiredness. This should make learning a new task more difficult during tiredness. Furthermore, the sleep debt accumulated during an extended awake period should be proportional to the fraction of that period during which synaptic changes were permitted. This could be evaluated by monitoring the ability to learn during a period of sleep deprivation. It is also expected that REM time should increase when it is anticipated that the day-to-day environment will soon become less predictable and less well typified by awake life. This could be tested by studying the sleep patterns of people about to alter their life circumstances.

We have already suggested that the phylogenetic consequences of our theory would explain the high encephalization of some cetaceans relative to their apparent cognitive abilities. Sleep is common to much of the animal kingdom and, although ranking learning ability or intelligence across species is a somewhat subjective enterprise, it does not appear that the amount of REM sleep is correlated with intelligence. We predict that further evidence for our theory will be revealed by a closer examination of sleep patterns, encephalization, cognitive ability, and environmental unpredictability. There is also a developmental implication: total sleep time should correlate with brain growth rather than body growth.

Thanks to Jack D. Cowan for helpful discussion and motivation, and to Stuart A. Tilbury for detailed helpful comments and research. Supported by Science Foundation Ireland grants 00/PI.1/C067 (BAP) and 06/RFP/BIM020 (CJH).

References

D. H. Ackley, G. E. Hinton, and T. J. Sejnowski. A learning algorithm for Boltzmann Machines. *Cognitive Science*, 9:147–69, 1985.

S. R. Allen, J. Oswald, S. Lewis, and J. Tagney. The effect of distorted visual input on sleep. *International Journal of Psychophysiology*, 9:498–504, 1972.

L. B. Almeida. A learning rule for asynchronous perceptrons with feedback in a combinatorial environment. In M. Caudill and C. Butler, editors, *IEEE First International Conference on Neural Networks*, volume 2, pages 609–18, San Diego, CA, June 21–24 1987.

A. F. Atiya and Y. Abu-Mostafa. A method for the associative storage of analog vectors. In *Advances in Neural Information Processing Systems 2*, pages 590–5. Morgan Kaufmann, 1990.

P. Bak. *The Science of Self-Organized Criticality*. Oxford University Press, 1997.

C. W. Bazil. Sleep and epilepsy. *Seminars in Neurology*, 22(3):321–7, 2002.

J. M. Beggs and D. Plenz. Neuronal avalanches in cultured slices of neocortex. *Journal of Neuroscience*, 23(35):11167–77, 2003.

J. M. Beggs and D. Plenz. Neuronal avalanches are diverse and precise activity patterns that are stable for many hours in cortical slice cultures. *Journal of Neuroscience*, 24(22):5216–29, 2004.

D. R. Chialvo. Critical brain networks. *Physica A*, 340(4):756, 2004. URL `http://arxiv.org/abs/cond-mat/0402538`. Contribution to the Niels Bohr Summer Institute on Complexity and Criticality (2003); in Per Bak Memorial Issue of *Physica A*.

C. Cirelli and G. Tononi. Differential expression of plasticity-related genes in waking and sleep and their regulation by the noradrenergic system. *Journal of Neuroscience*, 20(24):9187–94, 2000.

C. Cirelli, C. Gutierrez, and G. Tononi. Extensive and divergent effects of sleep and wakefulness on brain gene expression. *Neuron*, 41:35–43, 2004.

A. Coenen. Where is the classic interference theory for sleep and memory. *Behavioral and Brain Sciences*, 28:67–8, 2005.

R. M. Cohen, D. Pickar, D. Garnett, S. Lipper, J. C. Gillin, and D. L. Murphy. REM sleep suppression induced by selective monoamine oxidase inhibitors. *Psychopharmacology*, 78(2):137–40, Oct. 1982. doi: 10.1007/BF00432251. URL `http://www.springerlink.com/content/w7068p5747xxx575/`.

F. Crick and C. Koch. Constraints on cortical and thalamic projections: the no-strong-loops hypothesis. *Nature*, 391(6664):245–50, 1998.

F. Crick and G. Mitchison. The function of dream sleep. *Nature*, 304:111–4, 1983.

P. Dayan and G. E. Hinton. Varieties of Helmholtz machine. *Neural Networks*, 9(8):1385–403, 1996.

P. Dayan, G. E. Hinton, R. M. Neal, and R. S. Zemel. The Helmholtz machine. *Neural Computation*, 7(5):889–904, Sept. 1995.

S. Derégnaucourt, P. P. Mitra, O. Fehér1, C. Pytte, and O. Tchernichovski. How sleep affects the developmental learning of bird song. *Nature*, 433(7027):710–6, 17 Feb. 2005. doi: 10.1038/nature03275.

A. Destexhe, S. W. Hughes, M. Rudolph, and V. Crunelli. Are corticothalamic 'up' states fragments of wakefulness? *Trends in Neuroscience*, 30(7):334–42, 2007.

D. S. Dinner and H. O. Lüders, editors. *Epilepsy and Sleep: Physiological and Clinical Relationships*. Academic Press, 1st edition, Jan. 2001. ISBN 978-0122167706.

R. Feynman. Forces in molecules. *Physical Review*, 56(340), 1939. URL `http://link.aps.org/abstract/PR/v56/p340`.

L. A. Finelli and T. J. Sejnowski. What is consolidated during sleep-dependent motor skill learning? *Behavioral and Brain Sciences*, 28(1):70–1, 2005.

S. Fischer, M. Hallschmid, A. L. Elsner, and J. Born. Sleep forms memory for finger skills. *Proceedings of the National Academy of Sciences of the United States of America*, 99:11987–91, 2002.

W. Fishbein and B. M. Gutwein. Paradoxical sleep and memory storage processes. *Behavioral Biology*, 19:425–64, 1977.

D. J. Foster and M. A. Wilson. Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature*, 440(7084):680–3, 2006.

A. C. Gallup and G. G. Gallup, Jr. Yawning as a brain cooling mechanism: Nasal breathing and forehead cooling diminish the incidence of contagious yawning. *Evolutionary Psychology*, 5:92–101, 2007.

A. Georgotas, B. Reisberg, and S. Ferris. First results on the effects of MAO inhibition on cognitive functioning in elderly depressed patients. *Archives of Gerontology and Geriatrics*, 2:249–63, 1983.

E. Greenfield and H. Lecar. Mutual information in a dilute, asymmetric neural network model. *Physical Review E*, 63(4):041905, Apr. 2001. URL `http://link.aps.org/abstract/PRE/v63/e041905`.

B. M. Gutwein and W. Fishbein. Paradoxical sleep and memory (I): Selective alterations following enriched and impoverished environmental rearing. *Brain Research Bulletin*, 5:9–12, 1980.

E. Hartmann. *The Functions of Sleep.* Yale University Press, New Haven, CT, 1973.

E. Hartmann and V. Brewer. When is more or less sleep required? A study of variable sleepers. *Comparitive Psychiatry*, 17(2):275–84, Mar.–Apr. 1976.

G. Hinton, P. Dayan, B. Frey, and R. Neal. The wake-sleep algorithm for unsupervised neural networks. *Science*, 268:1158–61, 1995.

G. E. Hinton. Deterministic Boltzmann learning performs steepest descent in weight-space. *Neural Computation*, 1(1):143–50, 1989.

C. E. Ho, B. Pesaran, M. S. Fee, and P. P. Mitra. Characterisation of the structure and variability of zebrafinch song elements. *5th Joint Symposium on Neural Computation Proceedings*, pages 76–83, 1998.

S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–80, 1997.

J. J. Hopfield and C. D. Brody. What is a moment? transient synchrony as a collective mechanism for spatiotemporal integration. *Proceedings of the National Academy of Sciences of the United States of America*, 98:1282–7, 2001.

J. J. Hopfield, D. I. Feinstein, and R. G. Palmer. 'Unlearning' has a stabilizing effect in collective memories. *Nature*, 304(5922):158–9, July 14 1983.

J. A. Horne and M. J. McGrath. The consolidation hypothesis for REM sleep function: Stress and other confounding factors—a review. *Biological Psychology*, 18:165–84, 1984.

H. Jaeger and H. Haas. Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. *Science*, 304:78–80, 2004.

H. J. Jerison. *Evolution of the Brain and Intelligence.* Academic Press, 1973.

M. I. Jordan. Attractor dynamics and parallelism in a connectionist sequential machine. In *Proceedings of Ninth Annual Conference of the Cognitive Science Society*, pages 531–46. Lawrence Erlbaum Associates, 1986.

V. S. Kesarev. The inferior brain of the dolphin. *Soviet Sci. Rev.*, 2:52–8, 1971.

S. J. Kolb and B. Litt. Management of epilepsy and comorbid disorders in the emergency room and intensive care unit. In A. B. Ettinger and O. Devinsky, editors, *Managing Epilepsy and Co-Existing Disorders*, pages 515–35. Butterworth-Heinemann, 2002.

P. Lavie, H. Pratt, B. Scharf, R. Peled, and J. Brown. Localized pontine lesion, nearly absence of REM sleep. *Neurology*, 34:118, 1984.

K. Lehnertz and C. E. Elger. Can epileptic seizures be predicted? evidence from nonlinear time series analysis of brain electrical activity. *Physical Review Letters*, 80:5019–22, 1998.

P. S. Low, S. S. Shank, T. J. Sejnowski, and D. Margoliash. Mammalian-like features of sleep structure in zebra-finches. *Proceedings of the National Academy of Sciences of the United States of America*, 105(26):9081–6, 2008.

O. I. Lyamin, P. R. Manger, L. M. Mukhametova, J. M. Seigel, and O. V. Shpak. Rest and activity states in a gray whale. *Journal of Sleep Research*, 9:261, 2000.

W. Maass, T. Natschläger, and H. Markram. Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Computation*, 14:2531–2560, 2002.

E. MacPhail. *Brain and Intelligence in Vertebrates.* Clarendon Press, Oxford, 1982.

P. Maquet. The role of sleep in learning and memory. *Science*, 294(5544):1048–52, 2001.

M. Massimini, F. Ferrarelli, R. Huber, S. K. Esser, H. Singh, and G. Tononi. Breakdown of cortical effective connectivity during sleep. *Science*, 309: 2228–32, 2005.

M. A. Menza, K. R. Kaufman, and A. Castellanos. Modafinil augmentation of antidepressant treatment in depression. *J Clin Psychiatry*, 61(5):378–81, 2000.

L. M. Mukhametov. Sleep in marine mammals. *Experimental Brain Research*, 8:227–38, 1984.

L. M. Mukhametov. The absence of paradoxical sleep in dolphins. In W. P. Koella, F. Obal, H. Schulz, and P. Visser, editors, *Sleep 86*, pages 154–6, New York, 1988. Gustav Fischer.

L. M. Mukhametov. Unihemispheric slow wave sleep in the brain of dolphins and seals. In S. Inoué and a. A. Borbély, editors, *Endogenous Sleep Substances and Sleep Regulation*, pages 67–75, Tokyo, 1985. Japan Scientific Societies Press.

E. J. Mulder, W. H. Linssen, J. Passchier, J. F. Orlebeke, and E. J. de Geus. Interictal and postictal cognitive changes in migraine. *Cephalalgia*, 19(6): 541, 1999.

B. A. Pearlmutter. Gradient calculation for dynamic recurrent neural networks: a survey. *IEEE Transactions on Neural Networks*, 6(5):1212–28, 1995. doi: 10.1109/72.410363.

F. Pineda. Generalization of back-propagation to recurrent neural networks. *Physical Review Letters*, 19(59):2229–32, 1987.

S. Renals and R. Rohwer. A study of network dynamics. *Journal of Statistical Physics*, 58:825–48, June 1990.

J. J. Ross. Neurological findings after prolonged sleep deprivation. *Arch Neurol*, 12:399–403, 1965.

T. C. Roth, II, J. A. Lesku, C. J. Amlaner, and S. L. Lima. A phylogenetic analysis of the correlates of sleep in birds. *Journal of Sleep Research*, 15:395–402, 2006.

T. J. Sejnowski. Sleep and memory. *Current Biology*, 5(8):832–4, Aug. 1 1995.

T. J. Sejnowski and A. Destexhe. Why do we sleep? *Brain Research*, 886(1-2):208–23, Dec. 15 2000.

J. T. Shurley, E. A. Serafetinides, R. E. Brooks, R. Elsner, and D. W. Kenney. Sleep in cetaceans: I. the pilot whale, globicephala scammoni. *International Journal of Psychophysiology*, 6:230, 1969.

J. M. Siegel. The incredible, shrinking sleep-learning connection. *Behavioral and Brain Sciences*, 28:82–3, 2005.

J. M. Siegel. Do all animals sleep? *Trends in Neuroscience*, 2008. In press.

P. Y. Simard, J. P. Rayzs, and B. Victorri. A stability constraint for fixed points in recurrent networks. Technical report, Department of Computer Science, University of Rochester, 1990.

P. Y. Simard, J. P. Rayzs, and B. Victorri. Shaping the state space landscape in recurrent networks. In *Advances in Neural Information Processing Systems 3*, pages 105–12. Morgan Kaufmann, 1991.

C. Smith. Sleep states and memory processes in humans: procedural versus declarative memory systems. *Sleep Medicine Reviews*, 5(6):491–506, 2001.

O. Tchernichovski, F. Nottebohma, C. E. Ho, B. Pesaranb, and P. P. Mitra. A procedure for an automated measurement of song similarity. *Behaviour*, 59:1167–76, 2000.

G. Tononi and C. Cirelli. Sleep function and synaptic homeostasis. *Sleep Med. Rev.*, 10:49–62, 2006.

A. M. Turing. Computing machinery and intelligence. *Mind*, 49:433–60, 1950.

R. P. Vertes and J. M. Siegel. Time for the sleep community to take a critical look at the purported role of sleep in memory processing. *Sleep*, 28(12): 1228–9, Dec. 2005.

G. W. Vogel. An alternative view of the neurobiology of dreaming. *American Journal of Psychiatry*, 135: 1531–5, 1978.

M. P. Walker. A refined model of sleep and the time course of memory formation. *Behavioral and Brain Sciences*, 28(1):51–64, 2005.

M. P. Walker, T. Brakefield, A. Morgan, J. A. Hobson, and R. Stickgold. Practice with sleep makes perfect: Sleep-dependent motor skill learning. *Neuron*, 35:205–11, 2002.

A. Wirz-Justice and R. H. V. den Hoofdakker. Sleep deprivation in depression: What do we know, where do we go? *Biol Psychiatry*, 46:445–53, 1999.

S. Zapperi, I. K. Baekgaard, and H. E. Stanley. Self-organized branching processes: Mean-field theory for avalanches. *Physical Review Letters*, 75:4071–4, 1995.