Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Physiology - Paris 98 (2004) 507–529
Journal of Physiology Paris
www.elsevier.com/locate/jphysparis

Charting epilepsy by searching for intelligence in network space with the help of evolving autonomous agents

Elan L. Ohayon [a,b,*], Stiliyan Kalitzin [b], Piotr Suffczynski [b,c], Frank Y. Jin [d], Paul W. Tsang [d], Donald S. Borrett [e], W. McIntyre Burnham [a], Hon C. Kwan [d]

[a] *University of Toronto Epilepsy Research Program, Institute of Medical Science, Medical Sciences Building, Room 4306, University of Toronto, Toronto, Ont., Canada M5S 1A8*
[b] *Stichting Epilepsie Instellingen Nederland, Heemstede, The Netherlands*
[c] *Laboratory of Medical Physics, Warsaw University, Warsaw, Poland*
[d] *Department of Physiology, University of Toronto, Toronto, Canada*
[e] *Division of Neurology, Toronto East General Hospital, Toronto, Canada*

Abstract

The problem of demarcating neural network space is formidable. A simple fully connected recurrent network of five units (binary activations, synaptic weight resolution of 10) has $3.2 * 10^{26}$ possible initial states. The problem increases drastically with scaling. Here we consider three complementary approaches to help direct the exploration to distinguish epileptic from healthy networks. {1} First, we perform a gross mapping of the space of five-unit continuous recurrent networks using randomized weights and initial activations. The majority of weight patterns (>70%) were found to result in neural assemblies exhibiting periodic limit-cycle oscillatory behavior. {2} Next we examine the activation space of non-periodic networks demonstrating that the emergence of paroxysmal activity does not require changes in connectivity. {3} The next challenge is to focus the search of network space to identify networks with more complex dynamics. Here we rely on a major available indicator critical to clinical assessment but largely ignored by epilepsy modelers, namely: behavioral states. To this end, we connected the above network layout to an external robot in which interactive states were evolved. The first random generation showed a distribution in line with approach {1}. That is, the predominate phenotypes were fixed-point or oscillatory with seizure-like motor output. As evolution progressed the profile changed markedly. Within 20 generations the entire population was able to navigate a simple environment with all individuals exhibiting multiply-stable behaviors with no cases of default locked limit-cycle oscillatory motor behavior. The resultant population may thus afford us a view of the architectural principles demarcating healthy biological networks from the pathological. The approach has an advantage over other epilepsy modeling techniques in providing a way to clarify whether observed dynamics or suggested therapies are pointing to computational viability or dead space.

Keywords: Seizures; Autism; Synchrony; Graph theory; Close return; Recurrent neural network; Robot; Artificial life; Computer modeling; Genetic algorithm; Self-organization

1. Introduction

1.1. Networks and behavior

In framing the relation of network activity to behavior there are several important approaches and conceptual breakthroughs on which a theoretical neuroscientist might depend. In *The Organization of Behavior*, Hebb forwarded mechanisms by which persistent activity and use-dependent

* Corresponding author. Address: University of Toronto Epilepsy Research Program, Institute of Medical Science, Medical Sciences Building, Room 4306, University of Toronto, Toronto, Ont., Canada M5S 1A8. Tel.: +1 416 978 6381; fax: +1 416 971 2433.
E-mail address: ohayon@chass.utoronto.ca (E.L. Ohayon).

0928-4257/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jphysparis.2005.09.018

modifications could underwrite memory and behavior [22]. The connectionist modeling movement formalized algorithms that demonstrated ways in which networks could perform a range of categorization and behavioral tasks through algorithmic alterations in connection weights [23,60] However, recurrence was a central theme for Hebb and is undoubtedly a prominent feature of brain activity and whereas Hebb envisioned reverberating circuits these initial connectionist models were predominantly feed forward. Even those models that astutely incorporated recurrent connections and feedback between nonlinear units, nonetheless relied on equilibrium, with activity settling at fixed points, as a means of representation [24,28]. Certainly the importance of dynamics in physiology, network theory and self-organization has long been recognized [16,18,55,78]. Significant progress has been made in incorporating the dynamics of networks into connectionism, neuroscience and the cognitive sciences [14,54,88]. Nonetheless, the fundamental principles relating network structure to the dynamics of neural activity remain elusive. How these architectural principles and the emergent activity translate to behavior is even more confounding.

1.2. Networks and dynamics

The challenge then is to relate the structure of networks to the dynamics they can support. A recent indication that some of the most basic questions regarding network architecture have yet to be answered is the extraordinary increase in attention to general network theory [1,3,44,83]. Grounded in graph theory, these studies have brought to light just how fundamental network connectivity is to the functioning of so many of the systems that surround us. Although some of this work has considered dynamics [68,71] it is more often the case that network connectivity and network behavior are conflated. Most studies to date relate to the statistics of network linking and the dynamics of structure change rather than the dynamics of network activity. It is generally assumed that activity patterns will follow apparent network architecture and that current measures of structure will automatically inform us about functionality. Yet, the complexity of the patterns of activity carried in neural networks and the behaviors they support may stand in a non-trivial relation to the underlying network architecture. As will be demonstrated in this paper, complex network dynamics can occur in even small networks long before issues of large-scale connectivity arise.

The fact that the connection between activity, architecture and behavioral space remains obscure is not altogether surprising. Characterizing an ensemble's behavior and predicting its dynamical repertoire from structure is particularly difficult if not analytically intractable. In theoretical neuroscience analytical approaches to non-linear neural networks are often restricted to two cells [15]. Numerical approaches can overcome some limitations as seen in their application to systems like cellular automata [34,87]. But even these cases are often single-dimension and highly discretized systems that can take years to fully explore. Moreover, both cellular automata and recent mainstream network theorists tend to work with binary connectivity. When continuous connectivity is considered, as is the case of neural systems, a brute numerical approach becomes even less feasible given present computational resources. For example, a simple recurrent network of five interconnected units has $3.2 * 10^{26}$ possible initial states.[1] Iterating even such a simple model until all initial states are tested (1 ms per test) for all possible networks would take approximately $3.2 * 10^{25}$ s, or several million times the current estimates of the age of the universe (10–15 billion years). These estimates are not comforting given that most neural models are much more complicated and that having a general sense of the network architecture space and its relation to activity is likely essential to an understanding of the behavior of neural networks.

The dynamics of neural activity and the changing of relations of weights in complex networks underlie some of the most fundamental questions we need to answer if we are to understand activity in the nervous system with respect to perception, learning and autonomous behavior. For example: What is the likelihood that a network's activity will die out and settle on a fixed point? How likely or prone is a network to enter a limit cycle and become locked into periodicity? What is the full taxonomy of complex network behaviors and where are the various categories placed in the space of all possible networks? What is the connection between network structure, network dynamics, and an organism's interaction with the world? What role do these various types of dynamics play in physiological systems?

1.3. Dynamics and epilepsy

Answers to these basic questions may also underlie some of the most intractable clinical problems facing neuroscience. When network dynamics go wrong the effects can be devastating. Epilepsy is perhaps the paradigmatic case in which changes in network dynamics can have acute effects on a gamut of cognitive and behavioral phenomena. A disorder that is estimated to affect 1–3% of the population (varying by social and geographic position), epilepsy can be brought on by perceptual triggers, physiological stimuli, stress, physical trauma, pharmacological and genetic factors. Approximately 60–70% of cases are idio-

[1] Calculation of number of possible network evolutions for a five unit network ($n = 5$): *Network space*: there are 25 synapses (n^2). Assuming a synaptic weight resolution (w) of 10 there are 10^{25} possible network configurations (w^{n^2}); *Activation space*: Assuming even binary activation each network can have 32 (2^n) possible activation combinations. The number of network and activation combinations gives us a total of $3.2 * 10^{26}$ possible initial states. *Behavioral space*: Assuming the dynamics of a network can be assessed within 100 iterations, the total number of computational iterations needed to assess all the possible networks will be $3.2 * 10^{28}$.

pathic or cryptogenic, that is to say having no known or identified cause, and an estimated 20% of cases are currently intractable [8,21,65]. Yet despite a range of symptoms, causes and levels of responsiveness to treatment, the unifying, almost definitional, factor seems to be changes in neural dynamics.

Recent reviews by Lopes da Silva et al. [38,39] provide an overview of epilepsy as a dynamical disease and itemize several ways in which the transitions might occur including bifurcations in the system dynamics and switching between preexisting attractors [72]. To answer how phenomena such as hypersynchronization and broad fluctuations in activity come about we need to understand the relation of these dynamical patterns to the underlying neural structures. In recognizing the connection with dynamics it follows that epilepsy may also provide a general portal into the dynamics of the brain and the mechanism underlying the non-seizure conditions, that is behavior.

1.4. Epilepsy and networks

Much of the search for fundamental mechanisms of epilepsy assumes that the culprits are pathological cell properties (including local synaptic properties) or an imbalance in cell types [11,45,67]. An alternative to this cell-centric "epileptic neuron hypothesis" is the "epileptic aggregate" theory. The possibility thus arises that networks with normal cells might exhibit epileptic phenomena due to malformed connectivity. Outstanding questions relating to the mechanisms that bring about these massive dynamical changes in epilepsy are analogous, if not identical, to those posed earlier about networks in general. The intractability of the network activity question and epilepsy are strongly related and both depend on linking architecture and activity. Whether in graph theory or clinical epilepsy, the architecture, dynamics and emergent functional behavior of networks are all closely interwoven.

1.5. Embodied modeling

The connection between network structure, activity dynamics and behavior is not always obvious. Although many models incorporate extensive physiological detail it is difficult to know whether they capture the most general system dynamics or are just a particular correlation to observed brain dynamics that the field or given experimenter find interesting. For example, are occurrences of synchrony computationally necessary for behavior or simply a correlated transition? To this end, it is a major impediment that, to date, major oscillatory ensemble models have been both figuratively and literally disconnected from the world [74]. Recently there has been a growing recognition by a range of modelers for the need to relate neural dynamics to behavior and the value of doing so in embodied modeling. Important progress is now being made in neuroscience-related studies of robots (for an overview see [46,61,84]). The interrelation of dynamical and behavioral indicators of seizures suggests that such models might similarly help relate inquiries into the fundamental aspects of system dynamics to epilepsy. Given this connection between embodiment, behavior and seizures it is not a coincidence that the earliest pioneers of cybernetics and robots (hereafter termed autonomous agents) explicitly had interest in neuroscience and epilepsy [80–82,86]. Making the connection between activity and embodied behavior is essential to understanding epilepsy. It is important to remember that as much as diagnosis of epilepsy relies on biophysical measures such as electroencephalograms and recordings from implanted electrodes, ultimately it is behavioral indicators that are at the heart of the clinical diagnosis and as such are the canonical test for any model or cure.

1.6. Charting network space

The purpose of this paper is to explore techniques for addressing the relation of recurrent architectures to neural activity and ultimately behavior. In particular, we focus on relating structure to a taxonomy of network activity patterns in standalone and embodiment recurrent network models. In so doing, we ask how regions of network connectivity space might be charted and related to both healthy behavior and pathological phenomena such as epilepsy. This is the first work in embodied network modeling of epilepsy and more generally the escape from pathology in autonomous agents.

2. Methods

2.1. Recurrent neural network

Both the standalone and embodied modeling described in this paper used fully interconnected recurrent neural networks including unit self-feedback (see Fig. 1, dotted box). The recurrent networks were composed of five units, with each unit receiving five input connections ($w_0 \ldots, w_j$) for a total of 25 network connections. The connection strength between units was represented by a weight that ranges from -3 to $+3$ with negative values representing inhibition. The total input (E_i) for a unit is the weighted sum of the activations of the corresponding afferent units (S_j) such that:

$$E_i = \sum_j w_{ij} S_j \tag{1}$$

The sum input (E_i) is then fed through a non-linear activation function which yields the new activation value (S_i) for that unit (For review see: [5,15,23,46,60]). A unit's activation is also the output of the unit for the purpose of the next iteration. We tested both hyperbolic tangent (sigmoid) and radial basis functions (RBF) as activation functions:

$$S(E_i) = e^{(-E_i)^2} \quad \text{radial basis activation function} \tag{2.1}$$

$$S(E_i) = \tanh(E_i) \quad \text{sigmoid activation function} \tag{2.2}$$

In the first set of experiments we generated networks with all connections randomized (weight values ranging from −3 to 3), random initial unit activity levels (−1 to 1 for the sigmoid networks; 0–1 for RBF) and then iterated network activity for 1000 time steps.

2.2. Naive Hebbian plasticity

We also performed some preliminary studies of the effects of plasticity on dynamics using a simple Hebbian [22] rule:

$$w_{ij} = w_{ij} + ((S_i S_j) w_{ij} r) \quad \text{naive Hebbian plasticity} \tag{2.3}$$

For these plastic networks, the change in weights is a simple function of the activity of the two connected cells modulated by a rate factor (r). Network weights were left either to grow indefinitely or constrained within bounds. It is important to note that in the sections of this study concerned with the categorization of random networks, no plasticity was involved and weights were held steady throughout the run. In the embodied modeling portion weights were changed through a genetic algorithm but held steady for any particular individual.

2.3. Categorization of network dynamics

The activity of networks was categorized using a variation of the close returns algorithm [42,10,32,47]. In the current implementation the final state of all units in the network was compared to every other state in the evolution to see if the value returned within a given range. This window (ε) was set as a fraction of the difference between maximal and minimal possible activation values (e.g., 1% of $|\text{Max} - \text{Min}|$). If all units in a network returned precisely to the same state in two successive iterations (i.e., absolute difference for all units = 0) then the network had settled on a "fixed point" and would remain frozen in that state. If all units simultaneously returned to the same states in iterations separated by an interval p then the network would oscillate endlessly and was said to settle on a "limit cycle" with period p and frequency $f = (1/p)$. An occurrence in which the absolute difference in activation values ($|E_{i(\text{final})} - E_{i(t)}|$) dropped below the window threshold (ε) but not to 0 was termed a close return. For those cases in which activations did not return precisely to the same state but nevertheless saw all units simultaneously achieve a return within ε of the final iteration values, the network was categorized as exhibiting "close returns" (unstable periodic orbits). These "close return" networks were characterized by repetitive but not precisely repeating oscillations. Finally, networks that never returned within the window were marked as "unspecified" or "turbulent" and were characterized by behavior that generally appeared random.

2.4. Embodied modeling

For the embodied modeling portion of the experiment we connected the recurrent networks to a mobile robot and evolved network structures that could perform a simple obstacle avoidance task in the world.

2.4.1. Autonomous agent network architecture

The network portion of the model was analogous to a central nervous system (CNS) and the attached robot the body. The network had three layers including (i) an input layer comprising of eight feed-forward units each corresponding to a sensory transducer, (ii) a computational recurrent layer with five fully recurrent computational units and (iii) a motor output layer consisting of two motor output units. Each recurrent unit thus received eight sensory inputs in addition to the five recurrent inputs described earlier in the case of standalone networks. Activity in the network responded in real-time to the sensory input and recurrent activity. Output values controlled robot wheel rotation direction and speed. Fig. 1 is an illustration of the network and main robot body features.

2.4.2. Robot chassis and sensory transducers

For the robot body we used a modified version of a commercial miniature robot (Descartes model, Diversified Enterprises, Santa Barbara, CA, USA) equipped with two motors that rotated two independent wheels in the clockwise and counter-clockwise directions. The robot body was round and wheel orientation was fixed. Turns were accomplished by asymmetric rotation speeds. The robot was tethered to a personal computer running the recurrent neural network program. All network processing was done on the attached personal computer. A processor on board the robot was used solely for sensory and motor communication. The incoming sensory information and outgoing wheel instructions were relayed via a serial port. The communication and network modeling software were developed in our lab using Labview (National Instruments, Austin, TX, USA). Wheel speed update and sensory readings were performed in the same cycle. Network activity was computed online with an average interface cycle time—including sensory input, neural network computation and command output—of approximately 80 ms (~12.5 Hz). The limiting factor in the update was the serial communication.

Sensory information was relayed via eight transducers (right side of Fig. 1) including: two bumper sensors (LB, RB), four cadmium sulfide photocell sensors (PS1–PS4), and infrared encoders to monitor wheel rotation (RW, LW). The bumpers were located on each side of the unit and registered binary tactile perception. Front contact was indicated when the left and right were simultaneously triggered. There was a small tactile "blind-spot" in the rear. The four photosensitive detectors were located on the front, rear, right front and left-front orientations. Four light emitting diodes (LEDs) matched the photosensitive

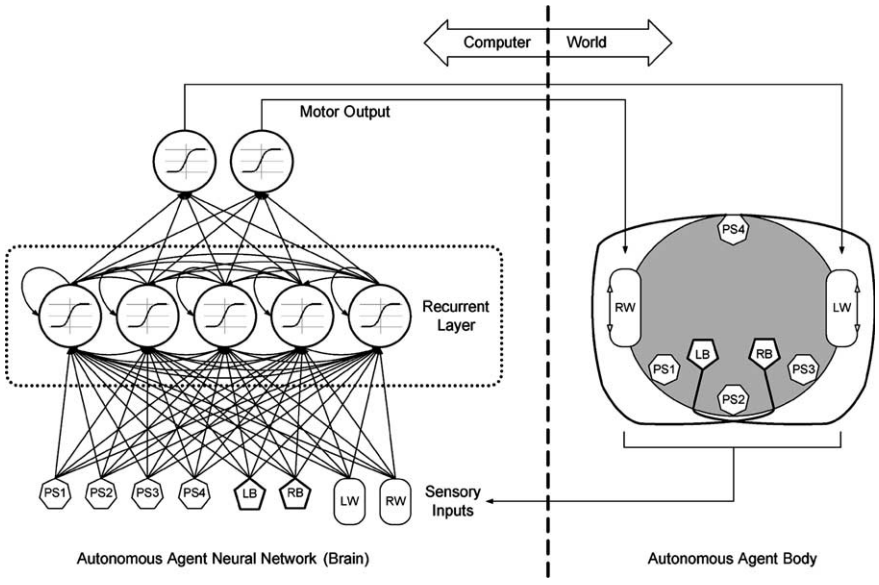

Fig. 1. Autonomous agent network and body diagrams. The left side of the figure is a schematic of the neural network portion of the autonomous agent model run on the computer. The right side of the figure is a schematic of the robot body placed in the world. The central recurrent layer of the network, indicated by the dotted box, corresponds to the network used in the standalone models of this paper. The sensory transducers on the body include right and left bumper sensors (RB; LB), four photoreceptors (PS1–PS4) and wheel rotational encoders (RW; LW). The area pointing down with the highest density of photoreceptors and where the bumpers meet is considered the front. The eight sensory transducers relay information to the network's corresponding sensory inputs. These are fully connected to the recurrent layer. The five recurrent units feed back onto each other as well as being fully connected to two motor output units. Unit activity is determined by taking each input to the unit and multiplying it by the respective weight. This weighted sum is then fed through a sigmoidal function which in turn yields the activation value and output for that unit. The motor output units set the speed for the two independently moving wheels. Network output was a real-time response to sensory input and internal recurrent activity.

detector position. The photoreceptors thus picked up the reflection of the LEDs and were set to return a value ranging from 0 to 255 as the robot approached an obstacle. The wheel encoders are analogous to proprioceptors and monitored rotational distance, returning a value between 0 and 65,535. All sensory readings were normalized to values between 0 and 1 and each of the five network processing units received information from all sensory inputs.

2.4.3. Learning environment and evolution

The robot body was placed in an arena with obstacles. Fig. 2 illustrates the experimental setup. Each individual was represented by a different network weight matrix and was assessed in the arena for times ranging between 20 s and 2 min (or approximately 250–1500 activity iterations). All sensory and neural activity was recorded. Evolution corresponded to changes in weights and was accomplished using a genetic algorithm.

2.4.4. Genetic algorithm

We used a genetic algorithm to evolve autonomous agents so that they could navigate the environment (for a review of genetic algorithms theory see [26]). In our implementation, the genetic code for an individual is its weight matrix and the evolutionary process acted directly on network structure [46]. Initial weights for the first 20 individuals were randomly generated. After all individuals in a given generation were tested in the arena, fitness scores were used to determine the individual network structures for the next generation. Selection was made using the roulette wheel method. According to this method the probability of a network's architecture (its set of weights) being passed on to the next generation is proportional to its behavioral success as compared to the sum of all other fitness scores across the generation. Network weights also underwent a mutation process which consisted of adding noise to each weight from an inverse normal random distribution with a mean of 0 and SD = 0.1. For each generation 10% of the most successful networks structures were automatically passed on to the next generation without alteration (elitism). The process was repeated until all individuals in a population could move and recover from collisions by changing direction. Although we initially tested other genetic mechanisms such as crossover, this noise mutation and performance-based selection proved sufficiently effective in generating behavior for the purposes of this study.

2.4.5. Fitness function

The fitness scores used in the genetic algorithm were calculated on-line with wheel rotation distance, direction of movement and collision all being considered. Instantaneous fitness was thus a function of wheel rotation direction:

$$\text{fitness}_t = \left(\frac{|(D_{L(t)} + D_{R(t)})|}{1 + |(D_{L(t)} - D_{R(t)})|} + \frac{|(D_{L(t)} - D_{R(t)})|}{1 + |(D_{L(t)} + D_{R(t)})|} \right) * \text{rf} \tag{3.1}$$

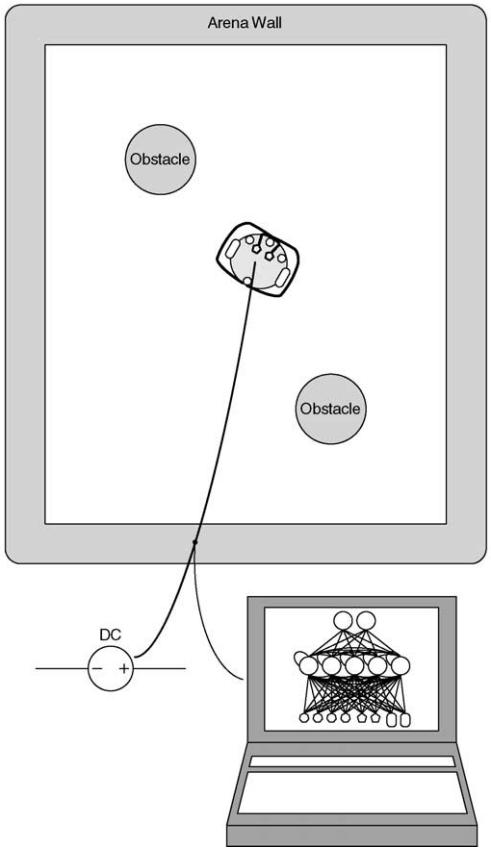

Fig. 2. Autonomous agent experimental setup. The robot was connected to a computer running the model neural network and placed in an arena with various obstacles. The sensory input and motor output were communicated via a serial line which tethered the robot to the computer. Attached via the tether was also a DC power source that supplied power to the two motors driving the wheels.

D_L and D_R are left and right wheel distances covered in the given iteration. The first ratio compares total distance $|(L+R)|$ moved with the difference between wheels $|(L-R)|$ thereby rewarding the most points for fast and balanced moves either forward or backward. The second ratio rewards the fast but opposite wheel movements. Straight movements and sharp turns thereby contributed most to fitness. The reward factor (rf) encouraged forward movements, where

$$\begin{aligned} &\text{If } (D_{L(t)} * D_{R(t)}) > 0 \quad \text{Then rf} = 3 \\ &\qquad\qquad\qquad\quad \text{Else rf} = 1 \end{aligned} \tag{3.2}$$

The directional component of the fitness function is shown in Fig. 3. The fitness measure was also adjusted to reflect a collision penalty (cp) raised to the power of number of collisions (nc):

$$\text{fitness}_{\text{cumulative}} = (\text{cp}^{\text{nc}}) * \sum_t \text{fitness}_t \tag{3.3}$$

Several other fitness measures were tested. We settled on the above function because of its efficiency in promoting forward movement and turning while discouraging collisions with walls and mid-arena obstacles.

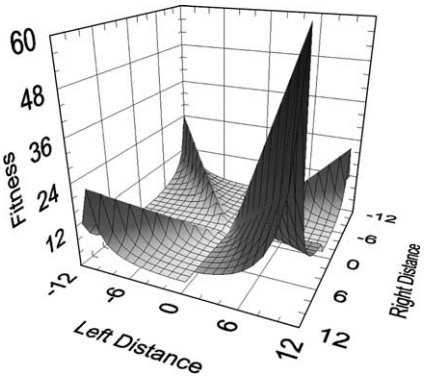

Fig. 3. Fitness function used in evolutionary algorithm. Fitness was a function of left and right distance traveled in wheel encoder units. Wheel movements that enhance forward or backward movement or turning in an efficient manner were rewarded proportional to speed and coordination. Movement with both wheels moving forward had the reward increased by a factor of 3. The function also has an additional dimension (not shown) in which fitness is reduced for instances of collision.

3. Results

3.1. Distribution of dynamics in populations of random networks

The first part of the study was aimed at establishing the distribution of dynamics in random networks of a given size and given intrinsic unit properties. In other words, to get a sense for how variable the distribution of dynamics might be if only network connectivity was varied. As a first step we generated 1000 random networks. Each of these five-unit, fully recurrent, networks was started with random activity conditions and run for 1000 iterations. We then categorized the networks according to the close returns algorithm described in the methods section.

Fig. 4 shows examples of (i) overlapping traces and (ii) intensity plots of activity in sigmoid activation function networks. The randomly connected networks' activity dynamics were categorized as (a) fixed point, (b) periodic oscillations, (c) close returns or (d) turbulent. We repeated the experiment for RBF functions.

Under both activation functions the majority of weight patterns were found to result in neural assemblies with limit-cycle oscillatory behavior ($71.1 \pm 1.4\%$ for sigmoid; $43.9 \pm 1.6\%$ for RBF). Table 1 summarizes the distribution of dynamics for both the sigmoid and RBF populations.

3.2. Distribution of limit cycle periods

Next we examined the distribution of the periods for those networks categorized as having limit cycle dynamics. Observed periods were as long as 90 iterations for the sigmoid networks and as high as 135 iterations for the RBF networks. The large outer graph in Fig. 5 shows this distribution of limit cycle periods for the sigmoid networks population. Over 95% of sigmoid networks exhibiting limit cycles had periods between 2 and 24 iterations. Though this

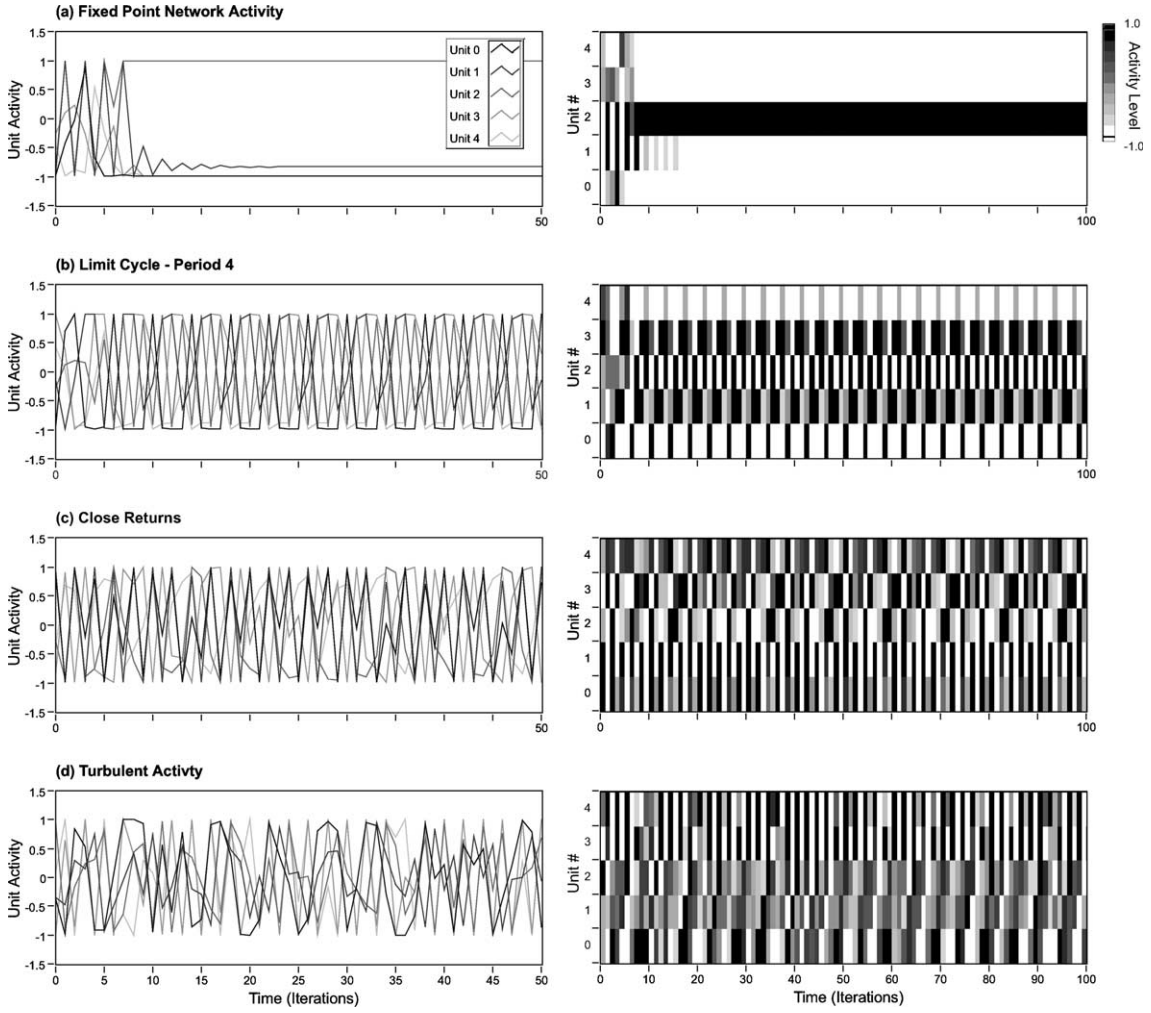

Fig. 4. Activity-dynamics categories in random networks. The left column panels are superimposed activity traces of all units in a given network. In the right column graphs each unit occupies a separate row and activity is represented by intensity (z scale). The left-hand traces are for the first 50 iterations to highlight the initial phase of the run and settling phenomena. The intensity graphs show the first 100 iterations for the same run and highlight any pattern that might eventually develop. The trace and intensity charts of a given row correspond to a dynamic category: (a) Fixed point: all units settle on a fixed value; (b) Limit cycle—period 4: the units eventually settle onto a periodic pattern, in which unit activation values repeat every forth iteration; (c) Close returns: The network has a pattern that comes close to repeating but never repeats perfectly. The close to periodic nature of the pattern can be discerned in the later portion of the intensity graph; (d) Turbulent activity: Networks with these dynamics do not have a readably identifiable activity pattern.

Table 1
Distribution of network activity-dynamics by category network activity was categorized into four types of dynamics

Activation function	Dynamics category	Count	% of Population	SE%
Sigmoid	Fixed point	100	10.00	0.95
	Limit cycle	711	71.10	1.43
	Close returns	134	13.40	1.08
	Turbulent/other	55	5.50	0.72
	Total	1000	100.00	
RBF	Fixed point	27	2.70	0.51
	Limit cycle	439	43.90	1.57
	Close returns	185	18.50	1.23
	Turbulent/other	349	34.90	1.51
	Total	1000	100.00	

This table shows the distribution across the categories as a percentage of population for sigmoid and radial basis activation functions. Each population consisted of 1000 networks. ($\mathrm{SE}\% = \sqrt{(p(100-p)/n)}$).

steep drop with a long tail may look like an x^{-a} power law relation, closer examination reveals a preference for certain periods and an interleaving in the probability of odd and even periods.

3.3. Dynamics and plasticity

We also performed preliminary tests to explore changes in activity distribution resulting from the introduction of a naive Hebbian algorithm. The addition of dynamic changes in weights found 100% of networks converging onto fixed point activity within 1000 iterations. Although these plastic networks all settled to fixed points, they tended to sweep through several activity regimes in a manner not seen in the non-plastic counterparts. For example, the activity records for plastic networks thus included chirp-like frequency fluctuations—a phenomenon never

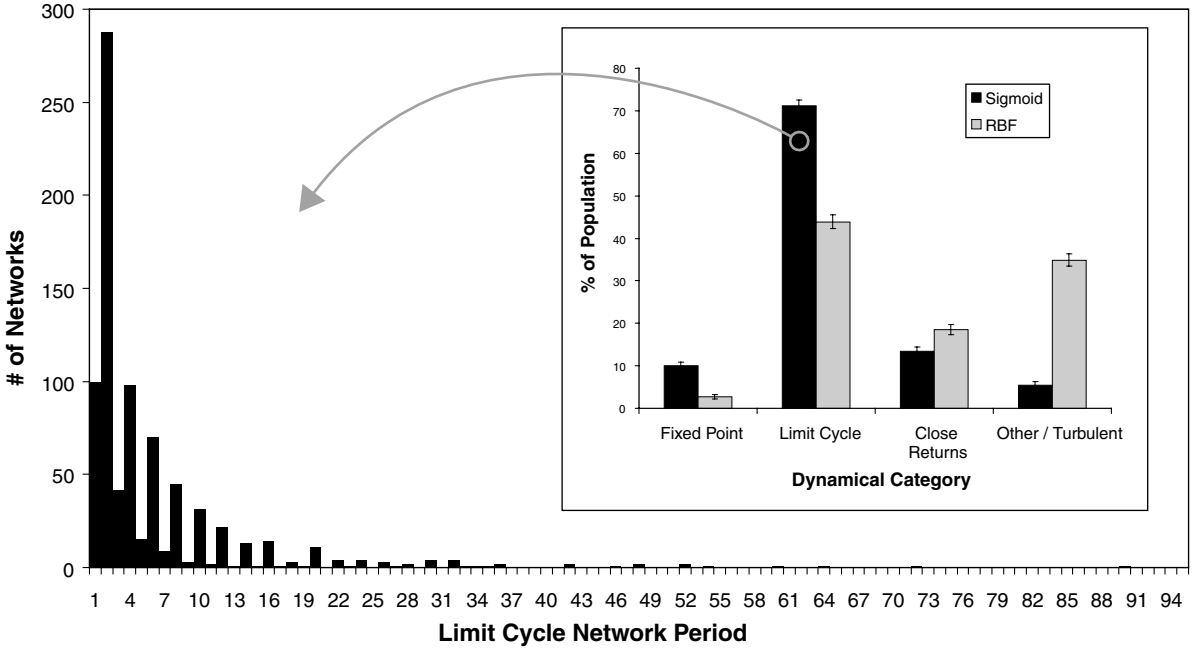

Fig. 5. Distribution of dynamics in random networks. Insert graph summarizes the distribution of dynamic categories in the sigmoid and RBF random network populations. In both cases the majority of networks exhibited limit-cycle dynamics. The outer graph is a breakdown of the limit cycle category in the sigmoid population. Specifically, this histogram shows the number of networks exhibiting each limit cycle period (in iterations). Fixed point networks are included as period 1. Most networks oscillated with a period of 2 iterations and 95% of networks exhibited periods between 2 and 24 iterations. A small proportion of networks showed long periods (>24 iterations) with a maximum recorded period of 90 iterations.

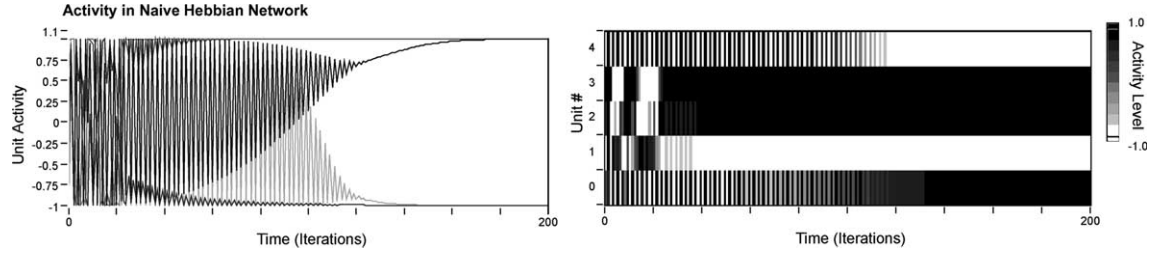

Fig. 6. Activity dynamics of a plastic network. Networks were run with the addition of a naive Hebbian plasticity rule. The ongoing changes in weights often resulted in networks covering a range of activity patterns. These networks quickly tended to synchronize and ultimately all settled to fixed point dynamics. The superimposed activity traces in the left panel highlight an example of a network's progression from turbulent activity through synchrony and ultimate convergence to a fixed point all within 200 iterations. The right panel clarifies the progression of activity for the individual units and highlights the transitional synchrony between the first and fifth unit for the same data.

observed in the static connectivity networks. Fig. 6 illustrates an example of the tendency of plastic networks with initial close returns or turbulent activity to quickly fall into a synchronous state and then settle to a final fixed point. The settling of activity onto fixed points was a direct result of the naive Hebbian algorithms tendency to push connections to extremes such that the weight growth eventually saturated the input in either the excitatory (+) or inhibitory (−) directions.

3.4. Complex network dynamics and intermittency

As a next step in understanding complex network dynamics we relied on the categorization algorithm to focus our search. The initial review of the non-plastic populations categorized as having close returns or turbulent activity included networks exhibiting dynamics that may hold important relation to those seen in the biological counterparts. Fig. 7 illustrates the activity of an RBF network with units that displayed an intermittent activity pattern, switching between a laminar quiescent phase and turbulent activity. This intermittency [53,63,70] is highly reminiscent of paroxysmal activity patterns seen in humans and other organisms. Variants of this intermittency patterns were also seen in sigmoid networks categorized as having close returns or turbulent activity. The variant included continuous switches between turbulent behavior and synchronized activity.

Note that the unit properties and connectivity were not plastic and did not change over time. The changes in acti-

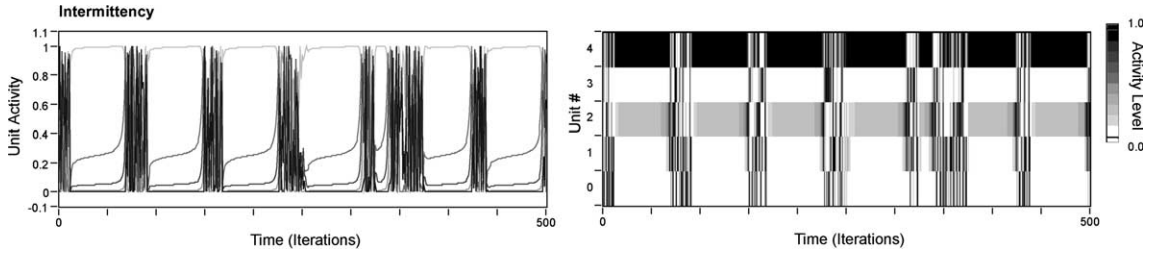

Fig. 7. Intermittent activity in a recurrent network. The left panel shows the superimposed activity traces of an intermittent network. The chart highlights the paroxysmal nature of the switch between turbulent and laminar phases. The right panel is an intensity graph of the same data showing the relation of activity between the units. Note that notwithstanding the fact that weights were not changed during the simulation there are clear changes in the dynamical output. This particular network thus demonstrates that the emergence of intermittent paroxysmal activity need not require changes in synaptic connectivity.

vation phases seen in these intermittent networks were completely due to the propagation of activity. The categorization thus proved to be an important technique for quickly demarcating complex dynamics and opens the possibility for identifying sub-classes of dynamics that can be potentially correlated with network structure. A detailed analysis of the statistical properties of this intermittency is presented in a forthcoming paper [51].

3.5. Evolution of behavior in autonomous agents

The third part of this study was designed to further focus the search of network space to those complex dynamics that relate to behavior. To this end, we equipped the same type of recurrent networks employed in the earlier sections with sensory inputs and motor outputs. We then evolved these networks to perform simple roaming and avoidance tasks. Here we report on the evolution of dynamics and behavior over generations in one such lineage.

3.5.1. Fitness evolution over generations

The first random embodied generations of 20 individuals showed a distribution of dynamics analogous to that seen in the disconnected random networks. The predominate neural activity patterns were fixed point and limit cycles with the addition of environmental noise. These neural dynamics were reflected in behavioral phenotypes in which individuals did not move or exhibited highly oscillatory activity of varying amplitudes with little to no motion across the arena. No individual in this first generation showed behavior that even remotely looked like roaming and avoidance.

Once the first generation was assessed, we began to apply the genetic algorithm described in the methods section. Panel (a) in Fig. 8 shows the evolutionary progression and rise in fitness measure over all individuals (Fig. 8a). Panel (b) plots the mean and standard deviation for these fitness measures across generations. Also included in (b) are the maximal and minimal fitness scores attained in a given generation. Note that the progression was not always

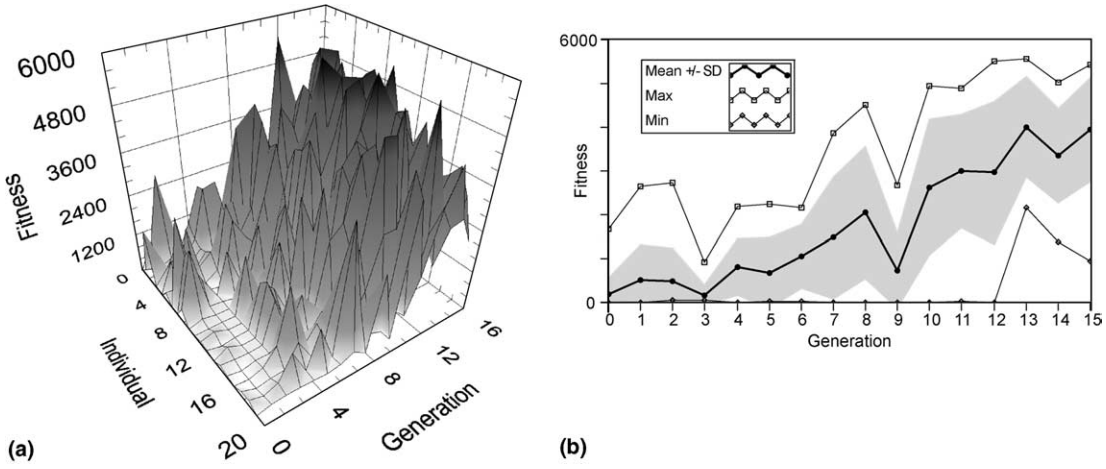

Fig. 8. Evolution in fitness of recurrent networks over generations. The left graph (a) illustrates the evolution of fitness in a population over 15 generations (X axis). Each generation includes 20 new individuals (Y axis). Fitness levels (Z axis) improved markedly over the course of the evolution. Fitness did not always rise from one generation to the next. The adjacent panel (b) charts the mean with standard deviation (gray area) across generations. The plot also includes the highest (max) and lowest (min) individual fitness scores achieved in a given generation. The entire population evolved out of the seizure within 14 generations from the initial set of 20 random individuals.

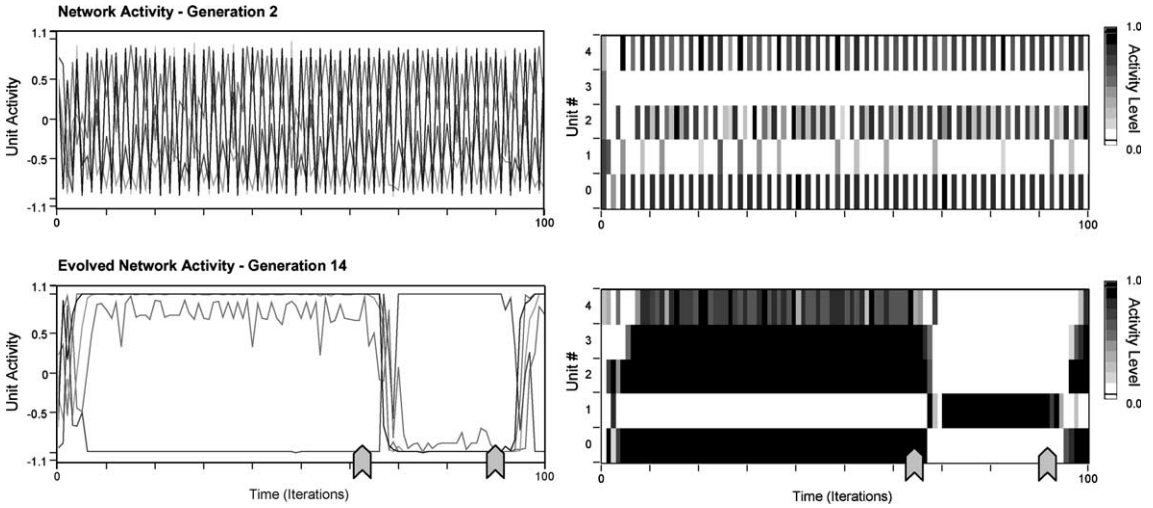

Fig. 9. Evolution of dynamics in autonomous agents. The left column panels are superimposed activity traces and the right panels are stacked intensity plots of recurrent unit activity. The graphs contrast a second generation's highly periodic neural activity (top panels) with the more complex activity seen in a fourteenth generation network (bottom panels). Chevrons on the bottom graphs of evolved network activity mark the time at which collisions are registered by a bumper. These collisions were followed by network reactions leading to behavioral changes in movement direction. In this case, the system's solution to extricating itself from obstacles is through reversal of all unit activity resulting in a physical reversal of the robot body.

in the direction of improvement as defined by the fitness function. On the whole though, as evolution progressed, the profile changed markedly as compared to the first generation. Within 14 generations the entire population was able to navigate the simple environment with almost all individuals exhibiting multiply-stable behaviors with few or no cases of default limit-cycle oscillatory motor behavior.

Examination of the various transducers and their correlation with behavior showed that the recurrent network relied chiefly on the bumpers and wheel encoders to navigate the environment and avoid obstacles. Although we ran the experiments in a dark environment, the ambient light from the LEDs biased the photoreceptors such that the change registered in approaching an obstacle was not sufficient to affect the network. This limitation could have been corrected by further normalizing the photoreceptor signal. However, in order to simplify the experiment and its interpretation, we opted to eliminate the photoreceptor input. Data reported in the results are from runs in which input is restricted to wheel encoders and bumpers.

The low mean fitness (189.1 ± 368.4) of the first generation was approximately 10% of the mean fitness later achieved by the evolved population (3973.8 ± 1130.3). The low first generation fitness corresponded closely to the motor movements reminiscent of seizures. The successful navigation seen in the evolved networks consisted of steady forward body motion with clear reversal and change of directions upon encountering a wall or mid-arena obstacle. The increase in these behaviors corresponded with a rise in fitness. Occasionally, following mutation, an individual or two in a given generation would fail to move or move only in reverse. The particulars of these behaviors usually dropped off by the next generation due to low probability of the responsible weight pattern being passed on by the evolutionary algorithm.

3.5.2. Evolving network activity dynamics in autonomous agents

The rise in fitness function and evolution of behavior was accompanied by an evolution in dynamics. The dramatic change in activity patterns is demonstrated in Fig. 9 which compares examples of neural activity traces of an individual from an early generation (generation 2) and a network with evolved behavior (generation 14). Whereas the networks of generation 0 exhibited mostly large amplitude, highly oscillatory activity, the activity in evolved networks was generally complex exhibiting both a range of frequencies and multistable responses as they interacted with the environment. The evolutionary algorithm thus reliably searched an exceedingly large network space and within 14 generations was able to focus on networks that not only exhibited complex dynamics but that could also accomplish a behavioral task.

4. Discussion

4.1. Network space and epilepsy as default

Epilepsy in its many forms is the archetypical dynamical ailment [38,39,72]. In trying to establish the underlying factors that lead to this condition much of the literature has focused on intrinsic changes to cells or the changes in cell type distributions [11,45,67]. However, the network models examined in this study show that when intrinsic properties are held fixed a broad range of network behaviors are still accessible through connectivity changes even in the sim-

plest of networks. The fact that connectivity patterns can underlie phenomena such as hyperexcited synchrony may not be surprising, but the ease at which they were found through random searches is surely notable. In particular, the suggestion that this class of networks and concomitant activity patterns may turn out to be the statistical norm and default scenario is disquieting. In the case of these deterministic non-plastic nets, once a network falls into a periodic state there is no way for it to break out. Moreover, some of the networks categorized as having nearly periodic orbits may have dropped into periodic limit cycle activity if allowed to run beyond the allotted 1000 iteration. As such, the number of limit cycle networks may have actually been underestimated. These observations turn the question of what causes epilepsy on its head. Instead of asking how epilepsy comes about they compel us to ask how recurrent neural ensembles ever manage to avoid this ubiquitous synchrony in the first place. That is, why are we not all epileptic, all the time?

4.2. Unbiased exploration of the relation of architecture to dynamics

In the introduction we drew attention to the analytical intractability and computational barriers hampering the establishment of a connection between the architecture of networks and the dynamics they exhibit. Although the models we present are simple and future work will require exploration of more complex models, the ability to study the distribution of major categories of dynamical behaviors gives some insight into the propensity of networks to get locked into limit cycles and offers a measure by which to examine the effects of network connectivity and intrinsic changes. An important feature of the network sampling approach presented in Section 3.1 is that it distinguishes itself from models based on known anatomical structures by exploring connectivity in a less biased manner. Rather than make assumptions regarding connections structure we search the space randomly. This approach does not preclude commencing with more restricted anatomical assumptions and indeed current work is exploring the dynamical propensities of spatial networks [52]. However, allowing for all possible connections proved productive in that we were able to establish some general distributions in the population despite the large number of possible network structures thereby providing a sense of the terrain and challenges biological networks face in evolution and development.

The findings on intermittency in these simple networks, as presented in Section 3.4, further underscores how such research might help elucidate the workings of real neural networks in their relation to activity and behavior. For example, the model clearly demonstrates that intermittent dynamics need not require changes to the underlying network architecture nor intrinsic unit properties. To think that such dramatic changes in activity can be supported by network activity propagation alone is highly intriguing and has direct implications to physiological recordings. For example, it is often assumed that paroxysm might be a reflection of changing connectivity (plasticity) or fundamental shifts in intrinsic cell properties, however this case is a potent counter-example that undermines the certainty of such assumptions [51].

4.3. Recurrence and universality in network modeling

Early connectionist models were mostly feed forward. Even those networks that incorporated recurrence had equilibrium as the general goal, in the sense that network activity was meant to settle on a fixed point [24,28]. In order to account for continuous interaction with the environment our network modeling focused on incorporating recurrence and achieving persistent activity [14,46,54,78]. The exclusion of cellular details to achieve computational simplicity may seem an ill-advised omission that undermines the applicability of any conclusions to the true workings of the biological network. However, the primary reason for simplification was not to speed computation or due to doubting the considerable effect that changes in intrinsic properties can have but rather to demonstrate the universality and generality of principles. Connectionist modeling of the sort applied in this study helps illustrate that some of the most salient questions facing neuroscientists can be addressed at this higher level. This sort of connectionist modeling thus allows for the exploration of the most fundamental parameters relating to neural dynamics. Furthermore, in abstracting to high-level features (inhibition, excitation, connectivity patterns, etc.) we specifically allow for the certainty that there are important mechanisms at the cellular level that have yet to be discovered.

4.4. Close return and the taxonomy of activation patterns

We chose a variant of close return analysis as a categorization algorithm primarily because it is well suited to quantify unstable periodic orbits and can distinguish them from limit cycles and turbulence in time series [42]. There are indications that it can be effective in analyzing normal and epileptiform EEG [10,32,47]. The technique was also found to correspond well to visual categorization and was computationally efficient as compared to other nonlinear dynamic approaches and Fourier-based methods. The intermittency case of Section 3.4 shows that these dynamics have yet to be classified into many related subcategories. The selection of close return for the analysis of network dynamics does not preclude the application of other dynamical categorization methods [35,51,53,63,70].

4.5. Categorization and distribution of dynamics in population

The difference between the activity of any two networks in a given activation function class was entirely attributable to differences in connectivity and initial states. The

distributions graphed in the Fig. 5 (insert) demonstrated that dynamics could be quickly parsed out even in the face of a combinatorial explosion and that the area taken up by each category could be reliably estimated. The findings imply that the behavior of a random network of a given size and activation function is, at the very least, statistically predictable. The reliability in which activity could be categorized suggests that the method can serve as a foundation for searching for correlations between activation patterns and gross architectural features.

4.6. Distribution of attractors within networks and considerations of multi-stability

Given the size of the sample, the predominance of limit cycles in the population was well established. However, there is the further issue of the number and size of attractors *per* network. It is important to note that the categorizations were performed across populations but assessed only once for each network. For example, reflection on the periodic sigmoid networks will show that all such networks have at least one additional attractor, namely the trivial fixed point case when all units are set to 0 as an initial condition. How many more attractors do these random networks typically have? If they have multiple attractors what is the distribution of the dynamic categories? How large are the basins of these attractors? In other words, how likely is it for activity to fall into any given attractor?

We performed some preliminary investigation of this question using all possible combinations of extreme activation (i.e., units were set to either max or min values) for the intermittency network of Section 3.4. For this particular case and set of initial conditions we were able to demonstrate that the network always returned to the same general activity pattern while not repeating itself. Thus, even if testing all initial states proves impractical either by analytical methods, linearization techniques or brute numerical approaches the question of the distribution of categories of dynamics for a given network is analogous to that of the distribution in a population and so is also approachable by the sampling techniques applied in this study. That is, just as we can sample and search the population for the distribution of dynamic categories, each network can similarly be explored through geometric, random or extreme seeding of initial states as well as through targeted searches such as genetic and learning algorithms.

This question of attractors becomes particularly important if we are interested in the formation of network properties during evolution, development, learning and pathogenesis. For example, it is possible that evolution in the embodied networks resulted in networks with an increase in the *size* of already existing basins of attraction. It is also possible that there was an increase in the *number* of attractors or change in the types of attractors. Analogously, we can inquire whether epileptic regimes are already present in the normal brain such that epileptogenesis merely increases the range of activations that will trigger a seizure or whether epileptogenic events actually introduce completely new attractors. Answering these questions and establishing whether the process might be reversed could mean the difference between providing temporary relief from seizures and finding the route to permanent cures.

4.7. Scale-free and clustered distributions—implication to network theory

The application of continuous networks versus binary networks may help bridge the gap between current network modeling trends and neuroscience. Analysis of the recurrent models also demonstrated that networks exhibiting complex dynamics and many of the critical questions listed in the introduction emerge in even these small scale systems. One need not focus on large "scale-free" networks of the size of the internet or human populations before complex and possibly scale-free activity is generated [1,3]. Moreover, these simple networks may help expose new distributions and principles that have yet to be identified in general network theory. For example, the distribution of limit cycle periods illustrated in Fig. 5 seems at first to be of a traditional scale-free form, but on closer inspection is shown to be highly non-monotonic. Preference in certain period lengths and the interleaving of probabilities alerts us to the fact that there are fundamental questions relating to limit cycle periodicity that cannot be simply dismissed in favor of focusing on the larger trend.

The peculiar distribution of limit cycles also highlights how important trends might go unnoticed in real data. The luxury of modeling with difference equations in the absence of noise helped expose this non-monotonic drop in the period length distribution. The interleaving of even and odd periods in and of itself is suggestive of the possibility that a given structure may entail predispositions to certain activation patterns and that periodicity of a network is likely closely related to topological features of connectivity [52,68]. This fascinating phenomenon suggests that the binning of noisy biological data could obscure important relations between structure and activity including inherent inclinations to be attracted to specific periods. If the clustering in certain periods and other types of patterning can be related to the architectural underpinnings this could be extremely important for identifying the mechanisms of frequency banding seen in biological networks and the shifts they undergo in pathological states [48].

4.8. Hebbian nightmares and searches with naive plasticity

As complex dynamical systems go, neural networks have features such as complicated connectivity patterns and recurrence that make them particularly difficult to understand and predict even when compared to the most complex physical counterparts (chemical lattices, percolation systems, spin glasses, etc.). Moreover, the addition of use-dependent plasticity can make the study of these

already complex systems seem absolutely daunting. For example, in assessing the dynamics of a five unit network, the introduction of plasticity necessitates the additional tracking of 25 weights thereby changing the case from a five-dimensional problem to one with 30 dimensions. Notwithstanding the difficulty, the fact that networks learn and change as they interact with the world cannot be ignored. Are the activity distributions seen in the static architecture case just as true for networks with plasticity or might plasticity hold the key to eliminating the predisposition to entering limit cycles?

Unfortunately, the preliminary tests with naive Hebbian learning in weights strengthened by connectivity showed increased inclination to oscillation and fixed point dynamics (Section 3.3). This propensity to fall within one of these two dynamical regimes is in line with previous studies pertaining to the dynamics of weights [29]. Thus, if synchrony of activity is a problem facing the brain, then the addition of plasticity in a naive Hebbian form seems only to markedly worsen the problem. At the very least, the restricted dynamics that resulted from the present addition of plasticity suggests that there must be more to the application of plasticity than accounted for by this overly simplified learning rule.

Despite the inability of this form of plasticity to avoid pathological dynamics, these plastic models may still be helpful in illuminating network space and significantly decreasing the time for exploration of possible networks. The self-organization seen in the plastic network described in Section 3.3 shows that simple Hebbian plasticity can skew the final distribution to limit cycles and ultimately fixed point activity. Networks with this simple form of plasticity could not reliably or persistently exhibit complex dynamics. Ironically, this failing is extremely informative. Specifically, the approach suggests that the probability of limit cycle and fixed-point behaviors may increase as weights are saturated. Most importantly, instead of the added dimensions expanding the time needed for exploration we see that plasticity can hasten the convergence in a given space. Thus, rather than compounding difficulties this naive plasticity, and plasticity in general, may be a way of expediting the charting of network space. The networks found in this manner can act as comparators for random and evolved embodied systems when studying network architecture.

4.9. Intermittency and uncharted dynamics

The random sampling and observations of Sections 3.1 and 3.4 show that the dynamical features of networks can help direct the exploration prior to even considering what their function may be. The intermittent network presented in Section 3.4 is a particularly important example in that it demonstrates how a simple network can exhibit emergent complex and possibly unexpected behavior without the dynamics being explicitly built into the workings of the constituent units. The intermittency seen in paroxysms may be similarly driven by activity reverberations originating in the network structure thus suggesting a completely new path to epilepsy beyond those already posited [38,39,72]. These models thus provide an opening for considering the endogenous mechanisms for transition in dynamics in both pathological and healthy conditions. Specifically, the model can generate shifts in dynamics without having to assume extraneous factors such as noise, targeted environmental interference or entering into the endless regress of postulating secondary neural populations responsible for driving the network. These autonomous transitions might also elucidate mechanisms related to behavioral shifts in healthy organisms. The relation of these intermittent networks to epilepsy and attention is further explored in a forthcoming paper [51].

4.10. Charting small but critical constellations in parameter space

In studying the network activation patterns we began by asking what range of dynamical behaviors would be found if the entire space was searched? The fact that simple networks can show a predictable distribution of dynamical behaviors suggests that the same assessment might be made in more complex models. As we move from the common case of periodicity and fixed points to focusing on intermittency and evolved networks, new questions arise: How peculiar are the behaviors presented? What are the implications to the biological case?

Given that a system with fixed unit properties and a relatively limited range of network connectivity parameters could, when searched, exhibit such a broad range of dynamics implies that when complex biophysical network models are presented they cannot be taken at face value simply because their units capture biological detail. The lesson here is that extensive sweeping of connectivity parameters (geometry, spatial distribution, homogeneity, etc.) may always be in order.

Though perhaps a given set of dynamics may appear peripheral to the statistically normal workings of networks of a given configuration it is important to remember that even rare patterns might be important for understanding pathology and health. Epilepsy is a relatively common ailment that can be devastating, yet it affects only 1% of the population and for the most part exhibits itself only a small portion of the time. This suggests that a pattern of clinical significance may be exhibited in only 1% of the networks. Conversely, dynamics that appear rare in random networks or designed models may prove to be the norm in biology. To deduce the applicability of a model based on the dynamics most prevalent within an expected parameter range may result in overlooking important behaviors. For this reason, the use of a search paradigm like an evolutionary algorithm in autonomous agent modeling can be essential. By following the path simulated evolution or learning algorithms trace to these small constellations of complex networks, we may be able to target the essential

factors and reveal solutions biological evolution has found in the past. Pioneering work in this field has shown that this approach to modeling can find solutions that elude more traditional forms of investigation [61,84].

At first glance, the predominance of epilepsy seen in the initial random networks might be taken as evidence that the specifics of the present model are too far removed from biology, and that the over-abstraction of unit properties is at fault. By this account successful behavior exhibited in the evolved networks is simply a small and skewed subset on the fringes of the distribution. This assumption however neglects to consider that, through learning and evolution, biological systems may similarly converge on very small constellations of useful network structures in a space occupied mostly by ineffective networks. Thus, the predominance of epilepsy in these models may not be due to an absence of critical intrinsic factors but rather the reflection of a real connectivity challenge that extends into the biological realm.

4.11. Dynamics and the world: spatio-temporal consideration in traditional, virtual world and embodied neural modeling

Though time and other elements of the model were discretized in this study and not specifically matched to biological measures they were not arbitrary. One of the most interesting consequences of running a model in real time and a real environment is the manner in which a network becomes attuned to its body. The importance of this became apparent when we increased transmission and processing rates only to find that these brain–body changes could profoundly affect network behaviors, which then required retraining (results not shown). This is an interesting phenomenon which one would not normally encounter under traditional forms of modeling, where time scales may appear closely linked to biological measures but can actually prove arbitrary with respect to the environment.

Given that actual brain–world interactions are not incorporated into most neurophysiological modeling it is not surprising that a great many models that are said to accomplish a given computation cannot be translated into physical world performance. The cost of not considering embodiment has figured prominently in the critique of traditional Artificial Intelligence and has been implicated in the field's stagnation on several fronts [4,13,79]. The difference that embodiment makes in Artificial Intelligence and Robotic modeling [46,61,84] may similarly be extended to basic computational neuroscience.

Whether virtually or actually embodied, the evolution of networks in a world can connect dynamics to behavior and impart meaning to computation. Modeling in embodied systems, however, has the disadvantage of being considerably slower than running traditional network models of comparable size. This is particularly true of physical embodiment. A virtual world version of this experiment may have provided similar information with the advantages of speed, technical simplicity and increased experimental control over the environment. So why bother with actual embodiment?

It is not a coincidence that the capabilities of models raised and tested in virtual worlds seem to eclipse the performance of corporeally embodied models. The embodied models in Section 3.5 responded to physical elements in the world, be they transmission rates, physical characteristics of the body (e.g., body shape, size, wheel rotation features) or properties of the environment (e.g., friction and gravity). Direct interaction with physical objects and explicit consideration of real world parameters is a much more challenging task than maneuvering in a streamlined and idealized version of an environment. In applying spatio-temporal simplification, factors that are clearly germane and informative to the biological system may be wrongly excluded as redundant or inadvertently conflated into the workings of the network models. It is for this reason that performance of networks raised in streamlined worlds may not always be portable to the embodied case. The fact that the embodied experiments did not build in such assumptions about the surrounding world can help increase the confidence in the findings. For a review of the reasons physically embodied modeling may be qualitatively different from virtual implementations see [46,84].

4.12. Evolution of dynamics, behaviorally relevant activity and autonomy

The aim of the embodied sections of this study was to go beyond gross categorization and to focus the search of network space to networks with dynamics that might relate directly to behavior. This approach is analogous to clinical diagnoses that do not limit themselves to mere study of physiological recordings and that consider behavioral states as an essential if not central indicator of an assessment.

The randomly connected networks predominately exhibited what might be considered pathological behavior. That is, continuous shaking or complete inactivity. Through interaction with the world the genetic algorithm was able to evolve networks away from this motor behavior. The activity traces following the collision time markers on Fig. 9 show how an evolved network responded to tactile input. As part of the dynamic response the network undergoes a brief episode of turbulent activity followed by a relaxation into new activity ranges. This flip in activity correlates to the change in wheel direction. The solution settled on by this evolutionary run thus appears to be a parsing of the activity space such that the network flips to appropriate nearly periodic attractors as a response to tactile input.

The degree to which these found strategies can be said to be autonomous is certainly debatable. The task was highly simplified and it is consequently not surprising that the evolved behavior is fundamentally reactive. These reactive behaviors are akin to those seen in Braitenberg vehicles [7]

and could in fact be accomplished with as few as two neurons [81,82].

The evolution of networks in a world is only a first step in the exploration of behavior. The next step will require identifying the elements necessary to go beyond simple reactive behavior. For example, is an increase in the number of units, transducers or actuators necessary? Is there a need to increase the complexity of the unit behavior or connectivity? Do the genetic code, evolutionary algorithm or fitness functions need to be made more sophisticated? Is there a need to boost the number of individuals and evolutionary generations? Does plasticity need to be incorporated? Is opening the range of possible behaviors by enriching the environment and removing the tether essential? Do autonomous agents need to be coevolved, requiring interaction with other intelligent systems? All the above seem to be obvious candidates for forwarding the search. What else might be needed for truly autonomous intelligent behavior? It is in the identification of factors required for changing the structure of networks so that they can interact intelligently with the world that will help demarcate the architectural principles that underlie healthy versus pathological dynamics.

4.13. Embodied neurophysiological modeling of behavior and pathology

Aside from the understandable need to expand the complexity of the model, even these preliminary recordings of activity in simple embodied versions can already present novel and revealing perspectives of epilepsy and pathology. On a technical level, the approach offers several advantages over traditional biological recordings including: (i) Full access to the activity levels across the entire network. (ii) Full access to network connectivity weight. (iii) The ability to repeat experiments with identical network structures over differing circumstances. (iv) Access to complete populations through evolution. (v) Allows for real-time interaction with any element of the model.

The main advantage compared to traditional computational models is the access to a physically instantiated behavioral measure. The fact that many physiological and computational models are both figuratively and literally disconnected from the world is a considerable impediment to the interpretation of their applicability. Although such models may claim kinship to reality through reductionist detail we do not know whether they capture the relevant dynamics. If a model's unit dynamics seem isomorphic to that of a real neuron will the system translate to behavioral competence? A central role of the neural system is to interact with the environment; to process data. Yet so few computational physiological models do so. The prevalent strategy is to attempt and reconstruct the observed dynamics of cell firings. The most thoughtful of these projects speculate on the essential aspects of the coding before considering how they might be generated by a network. But even in cases where the computational aspects are explicitly considered, performance is still measured by replicating the purported salient part of the dynamics in the absence of working proof. The limitations of this approach become apparent when behavioral measures are required.

Once a corporeal demonstration is demanded and a system's behavior is considered, central conjectures such as the step from synchrony to intelligent processing are not as obvious as supposed [20,66]. The embodied experiment of Section 3.5 suggests that (a) the highly oscillatory locked-in limit cycle activity is likely to be a problem, (b) that solutions to a behavioral task may lie in the direction of escape from synchrony and (c) that rather than synchrony, it is complex patterns of changes in firing frequencies, phases and amplitudes that may characterize behaviorally successful solutions. In any event, the experiment allowed for extension of model testing beyond the ability to mimic patterns pre-selected by an experimenter and instead its success can now be measured in terms of the outward behavior of the system. It is entirely likely that there are complex yet elegant dynamical patterns that are biologically relevant but have yet to be recognized by physiologists due to limitations in current imaging and recording techniques or due to an inability to imagine dynamical strategies beyond synchrony. Rather than just focusing on the most obvious dynamical features seen in biological systems, the evolutionary approach can help guide us to the candidate neural structures and dynamics that may have important computational applicability.

4.14. System-based oscillations do not require pacers or central clocks

Synchronizations seem to be everywhere in complex biological recordings. It is not surprising that synchrony has been so closely connected to higher cognitive functions such as perception, attention, action, and consciousness [20,66]. Moreover, because these oscillations can be incredibly precise, often the intuition is that there must be pacer cells or other mechanisms external to the network in question that keep an ensemble synchronized [36,37].

The simple network models presented here help underscore the role of network connectivity in oscillations and synchrony. Our use of the term synchrony is somewhat broad and is intended to cover the range of periodic phase-locked periodic activity [33] observed in the random, plastic and evolutionary networks. This definition of course subsumes networks in which the activity of all units was fully coherent. Note that if the concerns regarding the ubiquity and dysfunction of synchrony are valid even under this broader definition the problem becomes all the more intractable in the more precise zero phase delay cases.

Indeed, the first set of experiments clearly demonstrated that achieving synchrony is not as rare or difficult as it may seem, and that self-organization of activity is quite ubiquitous in these networks. These observations are in line with the recent findings in network and synchronization studies

[69,71]. Careful consideration shows that synchronous periodicity—and more importantly its avoidance—has been a major longstanding theme in the Artificial Life studies [17,34,87]. Incredibly, Alan Turing in a little known paper did some amazingly forward thinking work on binary recurrent networks [77,73]. Little recognized at the time and dismissed, the paper was a prescient pre-shadowing of connectionism. Although dynamics and their categorization were not explicitly Turing's concern, these clearly are a factor. Recently Turing's models have been materialized using computers and clearly show networks that fall into the synchronous and chaotic activity [73].

Obviously, the propensity of any given network structure to fall into synchrony and the distribution of periodic networks in a population can change depending on the intrinsic activation function. This is clear from the differences in the distribution seen in RBF and SAF populations as illustrated in the insert to Fig. 5. However, the results show that limit-cycle predominance was preserved even under this radical change in activation functions. Both cases clearly establish that even if intrinsic mechanisms are held steady, changes in dynamic patterns can be a collective network phenomenon. The present exploration of recurrent networks not only showed that these oscillatory dynamics can be common regimes but also that connections alone can drive oscillations at a range of frequencies independent of individual cell predisposition. Fig. 5 underscores the fact that these periods can far exceed the time scales of the units in a network or the update rate of the model. Although this may seem trivial to connectionist modelers, it is not clear that the implications are fully recognized by neurophysiologists. Oscillating biological ensembles may under certain circumstances be driven by pacer cells, but these models demonstrate that no central pacers are *necessary* and the search for such entities may be in vain or misdirected. The implications to pathology and behavior are obvious. Alterations of connectivity are at least *sufficient* to bring about the massive changes in dynamics that underlie both seizures and intelligent behavior.

4.15. The ubiquity of synchrony and the escape to complexity

Synchronous assemblies may actually be a hindrance rather than a prerequisite for action. The standalone networks and the first-generation embodied models show that networks will produce synchrony quite readily, that this synchrony does not seem to contribute to behavior and that it can hamper interaction with the environment. Synchrony may also exact various physiological and metabolic tolls. In a neural system these may be aggravated with the addition of certain forms of plasticity [55]. The emerging theme from the standalone, plastic and evolved models is that it is not the achievement of orderly synchrony that is the challenge but rather the escape from its engulfing ubiquity. One might say with respect to epilepsy that ''the disorder is in the order''.

It has long been postulated that synchrony might be associated with pathology [18]. Complex systems may be able to—and may need to—accomplish information processing without binding in the sense of synchronous subpopulations. From this alternate perspective the question of synchrony and how systems find escape from its grasp becomes compelling. How does processing take place without falling into synchrony? The ubiquity of synchrony suggests that the solution is not simple.

This study suggests that even though the a priori chances may be stacked against organisms, an evolving interaction with the world offers more hopeful odds of achieving autonomous dynamics. Here we demonstrate this principle in the evolution of small networks. Interestingly, there is also evidence that such changes take place in development. There are striking reports in the developmental and epilepsy literature that suggests the earliest forms of activity seen in the developing embryo are perilously ictal in character, that the neonatal brain is more susceptible to seizures and that crucial stages in development are accompanied by changes in these dynamic profiles [27,6,59].

4.16. Embodiment and the unbinding problem

The evolutionary sections of this study directly address the problem of trying to relate the escape from periodic dynamics to behavior. If a system is independent of the world then no interaction with the world is possible and the tendency to lock in synchrony can take over. If a system wholly relies on the environment to pull out of synchrony then the system may simply become captive to the dynamics of the world. Extending the preceding synchrony discussion to the cognitive realm, the question thus shifts from how representations get temporally bonded across a network (the binding problem) [20,66] to how unit activities may become sufficiently independent from the world and each other so that they can represent and act (the unbinding problem).

Even if the intrinsic dynamics of cells are such that they can underwrite intelligent behavior, a higher order of self organization, beyond simple synchrony, was required. Achieving the correct connectivity for intelligent behavior necessitated learning or evolution and the context of a world. The random networks suggested that death and seizures occupy the largest part of parametric space even for normal cells. As the intermittency case showed (Section 3.4), limiting the search to networks categorized under more complex dynamics and systematically looking at network activity patterns can be illuminating. Such a search can help identify basic mechanisms for escaping synchrony and autonomously switching between dynamical regimes [51]. And yet, though the intermittency network may elucidate mechanism, the relationship to behavior is still extrapolated, a conjecture. It is for this reason we needed ways to explore the sizable space to find architectures that can lead not only to more

complex activity patterns but that could also be directly related to behavior.

The ability to embody networks fundamentally changes the research landscape. These models allowed us to see that networks of very few units are sufficient to control elementary navigation. Once embodied, the dynamics observed in the random networks could immediately be assessed in relation to behavior. Particular dynamics could be seen as closely related to what we might regard as pathology in the biological counterpart. Fixed point activity corresponds to immobility, or death. The highly prevalent limit cycles corresponded to small shakes reminiscent of seizures. The first generation of embodied networks thus demonstrated that: (i) the world alone does not change the dynamics of these networks, (ii) that synchrony and self-organization of activity far from being peculiar are highly probable and possibly even problematic, thereby questioning the hope that simple synchrony will account for the brain's representational powers, (iii) the probability that a network will interact with the world by default is vanishingly small.

This is where the findings of Section 3.5 show the richness and tangible promise of embodied models as compared to the techniques that separate networks from action. The successful escape from synchrony in these models can help expand our conceptual grasp of how dynamic representation takes place in non-synchronized (unbound) systems.

4.17. Charting epilepsy by searching for intelligent behavior

This paper began by inquiring about the full landscape of network behaviors. We inquired as to the likelihood that a network's activity would die out (settle on a fixed point) and the likelihood that it would enter and get locked into a periodic limit cycle. We asked what other sorts of dynamics networks might exhibit, how these dynamics might relate to interactions with the world and the role these various types of dynamics might play in physiological systems. Most generally, we wondered about the connection between network structure, network dynamics, and an organism's interaction with the world.

The attempt to answer these questions began with general experimental observations regarding the ubiquity of oscillations (Section 3.1) and their worsening under simple plasticity (Section 3.3). The observations of Section 3.4 showed how certain types of more complex epilepsy-related phenomena such as interictal paroxysmal events might arise and signaled the possibility of heterogeneous dynamics. Section 3.5 illustrated how synchronous dynamics can drive pathology and linked ictal neuronal activity to motor output. This section also explored ways leading out of the dominant, purely repetitive, regimes and searched for networks that could interact with the world. This exploration of network architectures is what was meant by "charting epilepsy by searching for intelligence in network space". The first sections are intended to give a sense for the vastness of network space—the sheer number of possible networks and the types of behavior that a random search might uncover. They also clarified that finding the limits of these vast zones occupied by fixed points and limit cycles would require assistance. This is where the evolution of intelligent behavior helped demarcate the borders of seizure territory, and directed our search in non-epileptic space. To find a system that can interact with the environment is to find a system that has escaped locked-in periodic cycles (embodied model) or seizures (biological counterpart).

4.18. Evolving out of epilepsy and into the world

Our ability to actively interface with the world is both a reflection and direct consequence of internal network dynamics. A dynamic ailment might manifest itself as a complete disconnection from the world, or, at the other extreme, as a system that is entirely driven by the world and cannot free itself from the effects of external inputs. In this regard, the embodied behavior of the first random networks falls within the former type of pathology. Not one of the naive networks in Section 3.5 accomplished much beyond freezing or shaking behavior. With the initiation of evolution, the emergence of interaction with the world became detectable even in early generations with individuals showing some shaking in a forward direction and dashes toward arena walls. Within only 14 generations the evolutionary algorithms had found network structures that could do something reminiscent of behavior. These changes in behavior so clearly reflected in fitness measures (Fig. 8) were sustained by corresponding transformation in activity dynamics (Fig. 9). This multifaceted shift was generally from homogeneous narrowband activity to that with multiple-time scales and/or multi-stability.

The application of a genetic algorithm in Section 3.5 thus not only focused the behavioral search but could prove to be biologically relevant in helping explore the means by which evolution and learning find appropriate solutions in the incredibly large space of possible networks. Even with this simple task there seemed to be a combination of changes in both the apparent number of attractors and spectral characteristics of network activity. Certainly many of the evolved networks exhibited a move away from characteristic scales that was pronounced in the random networks. Similarly, the embodied networks showed how the activities could shift following a tactile input, suggesting the introduction of new attractors or expansion of existing basins. It is unclear which of these changes—the move from synchrony per se or the ability to switch between multiple attractors—best characterized the evolutionarily induced navigation strategies. An increase in task complexity and a careful analysis of these attractors could reveal the tactics the healthy brain employs in its interaction with the world. Autonomous mechanisms of transition are important in explaining how agents choose, attain autonomy and are able to avoid lockdown. As such, being

able to distinguish the size, character, transient nature and the number of attractors might help elucidate how neural ensembles perform transitions in the healthy condition.

Understanding the nature of signal heterogeneity and uncovering the mechanisms for switching between dynamical regimes is also particularly critical for understanding the pathological conditions and its termination. In generation 0, the activity was continuous and homogeneous. In contrast, the evolved system exhibited intermittent and heterogeneous patterns. This durational and spectral variability is comparable to inter-ictal phenomena in epileptiform recordings and could be suggestive of the general mechanism involved in pulling out of seizures. Studying these dynamics in an embodied model could help distinguish normal from pathological population dynamics. Disambiguating normal oscillations from pathology is not a trivial problem. Sleep signals, alpha waves and ictal discharges are all highly oscillatory. Are sleep and alpha oscillations closer to seizures than normal awake recordings? [12]. Perhaps the difference is one of degrees, for example: (a) amplitude of activities, (b) number of cells involved, (c) duration of the state, (d) frequency of firing and (e) synchrony of elements may all be metrics differentiating healthy from pathological oscillations. In our experiments, measures of activity dynamics varied depending on whether the agents exhibited movement or were behaviorally detached. The observations also suggest that there may exist more subtle, fundamental and qualitative differences in dynamics, such as shifts from periodic limit cycles to close returns that could be critical for the emergence and stability of intelligent interactive behavior. Here, the dynamical systems literature's use of the prefix "unstable" in "unstable periodic orbits" is somewhat unfortunate; for although the term may be accurate and descriptive from a dynamical systems perspective it may be misleading when applied in the behavioral or cognitive context. It may very well be that it is these non-limit-cycle oscillating networks that are the most immune to pathological synchrony, most robust in the face of ambient noise and the most stable when considered in the context of cognitive performance and behavior in the world. It is for this reason we chose to categorize such networks as exhibiting "close returns" rather than the more broadly used "unstable periodic orbits".

4.19. The role of world input and the consequences of disconnection and connection

Are the changes in behavior seen in the embodied networks really due to a change in architecture or are the dynamics simply a consequence of the addition of world input? The change in fitness values and observed behavior seen over evolution (Section 3.5) demonstrated that although the environment may be *necessary* it is certainly not *sufficient*. The mere linking of networks to sensors and actuators did not miraculously ameliorate network periodicity or result in intelligent behavior. Random networks remained pathological despite connectivity. The evolutionary transformation implied that the escape from epilepsy required not only a connection with the environment but also the introduction of network structures that allow for dynamic processing of the input.

A strong predisposition to fall into periodic activity clearly interfered with a meaningful connection to the environment. However, a causal connection may also exist in the reverse direction. That is, it seems reasonable to postulate that the degree of connection to the environment can have an effect on periodicity. Certainly, there is evidence supporting the idea that decreasing connection with the environment can correlate with an increase in the risk of seizures. After all, certain types of seizures can increase during sleep [12] and though the causal route is unclear, there are links between certain forms of autism and seizures [58,75,76]. As such, it is important to establish the role that impaired connections to the environment has in epilepsy and vice versa. The connection to the world studied in the embodied networks in Section 3.5 thus goes beyond studying evolutionary considerations and offers new ways to explore the relation between connections with the world and resultant dynamics. Rather than depriving a system of connectivity and testing for loss of stability (e.g., as is done in in-vitro slice models), we can now proceed by increasing a system's connectivity to the world and exploring robustness directly via behavioral responsiveness.

4.20. Beyond evolutionary perspectives: learning and self-organization

Caution should be taken to not over-interpret the application of a genetic algorithm and the analogy to biological evolution. We applied the technique primarily as a way to explore network space. Undoubtedly, the computational version shares important features with the mechanisms of biological evolution but in many aspects it is far removed from the genetic processes that occur in organic systems. For one thing, it is unlikely that biological genomes code detailed network structure directly. Secondly, even by computational modeling standards the approach was particularly elementary. Although we initially explored mechanisms such as crossover as well as a range of fitness functions, for simplicity and speed we chose a much more targeted and limited evolutionary algorithm. Our concern here was not to explore evolutionary algorithms but rather to quickly find a way to connect network architecture, activity dynamics and behavior in a manner that would be transparent and straightforward. Extending the genetic algorithm to consider more complex coding schemes for network structure might enrich the search and help automate the promotion or identification of specific architectural elements [68]. More sophisticated coding might also make the analogy with biological evolution more precise.

Nonetheless, being overly concerned with relating the genetic algorithm of Section 3.5 to the details of natural counterpart might overshadow important general princi-

ples. Specifically, the genetic algorithm at hand could easily be recast as a learning algorithm. The genetic search for behavior only relied on rewarding the most beneficial of the random weight changes. There is no reason to believe that the process might not be successfully repeated so that weights are changed within individuals. Movement and collision feedback could act as reward and discouragement in an online learning algorithm. Comparison and integration of learned schemes and corresponding network structures might be implemented in a learning algorithm in a manner analogous to the genetic crossover. Moreover, there is nothing to preclude the simultaneous running of both forms of weight modification; learning and genetic algorithms are not mutually exclusive. Although in most of our experiments the weights for a given network were frozen and changed only between generations, the synergistic combination of evolution and appropriate learning algorithm might result in the Baldwin effect [2], speeding up the search or even finding new solutions [2,25].

Most importantly, focusing exclusively on competitive evolution could have more profound limitations. In thinking in evolutionary fitness terms we may fail to consider distinct factors complementary, orthogonal and even opposite to Darwinian mechanisms. Certainly, the development of behavior in our models suggests that a Darwinian-like process can promote an escape from overly ordered dynamics. However, the very presence of high order is intriguing. It suggests that we must keep in mind issues of symmetry and symmetry breaking so fundamental to life. This perspective suggests that even random network structures have high ordering coefficients and that complexity and self-organization (rather than just competition among individuals) plays an important role in the rise (and demise) of symmetrical patterns in pathology and action [30,31].

4.21. Clinical implications and potential therapies

Thinking beyond genetic algorithms is particularly important for interpreting how synchrony in these models might relate to the susceptibility of certain individuals to seizures. Focusing on learning, development and self-organization could also be critical to extending these findings to potential clinical treatment. By considering the shift from monolithically oscillatory systems to complex dynamics as a process that can be driven by learning and that can occur within the life cycle of an organism, immediately suggests the connection between these models and the biological evidence [6,27]. By thinking beyond evolutionary time scales of millions of years to the individual life cycle, the prospect of epilepsy as default can be seen as a challenge that every neural system must—and can—overcome. Accordingly, if the process is revisited in normal development it follows that the problem can be solved within a single network and in short time frames.

If individual developing networks can rearrange themselves quickly such that they move out of the epileptic regions of network space and into behavioral domains there is hope that the process can be retriggered later in life and replicated in clinical application. To this end, it is important to consider what is taking place in the embodied model. Although the developmental literature focuses on the GABAergic system and the balance between inhibitory and excitatory elements [6,27], the embodied model suggests that considering issues of self-organization and more complex architectural factors will be critical to understanding how the developing systems escape synchrony. Simply turning up inhibition may not be sufficient and might even be counter-productive. Are there endogenous mechanisms that are sensitive to quiescence (fixed points) and oscillations (periodic limit cycles)? Are there intrinsic biological mechanisms that act as dynamical fitness functions and change connectivity so as to encourage more complex activity? What role do major structural changes brought on by neurogenesis and cell death play in epileptogenesis and its reversal [64]? How does the susceptibility to seizure dynamics relate to behavioral activity [43,62,49,50]?

One way of studying the mechanisms that cause network structures to revert and become vulnerable to synchronous periodicity is by introducing environmental conditions that trigger these changes, thereby bringing on epilepsy in the post-developmental network. This is precisely the approach taken by artificially inducing seizures by kindling [19,40,56,57,43,64]. Conversely, a complementary approach to studying these mechanisms is by attempting to introduce normal activity in networks that are by default overtaken by synchrony. This is the approach taken in the evolutionary modeling sections presented in this paper. The implicit assumption is that there are structural generalities underlying oscillations to be found in random networks and that identifying these patterns will eventually elucidate the corresponding mechanisms that affect connectivity and increase vulnerability in the biological systems [49,50]. Moreover, if genetic or environmental factors can result in changes to connectivity in a particular direction perhaps there are treatments analogous to the network alterations effected in Section 3.5 that can pull systems out of the epileptic regions and back to behavioral clusters. Even if the initial architecture is established by genetic predisposition or noxious events, there is nothing that precludes the possibility of an environmental cure. For example, learning mechanisms in conjunction with the appropriate stimuli could theoretically accomplish the changes observed in the embodied models.

The connection to the biological case is conjectural and it would be foolish and dangerous to proceed to any clinical conclusions based on these preliminary explorations. The point here is not to forward immediate therapies but to suggest ways of exploring connections between network architecture and seizures that would be impossible by other means. Mapping out detailed network architectures, let alone the subtle connectivity changes, in biological preparations is extremely difficult. Most computer models that focus on intrinsic cell properties do not explore

connectivity parameters and so also cannot begin to test how network structure might drive seizures. Even those computational models that do consider connectivity often neglect the connection to behavior. The embodied recurrent network approach presented in this paper overcomes these obstacles by both focusing on network structure and by providing a way to clarify whether observed dynamics or suggested therapies will result in intelligent behavior (computational viability) or pathological conditions (limit cycles or fixed point regions). The evolution out of periodic synchronous activity in the models is analogous to known quenching phenomena [85] with the added advantage that it offers a practical way of exploring the effects of environment on connectivity and developing ways of implementing such changes in the adult brain by tapping into learning mechanisms.

4.22. Cure-oriented modeling, cybernetics and epilepsy

Biological models generally focus on replicating pathology in the hopes of later finding a cure. Such modeling can become dominated by capturing the relevant pathology rather than escaping from the condition. The embodied modeling presented here focuses on increased interaction with the world and so has a fundamentally different outlook. Instead of modeling pathology and applying control, we focused on modeling the escape from pathology and how this might be assessed in the context of the world and behavior.

It is no coincidence that some of the founders of cybernetics had a direct interest in electroencephalography and epilepsy [80–82,86]. As the prototypical dynamical disease these pioneers surely knew that condition might affect any constructed system and might have hoped that their research would alleviate seizures in humans. However, it is worth noting that cybernetics—by definition—is also dominated by notions of control. This perspective may have tainted decades of robotic theory and possibly the neurosciences. The cybernetic outlook is often one in which periodic activity and computation by synchrony seems controlled and an appropriate design goal. Yet, the observations of this paper suggest that the fundamental objective of organic systems might be precisely the opposite; that the escape from the all-encompassing control of synchrony is the true challenge.

This study thus suggests that creative cybernetic systems are likely to face the same challenges as their biological counterparts. In the past, the connection between robots and medicine has mostly been limited to applications such as assistance in navigation of surgical instruments. Here we explore a more direct and profound connection between robots and pathology. That is, that neurally-inspired synthetic autonomous agents may be just as prone to the problem of seizures. The human need to generate activity that supports interactions with the environment while ensuring autonomy thus similarly extends to these systems. As such, any solution found through these models might go beyond humans and epilepsy and might apply to embodied computational systems.

4.23. Conclusions, open questions and future directions

A central aim of the preceding Ladislav Tauc Conference in Neurobiology and these proceedings was to explore neuronal assemblies and cybernetics. In our approach to the question of how cell assemblies interface with the world we attempted to ask the most basic of questions with the most abstract neural models. This paper focuses on the relation between structure, activity, and behavior. Rather than focusing on a particular structure it assumes only nonlinear recurrence [28] and attempts to provide ways of searching these large spaces for structures that can underlie healthy activity and relate complex dynamics to behavior in embodied systems. We suggested that the charting of this space can be accomplished through: (1) random searches of network structures, (2) categorization via dynamics, (3) subcategorization through exploration of found complex dynamical sets, (4) embodied modeling and searching for intelligent behavior using learning or evolutionary algorithms.

This line of research leads to many open questions. Already mentioned is the need to explore learning algorithms. We need to establish how distributions of dynamic categories vary with network size and various statistical changes in weight distributions. The full effects of noise have yet to be explored in detail. The categorization and subcategorization of networks could be accomplished through a range of dynamical tests. The embodied modeling could include more complex tasks, larger networks, spatially defined structures, time delayed connectivity, asynchronous updates, more sophisticated genetic coding schemes, more elaborate evolutionary algorithms and the addition of parallel learning. All of these might help identify universal elements in architecture that underlie pathological vs. healthy behavior. Analysis of the random and evolved network structure would benefit from the various tools being developed by current network theory [1,3,41,44,68,71,83].

Notwithstanding future modifications, the exploration of even these simple recurrent neural networks and embodied models provided a way to study the theoretical underpinning of movement disorders and the significance of the environment in the initiation and amelioration of seizures. This paper is fundamentally a proof-of-concept that such searches can yield results both in uncovering new mechanisms for epileptic activity (Section 3.4) and generating dynamical non-stationary and non-limit cycle networks that can support persistent activity while interacting with the world (Section 3.5). The resultant population may thus afford a view of the architectural principles demarcating healthy biological networks from the pathological. What is interesting and often neglected in autonomous agent modeling is the analysis of the dynamics of the underlying neural systems, especially at the early stages and through development. This conjoining of dynamical systems analy-

sis and studying development in embodied models may help provide understanding of the progressive escape from pathology in human as well as robotic systems.

4.24. Closing remarks: robots, epilepsy and the ethics of autonomous agents

The connection between epilepsy and artificial life is at least as old as the very coining of the term "robot" [9]. A key scene in "R.U.R." (Rossum's Universal Robots), the 1923 allegorical play by Karel Čapek that originated the term, has a central human character, Helena, witnessing a robot having a seizure. The Psychologist-in-Chief dismisses this activity: "*Occasionally they seem somehow to go off their heads. Something like epilepsy, you know. We call it Robot's cramp... It's evidently some breakdown in the mechanism*." The Chief Engineer and Head of the Physiological Department concur. The General Manager asserts: "*A flaw in the works. It'll have to be removed*." But Helena—clearly conveying Čapek's sympathies and outlook—disagrees and cautions that this supposed flaw may be an important first indication of life: "*Perhaps it's just a sign that there's a struggle. Oh, if you could infuse them with it*." [9].

Čapek's fictional, symbolic, observations are prescient cautions worth heeding on many levels. It would be simple to presume that the initial seizure-like interaction of our models with the environment are nothing but pathology—but it would be much more accurate and astute to recognize that these reverberations are also the basis for interaction with the world. The evolved model's ability to interact with the environment is premised on the initial existence of reverberations which are evolved into interactive behavior. Those initial artificial reverberations—as perhaps is the case with biological seizures—may be viewed as the transitional territory between fixed-point, comatose or dead systems at equilibrium and systems that can actively but independently interact with the world.

It is notable that social and ethical issues relating to cyborgs constituted another important theme of the present conference. In its most universal reading, R.U.R. is a cautionary tale outlining the cost of ignoring the freedom of autonomous systems and the responsibility associated with introducing such systems into the world. This message should not be taken lightly. The hope is that such research will help cure epilepsy in humans just as much as it might help alleviate pathology and ensure autonomy for any new denizens of this world.

Acknowledgements

The authors are grateful for the support of the Institute of Neurosciences, Mental Health and Addiction of the Canadian Institutes of Health Research (CIHR) and the Dutch Epilepsy Clinics Foundation (SEIN).

References

[1] R. Albert, A.-L. Barabási, Statistical mechanics of complex networks, Reviews of Modern Physics 74 (2002) 47–97.
[2] J.M. Baldwin, A new factor in evolution, American Naturalist 30 (354) (1896) 441–451.
[3] A.-L. Barabási, Linked: The New Science of Networks, Perseus Publishing, Cambridge, MA, 2002.
[4] W. Bechtel, A. Abrahamsen, Connectionism and the Mind: An Introduction to Parallel Processing, Basil Blackwell, Cambridge, MA, 1991.
[5] W. Bechtel, A. Abrahamsen, Connectionism and the Mind, Parallel Processing, Dynamics, and Evolution in Networks, second ed., Basil Blackwell, Cambridge, MA, 2002.
[6] Y. Ben-Ari, Excitatory actions of GABA during development: the nature of the nurture, in: Sixth IBRO World Congress of Neuroscience, 2003.
[7] V. Braitenberg, Vehicles: Experiments in Synthetic Psychology, MIT Press, Cambridge, MA, 1984.
[8] W.M. Burnham, P.L. Carlen, P.A. Hwang (Eds.), Intractable Seizures: Diagnosis, Treatment and Prevention. Advances in Experimental Medicine and Biology, vol. 497, Kluwer Academic/Plenum, New York, NY, 2002.
[9] K. Čapek, R.U.R. (Rossum's Universal Robots), Oxford University Press, Oxford, 1923.
[10] H.Y. Chun, E.L. Ohayon, H.C. Kwan, W.M. Burnham, P.A. Hwang, Comparison of close return and Fourier detection of unstable periodic events in interictal and ictal EEG, Eastern Association of Electroencephalographers, 57th Annual Meeting 2003, Clinical Neurophysiology 115 (1) (2004) 249–250.
[11] M.A. Dichter, Basic mechanisms of epilepsy: targets for therapeutic intervention, Epilepsia 38 (Suppl. 9) (1997) S2–S6.
[12] D.S. Dinner, H.O. Lüders, Epilepsy and Sleep: Physiological and Clinical Relationships, Academic Press, San Diego, CA, 2001.
[13] H.L. Dreyfus, What Computers Still Can't Do, MIT Press, Cambridge, MA, 1992.
[14] J. Elman, Development: it's about time, Developmental Science 6 (4) (2003) 430–433.
[15] B. Ermentrout, Phase-plane analysis of neural activity, in: M.A. Arbib (Ed.), The Handbook of Brain Theory and Neural Networks, MIT Press, Cambridge, MA, 1995, pp. 339–343.
[16] W.J. Freeman, Mass Action in the Nervous System, Academic Press, New York, 1975.
[17] M. Gardner, The fantastic combinations of John Conway's new solitaire game life, Scientific American 223 (1970) 120–123.
[18] L. Glass, Synchronization and rhythmic processes in physiology, Nature 410 (2001) 277–284.
[19] G.V. Goddard, Development of epileptic seizures through brain stimulation at low intensity, Nature 214 (1967) 1020–1021.
[20] C.M. Gray, The temporal correlation hypothesis of visual feature integration: still alive and well, Neuron 24 (1999) 31–47.
[21] A. Guberman, J. Bruni, Essentials of Clinical Epilepsy, second ed., Butterworth Heinemann, Boston, MA, 1999.
[22] D.O. Hebb, The Organization of Behavior, Wiley, New York, 1949.
[23] J. Hertz, A. Krogh, R.G. Palmer, Introduction to the Theory of Neural Computation, Addison-Wesley, Redwood City, CA, 1991.
[24] G.E. Hinton, T.J. Sejnowski, Learning and relearning in Boltzmann machines, in: D.E. Rumelhart, J.L. McClelland (Eds.), Parallel Distributed Processing: Explorations in the Microstructure of Cognition, vol. 1, MIT Press, Cambridge, MA, 1986, pp. 282–314.
[25] G.E. Hinton, S.J. Nowlan, How learning can guide evolution, Complex Systems 1 (1987) 495–502.
[26] J.H. Holland, Adaptation in Natural and Artificial Systems, 1992 ed., MIT Press, Cambridge, MA, 1992.
[27] G.L. Holmes, R. Khazipov, Y. Ben-Ari, New concepts in neonatal seizures, NeuroReport 13 (1) (2002) A3–A8.

[28] J.J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, Proceedings of the National Academy of Sciences of the USA 79 (1982) 2554–2558.
[29] S. Kalitzin, B.W. van Dijk, H. Spekreijse, Self-organized dynamics in plastic neural networks: bistability and coherence, Biological Cybernetics 83 (2000) 139–150.
[30] S.A. Kauffman, Emergent Properties in Random Complex Automata, Physica D 10 (1–2) (1984) 145–156.
[31] S.A. Kauffman, The Origins of Order: Self-Organization and Selection in Evolution, Oxford University Press, New York, NY, 1993.
[32] H.C. Kwan, E.L. Ohayon, P.A. Hwang, W.M. Burnham, Identification of close returns in human epileptiform EEG, Society for Neuroscience Abstracts, 2002, p. 793.1.
[33] J.P. Lachaux, E. Rodriguez, J. Martinerie, F.J. Varela, Measuring phase synchrony in brain signals, Human Brain Mapping 8 (1999) 194–208.
[34] C.G. Langton, Life at the edge of chaos, in: C.G. Langton, C. Taylor, J.D. Farmer, S. Rasmussen (Eds.), Artificial Life II: Santa Fe Institute Studies in the Sciences of Complexity, vol. X, Wesley, Redwood City, 1991.
[35] B. Litt, J. Echauz, Prediction of epileptic seizures, The Lancet Neurology 1 (2002) 22–30.
[36] R.R. Llinás, U. Ribary, Rostrocaudal scan in human brain: a global characteristic of the 40-Hz response during sensory input, in: T.H. Bullock, E. Basar (Eds.), Induced Rhythms in the Brain, Birkhäuser, Boston, MA, 1992.
[37] R.R. Llinás, U. Ribary, Coherent 40 Hz oscillation characterizes dream state in humans, Proceedings of the National Academy of Science USA 90 (1993) 2078–2081.
[38] F.H. Lopes da Silva, W. Blanes, S.N. Kalitzin, J. Parra, P. Suffczynski, D.N. Velis, Dynamical diseases of brain systems: different routes to epileptic seizures, IEEE Transactions of the Biomedical Engineering 50 (5) (2003) 540–548.
[39] F.H. Lopes da Silva, W. Blanes, S.N. Kalitzin, J. Parra, P. Suffczynski, D.N. Velis, Epilepsies as dynamical diseases of brain systems: basic models of the transition between normal and epileptic activity, Epilepsia 44 (Suppl. 12) (2003) 72–83.
[40] J.O. McNamara, Pursuit of the mechanisms of kindling, Trends in Neuroscience 11 (1) (1988) 33–36.
[41] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, U. Alon, Network motifs: simple building blocks of complex networks, Science 298 (2002) 824–827.
[42] G.B. Mindlin, R. Gilmore, Topological analysis and synthesis of chaotic time series, Physica D 58 (1–4) (1992) 229–242.
[43] P. Murphy, W.M. Burnham, The effect of kindled seizures on the locomotory behavior of long-Evans rats, Experimental Neurology 180 (2003) 87–91.
[44] M.E.J. Newman, The structure and function of complex networks, SIAM Review 45 (2) (2003) 167–256.
[45] J.L. Noebels, The biology of epilepsy genes, Annual Review of Neuroscience 26 (2003) 599–625.
[46] S. Nolfi, D. Floreano, Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines, MIT Press, Cambridge, MA, 2000.
[47] E.L. Ohayon, R. Liu, W.M. Burnham, H.C. Kwan, Characterization of human EEG by close return, Society for Neuroscience Abstracts, 2001, p. 83.9.
[48] E.L. Ohayon, A.J. Mendonça, P.W. Tsang, H.C. Kwan, W.M. Burnham, High frequency transients in kindling identified using a calibrated amplitude wavelet transform (CAWT), Eastern Association of Electroencephalographers, 56th Annual Meeting 2002, Clinical Neurophysiology 115 (1) (2004) 247–248.
[49] E.L. Ohayon, H.C. Kwan, W.M. Burnham, Ictal pathology in autonomous agents: epilepsy in robots, Society for Neuroscience Abstracts, 2002, p. 414.12.
[50] E.L. Ohayon, F.Y. Jin, D.S. Borrett, W.M. Burnham, H.C. Kwan, Reduction in seizure susceptibility through evolution in embodied neural models, 6th International Brain Research Organization World Congress of Neuroscience Abstract Book (2003) 477.
[51] E.L. Ohayon, H.C. Kwan, W.M. Burnham, P. Suffczynski, S. Kalitzin, Emergent complex patterns in autonomous distributed systems: mechanisms for attention recovery and relation to models of clinical epilepsy, in: Proceedings of IEEE SMC (2004) 2066–2072.
[52] E.L. Ohayon, H.C. Kwan, P.W. Tsang, D.S. Borrett, M. Burnham, I. Timofeev, M. Steriade, T.J. Sejnowski, M. Bazhenov, The effects of lesions on dynamics in network models of epilepsy, Society for Neuroscience Abstracts, 2004, p. 228.18.
[53] H.O. Peitgen, H. Juergens, D. Saupe, Chaos and Fractals: New Frontiers of Science, Springer-Verlag, New York, 1992.
[54] R. Port, T.J. van Gelder, Mind as Motion: Explorations in the Dynamics of Cognition, MIT Press, Cambridge, MA, 1995.
[55] A.D. Priesol, D.S. Borrett, H.C. Kwan, Dynamics of a chaotic neural network in response to a sustained stimulus, Technical Report # RBCV-TR-91-38, Research in Biological and Computational Vision, University of Toronto, Toronto, 1991.
[56] R.J. Racine, Modification of seizure activity by electrical stimulation. II. Motor seizure, Electroencephalography and Clinical Neurophysiology 32 (3) (1972) 281–294.
[57] R. Racine, V. Okujava, S. Chipashvili, Modification of seizure activity by electrical stimulation. III. Mechanisms, Electroencephalography and Clinical Neurophysiology 32 (3) (1972) 295–299.
[58] I. Rapin, Autism, New England Journal of Medicine 337 (2) (1997) 97–104.
[59] J.M. Rennie, G.B. Boylan, Neonatal seizures and their treatment, Current Opinion in Neurology 16 (2003) 177–181.
[60] D.E. Rumelhart, J.L. McClelland (Eds.), Parallel Distributed Processing: Explorations in the Microstructure of Cognition, vols. 1–2, MIT Press, Cambridge, MA, 1986.
[61] E. Ruppin, Evolutionary autonomous agents: a neuroscience perspective, Nature Reviews Neuroscience 3 (2) (2002) 132–141.
[62] J. Schoenfeld, M. Seidenberg, A. Woodard, K. Hecox, C. Inglese, K. Mack, B. Hermann, Neuropsychological and behavioral status of children with complex partial seizures, Developmental Medicine and Child Neurology 41 (1999) 724–732.
[63] H.G. Schuster, Deterministic Chaos: An Introduction, second ed., John Wiley & Sons, New York, 1988.
[64] B.W. Scott, S. Wang, W.M. Burnham, U. DeBoni, J.M. Wojtowicz, Kindling-induced neurogenesis in the Dentate Gyrus of the rat, Neuroscience Letters 248 (1998) 73–76.
[65] B.F. Shneker, N.B. Fountain, Epilepsy Disease-a-Month 49 (7) (2003) 426–478.
[66] W. Singer, Neuronal synchrony: a versatile code for the definition of relations? Neuron 24 (1999) 49–65.
[67] O.K. Steinlein, Genetic mechanisms that underlie epilepsy, Nature Reviews Neuroscience 5 (2004) 400–408.
[68] I. Stewart, Networking opportunity, Nature 427 (2004) 601–604.
[69] S.H. Strogatz, I. Stewart, Coupled oscillators and biological synchronization, Scientific American 269 (6) (1993) 102–109.
[70] S.H. Strogatz, Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering, Perseus Books, Cambridge, MA, 1994.
[71] S.H. Strogatz, Exploring complex networks, Nature 410 (2001) 268–276.
[72] P. Suffczynski, S. Kalitzin, F.H. Lopes da Silva, Dynamics of non-convulsive epileptic phenomena modeled by a bistable neuronal network, Neuroscience 126 (2) (2004) 467–484.
[73] C. Teuscher, Turing's Connectionism: An Investigation of Neural Network Architectures, Springer-Verlag, London, 2002.
[74] R.D. Traub, J.G.R. Jefferys, M.A. Whittington, Fast Oscillations in Cortical Circuits, MIT Press, Cambridge, MA, 1999.
[75] R.F. Tuchman, I. Rapin, Regression in pervasive developmental disorders: seizures and epileptiform electroencephalogram correlates, Pediatrics 99 (4) (1997) 560–566.

[76] R.F. Tuchman, I. Rapin, Epilepsy in autism, The Lancet Neurology 1 (6) (2002) 352–358.
[77] A.M. Turing, Intelligent machinery (National Physical Laboratory Report, 1948), in: D.C. Ince (Ed.), Collected Works of A.M. Turing: Mechanical Intelligence, Elsevier, Amsterdam, 1992.
[78] T.J. van Gelder, The dynamical hypothesis in cognitive science, Behavioral and Brain Sciences 21 (1998) 615–665.
[79] F.J. Varela, E. Thompson, E. Rosch, The Embodied Mind: Cognitive Science and Human Experience, MIT Press, Cambridge, MA, 1993.
[80] W.G. Walter, The location of cerebral tumours by electro-encephalography, The Lancet 228 (5893) (1936) 305–308.
[81] W.G. Walter, An imitation of life, Scientific American May (1950) 42–45.
[82] W.G. Walter, A machine that learns, Scientific American August (1951) 60–63.
[83] D.J. Watts, S.H. Strogatz, Collective dynamics of 'small-world' networks, Nature 393 (1998) 440–442.
[84] B. Webb, Robots in invertebrate neuroscience, Nature 417 (2002) 359–363.
[85] S.R.B. Weiss, A. Eidsath, X.-L. Li, T. Heynen, R.M. Post, Quenching revisited: low level direct current inhibits Amygdala-kindled seizures, Experimental Neurology 154 (1998) 185–192.
[86] N. Wiener, Cybernetics: Or Control and Communication in the Animal and the Machine, second ed., MIT Press, Cambridge, MA, 1961.
[87] S. Wolfram, Universality and complexity in cellular automata, Physica D 10 (1–2) (1984) 1–35.
[88] R.S. Yoon, D.S. Borrett, H.C. Kwan, Three-dimensional object recognition using a recurrent attractor neural network, IEEE Engineering in Medicine and Biology Society, 17th Annual Conference 1 (1995) 379–380.