Evolutionary Selection of Network Structure and Function

Larry Yaeger[1], Olaf Sporns[2], Steven Williams[1], Xin Shuai[1] and Sean Dougherty[3]

[1]School of Informatics and Computing and [2]Department of Psychological and Brain Sciences
Indiana University, Bloomington, IN 47408
[3]Open source contributor
larryy@indiana.edu

Abstract

We explore the relationship between evolved neural network structure and function, by applying graph theoretical tools to the analysis of the topology of artificial neural networks known to exhibit evolutionary increases in dynamical neural complexity. Our results suggest a synergistic convergence between network structures emerging due to physical constraints, such as wiring length and brain volume, and optimal network topologies evolved purely for function in the absence of physical constraints. We observe increases in clustering coefficients in concert with decreases in path lengths that together produce a driven evolutionary bias towards small-world networks relative to comparable networks in a passive null model. These small-world biases are exhibited during the same periods that evolution actively selects for increasing neural complexity (also during which the model's agents are behaviorally adapting to their environment), thus strengthening the association between small-world network structures and complex neural dynamics. We also introduce a new measure of path length in graphs, "normalized path length", that is better behaved than existing metrics for networks comprised of disjoint subgraphs and disconnected nodes, and a novel method of quantifying the degree of evolutionary selection for small world networks, "small-world bias".

Introduction

Dynamical processes in networks are unavoidably influenced by the networks' underlying topologies. As the study of networks has come to pervade all of science, a need has arisen to understand this relationship between the anatomical structure of networks and the dynamical functions they carry out (Strogatz, 2001).

Small-world properties have been shown (Watts and Strogatz, 1998) to characterize many networks of interest, including biological nervous systems. Small-world networks of Hodgkin-Huxley neurons have been shown (Lago-Fernández et al., 2000) to provide the best features of both random networks (fast system response) and regular networks (coherent oscillations). Small-world-ness has also been shown (Sporns et al., 2000) to be highly correlated with dynamical complexity in artificial neural networks evolved specifically for complexity. In the biological realm, cortical connection matrices for macaque visual cortex and rat cortex have been shown (Sporns et al., 2000) to exhibit both small world anatomical properties and high dynamical complexity.

It has been argued that physical constraints—evolutionary pressures to reduce overall wiring length (Mitchison, 1991; Cherniak, 1995) and to maximize connectivity while minimizing volume (Murre and Engelhardt, 1995)—might explain key aspects of biological brain connectivity. But it is unlikely that evolutionary pressure on wiring alone is responsible for the detailed patterns of connectivity seen in biological brains (Sporns et al., 2000). Thus one is led to ask how natural selection would act upon the topological characteristics of nervous systems in the absence of physical constraints, and whether such functional evolutionary pressures are opposed to, independent of, or aligned with physical evolutionary pressures.

In previous work using the Polyworld artificial life system (Yaeger, 1994) we have shown that when agents whose behaviors are controlled by a genetically prescribed artificial neural network are subject to natural selection, the networks' dynamical neural complexity increases over evolutionary time (Yaeger and Sporns, 2006), the networks' complexity will be actively selected for by evolution (Yaeger et al., 2008), and periods of neural complexity growth correspond to periods of behavioral adaptation of the agents to their environment (Yaeger, 2009).

We now seek to understand the underlying network topologies that give rise to this evolved functional complexity. Preliminary results for several graph theoretical metrics from one simulation suggested (Lizier et al., 2009) that evolutionary trends in Polyworld mirrored those in biological neural networks (and successfully related anatomical networks to inferred functional networks). We will more fully characterize those evolutionary trends, determine their robustness and statistical significance, quantify the small-world-ness of those trends, and confirm the role of natural selection (as opposed to random drift, in a "driven" vs. "passive" sense (McShea, 1996)) in the shaping of those trends. This allows us to characterize the relationship between evolutionary pressures on brain structure due to functional optimization vs. physical constraints.

Tools and Techniques

Polyworld

Polyworld is an ecosystem model in which the agents are controlled by artificial neural networks using a firing rate neuron model performing Hebbian learning at the synapses. The wiring diagrams of these networks are the primary subject of evolution in the system, through a genetic encoding of a generative model of network architectures. This genetic encoding describes the network topology in terms of a number of neural groups, containing a number of excitatory and inhibitory neurons, wired together with genetically determined connection densities, ordered-ness of connections, and learning rates. By eschewing any particular model of ontogenetic development, Polyworld avoids the biases inherent in such a model choice. Further, instead of evolving specific network topologies, Polyworld forces evolution to select for useful statistics of neural connectivity.

Vision, current energy level, and a randomly firing neuron are the inputs to the network. A suite of primitive behaviors (move, turn, eat, mate, attack, light, focus) are the outputs. All agent actions consume energy, which must be replenished by consuming food from the environment, or by killing and eating other agents. Normally there are per-neuron and per-synapse energy costs, but these have been eliminated for this study so as not to impose any pseudo-physical constraints on network topology. Survival and reproduction, variation and selection, are the only driving forces, so Polyworld acts as a model of natural selection, with no fitness function, rather than in the manner of a genetic algorithm (though that is possible, if desired).

In these experiments Polyworld is used to produce paired runs in which an initial, normal "driven" run is followed by a "passive", null-model run. (The terms driven and passive are used in the sense proposed by McShea (1996).) In the passive run, agents cannot reproduce or die on their own; rather, pairs are chosen for reproduction at random and individuals are killed at random to match the birth and death events of the original driven run, thus removing the effects of selection, while retaining population statistics and levels of genetic variation that are equivalent to those in the driven run. This allows the direct comparison of driven vs. passive, natural-selection vs. random-walk evolutionary trajectories. See (Yaeger et al., 2008; Yaeger, 2009) for more details.

The activation of every neuron at every time step for every agent is recorded to disk as simulations progress, as is the neural architecture of every agent. Thus we are able to study both the structure and the function of the evolved neural networks, under conditions in which either natural selection or increasing variance due to a random walk are holding sway.

The Polyworld source code and data analysis tools are available at http://sourceforge.net/projects/polyworld/ and instructions for installing and building Polyworld are at http://beanblossom.in.us/larryy/BuildingPolyworld.html.

Complexity

Our primary tool for analyzing neural dynamics is an information theoretic measure of neural complexity proposed by Tononi et al. (1994) and introduced in a simplified and more computationally tractable form in (Tononi et al., 1998). Referred to throughout as "complexity" (aka "TSE complexity", for the initials of its inventors), the measure captures a trade-off between integration (cooperation) and segregation (specialization) in any system of random variables, such as the temporal traces of our agents' neural activations. Maximally complex networks exhibit a high degree of both integration and segregation at multiple scales. Though not presented here, we have previously demonstrated (Yaeger, 2009) that complexity is actively selected for, in a driven, biased fashion, during periods of behavioral adaptation of the agents to their environment, which corresponds to approximately the first 7,000 time steps in these experiments. During this period complexity increases much more rapidly in the driven runs than in the passive runs, but once a "good enough" solution emerges and begins to propogate throughout the population, driven complexity plateaus, while passive complexity continues its random walk to higher values.

Graph Theoretical Metrics

For current purposes we are interested primarily in three graph theoretical metrics. Two of them—clustering coefficient and characteristic path length—were used by Watts and Strogatz (1998) to define and characterize small-world networks. The third is a quantitative means of characterizing the degree of small-world-ness exhibited by a network introduced by Humphries et al. (2006). Throughout we will talk about our neural networks as graphs, which can be described by the number of nodes (aka vertices or neurons) and the number of edges (aka links or synapses) that connect them.

Clustering coefficient (CC) is a local measure of cliquishness in a graph, and characterizes the degree to which a node's neighbors are likely to be neighbors of each other (where "neighbor" means a link exists between the nodes). In social networks this would be the degree to which friends of a common friend are likely to be friends of each other. It is defined at each node as the fraction of possible links between neighbors that are actually present in the graph, and defined for the entire network as the average of this fraction over all nodes in the graph.

Characteristic path length (CPL), also called average shortest path length, is a global measure of the average separation between all node pairs in a graph—an estimate of how far it is from any one node to another. The average distance to all other nodes is calculated for each node, and then averaged over all nodes.

Watts and Strogatz (1998) identified small-world networks by their combination of high clustering and low path length. By contrast, though regular lattice networks also exhibit high clustering, they typically have high path lengths,

since moving from one node to another requires the traversal of all intervening nodes and links. And while random graphs tend to have low path lengths, since any given node is only a few random hops away, they usually exhibit low clustering.

Small-world index (SWI) is a quantitative measure of small-world-ness introduced by Humphries et al. (2006). To calculate SWI one computes CC and CPL for the actual graph, CC and CPL for a corresponding random graph (or ensemble of random graphs as done here), and compares the ratios of actual to random measurements as follows:

$$\gamma = CC/\langle CC_r \rangle \tag{1}$$

$$\lambda = CPL/\langle CPL_r \rangle \tag{2}$$

$$s = \gamma/\lambda \tag{3}$$

where $\langle CC_r \rangle$ and $\langle CPL_r \rangle$ are the ensemble averages of CC and CPL over some number of random graphs having the same number of nodes and edges as the original graph, and s is the desired SWI.[1] SWI captures the degree to which clustering and path length in the actual, original graph vary, in the appropriate directions, from the values seen in comparable random graphs. The more small-world a network is, the greater its SWI will be above 1.0.

These metrics are most frequently applied to undirected graphs (a given edge connects in both directions), often with binary edges (either present or not). However, neural networks importantly have both weighted and directed edges. Fortunately these metrics extend straightforwardly to support the analysis of weighted, directed (WD) graphs, but their application to such networks has been less well characterized than for binary, undirected (BU) graphs and, indeed, there turn out to be some issues applying them to WD graphs. (Such as a greater prevalence of disconnected nodes in WD graphs.) Accordingly, we analyzed our networks treating them both as BU and WD graphs.

Neural network edge weights are also signed—positive for excitatory connections, negative for inhibitory connections. Unfortunately, few graph theoretical metrics extend well to signed graphs. So for these analyses we have made the less than desirable, but simple and common, approximation of using the absolute values of the network weights on the graph edges.

The fact that one of our key metrics, path length, is based on distances between nodes, yet our neural networks have weights, not distances, associated with their connections, presents another small conundrum. We again take the simplest, most common approach, and invert the weights to provide a distance measure. Thus a strong weight, which produces a strong influence, after inversion corresponds to a short distance. So nodes that strongly influence each other

are seen as close neighbors, while nodes that only weakly influence each other are seen as distant neighbors, and nodes that do not directly affect each other at all (have zero weight) are infinitely far apart (though they may be reachable indirectly, through other nodes and links). For our other fundamental metric, clustering coefficient, we use the original neural network weights on the edges.

A question also arises as to which neural network nodes to include in the graph being analyzed. One obvious answer is all of them. However, the sensory nodes have an unusual constraint—zero in-degree (no incoming connections)—and their activations are purely determined by what the agent senses in its environment rather than what happens within the neural network. Another answer, then, is the non-sensory neurons; i.e., all internal and output/behavioral neurons. In our complexity work we have referred to this set of non-sensory neurons as the "processing" neurons. Accordingly, we have carried out our graph theoretical analyses looking at both cases: all (A) neurons and processing (P) neurons.

Finally, especially early on in our simulations, some of the graphs are quite small and consist of multiple components (disconnected sub-graphs) and even contain disconnected neurons. It turns out that CPL behaves poorly and erratically in this situation. This is due to its treatment of inter-node distances between disconnected nodes as infinite. Thus path lengths are computed only within each disconnected subgraph and the metric can exhibit sudden large changes as subgraphs become connected or disconnected and shortest paths span much larger or smaller subsets of nodes.

A length metric proposed by Marchiori and Latora (2000), connectivity length (CL), uses inverted lengths to calculate the harmonic (rather than arithmetic) mean of average shortest path length, and better handles multiple components and disconnected nodes. However, by effectively including all those infinities (as zeroes), it can compress the distinctions between sparsely connected and disconnected graphs.

We therefore devised, and introduce here, a new length metric, normalized path length (NPL), that appears to be better behaved than either CPL or CL for the class of graphs we are analyzing, though it too has some quirks (a sensitivity to edge weights that makes it somewhat noisy in its WD form).

To calculate NPL, node pairs that have no path between them are assigned a maximum path length l_{max} defined as N/w_{max}, rather than infinity, where N is the number of nodes in the graph and w_{max} is the maximum possible synaptic weight in our neural networks. (For binary networks the greatest possible path length is $N - 1$, hence this value of N is one that cannot occur by any means other than disconnection.) Inverting to convert weight to distance, we also define a minimum path length l_{min}, which is just $1/w_{max}$. We then proceed to compute CPL normally, limiting path length to the defined maximum, and normalize first by subtracting the minimum path length and then dividing

---

[1]Humphries used a single random graph corresponding to each original graph, but there is sufficient variance in CC and CPL amongst graphs with the same numbers of nodes and links that we have chosen to use ensemble averages instead.

by the difference between the maximimum and minimum path lengths. Thus, in terms of CPL, NPL may be written as follows:

$$NPL = (CPL^* - l_{min})/(l_{max} - l_{min}) \qquad (4)$$

where CPL^* is a normally calculated CPL using l_{max} as the maximum possible distance between nodes. Or expressed in terms of path lengths:

$$NPL = \frac{\frac{\sum\limits_{\substack{i,j=1 \\ j \neq i}}^{N} min(l_{ij}, l_{max})}{N(N-1)} - l_{min}}{l_{max} - l_{min}} \qquad (5)$$

where l_{ij} is the shortest path from node j to node i. NPL is guaranteed to lie between 0.0, for a fully connected graph, and 1.0, for a fully disconnected graph (a collection of nodes with no links between them), and has proven to be well behaved for graphs with multiple components and disconnected nodes (as well as for the more commonly analyzed strongly connected graphs).

Since none of our three length metrics is "perfect" and NPL is entirely new, wherever a length metric is calculated or used, we examine all three, and refer in general to simply path length. Thus for each metric we treat the graph as consisting of either the A neurons or the P neurons and we treat the graph edges as being either BU or WD, and for length metrics we look at each of CPL, CL, and NPL. Different neuron sets, graph types, and length metrics usually agree on common trends, but do sometimes provide different insights into the algorithms and architectures. Unfortunately, due to space constraints we cannot show all variations of all metrics. A complete set of plots of these metrics may be obtained as supplementary material here: http://informatics.indiana.edu/larryy/alife12_sup.zip. The abbreviations defined here (CC, CPL, CL, NPL, SWI, A, P, BU, WD) and another new metric (SWB) defined later are consistently applied in these plots as well as this paper.

All graph theoretical metrics were calculated using our new C++ implementation (bct-cpp) of the Brain Connectivity Toolbox (BCT) MATLAB module (Rubinov and Sporns, 2010). The original BCT may be found at http://www.brain-connectivity-toolbox.net/ and bct-cpp may be found at http://code.google.com/p/bct-cpp/.

Simulations and Data Acquisition

A set of 10 paired simulations, differing only in initial random number seeds, were run in driven and passive modes; i.e., 20 simulations in all. Each simulation ran for 30,000 time steps (approximately 400 generations) with a population varying from 90 to 300 agents. Temporal traces of neural activations and structural descriptions of neural anatomies were recorded for all agents. Agents were assigned to temporal bins corresponding to 1,000 time steps, according to the time of their death.

This type of binning was necessary for our complexity studies, since an agent's neural complexity can only be accurately computed after the completion of its neural activation time series—its death. We have retained this binning in our graph theoretical analysis so we can directly compare structural and functional results.

Complexity and graph theoretical metrics were calculated for each agent and averaged to produce a population mean (and standard deviation) in each temporal bin, for each driven and passive run. In addition, for each agent's actual neural network, 10 graphs with an identical node count, edge count, and distribution of weights were generated randomly, and the means of the graph theoretical measures for these networks were used to characterize the structure of a random graph corresponding to each actual graph.

As configured for these runs, a maximum of 217 neurons and 45,584 edges were possible. Evolved neuron counts ranged from 12 to 187, with a mean of 56. Evolved edge counts ranged from 33 to 13,081, with a mean of 1,077. In all, over half a million evolved graphs were analyzed using 42 different metrics (counting metrics for different neuron sets and graph types as distinct), and in excess of five million random graphs were analyzed using 24 of those metrics.

Results and Discussion

Given that we know complexity increases over evolutionary time in Polyworld and is, in fact, actively selected for by evolution under certain conditions, our intention is to develop a better understanding of the structural characteristics that give rise to these complex network dynamics. To this end we start by examining clustering coefficient.

The various neuron sets and graph types tell much the same story for clustering coefficient, as represented by the P,WD results in Figure 1. Initially CC is actively selected for by evolution, as evidenced by the more rapid rate of increase in the driven runs than in the passive runs. But once a "good enough" solution emerges and spreads throughout the population, CC in the passive populations surpasses that in the driven populations. The period during which there exists a statistically significant bias for high CC in the driven runs is from about t=1000 to t=11000. This mimics but extends the trend previously observed in neural complexity (Yaeger et al., 2008), as complexity's period of statistically significant differences lasted only from about t=1000 to t=4000, and passive complexity caught up to driven complexity by about t=7000. The period of behavioral adaptation is approximately t=1000 to t=7000 (Yaeger, 2009).

A traditional means of looking for meaningful graph structure is to compare suitable graph theoretical metrics computed for one's actual graphs to the same metrics calculated for comparable random graphs. We examined driven
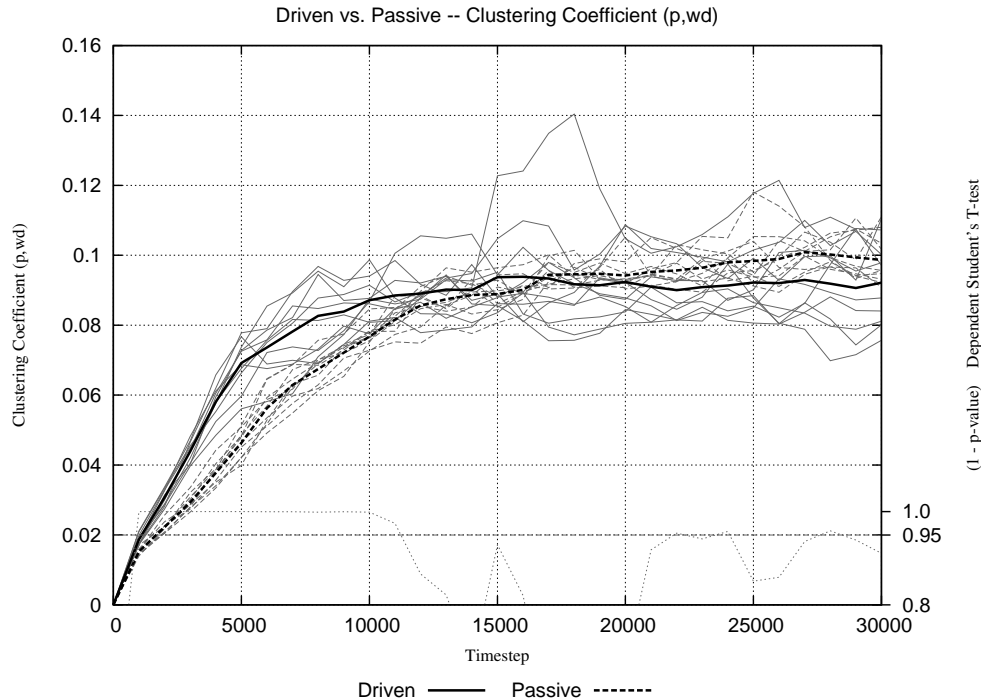
Figure 1: Driven *vs.* passive clustering coefficient as a function of time. Light solid lines show mean population CC for each driven run. Light dashed lines show mean population CC for each passive run. Heavy lines show meta-means of all ten runs for the corresponding line style. Light dotted line at bottom shows dependent $1-$p-value for a Student's T-test with typical $p > 0.05$ statistical significance indicated by the horizontal line at $p = 0.95$.

vs. random and passive vs. random CC, but do not include the results here due to space considerations. CC was substantially and statistically significantly greater in the actual evolved graphs than in the corresponding random graphs. Curiously, this difference was observed in passive vs. random as well as driven vs. random graphs, which we take as a warning that there is a bias present in our genetic encoding mechanism towards at least some degree of clustering. Given that the encoding expresses connectivity between groups of neurons, this seems reasonable. This result suggests that the differences we observe between driven and passive results may be lower than one might find with a completely unbiased encoding scheme. It also means we are probably better off focusing on driven vs. passive results than driven vs. random results, since the passive runs represent a more appropriate and tightly constrained null model than do the random graphs.

Turning to path length, the stories told by NPL and CL are very similar to each other and to that told by CC and complexity. CPL is less consistent, due to its previously discussed shortcomings, showing generally the same trends, but without much statistical significance in both WD analyses, large and greatly extended statistical significance in the P,BU analysis, and a result much like the other length metrics in the A,BU analysis. Figure 2, though somewhat noisy, shows the typical trends in path length, using NPL. Path length initially drops much more rapidly in the driven runs

than it does in the passive runs, but as that "good enough" solution becomes weakly stabilized in the driven runs, path length in the passive runs drops below that in the driven runs. In fact, path length in the passive runs drops nearly to the level seen in random graphs (not shown). The initial period of driven vs. passive statistical significance is from about t=1000 to t=7000, again corresponding well to the period of complexity growth and behavioral adaptation.

Thus we have seen that during the period of growth in the complexity of the agents' neural dynamics there is a corresponding, statistically significant growth in clustering coefficient and reduction in path length. High clustering coefficient and low path length are the defining characteristics of a small-world network. So our results are suggestive of a selective pressure towards small-world networks, and provide support for a correlation between small-world structure and complex function.

To investigate this trend towards small-world-ness, we turned to the small world index proposed by Humphries et al. (2006). As it was originally formulated, SWI is based on comparing CC and CPL in actual graphs vs. random graphs. However, given the problems previously discussed in applying CPL to our small, sparse, multi-component graphs with disconnected nodes, the standard version of SWI proved to be uninformative, displaying little consistency amongst the different neuron sets and graph types we analyzed and with sufficient noise to render some results un-
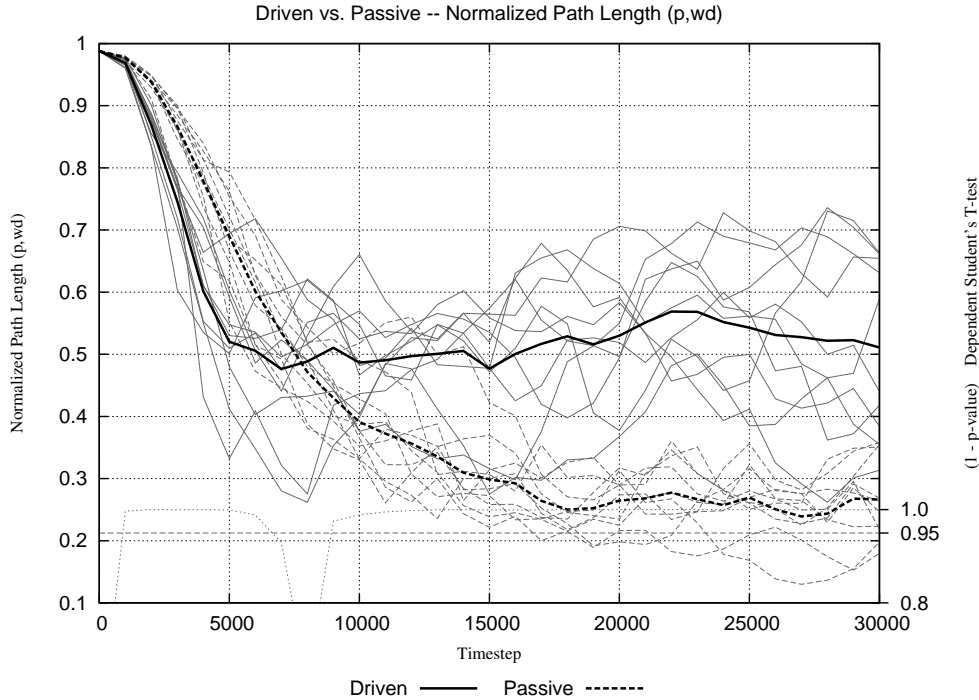
Figure 2: Driven *vs.* passive normalized path length as a function of time. Light solid lines show mean population NPL for each driven run. Light dashed lines show mean population NPL for each passive run. Heavy lines show meta-means of all ten runs for the corresponding line style. Light dotted line at bottom shows dependent $1-$p-value for a Student's T-test with typical $p > 0.05$ statistical significance indicated by the horizontal line at $p = 0.95$.

interpretable. So we developed alternative formulations of SWI, using our better behaved length metrics, CL and NPL. Curiously, some of the inconsistencies were present in these formulations as well.

We could have cherry-picked an SWI result based on NPL for the A neuron set and BU graph type that looks very much like we expected, with a statistically significantly higher growth rate in SWI for the driven runs compared to the passive runs. However, the P,WD version of this metric, even using NPL, actually reverses the roles of driven and passive (in a clear, although not significant fashion). We believe that the small and weakly connected character of our early nets are contributing to these difficulties, which explains why the problems are most exacerbated in the nets with the most limited set of connections (P,WD), but are not entirely satisfied with any of the explanations we have devised so far and feel this needs further investigation, which is why none of these results are included here (though they are all present in the supplemental materials).

The actual numerical values of all these different versions of SWI are greater than 1.0 for the driven runs, ranging from 1.5 to as much as 32.0, depending on the specific data and specific form of the metric, and the values are generally (though not always) greater for the driven runs than they are for the passive runs. So all we can really take away from the SWI analysis is that the evolved nets are small-world nets.

Given the difficulties and inconsistencies with SWI, we

sought to define a metric that would more directly capture and quantify the apparent bias towards high clustering and short path lengths evidenced in all of the raw clustering and path length data. To this end we have defined a new "small-world bias" (SWB) metric that takes its form from Humphries et al's SWI, but directly compares driven to passive—instead of actual to random—clustering and length metrics:

$$\gamma = \langle CC_{driven} \rangle / \langle CC_{passive} \rangle \qquad (6)$$

$$\lambda = \langle L_{driven} \rangle / \langle L_{passive} \rangle \qquad (7)$$

$$SWB = \gamma/\lambda \qquad (8)$$

where L can be any suitable length metric (such as CPL, CL, or NPL). The ensemble averages are taken over the usual population of agents expiring during a given temporal epoch. The numerator captures the degree to which a driven run favors high clustering relative to a passive run. The denominator captures the degree to which a driven run favors low path length relative to a passive run. Accordingly, when SWB exceeds 1.0, the driven run is at least slightly biased towards small world network characteristics relative to a passive run. It is not actually possible (because driven and passive graph sizes are different), but if one could calculate Humphries et al. (2006)'s SWI using the same random-graph basis for corresponding terms in SWI_{driven} and $SWI_{passive}$, then take their ratio, all the random-graph terms would cancel out and what one would be left with is SWB.
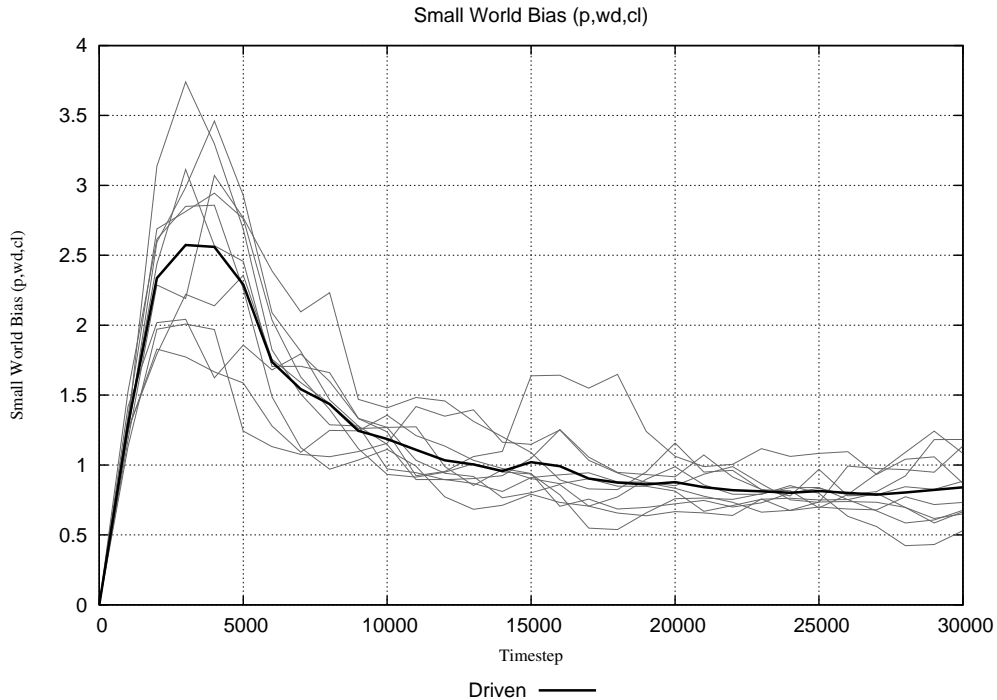
Figure 3: Small-world bias as a function of time. Where SWB is > 1.0, the driven run is exhibiting a bias towards small-world networks relative to the passive run.

Precise numerical values and periods of bias vary, but the resultant trends in SWB are remarkably consistent for both sets of neurons (A and P), both graph types (BU and WD), and all length metrics (CPL, CL, and NPL). Figure 3 shows SWB based on connectivity length for the processing neurons treated as weighted, directed graphs. There is a strong (> 1.5) bias towards small-world-ness from about t=2000 to t=7000, corresponding to the previously observed period of growth in neural complexity and behavioral adaptation to the environment. Once the agents have adapted to their environment evolutionary pressure on complexity diminishes, leading to the reduction in SWB at later times.

Conclusions

We have shown strong, reproducible evolutionary biases towards high clustering coefficients, short path lengths, and small-world-ness in driven runs subject to natural selection relative to passive runs in which natural selection is disabled. These structural, graph theoretical trends correspond to previously observed evolutionary trends in the dynamical complexity of neural function and behavioral adaptation of agents to their environment, thus strengthening the association between small-world-ness and complexity.

Short path lengths contribute to increased "integration" of neural function throughout the brain. Clustering can contribute to and is often evidence of increased "segregation" of specialized neural functions in the brain. It is this combination of increasing integration and segregation that produces the measured increases in dynamical neural complex-

ity (Tononi et al., 1994).

Our work demonstrates that even in the absence of physical constraints on wiring length and brain volume, evolution selects for small-world networks in order to enhance brain function. The resulting networks thus combine the predominantly local connectivity imposed by physical volume constraints (Murre and Engelhardt, 1995) with the short path lengths necessary to satisfy fast response time requirements (Lago-Fernández et al., 2000), despite a lack of physical constraints in their evolution. We suggest that humans (and all biological organisms with even modestly complex nervous systems) are the fortunate beneficiaries of these convergent and synergistic physical and functional constraints. Rather than physical constraints acting to limit brain function, our evidence suggests that physical constraints work in concert with evolutionary pressures to select neural topologies that foster more complex, adaptive behaviors.

Future Directions

There is one instance in which increases in clustering coefficient are not correlated with increasing neural segregation and complexity, which is progression towards a single large cluster. Since we do see correlated increases in neural complexity our clustering increases cannot be the result of network topologies approaching a single large cluster, however in the future we intend to look into modularity metrics that more directly address community structure. Our expectations are that structural modularity and functional complexity will be positively correlated. However, preliminary in-

consistent and contradictory results have led to the realization that standard measures of modularity, such as those due to Newman (2006) and Blondel et al. (2008), are not well suited to the types of networks generated early in our simulations and we believe values of these metrics are artificially elevated for such graphs. Further research is required to either develop better ways to characterize community structure in these networks or determine suitable subsets of these graphs to which the standard modularity metrics may reasonably be applied, perhaps only after having evolved beyond certain minimum size and connectivity constraints.

We further hope to identify more discriminating structural metrics, that will be reliably predictive of functional complexity. We also seek to improve upon our current technique of ignoring (by taking absolute values) what is likely to be a crucial distinction between the positive and negative weights associated with excitatory and inhibitory connections. One particular direction we intend to explore may address both aims at once, which is distributions of signed motifs. Network motifs, such as those advanced by Milo et al. (2004) and related to small-world properties and complexity by Sporns and Kötter (2004), are typically treated as unsigned, though there has been some discussion of small subsets of signed motifs in genetic transcription and other biological networks (Alon, 2007). Work by Kashtan and Alon (2005) demonstrates that modularity and motif distributions are sometimes correlated, but not uniquely so. We speculate that motif distributions may be more discriminating and predictive of functional complexity than modularity or the other metrics we have examined to date. We also expect that extending the standard 13 unsigned motifs to a corresponding 204 signed motifs will provide much greater discrimination, as well as greater relevance to neural networks.

Acknowledgements

References

Alon, U. (2007). Network motifs: theory and experimental approaches. *Nat Rev Genet*, 8:450–461.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *J. Stat. Mech.*, page P10008.

Cherniak, C. (1995). Neural component placement. *Trends in Neurosciences*, 18:522–527.

Humphries, M. D., Gurney, K., and Prescott, T. J. (2006). The brainstem reticular formation is a small-world, not scale-free, network. *Proc. R. Soc. B*, 273:503–511.

Kashtan, N. and Alon, U. (2005). Spontaneous evolution of modularity and network motifs. *Proc. Natl. Acad. Sci. U. S. A.*

Lago-Fernández, L. F., Huerta, R., Corbacho, F., and Sigüenza, J. A. (2000). Fast Response and Temporal Coherent Oscil-

lations in Small-World Networks. *Phys. Rev. Lett.*, 84:2758–2761.

Lizier, J. T., Piraveenan, M., Dany, P., Prokopenko, M., and Yaeger, L. S. (2009). Functional and Structural Topologies in Evolved Neural Networks. In Kampis, G. e. a., editor, *Proceedings of the Tenth European Conference on Artificial Life*. Springer Verlag, Heidelberg.

Marchiori, M. and Latora, V. (2000). Harmony in the Small-World. *Physica A*, 285(3-4):539–546.

McShea, D. W. (1996). Metazoan complexity and evolution: Is there a trend? *Evolution*, 50:477–492.

Milo, R., Itzkovitz, S., Kashtan, N., Levitt, R., Shen-Orr, S., Ayzenshtat, I., Sheffer, M., and Alon, U. (2004). Superfamilies of Evolved and Designed Networks. *Science*, 303:1538–1542.

Mitchison, G. (1991). Neuronal branching patterns and the economy of cortical wiring. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 245(1313):151–158.

Murre, J. M. J. and Engelhardt, D. P. F. (1995). The connectivity of the brain: multi-level quantitative analysis. *Biological Cybernetics*, 73:529–545.

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proc. Natl. Acad. Sci. U. S. A.*, 103:8577–8582.

Rubinov, M. and Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, In Press, Corrected Proof:–.

Sporns, O. and Kötter, R. (2004). Motifs in brain networks. *PLoS Biol*, 2(11):e369.

Sporns, O., Tononi, G., and Edelman, G. (2000). Theoretical Neuroanatomy: Relating Anatomical and Functional Connectivity in Graphs and Cortical Connection Matrices. *Cerebral Cortex*, 10:127–141.

Strogatz, S. H. (2001). Exploring complex networks. *Nature*, 410:268–276.

Tononi, G., Edelman, G., and Sporns, O. (1998). Complexity and coherency: integrating information in the brain. *Trends in Cognitive Sciences*, 2(12):474–484.

Tononi, G., Sporns, O., and Edelman, G. (1994). A measure for brain complexity: Relating functional segregation and integration in the nervous system. *Proc. Nat. Acad. Sci.*, 91:5033–5037.

Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442.

Yaeger, L. S. (1994). Computational Genetics, Physiology, Metabolism, Neural Systems, Learning, Vision, and Behavior or Polyworld: Life in a New Context. In Langton, C. G., editor, *Proceedings of the Artificial Life III Conference*, pages 263–298. Addison-Wesley, Reading, MA.

Yaeger, L. S. (2009). How evolution guides complexity. *HFSP*, 3(5):328–339.

Yaeger, L. S., Griffith, V., and Sporns, O. (2008). Passive and Driven Trends in the Evolution of Complexity. In Bullock, S. e. a., editor, *Proceedings of the Artificial Life XI Conference*, pages 725–732. MIT Press, Cambridge, MA.

Yaeger, L. S. and Sporns, O. (2006). Evolution of Neural Structure and Complexity in a Computational Ecology. In Rocha, L. e. a., editor, *Proceedings of the Artificial Life X Conference*, pages 330–336. MIT Press (Bradford Books), Cambridge, MA.