Four Correlates of Complex Behavioral Networks: Differentiation, Behavior, Connectivity, and Compartmentalization

Carving Networks at Their Joints

MARK A. CHANGIZI AND DARREN HE

Received July 31, 2003; revised February 11, 2004 and March 13, 2005; accepted March 13, 2005

Some of the most complex networks are those that (i) have been engineered under selective pressure (either economic or evolutionary), and (ii) are capable of eliciting network-level behaviors. Some examples are nervous systems, ant colonies, electronic circuits and computer software. Here we provide evidence that many such selected, behavioral *networks are similar in at least four respects. (1)* Differentiation: *Nodes of different types are used in a combinatorial fashion to build network* structures *through local connections, and networks accommodate more structure types via increasing the number of node types in the network (i.e., increasing differentiation), not via increasing the length of structures. (2)* Behavior: *Structures are themselves combined globally to implement* behaviors, *and networks accommodate a greater behavioral repertoire via increasing the number of lower-level behavior types (including structures), not via increasing the length of behaviors. (3)* Connectivity: *In order for structures in behavioral networks to combine with other structures within a fixed behavior length, the network must maintain an invariant network diameter, and this is accomplished via increasing network connectivity in larger networks. (4)* Compartmentalization: *Finally, for reasons of economical wiring, behavioral networks become increasingly parcellated. Special attention is given to nervous systems and computer software, but data from a variety of other behavioral selected networks are also provided, including ant colonies, electronic circuits, web sites and businesses. A general framework is introduced illuminating why behavioral selected networks share these four correlates. Because the four above features appear to apply to computer software as well as to biological networks, computer software provides a useful framework for comprehending the large-scale function and organization of biological networks. © 2005 Wiley Periodicals, Inc. Complexity 10: 13–40, 2005*

Key Words: differentiation; compartmentalization; behavior; connectivity; complexity; brain software

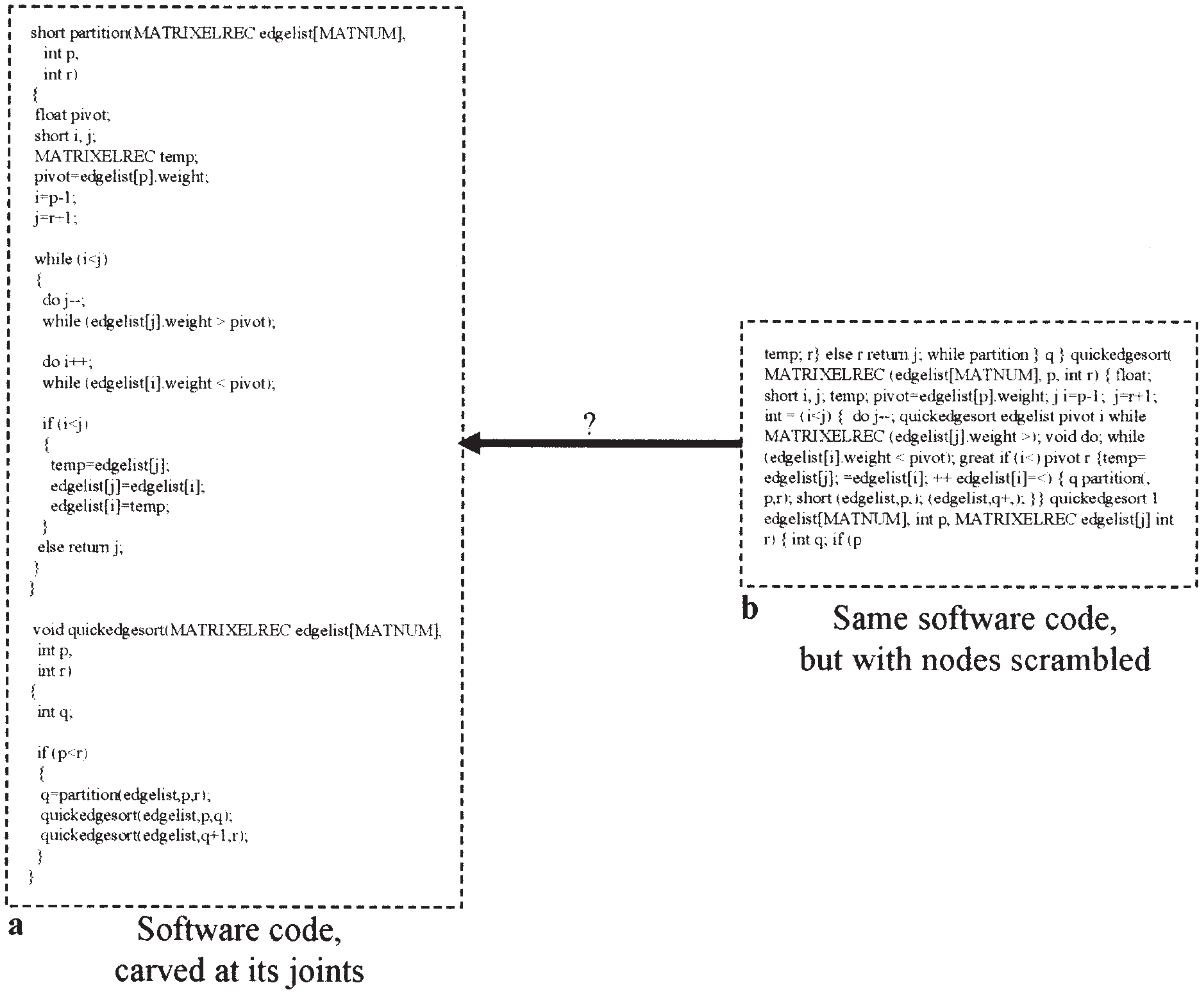
INTRODUCTION

Today we know what biological networks are made of, but it is fair to say that we have little idea how they work. For example, we know the one thousand cells of

Mark A. Changizi and Darren He are with the Sloan-Swartz Center for Theoretical Neurobiology, MC 139-74, Caltech, Pasadena, CA 91125. E-mail: changizi@caltech.edu; http://www.changizi.com

DOI 10.1002/cplx.20085

FIGURE 1



```
short partition(MATRIXELREC edgelist[MATNUM],
  int p,
  int r)
{
 float pivot;
 short i, j;
 MATRIXELREC temp;
 pivot=edgelist[p].weight;
 i=p-1;
 j=r+1;

 while (i<j)
 {
  do j--;
  while (edgelist[j].weight > pivot);

  do i++;
  while (edgelist[i].weight < pivot);

  if (i<j)
  {
   temp=edgelist[j];
   edgelist[j]=edgelist[i];
   edgelist[i]=temp;
  }
  else return j;
 }
}

void quickedgesort(MATRIXELREC edgelist[MATNUM],
 int p,
 int r)
{
 int q;

 if (p<r)
 {
 q=partition(edgelist,p,r);
 quickedgesort(edgelist,p,q);
 quickedgesort(edgelist,q+1,r);
 }
}
```

```
temp; r} else r return j; while partition } q } quickedgesort(
MATRIXELREC (edgelist[MATNUM], p, int r) { float;
short i, j; temp; pivot=edgelist[p].weight; j i=p-1; j=r+1;
int = (i<j) { do j--; quickedgesort edgelist pivot i while
MATRIXELREC (edgelist[j].weight >); void do; while
(edgelist[i].weight < pivot); great if (i<) pivot r {temp=
edgelist[j]; =edgelist[i]; ++ edgelist[i]=<) { q partition(,
p,r); short (edgelist,p,); (edgelist,q+,); }} quickedgesort 1
edgelist[MATNUM], int p, MATRIXELREC edgelist[j] int
r) { int q; if (p
```

Software code can be treated as a network, where the operators are the nodes, and program flow defines the edges. (This is commonly done in software engineering.) (a) Software code presented in the usual fashion, where the instructions, or lines of code, are readily apparent (as is the program flow). (b) The same software code, but where the operators have been scrambled (and the edges are presumed to still exist as before, so that the program still works). This is analogous to how biological networks appear to us, and one task of 21st century science is to parse the "scrambled network" into a translated, organized version where the basic functional structures (analogous to lines of code) are apparent.

the round worm *Caenorhabditis elegans* in and out, but we are in the dark when it comes to how they work together as a network whole. One reason biological networks are so perplexing is that they do not wear their organizational structure on their sleeves. In fact, they have a tendency to look like chunks of meat (e.g., brains), alphabet soup (e.g., protein networks), or mobs of scurrying individuals (e.g., ant colonies). This is in stark contrast to another kind of network, computer software code, where operators are the nodes, and program flow defines the edges. By "computer software," we refer to software actually found in "nature", i.e., actually engineered by people, as opposed to "any old" program in Platonic heaven. For any computable function there are infinitely many programs, or networks, that compute that function, but only an infinitesimal fraction of these programs will be economically organized. It is *this* class of economically organized programs that is the domain of computer software. Unlike the enigmatic biological networks, computer software *does* tend to wear its organizational structure on its sleeve—as seen, for example, in Figure 1(a)—and is, accordingly, much easier to comprehend. Consider, however, how mysterious computer software would become if the code were scrambled up, but with its program flow, or edges, unaffected, as seen in Figure 1(b). This is more akin to the way biological networks present themselves to us. Just as it is crucial for understanding computer software to have the code "carved at its joints," as in Figure 1(a), it will be crucial for understanding biological networks to carve the networks at their joints.

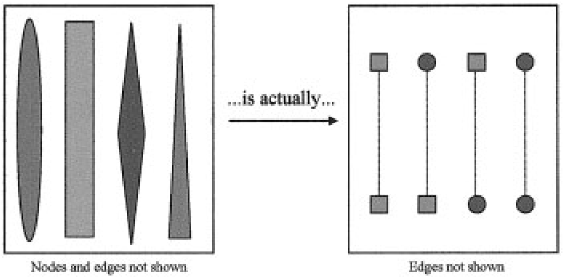
In this article we put forth a general theoretical framework for the organization of complex networks, with the idea that it will eventually aid in the parsing (or "carving at the joints") of these networks. The article is structured into four sections, each concerning one of four distinct features that appears to be shared among many complex, biological networks: (1) differentiation, (2) behavior, (3) connectivity, and (4) parcellation. Section 1 concerns *differentiation* and considers how network *structures*—the lowest-level behaviors—are put together from *nodes*. In Section 2 we look at how network-level *behaviors* are built out of combinations of lower-level behaviors in the network. We will see in Section 3 the manner in which *connectivity* and network diameter change in larger networks. Finally, in Section 4 we will see that, for reasons of economical well-connectedness, larger networks are expected to have a greater degree of *compartmentalization*. In Section 5 we summarize the relationships between these four "correlates of complex networks" and show how the drive for greater behavioral repertoire size leads to all the correlates.

Our framework is intended to apply to networks that (i) have been engineered (or "designed") under economic or evolutionary selective pressure (*selected*) and (ii) are capable of eliciting network-level behaviors (*behavioral*). Such "selected, behavioral networks" cover nervous systems, ant colonies, organisms (networks of cells), businesses (networks of employees), electronic circuits, web sites, and computer software, all of which we will cover. The framework does not apply to nonselected networks (like ecosystems or economies, where selection primarily acts at the level of the nodes) or nonbehavioral networks (like traditional Legos, or furniture). We give special attention to nervous systems, and to the mammalian neocortex in particular, and will demonstrate its similarity to many other selected, behavioral networks, including computer software. Not only does computer software have the same "correlates" as the biological networks, but *we will argue that computer software actually provides a powerful model for the largest-scale function and organization of complex networks*. This potentially has great significance for understanding the organization of biological networks, because we *know* how to think about computer software.

1. BUILDING STRUCTURES FROM NODES

Networks are composed of *nodes* (and edges), and nodes combine *locally* to build *structures* (Figure 2), which are the parts of networks responsible for the lowest-level, or primitive, behaviors. "Structures" are meant to refer to the largest objects nodes locally combine to make, and structures, in turn, behaviorally combine globally with other structures to instantiate higher-level behaviors, as we will discuss in Section 2. For example, neurons in the neocortex locally combine to make anatomical formations such as minicolumns, as well as modules (such as barrels or blobs), and these interact globally (via long-range white matter connections) with other such structures to instantiate neocortical behaviors (e.g., thoughts, representations, motor commands, etc.). In electronic circuits, electronic components combine to make simple functional circuits like an integrator [1], and these interact more globally to carry out circuit "behaviors," or device functions. And in computer software, operators combine to make instructions, or lines of code, and these globally interact to build software behaviors, or runs. Table 1 gives examples for nodes and (possible) structures in a number of kinds of network, including organisms, social insects, businesses, and in nonselected or nonbehavioral networks such as ecosystems and Legos.

FIGURE 2



Structures (primitive behaviors) in networks are built out of local combinations of nodes. Left: The rectangular box shows a network, inside of which are four distinct structures. The nodes and edges are not shown. Right: Structures are actually built out of nodes. Here, each structure is built from two nodes, where there are two node types in the network (■ and ●). The lines on the right indicate that the nodes work together as a structure; the lines are *not* the edges in the network, which are not shown. The illustration here treats the structures as nonoverlapping, but they could overlap; all that is required in the framework is that any "degree of overlap" does not itself tend to vary as a function of network size.

The question we ask in this section is this. In what manner are nodes employed to accommodate a greater number of structure types? There are broadly two different possible ways (Figure 3). Under the *universal language approach*, the set of node types is kept invariant, or universal, and the length of structures (i.e., the number of nodes in a structure) is increased (logarithmically with the number of structures, see Box 1). If this approach were followed, we would expect differentiation *not* to increase as a function of network size (see Box 2). Under the *invariant-length approach*, however, structure length is kept invariant, and the number of node types is increased (as a power law with the number of structures, see Box 1). If this approach were

TABLE 1

Some Example Hierarchical Relationships between Nodes and Structures

	Nodes	Structures (Lowest-Level Behaviors)
Behavioral, selected		
Nervous systems	Neurons	Basic circuit (minicolumns?)
Organisms	Cells	Basic cell combinations (organs?)
Social insects	Insects	Basic insect combinations (e.g., bridge?)
Businesses	Employees	Basic employee groups (committee?)
Electronic circuits	Components	Basic circuits (e.g., integrator?)
Computer software	Operators	Instructions (e.g., lines of code?)
Nonbehavioral, selected		
Legos	Lego pieces	Connections
Nonselected (nonbehavioral?)		
Ecosystems	Organism	Food chains

The first group of networks in the table are behavioral selected networks. Some nonbehavioral networks also may have structures, although they are not themselves used as symbols in higher-level behaviors. For example, Legos are nonbehavioral selected networks, and connections between pieces (i.e., pairs of pieces) may be the fundamental structure. (By "Legos" we refer to traditional Legos, not to the newer varieties that are computer-controlled and do have behaviors.) Some nonselected networks also may have structures, such as ecosystems (which are arguably nonbehavioral), where organisms interact combinatorially and make food chains (see Ref. 1).

followed, we would expect differentiation to increase as a function of network size (as a power law with positive exponent less than one, see Box 2).

In which of these manners *does* structural repertoire size increase in selected, behavioral networks? Figure 4 shows data for how differentiation scales with network size for a number of kinds of network, including two kinds of nervous network (neocortex and retina). In each case, differentiation increases with network size consistent with a power law, and thus the invariant-length approach is implicated, and the universal language approach can be rejected. This is true even for electronic circuits and computer software where universal languages are known; it is not, then, the case that the scarcity of the universal language approach is due to the difficulty in obtaining universal languages.

Another way to distinguish between the universal-language and invariant-length approaches to the construction of structures from nodes is to actually measure the lengths of structures in networks of varying size, and see if their length remains invariant. The difficulty with this, however, is that it is not always easy to know what the structures (or primitive behavior-producing parts) are in a network. Plotting differentiation versus network size as in Figure 4 allowed us to test whether there are invariant-length structures indirectly, and we were able to remain ignorant about what exactly the structures are. In neocortex, one of the most plausible notions of what a structure might be (i.e., what a local group of neurons combine to make, such that that group acts as a lowest-level behavior in the global construction of higher-level behaviors) is the *minicolumn*, which is a group of on the order of 100 neurons spanning the thickness of the neocortex [2–6]. Minicolumns *are*, in fact, invariant in length [7], or, equivalently, it is known that gray matter thickness increases in larger brains at the same rate as that expected by the overall neuron density decrease [8–17]. Another neocortical candidate for a structure is the *module*, which refers to columns (not minicolumns), blobs, barrels, and stripes. Figure 5(a) shows that the number of neurons in a module does not vary as a function of brain size [16]. In computer software, structures are most plausibly instructions, which typically lie on a single line of code. Figure 5(b) presents data showing that the number of operators per line of code is invariant.

The fact that computer software possesses invariant-length lines of code may not be empirically surprising to many readers, for most of us have seen many programs of varying sizes, and lines of code seem to be roughly the same length independent of program size. That these other diverse networks have invariant-length structures (as power-law scaling of differentiation indicates), however, is not at all empirically obvious. Not only do power-law plots of differentiation versus network size tell us that structures are invariant in length, *they tell us roughly how large the structures are*, for the inverse of the slope measures the average length of the structures. Actually, it measures the number of degrees of freedom, or *combinatorial degree*, of structures (see Boxes 1 and 2). The combinatorial degree is a number greater than or equal to 1, where 1 would imply that nodes are not used combinatorially at all, and greater values mean that nodes are used more combinatorially. For example, if the number of node types scales as the $1/2$ power of network size, then that suggests a combinatorial degree of $1/(1/2) = 2$, or two degrees of freedom per structure. For computer software in Figure 4, the inverse of the slope is a little over two, which means there are a little over two degrees of freedom (intuitively, a length of a little over two) in the construction of instructions (or lines of code). See Ref. 1 for discussion of the combinatorial degrees for many of

FIGURE 3

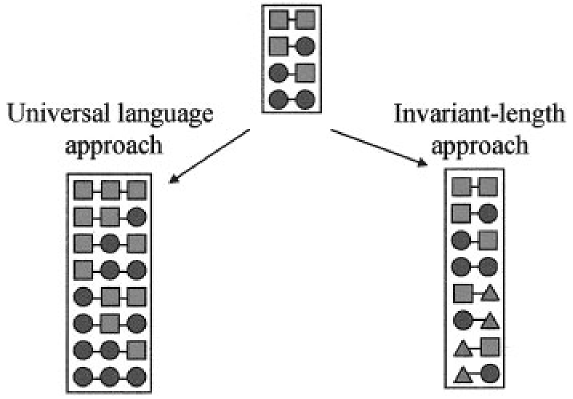


Illustration of the two main ways that structural repertoire size may be increased. Each of the three rectangles shown represents a network, where its nodes, but not its edges, are shown. The nodes come in multiple types. On a single row in each network are nodes involved in building a structure; the horizontal line connecting two nodes here indicates that the nodes interact as part of a structure. Consider the top, small network first, the one with 8 nodes coming in two types (● and ■). Supposing that its nodes can combine into structures of length two and that all such length-two combinations are allowed (or grammatical), there are four structures that can be built. How may a network accommodate twice this many structures? There are two broadly different ways. The first is the *universal language approach*: the box-network on the left has the same number of node types as the small network, but is able to string them into structures of length three, instead of just two, and it consequently can build the eight structures shown inside it. Structures are like computer instructions, or lines of code, and nodes are like operators (see Table 1). In this light, the universal language approach accepts the cost of longer lines of code (which takes up more space and may take more time to implement) for the payoff of a reduced number of node types to build. The second is the *invariant-length approach*: the network (the box on the right) has one new node type (▲), and is still confined to structures of length two. With this new node type the network can also build eight structures, as are shown row-by-row within it. In computer software terms, the invariant-length approach sacrifices on node-complexity in order to keep running-time and overall program size low. Note that all these ideas hold more generally, even when we abandon the simplification made for illustration-sake in this figure that all combinations for the allowed length are grammatical. See Boxes 1 and 2 for more details.

these networks; in Figure 4 the combinatorial degree, d, values are indicated for each of the plots.

In addition to leading to conclusions about invariant-length structures, another main, and connected, generalization we may make is that, because of the invariant-length structures, behavioral selected networks become increasingly differentiated as they enlarge. Others have studied the increase of differentiation in behavioral networks such as organisms [18–29], social insects [1, 30–32], businesses [1, 33–41], computer software [42–48], electronic circuits [1], as well on hierarchy more generally [49–52]. We see here in Figure 4 that increasing differentiation in larger networks appears to be common (possibly universal?) among selected behavioral networks, and more specifically, differentiation appears to increase as a power law. There also may be invariant length structures in nonbehavioral selected networks like Legos, and in nonselected networks like ecosystems and social organizations [53], and in each case differentiation increases with network size as well [1].

But *why* should selected behavioral networks have invariant-length structures? Size minimization is a fundamental principle in a variety of selected networks, including the brain [15, 54–73], vascular networks [74–82], and electronic circuits (e.g., [83, 84]). From Figure 3 (and see Box 2) we can see that network size scales up more slowly under the invariant-length approach than the universal-language approach. Accordingly, it is natural to hypothesize that the invariant-length approach is followed for this reason [1, 16, 52, 85]. In the case of computer software (and any other human-created network), structure length may also be constrained to be invariant because of comprehension limits of the programmer.

How general might these observations be? In the networks discussed above, the nodes are themselves complex objects in their own right, built from subsymbols. Might it be that the invariant-length approach tends to hold when the nodes are themselves complex, emergent objects, but that the universal-language approach holds at lower levels? For example, in electronic circuits, the nodes are often built from a small finite repertoire of microelectronic subcomponent types. In computer software, operators are built from a fixed repertoire of Latin letters. Cells (and neurons) are built from genes/proteins, and ultimately from four DNA bases. And, of course, all physical formations are built from on the order of 100 different atomic elements. Prima facie, then, it would seem that the universal language approach dominates at the lowest hierarchical levels. This conclusion, however, is not warranted, as we now explain.

The two possible approaches for how combinations are increased are defined relative to a designated pair of lower and upper level. To show that the universal language approach applies, one must show that the number of lower-level types, B, does not vary, over a wide range of upper-level types, C. *Furthermore*, to reject the possibility that the invariant-length approach applies—i.e., that the number of lower-level types is increasing—the range of upper-level types must be wide enough that one would expect to see an increase in the number of lower-level types if the invariant-length approach were being followed, given the likely magnitude of the power-law exponent relating the pair of levels.

For example, consider DNA as the lower level, and amino acids as the upper level. For all organisms, there are B = 4 lower-level types, and C = 20 upper-level types. That is, a log-log plot of B versus C for very many species would result in a plot with just one point, and nothing could therefore be

Box 1: Some Details about Combinatorial Systems

Simple Combinatorics

If there are 2 symbol types in a language, or combinatorial system (e.g., two lower-level behavior types, or perhaps two node types), and they are strung together into combinations of length 3, then there are obviously $2^3 = 8$ possible combinations. Letting S be the number of symbol types, L be the combination-length, and E be the number of combinations (or "expressions"), we have $E = S^L$. This, however, is insufficiently general, for (i) only some fraction α of the total possible combinations might be allowable, or grammatical, and (ii) there may be constraints, or correlations, between the co-occurrences of different symbol types and positions within the combination that reduce the effective length of combinations to $d = \beta L$, where β is a constant in the interval (0,1]. More generally, then, the equation relating S, L and E is $E = \alpha S(E)^{\beta L(E)}$, where S and L are written as functions of E to emphasize that they are not necessarily to be treated as constants. We may write this as follows:

$$E \sim S(E)^{d(E)}. \tag{1}$$

Combinatorial Degree

The exponent in Equation (1), $d = \beta L$, is the *combinatorial degree*, and it is important to understand its interpretation. When a combination has length L, there are a maximum of L potential degrees of freedom (or, a maximum base-S entropy of L). Not all of them, however, might be utilized in a language. For example, in English, if you place some arbitrary word at the beginning of a sentence, it typically severely constrains the set of probable words that might come next. Such constraints reduce the overall number of degrees of freedom to a value below L. The combinatorial degree, d, is the actual number of degrees of freedom (or base-S entropy) of combinations for a combinatorial system. Intuitively, it is the effective length of combinations, in that combinations scale *as if* they have lengths of d (rather than L). (When $L > d$, we say that the system is *redundant*.) For example, English sentences typically have a length of on the order of 20 or so words, but their combinatorial degree is only on the order of about 5 [85]. It is called "combinatorial degree" because it measures the "degree of combinatorialness" of a combinatorial system. The lowest possible combinatorial degree is 1, and in that case the combinatorial system is not combinatorial at all, for in this case, to double the number of combinations E requires doubling the number of symbol types S. Higher combinatorial degrees mean that the language is increasingly combinatorial. The combination length L puts an upper limit on the combinatorial degree. (See Ref. 1 for the relationship with Shannon's entropy.)

Increasing the Number of Possible Combinations

Figure 3 shows the two central ways for a kind of combinatorial system, or language, to increase the number of combinations allowed. The first is the *universal language approach*, where the number of symbol types remains invariant: $S \sim E^0$. To achieve more combinations, the combinatorial degree (and length) must be increased, and, solving for d in Equation (1), it must increase logarithmically: $d \sim \log(E)$. The "opposite" way to increase the number of combinations is the *invariant-length approach*, where, instead, the combinatorial degree (and length) remains invariant: $d \sim E^0$. To accommodate more combinations, the number of symbol types must be increased, and it must increase as a power law: $S \sim E^{1/d}$, where d is constant and ≥ 1. Thus, a log-log plot of S versus E gives a straight line, and the inverse of the slope is the combinatorial degree d. The following summarizes how the number of symbol types, S, and the combinatorial degree, d, scale as a function of the number of allowed combinations.

- *Universal language*: $S \sim E^0$ and $d \sim \log(E)$.
- *Invariant-length*: $S \sim E^{1/d}$ and $d \sim E^0$.

It should be recognized that there are a variety of in-between possibilities, where both d and S increase, such as where $S \sim \log(E)$, in which case Equation (1) leads to the sublogarithmic scaling relationship $d \sim \log(E)/\log(\log(E))$. However, in this article we treat the two approaches—the universal-language and invariant-length approaches—as two null hypotheses, and we will primarily aim to reject one or the other.

concluded concerning the possible scaling approaches. Consider, instead, choosing DNA again as the lower-level, but now choosing some much-higher level as the upper level, such as the number of cell types, E. Again, B = 4 for all life on Earth, but we will now have a range for the upper level, E, from 1 to on the order of 100. At first glance, this suggests the universal-language approach. However, to reject the invariant-length approach, one must recognize that, because the levels are so far removed from one another, the combinatorial degree will tend to be extremely high [52],

Box 2: Structures and Network Size

Combination Repertoire "Mass"

The *"mass" of a combination repertoire*, N_E, is the total number of symbols required to write down a list of all the (E many) combinations in the combination repertoire, or equivalently, the sum of the lengths of all the combinations. For example, for bird vocalization (see Section 2), a given species' song repertoire mass would be the sum of the lengths of all the different songs. For nodes combining into structures (see Section 1), the structure repertoire mass is the sum of the lengths of all the structures in the network. The combination repertoire mass, N_E, is just the number of combinations, E, times combination length, L. The following summarizes how the combination repertoire mass scales under the two main scaling possibilities (see Box 1), and how the number of symbol types, S, scales with mass (using what we know from Box 1 concerning how length scales with number of combinations, E).

- *Universal language*: $\mathrm{N_E} \sim \mathrm{Elog(E)}$ and $\mathrm{S} \sim \mathrm{N_E^0}$.
- *Invariant-length*: $N_E \sim E$ and $S \sim N_E^{1/d}$.

Network Size

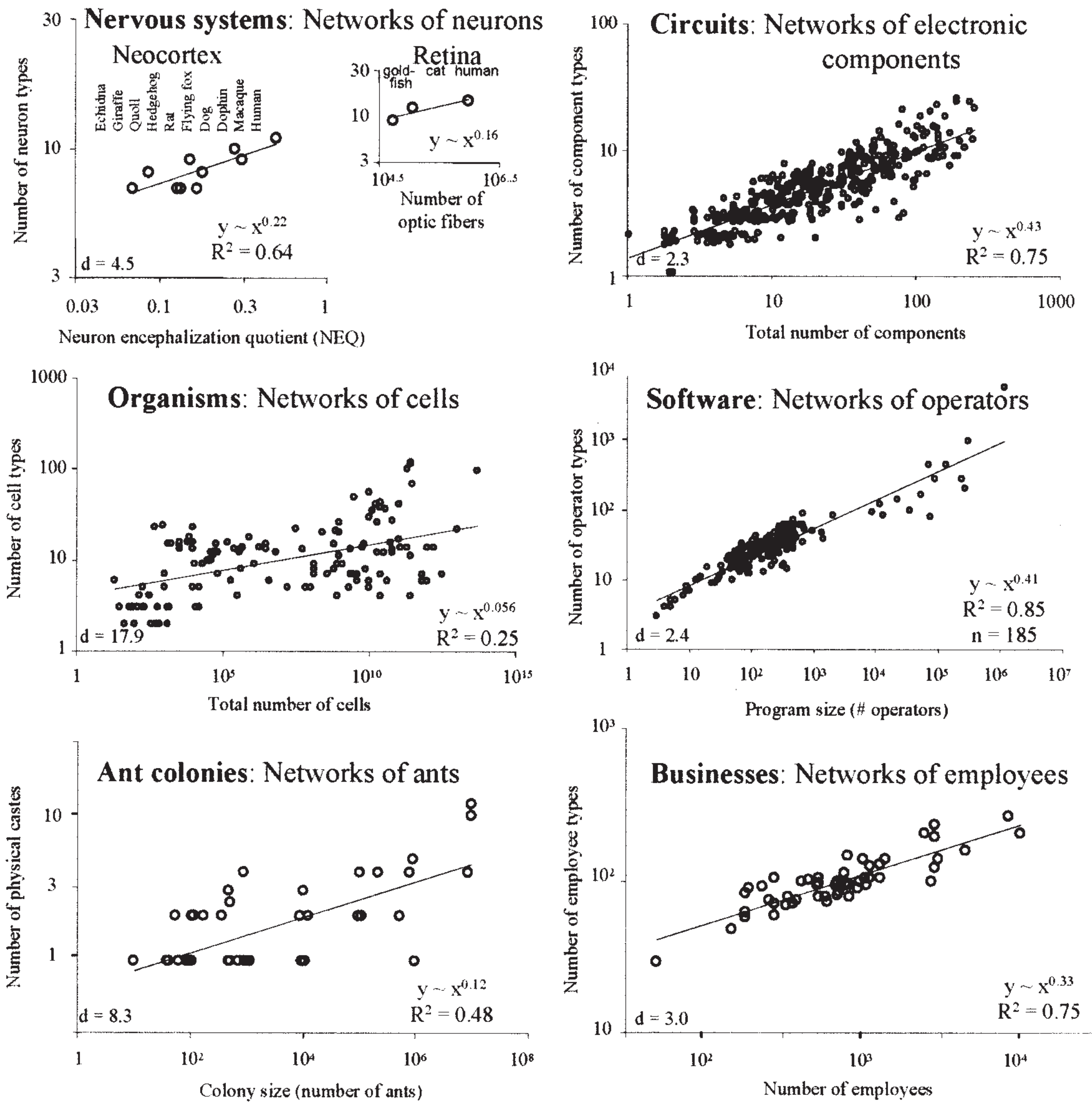
When the symbols are nodes, and the combinations are structures (as in Section 1), how does the combination repertoire mass, N_E, relate to the size of the network, N (i.e., the total number of nodes)? Nodes of different types combine *locally* to build structure types, and thus in networks with more structure types, the number of *copies*, or tokens, of each node type must increase (see Figure 2). Structures, however, interact *globally* with other structures in order to build behaviors (see Section 2), and it is therefore possible for the number of copies of each structure type to remain invariant as the number of structure types increases. This is, in fact, what one might expect for selected networks, where there is selective pressure to minimize the overall size of the network (see Refs. 1 and 85, and see the discussion in Section 1), and in this case network size, N, is driven primarily by combination repertoire mass, N_E, so that $N \sim N_E$. This has been called the "Minimal N Hypothesis" [1]. If the invariant-length approach applies for nodes combining to make structures—which is *also* expected under a size-minimization desideratum—then it would follow that $N \sim E$, i.e., network size would scale directly proportional to the number of structure types. And, the inverse slope of a log-log plot of the number of node types, S, versus network size, N, would be a measure of the combinatorial degree relating S and E. Unless otherwise stated in the text, our default operational assumption will be that $N \sim N_E$ for selected networks. That is, we will assume that network size, N, is primarily driven by structure repertoire mass, N_E, and given this, the following summarizes how network size relates to the number of structure types, E, and also how the number of node types, S, scales with network size, under the two possible scaling approaches.

- *Universal language*: $N \sim E\log(E)$ and $S \sim N^0$.
- *Invariant-length*: $N \sim E$ and $S \sim iN^{1/d}$.

Our central observations will rely only upon the weaker assumption that N and N_E are related by some power law.

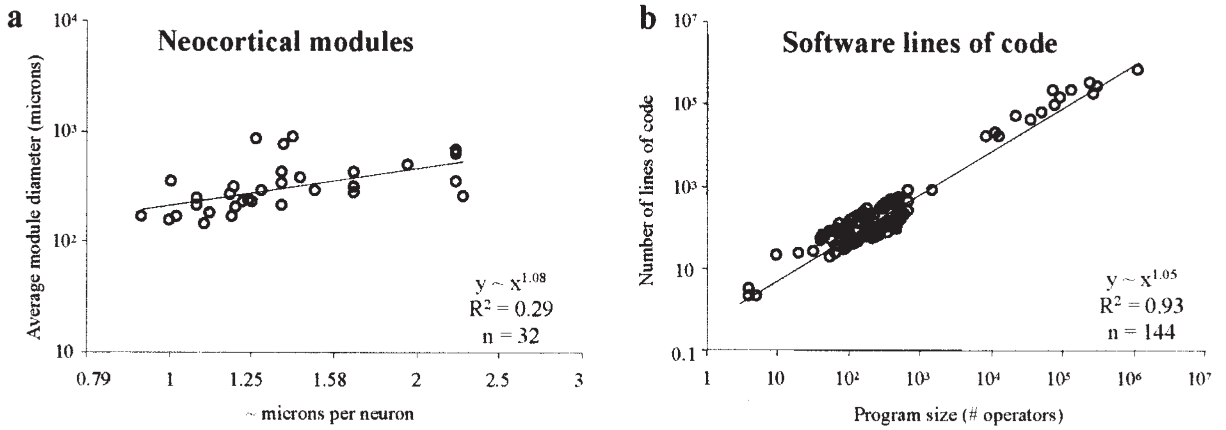
and the invariant-length approach would expect an accordingly very small, possibly negligible, increase in the number of lower-level types. As a toy example, consider the following four levels: let B refer to the number of DNA base types, C the number of amino acid types, D the number of protein types, and E the number of cell types. What are the power-law exponents relating each adjacent pair of levels? Given the actual number of bases and amino acids (i.e., 4 and 20), B and C are related as $\mathrm{C} = \mathrm{B}^{2.16}$, i.e., an estimated combinatorial degree of $\mathrm{b} \approx 2$ between that pair of levels. We have no data for the combinatorial degree, c, relating levels C (amino acids) and D (proteins), and will momentarily leave it variable. We will approximate the combinatorial degree between levels D and E by the combinatorial degree relating the number of genes and the number of cell types, which is approximately on the order of 3, or $\mathrm{d} \approx 3$ [1]. Consider now the *total* combinatorial degree relating B and E. By combining the previous scaling equations for the adjacent levels, we have $\mathrm{E} \sim \mathrm{D^d} \sim [\mathrm{C^c}]^{\mathrm{d}} \sim [(\mathrm{B^b})^{\mathrm{c}}]^{\mathrm{d}} = \mathrm{B^{bcd}}$; i.e., $\mathrm{E} \sim \mathrm{B^{bcd}}$. One can see that the combinatorial degree between far-removed levels—here between B (base types) and E (cell types)—is the *product* of the combinatorial degrees for all the adjacent pairs in between. In particular, with $\mathrm{b} \approx 2$ and $\mathrm{d} \approx 3$, we have $\mathrm{E} \sim \mathrm{B^{6c}}$. The combinatorial degree, c, relating levels C and D—amino acids and proteins—is plausibly quite high, but even if it were only 5, we would have $\mathrm{E} \sim \mathrm{B}^{30}$, or $\mathrm{B} \sim \mathrm{E}^{1/30}$. The number of cell types, E, ranges over a factor of about 100, and from this we would expect the number of base types, B, to range over a factor of $(100)^{1/30}$, or only 1.16. Therefore, given the variation that exists on Earth, it is not possible to determine whether DNA follows a universal-language or an invariant-length approach. What one would

FIGURE 4



How greater structural repertoire size is accommodated in a variety of behavioral networks. Each plot is a log-log (base 10) plot of number of node types versus network size. The moral here is that the number of node types (i.e., network differentiation) *increases* as network size and the number of structures increases, and thus the universal language approach is not followed. Instead, the invariant-length approach appears to be followed, suggesting that there are invariant-length structures (think: invariant-length lines of code) in each of these networks. The data for neocortex comes from Ref. 132 and Ref. 133, where we counted all distinct types of neurons they found using immunoreactive and morphological criteria. Neocortex network size was measured as the number of neurons corrected for the body size of an animal, called the neuron encephalization quotient, NEQ (computed as the 2/3 power of the traditional encephalization quotient, which is brain volume divided by body mass to the 3/4 power, see Ref. 1). NEQ was used rather than the number of neurons because it is known that encephalization quotients correlate well with measures of behavioral complexity [see Figure 9(a)], but that brain size does not [16, 111, 134]. Plotting the number of neuron types against total neuron number (not shown here), uncorrected for body size, leads to a lower slope (0.04) and correlation ($R^2 = 0.38$), but still significant (df = 8, $t = 2.07$, $p < 0.05$ for one-way *t*-test). The data for retina are from [135–137] (counts do not include epithelial or glial cells). Each point in the plot for organisms is a species, for more than 30 phyla, and the data are from Ref. 20. See Ref. 1 for more information about the other plots, except for software, which is new here, and was obtained from data in [42–48]. (Note that Halstead's theories [42] are not well-founded, but using his data is appropriate.) The business plot is for employment insurance companies [38]. Best-fit power law and correlation are shown. Also shown are the combinatorial degree values, *d*, for the systems, which is the inverse of the best-fit exponent, and measures how combinatorially the node types are used to build structures.

FIGURE 5



Evidence of invariant length structures in nervous systems and software. Figure 4 shows that differentiation (i.e., the number of node types) scales as a power law with network size, which implies the invariant-length approach. Another way to test that networks follow the invariant-length approach is to see if the structures in that kind of network have invariant length. This has the disadvantage of requiring us to hypothesize about what the structures might specifically be (whereas the power law plots of Figure 4 allow us to conclude "invariant-length" without actually knowing what the structures exactly are). (a) In neocortex, one of the more plausible ideas as to what a structure might be is the *minicolumn*. The number of neurons in a minicolumn (i.e., minicolumn length) indeed appears to be independent of brain size (see Ref. 7), and this can also be concluded from the fact that neocortical thickness scales as the 1/9 power of gray matter volume, just fast enough to counteract the linear density decrease of $-1/9$ (see Ref. 15). Another plausible idea as to what a neocortical structure might be is the *module*, which refers to columns (not minicolumns), barrels, blobs, stripes, and other anatomical features. (Modules have many minicolumns in them.) The plot shows a log-log (base 10) plot of the average diameter of a module (measured in microns along the neocortex surface, and taken from Ref. 138) versus the inverse of linear neuron density. Linear neuron density is measured as the $-1/9$ power of brain volume (see Ref. 15). The idea here is this: Neuron density decreases in larger brains, so an invariant number of neurons per module would nevertheless mean a physically larger module. By plotting these two quantities, a slope of one would mean that modules are expanding in size only as much as would be expected due to the density decrease, and thus the number of neurons per module is actually invariant. The slope is, indeed, very close to one. (See also Ref. 16.) (b) A structure in computer software is something like an instruction, which is usually on a single line of code. If lines of code have invariant length, we expect that the number of lines of code should scale proportionally with network size. The plot here shows that, indeed, they scale proportionally. These data are from Refs. 44 and 48.

ideally require are alternative possible life-forms—ones that did not, as a matter of fact, come to be—where some utilize a much greater number of amino acid types than the actual 20, and measure whether they have any increase in the number of base types (invariant-length approach), or whether they, instead, have an increased codon length (universal language approach).

The argument we have made concerning DNA applies in an analogous fashion to the other kinds of networks where there seem to be, at first glance, a universal set of subcomponents (e.g., a fixed set of microelectronic subcomponents underlying all electrical component types). In most such cases, it is not possible to distinguish between the distinct scaling possibilities, and thus they do not provide counterexamples to the invariant-length approach. That is, at the lowest hierarchical levels in hierarchies with multiple levels above, both the universal-language and invariant-length approaches expect the number of lowest-level types to remain constant. From the results of this section, one might reasonably argue as follows: in the cases (i.e., pairs of levels) where we *are* able to test between the universal-language and invariant-length approaches, we find greater conformance to an invariant-length approach. This should lead us to favor this approach in those cases for which data do not exist.

We record our empirical findings in regards to structural repertoire size increase as the following "observation."

> **Observation 1:** Invariant-Length Structures and Power-Law Differentiation: Behavioral selected networks (at least many for which the nodes are themselves complex objects) increase their structural repertoire size by increasing the number of node types, not by increasing the length (nor the combinatorial degree) of structures (i.e., the number of nodes per structure).

The predominance of nonuniversal-language approaches we see here and in Section 2 may provide reasons for doubting Wolfram's Principle of Computational Equivalence [86], although there is some uncertainty in the intended meaning of the principle [87].

TABLE 2

Some Example Hierarchical Relationships for Behaviors, where the Lower-Level Behaviors Combine to Make Higher-Level Behaviors

	Lower-Level Behavior	Higher-Level Behavior
Mammalian behavior	Basic muscle actions	Simple ethogram behaviors
Bird vocalization	Syllables	Songs
Electronic devices	Basic button-press actions	Device functions
English language	Words	Sentences
Writing systems	Strokes	Characters
Computer software	Instructions (lines of code)	Runs

2. BUILDING BEHAVIORS FROM SIMPLER BEHAVIORS

Behaviors are often hierarchically organized, with lower-level behaviors combining to make higher-level behaviors, and these, in turn, combining to instantiate still-higher-level behaviors [88–100]. For example, mammalian behavior can be viewed as the interaction of multiple muscle activations of different types to carry out a whole-body behavior. In bird vocalization, syllables are put together to make songs. In computer software, instructions (each typically on a single line of code) sequentially cooperate to run a computation, which can be thought of as a computational behavior. And, in electronic devices (like a calculator, say), simple actions caused by a button-press combine to realize a device function (such as “3 + 2 = ”; see Table 2).

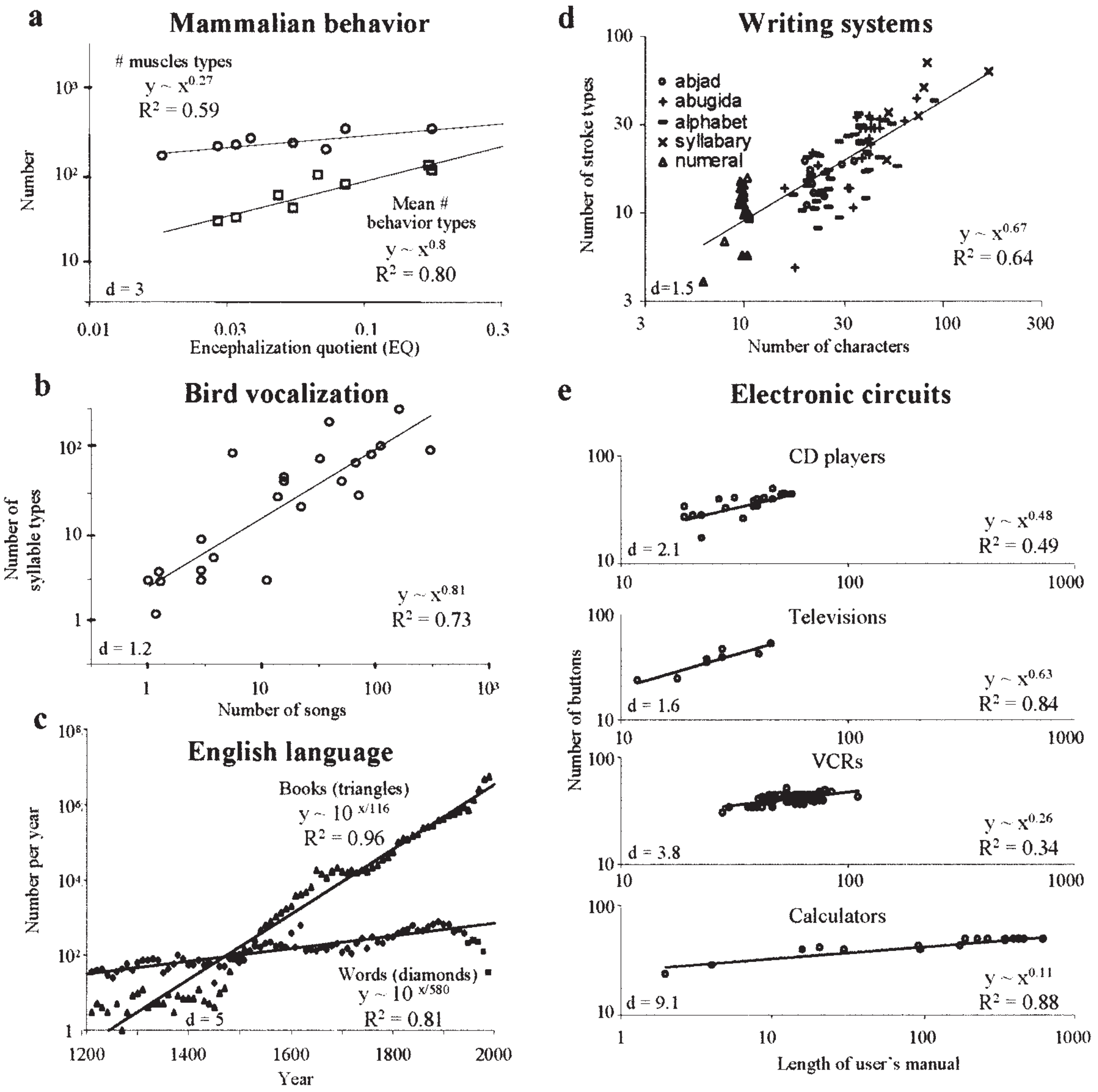
In this section we ask, In what manner are lower-level behaviors used to build a greater number of higher-level behaviors? There are broadly two different ways, analogous to our earlier discussion for nodes and structures. The first is the *universal language approach*, where the number of lower-level behavior types is kept fixed, or universal, and a greater number of higher-level behaviors is achieved via increasing the length of the higher-level behaviors (i.e., via increasing the number of lower-level behaviors per single higher-level behavior). Higher-level behavior length must, in fact, scale up as the logarithm of the number of higher-level behaviors for this case (see Box 1). The second is the *invariant length approach*, where the length of a higher-level behavior is kept fixed, or invariant, and a greater number of higher-level behaviors is accommodated via having a greater number of lower-level behavior *types* with which to build the higher-level behaviors. The number of lower-level behavior types must, in this case, increase as a power law with the number of higher-level behaviors, with a positive exponent less than or equal to one (see Box 1).

So, how *do* lower-level behaviors accommodate a greater number of higher-level behavior types? Figure 6 presents data for five kinds of lower-level/higher-level behavior pairs: (6a) Number of muscle types (i.e., number of muscle-level behavior types) versus number of behaviors as measured by ethologists, across mammals, (6b) number of bird syllable types versus number of songs in the repertoire, across a variety of species of song bird, (6c) vocabulary growth versus growth in the number of written sentences, in the English-speaking community over 800 years, (6d) number of stroke types versus number of characters, across writing systems over human history, and (6e) number of basic electronic device actions (measured as the number of button-press types, i.e., the number of buttons) versus the number of higher-level device functions (measured by the length of the user’s manual), across compact disc players, televisions, video casette recorders, and calculators. In each case, the number of lower-level behavior types increases as the number of higher-level behavior types is increased (Figure 6). In each case the increase is consistent with a power law, and thus the invariant-length approach is implicated, and we can reject the universal language approach.

How general are these results concerning lower and higher-level behaviors? For the behaviors discussed above (and in Figure 6), the lower-level behaviors are complex behaviors in their own right, built from even lower-level behaviors. Might it be that the invariant-length approach tends to hold when the behaviors are themselves higher-level behaviors, but that the universal language approach holds at the lowest behavioral levels? *First*, as we have learned from the discussion in Section 1, when there are many levels above a given lower-level, one must be cautious if one finds an unchanging number of lower-level types, for it is often the case that *both* the universal-language *and* invariant-length approaches expect a constant number of lower-level types. For example, present-day computers may possess microprocessors with a small, universal set of primitive behaviors, and this might lead one to think the universal-language approach applies. However, given the plethora of hierarchical levels in contemporary computation, as was the case for DNA, the invariant-length approach may well expect the lowest-level behavioral repertoire size to not increase over the actual range of computational behaviors found among computers today. *Second*, and as we now argue, the results from Section 1 tell us that the lowest behavioral levels *do* follow the invariant-length approach.

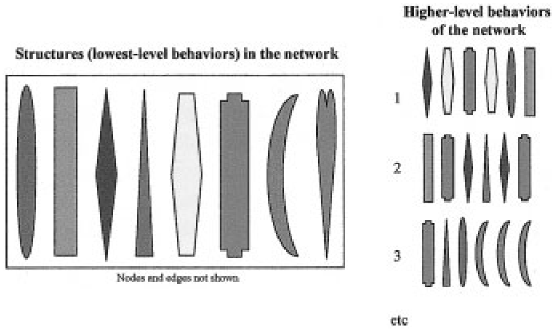
Structures—themselves the largest objects built locally from nodes, as discussed in Section 1—are defined as the primitive, or lowest-level, behavior-expressing parts of a network, which interact globally with one another to implement higher-level behaviors (see Figure 7). The way that structures interact to build higher-level things is different in kind than the way nodes interact to build structures, for structures interact *globally* with other structures, whereas

FIGURE 6



How lower-level behaviors accommodate a greater number of higher-level behaviors. The moral here is that the number of lower-level behavior types (think: symbols) *increases* as the higher-level behavioral repertoire size increases, and thus the universal language approach for behaviors can be rejected. Combinatorial degree values, *d*, are shown, which measures how combinatorially lower-level behaviors combine to build higher-level behaviors; *d* is the inverse of the best-fit power-law exponent of the number of lower-level behaviors versus the number of higher-level behaviors. (a) Log-log (base 10) plot of number of muscle types (e.g., triceps, sartorius) and number of ethobehavior types (simple whole-animal behaviors), each as a function of encephalization quotient (which is brain mass corrected for body mass). Each point represents a mammalian order (for data see Refs. 16 and 134). One can see that the number of muscle types (slowly) *increases* as behavioral complexity increases. (b) Log-log plot of number of syllable types versus number of songs in 23 species of bird (for data see Ref. 16). (c) Plots of the log growth of the English vocabulary (measured by the Oxford English Dictionary, Second Edition) and sentences (measured by the number of new books, via WorldCat) over the last 800 years. Vocabulary size grows disproportionately slowly compared to the number of sentences, with a combinatorial degree of approximately 5, consistent with comprehension limits. See Refs. 16 and 85 for full discussion. (d) Log-log plot of the number of stroke types versus number of characters in 91 scripts (including 18 numeral systems, 36 alphabets, 22 abugidas, 9 abjads, and 6 syllabaries) from human history. See Ref. 139 for full discussion. (e) Log-log plots of number of buttons versus number of user's manual pages, for four kinds of electronic device [52, 85]. Each button-press elicits a lower-level behavior from the electronic circuit network underlying the device, and combinations of these "symbols" are higher-level behaviors. Electronic devices like those shown here are useful kinds of electronic circuits for studying behaviors, for they come with user's manuals which aid us in estimating the number of higher-level behaviors (under the plausible assumption that doubling the number of device-functions tends to double the length of the user's manual).

FIGURE 7



Behaviors by networks are built out of global combinations of lower-level behaviors, and so on until the lowest level of primitive behaviors. Such primitive behaviors are elicited by parts of a network, or *structures*: a structure is a local combination of nodes that expresses a primitive behavior. Left: The rectangular box shows a network, inside of which are eight distinct structures. The nodes and edges are not shown. Right: These structures act as primitive behaviors in the global construction of network-level behaviors, and a list of such behaviors is shown, built from sequences of structures. (These behaviors may, in turn, be used as components to build yet higher-level behaviors.)

nodes interact *locally* to build structures. It is the global interaction that makes it intuitively reasonable to call the interactions "behaviors" of the network, and the structures the "primitive behaviors." This difference in kind also leads to a difference in that the number of copies of a given node type must increase as that node type is used in more structures, but the number of copies of a given structure type need not increase as that structure type is used in more behaviors. This is because one (or some fixed number of copies of the) structure type token can itself be used in many different higher-level behaviors, akin to the way that a single muscle, say a bicep, participates in a host of behaviors. Structures may accommodate increased higher-level behavioral repertoire sizes in the two possible manners discussed earlier, the universal-language and invariant-length approaches, and illustrated in Figure 8.

FIGURE 8

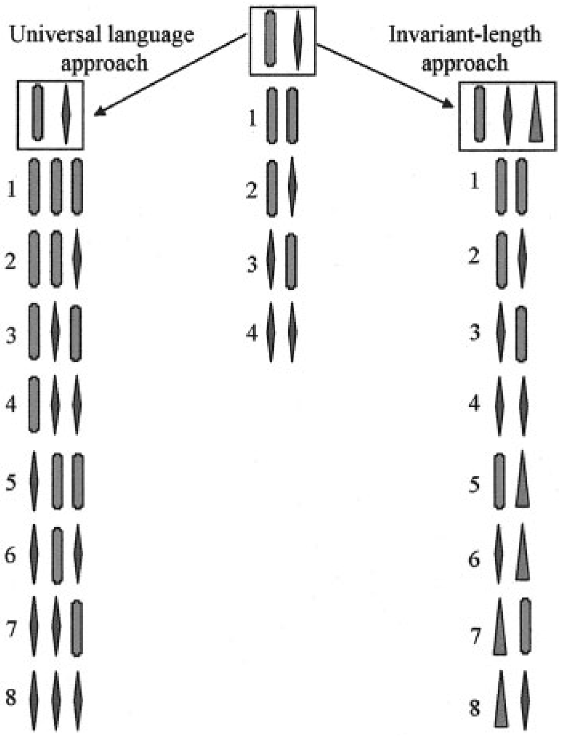
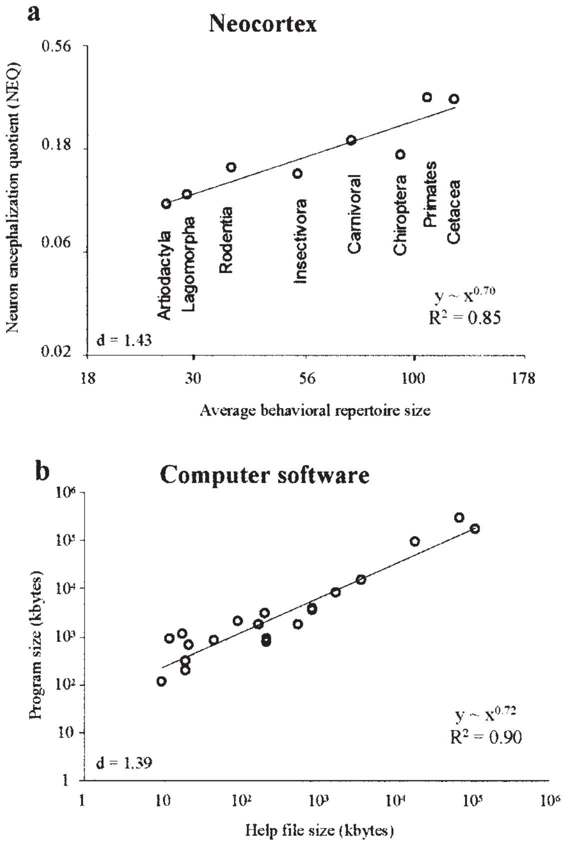


Illustration of the two main ways that structures may accommodate an increased behavioral repertoire size. At the top is a box depicting a network with two distinct structures. Supposing for simplicity that the network can build behaviors of length two (i.e., two structures per behavior), and that all such length-two combinations are allowed (or grammatical), its repertoire consists of the four behaviors shown below the box. How may a network accommodate twice this many behaviors? There are two broadly different ways. The first is the *universal language approach*: the box-network on the top left has the same two structures, but is able to string them into behaviors of length three, consequently having the eight behaviors shown below the box. Behaviors are like computer runs, and structures are like instructions, or lines of code (see Table 2), and in this light we may say that the universal language approach sacrifices on running-time (i.e., number of instructions per behavioral run) in order to keep the program (i.e., network) size to a minimum. The second is the *invariant-length approach*: the network (the box on the upper right) has one new kind of structure, and is still confined to behaviors of length two. With this new structure the network can also build eight behaviors, as shown below it. In computer software terms, the invariant-length approach sacrifices on size in order to keep running-time low. All these same ideas hold more generally, even when we abandon the simplification made for illustration sake in this figure that all combinations for the allowed length are grammatical (see Box 1).

How *do* structures in behavioral selected networks accommodate a greater number of higher-level behaviors? Our results from Section 1, concerning how node differentiation increases with network size, inform us that structures conform to the invariant-length approach. Here is why. The number of node types increases with network size (see Figure 4) and that must be in order to build a disproportionately increasing number of structure types (since nodes are used combinatorially). Or, equivalently, if the number of structure types did not scale up with network size in order to accommodate a greater number of higher-level behaviors, then we would not expect the number of node types to increase with network size. But the number of node types does increase with network size, meaning that the number of structure types must also be increasing. The invariant-length approach therefore appears to govern even lowest-level behaviors; that is, *structures follow the invariant-length approach when they globally combine to make higher-level behaviors*.

We may now more explicitly examine how the number of structure types scales with the number of higher-level network behaviors. Under the invariant-length approach for nodes combining into structures, network structures have invariant lengths, and thus network size scales in direct proportion to the number of structure types. (Network size would scale a bit more quickly than the number of structure types if the universal-language approach for nodes combining into structures were to hold, as discussed in Box 2.) If the invariant-length approach also holds for structures combining to make higher-level behaviors, then we expect network size to scale up as a function of the number of higher-level behaviors and to do so disproportionately slowly. The universal-language approach, on the other hand, would expect that network size does *not* increase as a function of the number of higher-level behaviors (see Figure 8). Figure 9 shows how network size scales with behavioral repertoire size, for neocortex [Figure 9(a)] and computer software [Figure 9(b)]. In Figure 9(a), neocortical network size (on the y axis) is measured as the number of neurons in neocortex, corrected for body size (i.e., the neuron encephalization quotient, see legend of Figure 4), and one can see that as behavioral repertoire size (along the x axis) increases, so does the neuron encephalization quotient (i.e., the measure of network size, along the y axis). Furthermore, also as expected, the neuron encephalization quotient increases disproportionately slowly as a function of the behavioral repertoire size (i.e., the exponent is <1). Neocortical structures, then, do not follow the universal-language approach in the construction of higher-level behaviors. Figure 9(b) shows data on how software size increases as a function of the size of the help file. As in electronic devices, help file size may be expected to scale more in line with the total behavioral repertoire size of the software, since it is this behavioral repertoire that needs to be explained to the user. The plot in Figure 9(b) shows that software size (y-axis) increases as a function of software behavioral repertoire size (x-axis), and also increases disproportionately slowly. Computer software, therefore, also does not follow a universal language approach in using structures to accommodate greater behavioral repertoire size.

FIGURE 9

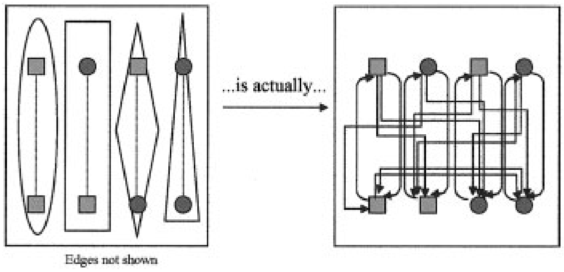


How structures accommodate greater behavioral repertoire size. In each case, network size (y axis) *increases* as a function of the number of higher-level behavior types (x axis), consistent with greater behavioral repertoire size being accommodated by a greater number of structure types (which requires greater network size). (a) Log-log (base 10) plot of neuron encephalization quotient (i.e., neocortical network size, corrected for body mass, see legend of Figure 4) versus number of behaviors (as measured within the ethology literature, see Refs. 16 and 134). We use neuron encephalization quotient here, rather than total neocortex network size, since ethological behavioral repertoire size correlates much better with neuron encephalization quotient than with brain size, the latter correlation which is $R^2 = 0.10$ (not shown), and is not significant (df $= 6$, $t = 0.82$, $p > 0.2$). Larger brain size may correlate with greater behavioral repertoire size as well, but where the behaviors are not the outwardly visible ones measured by ethologists, but, instead, "housekeeping" ones dealing with the animal's greater body mass. (b) Log-log plot of program size versus help file size for software found in a Windows operating system (EasySQLite, NotesSQL, Quicken 2001, Lotus 123, Winzip, PowerDVD, Encarta Encyclopedia, Encarta World Atlas, Encarta Dictionary, Rockwell, Siemens Wireless PC Card, Symantec LiveUpdate, EasyCalc, Office, Matlab, Intuit, Easy Button, Windows Media Player, Wordpad, Internet Explorer).

Let us sum up what we have learned thus far in this section. Figure 6 shows that the invariant-length approach applies to a wide variety of lower and upper behavioral levels, when the lower-level is (possibly) not the lowest level (the structures). We then argued that the fact that node differentiation increases in larger networks—i.e., the results from Figure 4 in Section 1—tells us that the invariant-length approach is also followed by the structures, the lowest-level behaviors, in their combining to build higher-level behaviors. This is because node differentiation would only increase in order to increase the number of structure types; therefore, the fact that node differentiation increases in larger networks is a signal that the number of structure types is increasing, something that would *not* occur if the universal-language approach applied to structures combining to build higher-level behaviors. In addition, we saw more explicitly in Figure 9 that the number of structure types increases (and disproportionately slowly) as higher-level behavioral repertoire size is increased. Together, these results suggest that behaviors are typically built using the invariant-length approach, even at the lowest behavioral levels, the structures. We record this as a second "observation."

> **Observation 2:** Invariant-Length Behaviors: Behavioral selected networks increase their higher-level behavioral repertoire size largely (or exclusively) by increasing the number of lower-level behaviors, not primarily by increasing the length (nor the combinatorial degree) of the higher-level behaviors. This appears to hold even for the lowest-level behaviors, or structures, in their combining to build higher-level behaviors.

FIGURE 10



Nodes in networks are connected via edges. Left: The rectangular box shows a network, inside of which are four distinct structures (the big shaped outlines), and one can see within them the nodes making them up. The edges are not shown. Right: Nodes interact with one another via edges, which up to this point, we have ignored.

This "behavioral nonuniversality" is a central driver of the organization of behavioral selected networks, as we will discuss in detail in Section 5, for it is because of this that (a) the number of structure types (the lowest-level behaviors) increases, and (b) behaviorally more complex networks become larger (see Figure 8). The fact that the number of structure types increases explains two organizational features. First, it explains why node differentiation increases: in order for nodes to build the greater structural repertoire using the invariant-length approach requires new node types (see Figure 3). Second, as we will see in Section 3, a greater number of structure types requires disproportionately quickly increasing the total number of edges between structures. Finally, the fact that behaviorally more complex networks become larger leads to greater compartmentalization, as we will see in Section 4.

Why, one might ask, do networks of a given kind need to increase the number of structure types (and thus increase network size) at all? Why not, instead, just build more and more high-level behaviors using a universal repertoire of structure types? Figure 8 depicts overall network size remaining invariant under the universal language approach, but this is, in fact, not possible in the long run, for reasons having to do with Kolgomorov complexity: as the repertoire of higher-level behaviors increases, eventually the minimum size program that can implement that repertoire increases in size. Although the network size cannot, then, remain invariant, it may nevertheless increase very significantly less quickly than network size under the invariant-length approach. What is costly about the universal language approach to building behaviors is that it leads to behavior lengths becoming increasingly long (see Figure 8), and, accordingly, taking longer to compute. *We propose that because behaviors in all selected networks are under time constraints, that sets an upper limit to behavior length; and this, in turn, leads to the invariant-length approach.* For example, when software engineers aim to increase the behavioral repertoire size for their computer software (i.e., increase the number of things their software can do), they are under running-time constraints; the behaviors of interest for their software are only those that may be run in some feasible number of steps. Such a limit on the number of steps is, in essence, a limit on the number of lower-level behaviors (i.e., instructions, the structures in programs) that can be part of a behavior (i.e., part of a run). This is why faster running-time algorithms tend to possess more distinct lines of code (more structure types) than slower algorithms: a faster running-time algorithm has a greater behavioral repertoire size (that is, given feasibility constraints the algorithm can carry out more behaviors), and, just like the other behavioral networks we have seen, the number of lower-level behavior types (instruction types) increases.

3. CONNECTIVITY AND NETWORK DIAMETER FOR BEHAVIORS

Thus far, our cartoon illustrations of behavioral networks (Figures 2 and 7) have not shown edges. But, of course, nodes must be interconnected in order for networks to work (Figure 10). Table 3 lists some example nodes, structures, and edges in several kinds of networks. There are two central reasons for needing edges. The first is that nodes must connect with one another locally in the assembly of structures. Because (as we have seen in Section 1) structures have invariant length, the number of edges per node need not increase in larger networks to maintain intra-structure connectivity, and we accordingly do not expect intra-structure connectivity to be a driving force on overall network connectivity. The second central reason for edges in behavioral networks is that structures must globally communicate with one another in order to act in a concerted fashion in the construction of a behavior. Since the number of distinct structures increases with network size (see Sections 1 and 2), there are potentially an ever-increasing number of structures with which a structure may need to interact. To understand connectivity in behavioral networks, it is therefore crucial that we understand how behavioral networks solve the problem of structure-structure communication.

The question, then, is how does structure-degree (the number of structures with which a structure connects) increase as structural repertoire size increases? Since structures are invariant in length, this question is equivalent to asking how structure-degree increases with network size. Figure 11 illustrates the three qualitatively different manners in which structure-degree may be increased. The first is the *invariant-degree approach*, where structure-degree does not vary as a function of network size. A consequence of this kind of connectivity scaling is that network diameter (which is the average distance between two structures, in terms of the number of edges that must be passed through to get from to the other) enlarges with network size (see Box 3). The "opposite" method of scaling connectivity is the *full-degree approach*, where structure-degree scales up proportionally with network size, so that, independent of network size, each structure connects to a fixed fraction of all the structures. This would lead to the lowest possible network diameter, and approximately equal to one for sufficiently large networks (supposing the network is random or small-world, see Box 3). The third mode of increasing connectivity is to take a middle-of-the-road approach between the previous two extremes, and we call this the *in-between-degree approach*. Here, structure-degree increases as a power law with network size, but disproportionately slowly. Under this approach, network diameter would be higher than in the full-degree approach, but would remain invariant (supposing, again, that the network is random or small-world).

What happens in real behavioral networks? Figure 12 shows data on how the total number of connections scales with the number of structures, for neocortex, electronic circuits, and web sites. Exponents of 1 in the plots would be the signature of the invariant-degree approach, exponents of 2 the signature of the full-degree approach, and exponents in between for the in-between-degree approach. In each of these cases connectivity scales as expected for the in-between-degree approach. The network diameter is, accordingly, expected to remain invariant. In particular, network diameters for these three networks may be very approximately 2 for neocortex [15], 1.7 for electronic circuits, and 4 for web sites.

We distill our empirical findings on how connectivity appears to scale with network size as the following observation.

> **Observation 3:** Invariant Network Diameter > 1: In behavioral selected networks, structure-degree (the number of edges per structure) increases as a power law with network size, but disproportionately slowly; and because behavioral networks are typically small-world (see Box 3), the network diameter is an invariant and greater than 1 (and approximately equal to the inverse of the exponent relating structure-degree to network size).

Scaling in this fashion would not generally be expected in nonselected networks. For example, in ecosystems, the number of individual animals with which any animal trophically interacts is probably independent of network size; intuitively, a lion will eat only so many prey in its lifetime, independent of ecosystem size. This is not to be confused with the fact that larger ecosystems have *species* with more connections in a food web; this is to confuse a network (the ecosystem) with its symbol-type-network (the food web) [see Box 3 and Ref. 1].

Why does connectivity follow the in-between-degree approach in behavioral networks? First, let us ask why behavioral networks do not follow the invariant-degree approach? Connection costs would be minimized under the invariant-degree approach, so why not conform to it? Actually, we are

TABLE 3

Some Example Nodes, Structures, and Edges in Behavioral Selected Networks

	Node	Structure	Edge
Nervous systems	Neuron	Minicolumn?, Module?	Axon
Electronic devices	Electronic component	Basic functional circuit	Wire
Computer software	Operator	Instruction, line of code	Program flow edge
Web sites	Word and pictures	Web page	Link

Web sites are behavioral networks: they are often designed such that sequential combinations of web pages satisfy the needs of the browser, and the web site behavior is the sequence of web pages "activated" when the web site is traversed by a browser. For web sites, we take the words and pictures to be the nodes, for they are the "smallest meaningful units." They locally combine to make up web pages, which are plausible candidate structures (or primitive behaviors). Web sites differ from other networks in that the nodes in other networks can have a degree greater than one. For example, neurons connect to many neurons, and neural structures connect to more structures by increasing the number of neurons to which a neuron connects. For web sites, however, each word on a web page can link to at most one other page. Greater structure-degree is possible by adding more and more "connection-nodes" to a page (i.e., a structure), nodes whose only job is to link. See also the legend of Figure 12.

FIGURE 11

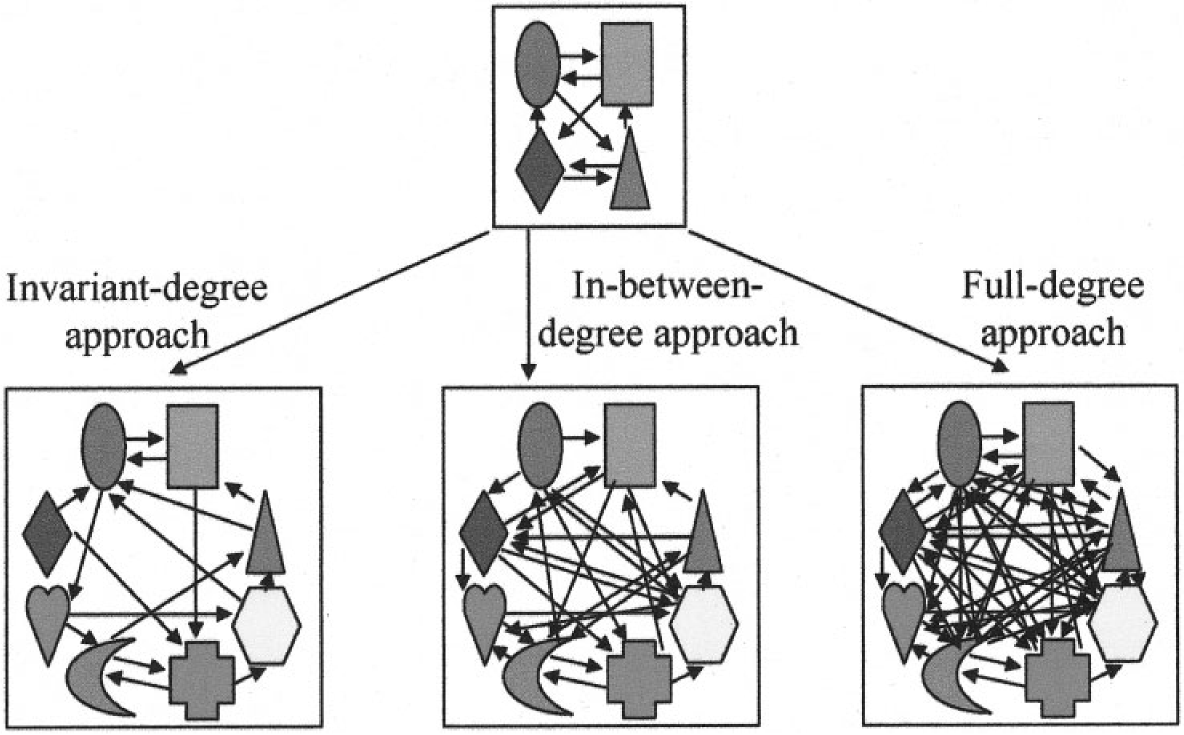


Illustration of the three qualitatively distinct ways in which structure-degree (the number of structures with which a structure connects) may increase with network size. Under the *invariant-degree approach*, structure-degree remains constant, and so the total number of edges scales proportionally with network size. If this possibility were true, network diameter (the average distance between two structures in the network, in terms of the number of edges that must be traversed) would increase proportionally with network size. Behaviors (combinations of structures) would have to become longer and longer (and more redundant) in larger (more behaviorally complex) networks (see Box 3). This approach thus achieves minimal wiring, but suffers the costs of an ever-increasing network diameter and consequently lengthier (longer run-time) behaviors. Under the *full-degree approach*, on the other hand, structure-degree scales proportionally with the number of structures, and total number of edges as the square. (Note that the "full-degree approach" does not require full connectivity; structure-degree must simply scale up as if fully connected.) If this possibility were followed, network diameter would be invariant and at a minimum, and approximately one in sufficiently large networks (see Box 3). This approach thus achieves minimal length (minimal run-time complexity) behaviors, but bears the cost of having the greatest possible number of edges (the number of edges increases as the square of network size). Finally, there is the *in-between-degree approach*, where structure-degree increases, but disproportionately slowly compared to the number of structures in the network, and as a power law (i.e., exponent between 0 and 1). The total number of connections, then, scales faster than the number of structures, but slower than as the square (i.e., an exponent between 1 and 2). If the networks are random or small-world, then the in-between-degree approach leads to network diameters that are invariant and approximately equal to the inverse of the exponent relating structure-degree to network size (see Box 3). This case provides a nice compromise between run-time behavior complexity (i.e., behavior length) and the burdens of network connectivity. For example if structure-degree were to scale as the square root of network size—which means *significantly* slower growth in connectivity than the full-degree approach—then for large networks, the diameter is invariant and low (approximately two). This is, in fact, how connectivity scales in neocortex [15]. Since we have seen that behavior length is invariant (see Section 2), we expect that the invariant-degree approach is *not* followed. Since we also might expect that the full-degree approach is exorbitantly costly in terms of connections, we expect behavioral networks to follow the in-between-degree approach.

already in a position to answer this on the basis of the behavioral generalization we have made earlier in Section 2. We saw then that, across networks of a given kind, behaviors appear to have invariant lengths (i.e., an invariant number of lowest-level behaviors, or structures). If connectivity scaled with invariant-degrees, then network diameter would get increasingly long in larger, more behavioral, networks. But larger network diameter has the consequence of making behavioral length progressively increase relative to the combinatorial degree (the number of degrees of freedom in the construction of a behavior from structures, see Boxes 1 and 3). (This is to say that the *redundancy* of behavior increases; see Box 3). To maintain an invariant combinatorial degree, behavior length would have to increase proportionally with network diameter. (Maintenance of an invariant length, at the expense of a decreasing combinatorial degree, could be maintained only until the combinatorial degree reaches one, at which point behavioral length must again increase.)

Box 3 : Some Details about Symbol-type Networks

Scaling of Connectivity, Network Diameter, and Small-Worlds

The *diameter* of a network is, over all pairs of nodes, the average number of edges that must be crossed to get from one node to the other. The *degree* of a node is the number of edges impinging on that node. If the average degree, δ, in a network does not increase with network size, N, then the network diameter will increase. If, at the other extreme, $\delta = N$, then diameter is one, no matter the network size. What happens in between these two extremes, where $\delta \sim N^v$, with $0 < v < 1$? If the network is random (where there is a fixed probability that two nodes are connected), the diameter $\Lambda \approx (\log N)/(\log\delta)$ [120] [assuming $N \gg \delta \gg \log N \gg 1$]. Since $N = c\delta^{1/v}$ (where c is a proportionality constant), we may write $\Lambda \approx [\log(c\delta^{1/v})]/[\log\delta] = (1/\nu) + (\log c)/(\log\delta)$. Since c is constant and δ is increasing in larger networks, for sufficiently large networks $\Lambda \rightarrow 1/v$. This is important, for network diameter can remain invariant despite the average degree scaling up disproportionately slowly with network size. For example, if $\delta \sim N^{1/2}$, then $\Lambda \rightarrow 2$ for sufficiently large networks. But networks of interest are rarely random, but very many kinds of network have been argued to be *small-world* (see [123]), where there are enough long-range shortcuts that the network diameter is approximately as low as that in a random network, despite retaining a highly organized topology. In the networks we study here, if connectivity scales up, we will presume the network is small-world, for nearly every complex network studied has been discovered to be small-world: for example, airport traffic [121], power grid [121], movie-actor collaborations [121], world wide web [122], acquaintances [121], the nervous system of *C. elegans* [123], citation networks [124–126], electronic circuits [127]. (In our discussion in the text in Section 3, we will primarily be interested in the degree of structures, not nodes, and the diameter between structures. The same points above apply, where structures are the nodes.)

Symbol-type Networks

A *symbol-type network* is any network where the vertices represent *types* of objects, or symbol types (rather than a traditional network, where there may be multiple vertices of the same type). One example kind of symbol-type network is an *ethogram*, which consists of a vertex for each of an animal's distinct behaviors—all at the same hierarchical level—and an arrow points from one behavior to another just in case the first tends to be followed by the next. (For example, if one treats the structures in networks as the vertices, then if each structure tends to be unique—at least in regard to scaling, such networks are symbol-type-networks.) Another kind of symbol-type-network is a food web, where each vertex represents a species, and an arrow points from one species to another just in case there are animals of the first species that sometimes eat animals of the second. Many kinds of symbol-type networks have also been found to be small-world networks.

Redundancy

Suppose that a symbol-type-network has S vertices, or symbol types, that δ is the average degree, and that symbol combinations are of length L. (Also suppose the symbol-type network is a small-world network.) How many possible combinations are there? Any of the S symbol types may interact with δ possible symbol types, which in turn may interact with δ others, and so on until L symbols have been put into the combination. Thus, there are $E \sim S\delta^{L-1}$ many possible combinations. If $\delta \sim S^v$, with v a constant in the interval [0, 1], then we may conclude that $E \sim S(S^v)^{L-1} = S^{vL-v+1}$. Recall from Box 1 that $E \sim S^d$, where d is the combinatorial degree. We may therefore set the exponents of the two previous relationships equal to one another, and we have $d = vL - v + 1$, or $v = (d-1)/(L-1)$. The exponent v (if $v > 0$) from the relationship $\delta \sim S^v$ is roughly the ratio of the combinatorial degree divided by the total potential combinatorial degree (or entropy divided by the total possible entropy). $1/v$ is, accordingly, a measure of the amount of *redundancy* in the combinatorial system. Recall (from earlier in this Box) that $1/v$ is also, for sufficiently large (small-world) networks, approximately the network diameter, and thus network diameter and redundancy are roughly equivalent. The reason, informally, is this: a diameter of Λ means that if one symbol type in the symbol-type network needs to interact with some other particular symbol type (for the purposes of some particular combination), then communication must proceed via $\sim\Lambda$ other symbol types. These scaling techniques were employed to argue that nodes are used combinatorially in metabolic networks (see Figure 8b of Ref. 1 from which one may compute that $v = 0.28 > 0$, a signature that a network is using its nodes combinatorially), and in electronic circuits (see Figure 2c of Ref. 1 from which one may compute that $v \approx 1 > 0$). They are also used in Section 3 of the current article, concerning the manner in which structures (or the primitive behaviors) interconnect for the purpose of combining to build behaviors.

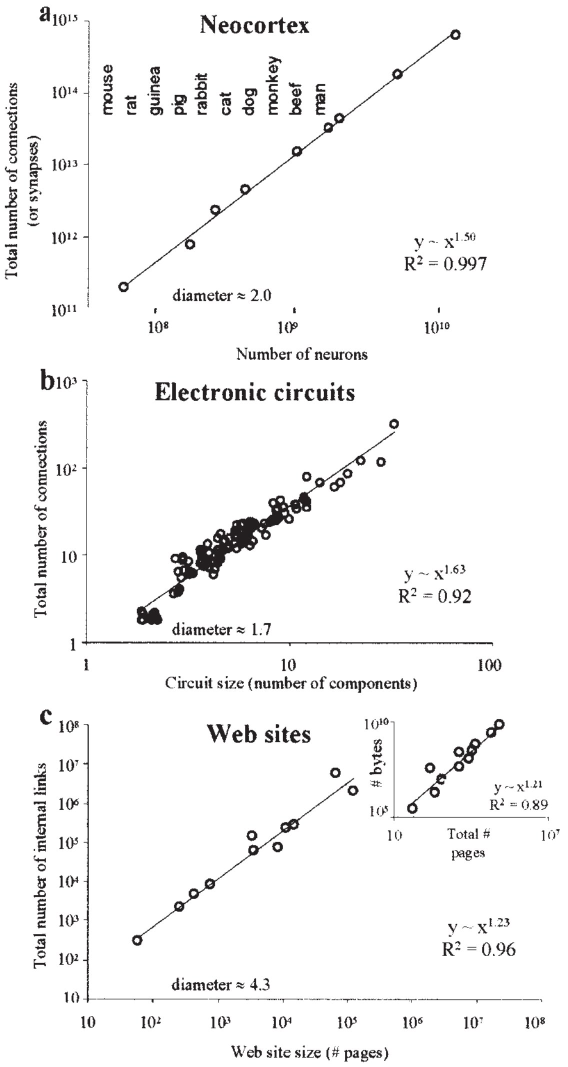
In short, structure-degree must increase at least as a power law with the total number of structures in order to maintain invariant length behaviors, and *this*, in turn, was for running-time minimization reasons. We accordingly expect computer software to have increasing structure-degree (i.e., an increasing typical number of program-flow edges per structure), but we do not currently possess structure-connectivity data for software; that is the subject of future work.

Why structure-degree does not scale up as in the full-degree approach is presumably because of the exorbitant connection costs (see earlier citations on volume optimization principles in networks in Section 1), for the total number of connections would then scale as the square of network size.

The above may help explain why behavioral networks follow the in-between-degree approach, but it leaves us in the dark as to why connectivity scales precisely as it does. Explaining the particular exponent relating connectivity and network size may depend on network-specific issues. For example, for neocortex, Changizi [15, 16, 72] put forth a theory of economical well-connectedness explaining why neuron-degree scales as the square root of network size (see the next section, and Figure 17), but the second of the two assumptions of the theory ("invariant area-infiltration") may not be expected to be satisfied in other kinds of behavioral networks. One conjecture for why web sites scale connectivity the way they do is that it results in a network diameter on the order of four (4.3), which may be due to an upper limit on how many links consumers are willing or able to wade through to get from any one part of a web site to another (perhaps for working memory reasons).

There are a number of consequences of an increasing structure-degree in networks, and they have been worked out in detail mostly for neocortex [15, 16, 72]. One consequence is that more and more space must be devoted to connection considerations. In neocortex, structure-degree increase is achieved by increasing the number of synapses per neuron [Figure 12(a)], and neuron density accordingly falls in larger brains [8, 11, 101–103]; namely neuron density scales as the −1/3 power of gray matter volume, or, equivalently, number of neurons scales as the 2/3 power of gray matter volume. From this we may conclude that gray matter volume (or mass) scales as the number of neurons to the power of 1.5, and this has an analogy in web sites, where the

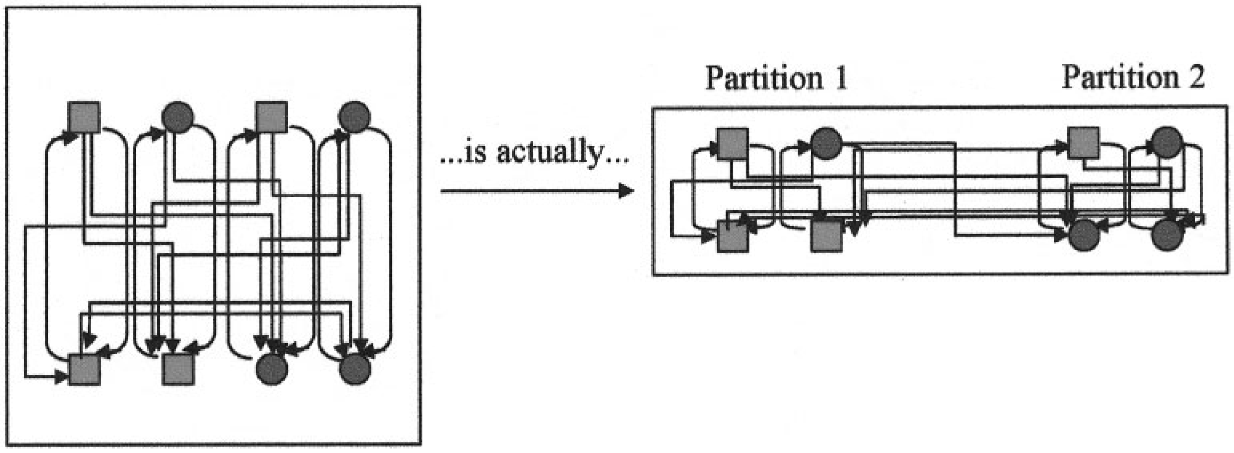
FIGURE 12



The manner in which connectivity scales with the number of structures, in three kinds of behavioral network: neocortex, electronic circuits, and web sites. Each follows the in-between-degree approach. (a) Log-log (base 10) plot of the total number of neocortical synapses versus neocortical network size (the latter which scales proportionally with the number of structures supposing neocortex has invariant-length structures, see Section 1). Data are computed using data from Ref. 101, who present data for how neuron density decreases in larger brains (namely, it decreases as the −1/3 power of brain volume). Since synapse density remains invariant [15, 140], a neuron density decrease corresponds to an increase in the number of synapses per neuron (neuron-degree). Since, as we see here, the total number of connections scales as the 3/2 power of network size, this corresponds to number of synapses per neuron scaling as the (3/2 − 1 =) 1/2 power of network size. If the neocortex is small-world, the network diameter will be approximately the inverse of this exponent, or 2 (see Ref. 15 and Box 3). (b) Log-log plot of the total number of connections versus network size, for electronic circuits found in Ref. 141. Because of overlapping data points, for visualization sake, each data point has been perturbed by adding a random jitter in the interval

FIGURE 12 Continued. [−0.05, 0.05] to the *x* and *y* values. The scaling exponent is ≈1.6, and so degree scales as the 0.6 power of network size. Network diameter here is approximately (1/0.6 =) 1.7. (c) Log-log plot of the total number of intra-web-site connections versus the number of web pages at the site, for 11 commercial web sites (plumtree, e4me, thesinc, corel, google, dell, att, palm, panasonic, sparkletts and sony), which we determined using a freeware web-crawler *Xenu*. The scaling exponent is about 1.23, and so degree scales as the 0.23 power of network size, corresponding to a network diameter of about 4 for web sites, independent of size. Since greater structure-degree is accommodated in web sites by increasing the number of "connection nodes" (rather than increasing node-degree as in most other networks), we expect that the average size of a web page (or structure) should increase in "mass" (bytes) due to these added "connection nodes." (The number of nonconnection nodes should remain invariant, since structural repertoire size increase in behavioral networks follows the invariant-length approach.) Thus, the overall mass in bytes of a web site should increase more quickly than the number of web pages (or structures), and, in particular, it should scale just fast enough to accommodate the added "connection nodes". The inset shows how the overall web site mass (no. of bytes) scales with number of web pages, and one can see that the scaling exponent is 1.20, near the earlier exponent, 1.23.

FIGURE 13



Behavioral networks often have distinct regions, or partitions, where intra-partition connections are made locally (and are thus less costly), and inter-partition connections are long-range (and costly). Left: A network, with its edges and differentiated nodes are shown. Also, the nodes are arranged here in columns, each column corresponding to a structure, as discussed in Figure 10. Right: The same network, but parcelled into two regions. Notice that the edges are shorter within the partitions, but longer between the partitions.

total "mass," in bytes, of a web site scales disproportionately quickly compared to the number of web pages [see inset of Figure 12(c)], namely with an exponent of 1.21, due to the increasing number of connections. It is because of this neocortical density decrease that neocortex surface area increases as the 8/9 power of gray matter volume [8–14], rather than the 2/3 power, which is why the neocortex becomes increasingly convoluted in larger brains. It is also what explains the slow increase in neocortex thickness and the linear dimensions of neocortical modules [see Figure 5(a)]. The increasing number of synapses per neuron also has the effect of increasing the physical diameters of axons and somas as the 1/9 power of gray matter volume [15, 17, 104], which can be explained [16] by neural vessel diameters conforming to Murray's Law [105–107], where the cube of a parent segment diameter is equal to the sum of the cubes of the daughter segment diameters [66, 108]. Finally, this physical increase in axon caliber, in turn, is critical in understanding why neocortical white matter scales disproportionately quickly as a function of gray matter volume, and namely, with exponent approximately 4/3 [13, 14, 109, 110]: if axon caliber somehow remained invariant, white matter volume would scale proportionally with gray matter volume [15, 16, 111]. We see, then, that connectivity increase drives many of the broad macro- and micro-features of neocortical anatomy, and that this connectivity increase is due, ultimately, to a pressure for invariant-length behaviors.

4. COMPARTMENTALIZATION

One feature found in behavioral networks, but not found in our latest cartoon illustration of networks (Figure 10), is that networks are often *partitioned* into multiple areas (or regions, or partitions) where intra-partition connections are made locally, and inter-partition connections made via long-range connections (Figure 13). Table 4 lists example cases of partitions (or parcellations) and partition-partition connections in several kinds of network. For example, partitions are commonly used in very-large-scale integrated (VLSI) design to aid in the reduction of wire (e.g., [84]): a circuit is partitioned into multiple subsets, each subset is placed in geographically separated regions, and any connections between the subsets become, in effect, long-range inter-partition connections. The neocortex also is partitioned into many areas, where neurons in the same area connect via local connections, and neurons in different areas connect via long-range, white matter axons. And people have proposed that perhaps neocortical areas, like in VLSI design, are driven by save-wire pressures [15, 16, 55–57, 59–62, 67, 72, 111–118]. Businesses also often appear to have multiple departments, or divisions.

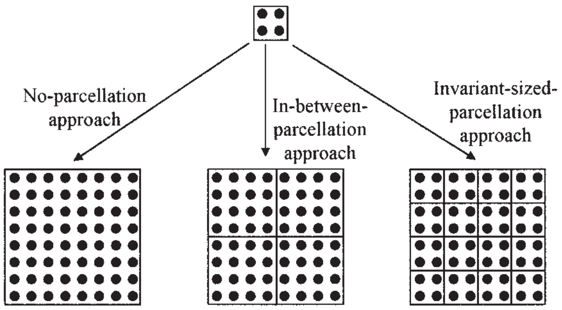
Our question here is, How does parcellation vary as a function of size in behavioral selected networks? Figure 14

TABLE 4

Some Example Kinds of Parcellation in Behavioral Networks

	Parcellation	Parcellation Connection
Neocortex	Area	White matter axons
Electronic devices	Partition	Long-range wires
Computer software	Module, or procedure	Procedure calls
University	Department	Contacts

FIGURE 14



The possible manners in which parcellation may scale with network size. The small network at the top has a size of four, and the networks below have been scaled up to a size of 64, or 16 times larger. Under the *no-parcellation approach*, there are no parcellations at any scale (left); or, more weakly, there are an invariant number of partitions. At the opposite extreme (right), the *invariant-sized-parcellation approach* leads to the number of partitions scaling proportionally with network size. In terms of scaling, partitions under this approach act as nodes themselves, and there is, in effect, just one large area; therefore, as far as scaling issues are concerned, these two seemingly opposite approaches are equivalent. Conceptually distinct from these, however, is the *in-between-parcellation approach* (middle), where the number of partitions scales up, but disproportionately slowly compared to network size.

shows three central possibilities. The first is the *no-parcellation approach,* where the number of partitions is invariant; or, intuitively, there are, as far as scaling issues are concerned, no partitions at all. The opposite extreme is the *invariant-sized-parcellation approach,* where the average size of a partition remains invariant. Although this scaling possibility may at first seem to be the opposite of the no-parcellation approach, they are actually identical in terms of scaling, since the partitions may now be treated as nodes. And, lastly, there is the *in-between-parcellation approach,* where the number of partitions increases with network size, but disproportionately slowly (and as a power law).

How does parcellation, in fact, scale in behavioral networks? Figure 15 shows parcellation data from three kinds of network: neocortex (areas), computer software (modules), and universities (departments). In each case, parcellation increases, but disproportionately slowly as a function of network size. The in-between-parcellation approach therefore applies in these cases. In particular, a power law applies in each case, and the exponent is near 1/2.

FIGURE 15

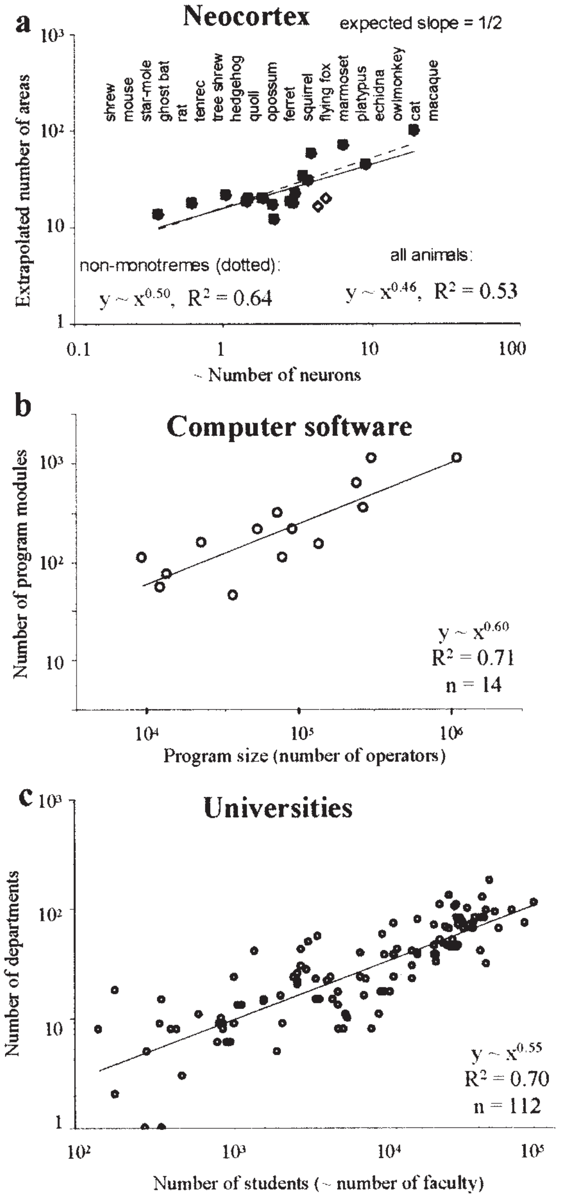
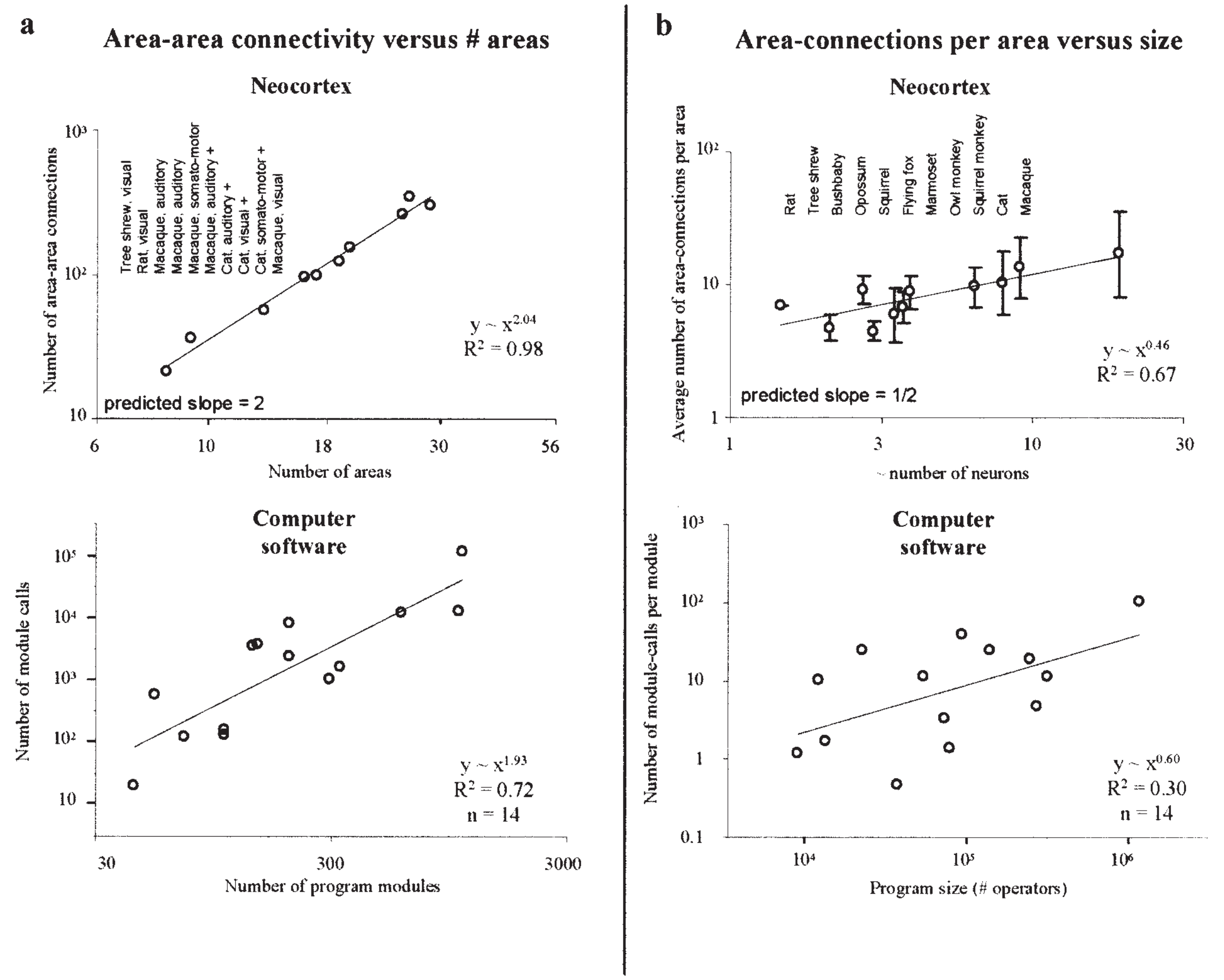


FIGURE 15 How parcellation varies as a function of network size in three kinds of network: neocortex, computer software, and universities. In each case parcellation scales up disproportionately slowly with network size, and namely nearly as a square-root law. (a) Log-log (base 10) plot of number of areas versus number of neurons in neocortex, for 19 mammals [72]. Parcellation scales as predicted by economical well-connectedness (see Figure 17). The area counts are computed via measuring the average sizes of areas, as a percentage of neocortex, and extrapolating the expected total number of areas. The *x*-axis here scales proportionally with number of neurons (and was computed as the 2/3 power of brain volume). Measurements are taken from sensory-motor areas from the Kaas-Krubitzer literature [142–151]. (b) Log-log plot of number of program modules versus program size for microprogramming data from Ref. 48. (c) Log-log plot of number of departments versus total number of faculty (measured as the number of students). Data are from Refs. 85 and 1. Number of departments was used in those articles as a measure of network differentiation, but we now believe departments are more analogous to partitions. A square-root law also appears to govern the growth in the number of departments as a function of university size for a single university (Duke) through time (not shown here, see Ref. 1).

FIGURE 16

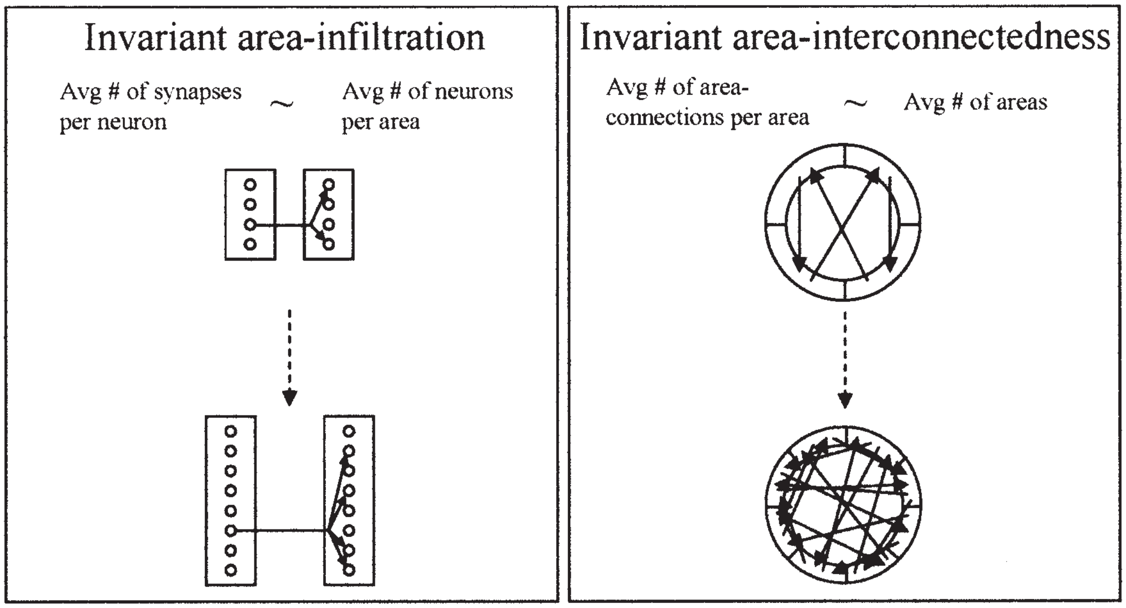


Data showing that area-area connectivity scales as if fully connected for neocortex and computer software. (a) Log-log (base 10) plots of the total number of area-area connections versus the total number of areas, for neocortex and computer software. Data for neocortex are from 10 sensory (and somato-motor) subnetworks [148, 152–156]. See Ref. 72. Economical well-connectedness (Figure 17) predicts a slope of 2, and one can see the slope is 2 here. Data for computer software are for programs written in microprogramming languages, and are from [48], where we use the number of procedure calls as a proportional proxy for module-module interconnectivity. (b) Log-log plots of the average area-degree versus the total network size (i.e., number of nodes, not number of areas). Data for neocortex are from sensory (and somato-motor) areas, and are compiled in Ref. 72. We expect (see Figure 17), and find, a slope of 1/2. Since the number of areas also scales as the 1/2 power of network size, area-degrees scale up as fast as possible. A similar phenomenon is found among the same set of computer programs as mentioned in (a).

A related question about parcellation is how partition-degree (the number of partitions with which a partition connects) scales with network size. We saw in the previous section that structure-degree increases with network size, but disproportionately slowly since full structure-structure connectivity would be prohibitive. Since the number of partitions increases disproportionately slowly as a function of network size, full partition-partition connectivity may be affordable. Analogous to the three possible manners for scaling structure connectivity (see Figure 11 again), there are three possible manners for scaling partition connectivity.

Figure 16 shows data on partition-partition connectivity for two kinds of network: neocortex and computer software. Figure 16(a) shows how the number of partition-partition connections scales as a function of the number of partitions, and for each kind of network the exponent is approximately 2. Recall that the analogous plot for structures (Figure 12) showed that the number of structure-structure connections scales as the number of structures to the power of much lower exponents (e.g., 1.5), which meant that structure-structure connectivity increased disproportionately slowly with network size. The scaling exponent of 2 for

FIGURE 17



Illustrations of the two invariance assumptions of *well-connectedness. Invariant area-infiltration*: In each pair of areas, a small pair (left) and large pair (right), one area infiltrates half of the neurons in the other area. *Invariant area-interconnectedness*: Two brains are shown, one with four cortical areas (left) and another with eight. They each have the same area-interconnectedness since each area in each brain is connected with 1/4 of the total number of areas.

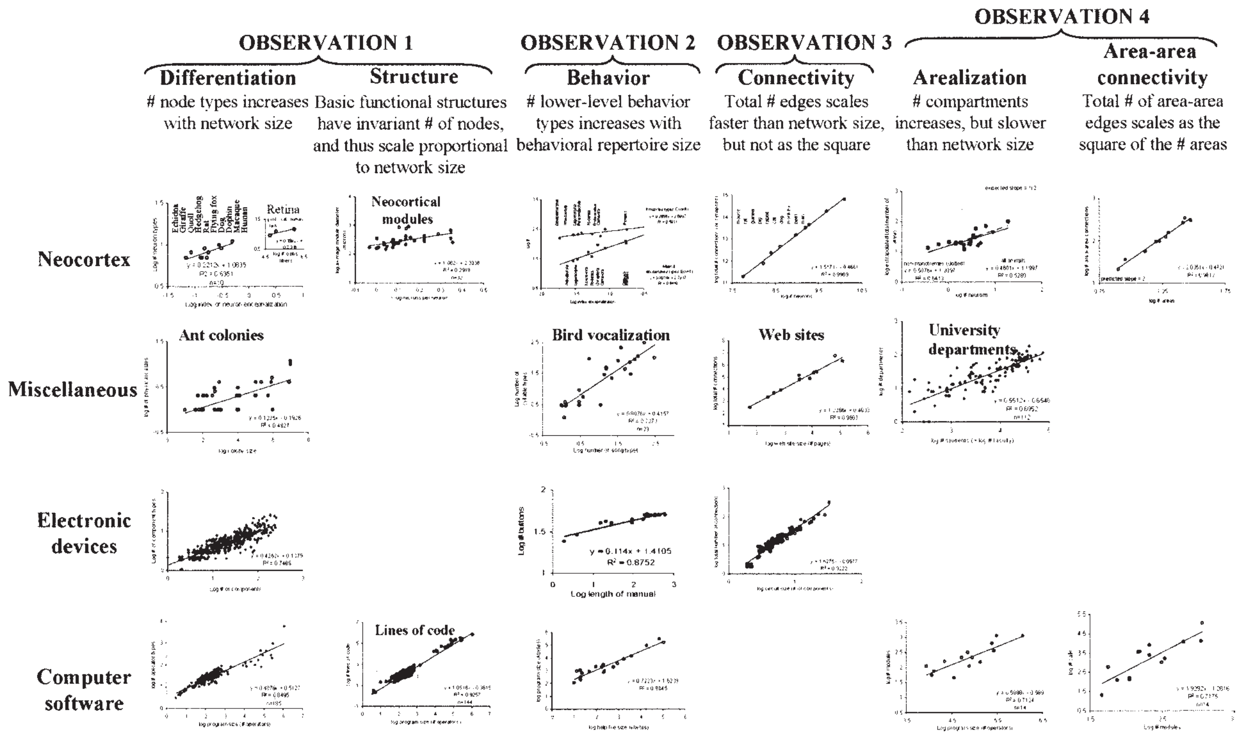
partition connectivity, however, means that partition connectivity scales up as quickly as possible: up to a constant factor, partition networks are fully connected, independent of network size. Figure 16(b) demonstrates this conclusion in a different fashion, showing data on how partition-degree increases with network size. In neocortex and the computer software data set, partition-degree increases as the square root of network size. Since parcellation also increases as the square root of network size (see Figure 15), partition-degree scales proportionally with the number of partitions, which is just to again conclude that, up to a constant factor, partition-networks maintain full-connectivity no matter the network size.

One way to describe these parcellation results is as follows: Behavioral networks appear to increase the number of partitions as fast as possible, subject to the constraint that areas are, up to a constant factor, fully connected. In order to conform to this—i.e., in order for a partition to connect to all the other partitions—each partition must have enough nodes within it to make the connections. Thus, the size of a partition cannot scale up less quickly than the number of partitions, lest full partition-connectivity be impossible to maintain. The fastest that the number of partitions may be increased, then, is proportionally with partition size, and since network size is the product of the number of partitions and partition size, parcellation scales as the square root of network size. Partitions conforming to this have been called *square-root compartments* [116, 119]. In this light, we record our empirical parcellation results as the following observation.

> **Observation 4:** Slow Parcellation Increase, and Full-Connectivity Parcellation: Larger behavioral selected networks are parcelled into more and more areas, with number of areas scaling as fast as possible subject to the constraint that areas are, up to a constant proportion, fully connected to one another; this leads to parcellation scaling as the square root of network size.

Why, though, might behavioral networks conform to this? That is, why does parcellation appear to scale up as quickly as possible subject to the constraint of full partition-partition connectivity (up to a constant factor)? We currently have no general answer to this question. We do, however, possess a theory specific to the neocortex where we can answer the question [15, 16, 72, 111]. The constraint that partition-networks scale up connectivity as quickly as possible we label *invariant area-interconnectedness* (Figure 17). Square-root compartments assume this. The second constraint is that when a partition connects to another partition, the neuron making the link synapses to an invariant fraction of the neurons in the partition. We call this *invariant area-infiltration* (Figure 17). If these two invariances are satisfied *economically*, i.e., using the least amount of wire, then one may derive square-root compartments (see [15, 16, 72, 111]). Areas scale up as quickly as possible (subject to the constraint of invari-

FIGURE 18



Summary of the correlates of behavioral networks, emphasizing the unifying power of the framework.

ant area-interconnectedness) because, that way, the size of areas scales up as slowly as possible, and this, in turn, allows the number of synapses per neuron, and the wiring volume supporting it, to scale up as slowly as possible [15, 16, 72, 111]. Other networks (computer software and businesses) may also follow a square root law [Figure 15(b and c)], but the explanation we just gave does not apply more generally to these other networks, since they are unlikely to satisfy invariant area-infiltration. Also, invariant area-infiltration is a necessary assumption in deriving that neuron-degree scales as the square root of network size [15, 16, 72, 111], something we found to be the case in Figure 12(a). Since other behavioral networks cannot be expected to satisfy invariant area-infiltration, structure-degree in other networks is not necessarily expected to scale as a square root law.

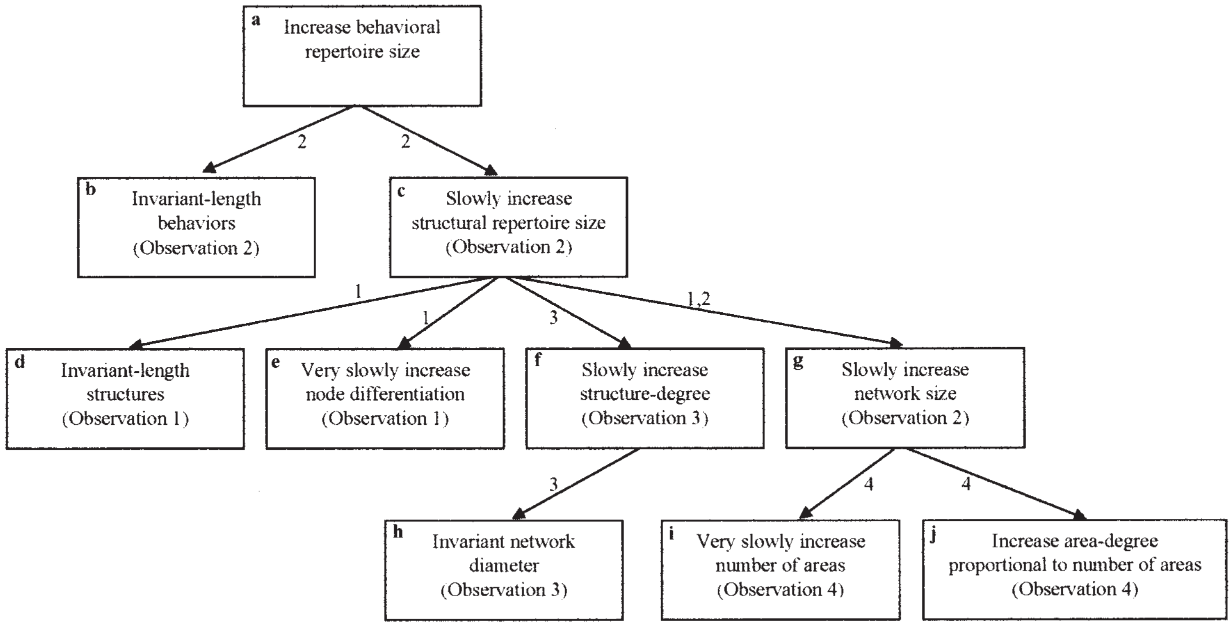
Parcellation, then, increases in larger behavioral selected networks and does so disproportionately slowly. Parcellation is generally thought to be driven, in part, by wire-optimization reasons, and we put forth one argument above, relevant for the neocortex specifically, that predicts parcellation to scale up as it does. We do not yet have a more general argument for why parcellation may increase in the manner it does for other networks, but it seems reasonable to suspect that wire-optimization reasons will be fundamental to any such explanation. If, indeed, wire-optimization generally underlies the scaling of parcellation, then we may attribute the ultimate cause for parcellation as follows: (a) parcellation increases for wire-optimization reasons, (b) pressure for wire-optimization is a central issue only because structure-degree increases in larger networks, (c) this, in turn, occurs in order to maintain an invariant network diameter, and (d) network diameter is kept invariant in order to keep behavioral running-times low. In this way we see how parcellation is something to expect in behavioral selected networks, but *not* necessarily in nonbehavioral or nonselected networks.

5. CONCLUSION

Our purpose here was to examine commonalities between behavioral selected networks of all kinds, and for neocortex and computer software more specifically. We have made four "observations" throughout the article, which we rewrite here in one place.

> **Observation 1:** Invariant-Length Structures and Power-Law Differentiation: Behavioral selected networks increase their structural repertoire size by increasing the number of node types, not by increasing the length of structures.

FIGURE 19



Correlates of behavioral networks, as discussed in this article, emphasizing the connections between the correlates. The arrows show the chains of reasoning, and the numbers on the arrows indicate the section of the article where the implication is discussed. In each case, the most relevant observation is noted. It is to be read like this: For example, starting at the top, it states that when behavioral repertoire size is increased, it results in two consequences: behavior lengths that remain invariant, and a slow increase in the number of lower-level, structure, types. This increase in the number of structure types, in turn, has the four consequences shown below it, and so on. See the discussion in Section 5 for more comment.

Observation 2: Invariant-Length Behaviors: Behavioral selected networks increase their behavioral repertoire size largely by increasing the number of lower-level behaviors, not primarily by increasing the length of behaviors. This appears to hold for even the lowest-level behaviors, or structures.

Observation 3: Invariant Network Diameter > 1: In behavioral selected networks, structure-degree (the number of edges per structure) increases as a power law with network size, but disproportionately slowly; and because behavioral networks are small-world, the network diameter is an invariant and greater than 1.

Observation 4: Slow Parcellation Increase, and Full-Connectivity Parcellation: Larger behavioral selected networks are parcelled into more and more areas, with number of areas scaling as fast as possible subject to the constraint that areas are, up to a constant proportion, fully connected to one another; this leads to parcellation scaling approximately as the square root of network size.

Concerning Observation 1, in our data the nodes are themselves complex, emergent, objects built from even lower-level objects, and, as discussed at the end of Section 1, it is an open question as to whether the observation applies to the lowest levels, such as DNA or fundamental microelectronic circuits. Figure 18 summarizes the main "correlates" of behavioral networks we have discussed, covering the four observations, and the wide variety of networks treated here. This figure emphasizes the unifying power of the framework.

Figure 19 summarizes many of the implications we have made throughout the article, showing how they all link together, and saying where in this article the implication was discussed. One can see from Figure 19 that all the correlates of behavioral networks we discussed appear to be consequences of a selected, behavioral network accommodating greater behavioral repertoire size. In particular, when a selected behavioral network increases its number of behaviors [Figure 19(a)], the lengths of the behaviors do not appear to increase [Figure 19(b)]. This may be due to selective pressure to keep the run-times of behaviors below some limit. Instead, the number of lower-level behavior types is increased [Figure 19(c)]. This also appears to hold for the lowest-level behavior types, the structure types [19(c)]. The number of structure types increases "slowly" compared to the behavioral repertoire size increase, because structures are used combinatorially to build behaviors. The fact that the number of structure types increases has at least four

TABLE 5

Summary of the Analogy between Computer Software and Neocortical Organization

Network	Node	Edge	Structure that Nodes Combine to Make	Behavior	Parcellation	Parcellation Connections
Computer software	Operator	Program flow edge	Instruction, line of code	Runs	Procedure	Procedure call
Neocortex	Neuron	Axon/dendrite	Minicolumn, module?	Behaviors/thoughts	Area	White matter axons

Note that networks of all kind, selected or not, may be treated as implementing computable functions [16, 128]; what is novel here is that the organization of actual selected software and actual behavioral selected networks appears to share so many commonalities. Analogies between programs and organisms have been made before. For example, artificial life [129] is an example field that, in a sense, makes a similar philosophical point; they model organisms in an ecosystem as programs competing against one another for resources. Their interest, however, has primarily concerned the self-organizing properties of competitive networks of these programs, rather than on the internal, syntactic program properties that are selected for. What research has been done on programs as models for biological structure has concentrated on genetic complexity [130, 131].

consequences, [19(d)] through [19(g)]. The lengths of structures do not appear to increase [19(d)], and, instead, the number of node types, or network differentiation, is increased very slowly [19(e)]. This may be due to selective pressure to minimize overall network size. This increase is "very slow" compared to the behavioral repertoire increase because nodes combine to build structures, which combine to build behaviors. The increasing structure type repertoire also leads to a slowly increasing structure-degree (the number of structures to which a structure connects) [19(f)]. This is in order to maintain an invariant network diameter [19(h)], for the purpose of invariant length behaviors. Greater structure repertoire size also means greater network size [19(g)], and *this* has the consequences, for wire-optimization reasons, that the number of areas slowly increases [19(i)] and the area-degree (the number of areas to which an area connects) increases in proportion to the number of areas [19(j)].

We have seen above that computer software is similar to the other behavioral selected networks, and to the neocortex in particular, in regards to our principal observations. But not only is computer software like other behavioral selected networks, but, more importantly, vice versa. *Our framework and data here suggest that behavioral selected networks of all kinds are like computer software in the respects summarized in Figures 18 and 19.* (See Table 5 for the full analogy between neocortex and computer software.) We believe this provides a powerful framework for thinking about complex behavioral, selected networks, because computer software is, of course, better understood than biological networks.

ACKNOWLEDGMENTS

I thank Rodrigo Quian Quiroga, Daniel McShea, Alex Bäcker, and two anonymous referees, for their comments. This work was supported in part by the Sloan-Swartz Foundation.

REFERENCES

1. Changizi, M.A.; McDannald M.A.; Widders, D. Scaling of differentiation in networks: Nervous systems, organisms, ant colonies, ecosystems, businesses, universities, cities, electronic circuits, and legos. J Theor Biol 2002, 218, 215–237.
2. Mountcastle, V.B. Modality and topographic properties of single neurons of cat's somatic sensory cortex. J Neurophysiol. 1957, 20, 408–434.
3. Mountcastle, V.B. The columnar organization of the neocortex. Brain 1997, 120, 701–722.
4. Tommerdahl, M.; Favorov, O.; Whitsel, B.L.; Nakhle, B. Minicolumnar activation patterns in cat and monkey S1 cortex. Cerebral Cortex 1993, 3, 399–411.
5. Peters, A. The organization of the primary visual cortex in the macaque. In: Cerebral Cortex, Vol. 10, Primary Visual Cortex of Primates; Peters, A.; Rockland, K.S., Eds.; Plenum; New York, 1994; pp 1–36.
6. Jones, E.G. Microcolumns in the cerebral cortex. Proc Natl Acad Sci USA 2000, 97, 5019–5021.
7. Rockel A.J.; Hiorns R.W.; Powell T.P.S. The basic uniformity in structure of the neocortex. Brain 1980, 103, 221–244.
8. Jerison, H.J. Allometry, brain size, cortical surface, and convolutedness. In: Primate Brain Evolution; Armstrong, E.; Falk, D., Eds.; Plenum Press: New York, 1982; pp 77–84.
9. Prothero, J.W.; Sundsten, J.W. Folding of the cerebral cortex in mammals. Brain Behav Evol 1985, 24, 152–167.
10. Prothero, J. Cortical scaling in mammals: A repeating units model. J Brain Res 1997, 38, 195–207.
11. Prothero, J. Scaling of cortical neuron density and white matter volume in mammals. J Brain Res 1997, 38, 513–524.
12. Hofman, M.A. Size and shape of the cerebral cortex in mammals. I. The cortical surface. Brain Behav Evol 1985, 27, 28–40.
13. Hofman, M.A. On the evolution and geoetry of the brain in mammals. Prog Neurobiol 1989, 32, 137–158.
14. Hofman, M.A. The fractal geometry of convoluted brains. J Hirnforschung 1991, 32, 103–111.
15. Changizi, M.A. Principles underlying mammalian neocortical scaling. Biol Cybern 2001, 84, 207–215.

16. Changizi, M.A. The Brain from 25,000 Feet: High Level Explorations of Brain Complexity, Perception, Induction and Vagueness; Kluwer Academic: Dordrecht, 2003.
17. Harrison, K.H.; Hof, P.R.; Wang, S.S.-H. Scaling laws in the mammalian neocortex. J Neurocytol 2002, 31, 289–298.
18. Bonner, J. T. The Evolution of Complexity; Princeton University Press: Princeton, 1988.
19. Bonner, J. T. Dividing the labour in cells and societies. Curr Sci 1993, 64, 459–466.
20. Bell, G.; Mooers, A.O. Size and complexity among multicellular organisms. Biol J Linnean Soc 1997, 60, 345–363.
21. Valentine, J.W.; Collins, A.G.; Meyer, C.P. Morphological complexity increase in metazoans. Paleobiology 1994, 20, 131–142.
22. Valentine, J.W.; May, C.L. Hierarchies in biology and paleontology. Paleobiology 1996, 22, 23–33.
23. McShea, D.W. Metazoan complexity and evolution: Is there a trend? Evolution 1996, 50, 477–492.
24. McShea, D.W. Functional complexity in organisms: parts as proxies. Biol Philos 2000, 15, 641–668.
25. McShea, D.W. Parts and integration: The consequences of hierarchy. In: Evolutionary Patterns: Growth, Form, and Tempo in the Fossil Record; Jackson, J.B.C.; Lidgard, S.; McKinney, F.K., Eds.; University of Chicago Press: Chicago, 2001.
26. McShea, D.W. The hierarchical structure of organisms: A scale and documentation of a trend in the maximum. Paleobiology 2001, 27, 405–423.
27. McShea, D.W. A complexity drain on cells in the evolution of multicellularity. Evolution 2001, 56, 441–452.
28. McShea, D.W.; Anderson, C. The remodularization of the organism. In: Modularity: Understanding the Development and Evolution of Complex Natural Systems; Callebaut, W., Raskin-Gutman, D., Eds.; MIT Press: Cambridge, MA. In press.
29. McShea, D.W.; Changizi, M.A. Three puzzles in hierarchical evolution. Integrative Comp Biol 2003, 43, 74–81.
30. Karsai, I., Wenzel, J.W. Productivity, individual-level and colony-level flexibility, and organization of work as consequences of colony size. Proc Natl Acad Sci USA 1998, 95, 8665–8669.
31. Bourke, A.F.G. Colony size, social complexity and reproductive conflict in social insects. J Evol Biol 1999, 12, 245–257.
32. Anderson, C.; McShea, D.W. Individual versus social complexity, with particular reference to ant colonies. Biol Rev 2001, 76, 211–237.
33. Simmel, G. The number of members as determining the sociological form of the group. Am J Sociol 1902, 8, 1–46, .
34. Caplow, T. Organizational size. Admin Sci Quart 1957, 1, 484–505.
35. Hall, R.H.; Haas, J.E.; Johnson, N.J. Organizational size, complexity, and formalization. Am Sociol Rev 1967, 32, 903–912.
36. Pugh, D.S.; Hickson, D.J.; Hinings, C.R.; Turner, C. Dimensions of organization structure. Admin Sci Quart 1968, 13, 65–105.
37. Blau, P.M. A formal theory of differentiation in organizations. Am Sociol Rev 1970, 35, 201–218.
38. Blau, P.M.; Schoenherr, P.M. The Structure of Organizations; Basic Books: New York, 1971.
39. Childers, G.W.; Mayhew, B.H., Jr.; Gray, L.N. System size and structural differentiation in military organizations: Testing a baseline model of division of labor. Am J Sociol 1971, 76, 813–830.
40. Kimberly, J.R. Organizational size and the structuralist perspective: A review, critique, and proposal. Admin Sci Quart 1976, 21, 570–597.
41. Slater, R.O. Organization size and differentiation. Res Sociol Organiz 1985, 4, 127–180.
42. Halstead, M.H. Elements of Software Science, Operating, and Programming Systems Series, Volume 7; Elsevier: New York, 1977.
43. Lassez, J.-L.; van der Knijff, D., Shepherd, J.; Lassez, C. A critical examination of software science. J Systems Software 1981, 2, 105–112.
44. Woodfield, S.N.; Shen, V.Y.; Dunsmore, H.E. A study of several metrics for programming effort. J Systems Software 1981, 2, 97–103.
45. Evangelist, W.M. Software complexity metric sensitivity to program structuring rules. J Systems Software 1983, 3, 231–243.
46. Card, D.N.; Agresti, W.W. Resolving the software science anomaly. J Systems Software 1987, 7, 29–35.
47. Kitchenham, B.; Pickard, L. Towards a constructive quality model. Part II: Statistical techniques for modelling software quality in the ESPRIT REQUEST project. Software Engineering J 1987, 2(4), 114–126.
48. Triantafyllos, G.; Vassiladis, S.; Kobrosly, W. Software metrics for the microcode of computer systems. J Systems Software 1994, 26, 221–232.
49. Simon, H.A. The architecture of complexity. Proc Am Philos Soc 1962, 106, 467–482.
50. Salthe, S.N. Evolving Hierarchical Systems; Columbia University Press: New York, 1985.
51. Pettersson, M. Complexity and Evolution; Cambridge University Press: Cambridge, 1996.
52. Changizi, M.A. Universal laws for hierarchical systems. Comments Theor Biol 2001, 6, 25–75.
53. Carneiro, R.L. On the relationship between size of population and complexity of social organization. Southwest J Anthropol 1967, 23, 234–243.
54. Cajal, S.R.y. Histology of the Nervous System, Vol. 1; Oxford University Press: New York, 1995.
55. Kaas, J.H. Why is brain size so important: design problems and solutions as neocortex gets bigger or smaller. Brain Mind 2000, 1, 7–23.
56. Cowey, A. Cortical maps and visual perception. The Grindley Memorial Lecture. Q J Exp Psychol 1979, 31, 1–17.
57. Cowey, A. Why are there so many visual areas? In: The Organization of the Cerebral Cortex; Schmitt, F.O.; Warden, F.G.; Adelman, G.; Dennis, S.G., Eds.; MIT Press: Cambridge, MA, 1981; pp 395–413.
58. Mead, C. Analog VLSI and Neural Systems; Addison-Wesley: Boston, 1989.
59. Durbin, R.; Mitchison, G.A. dimension reduction framework for understanding cortical maps. Nature 1990, 343, 644–647.
60. Mitchison, G. Neuronal branching patterns and the economy of cortical wiring. Proc R Soc Lond B 1991, 245, 151–158.
61. Mitchison, G. Axonal trees and cortical architecture. Trends Neurosci 1992, 15, 122–126.
62. Ringo, J.L. Neuronal interconnection as a function of brain size. Brain Behav Evol 1991, 38, 1–6.
63. Cherniak, C. Local optimization of neuron arbors. Biol Cybern 1992, 66, 503–510.
64. Cherniak, C. Component placement optimization in the brain. J Neurosci 1994, 14, 2418–2427.
65. Cherniak, C. Neural component placement. Trends Neurosci 1995, 18, 522–527.
66. Cherniak, C.; Changizi, M.A.; Kang, D. Large-scale optimization of neuron arbors. Phys Rev E 1999, 59, 6001–6009.
67. Jacobs, R.A.; Jordan, M.I. Computational consequences of a bias toward short connections. J Cogn Neurosci 1992, 4, 323–336.
68. Traverso, S.; Morchio, R.; Tamone, G. Neuronal growth and the Steiner problem. Rivista di Biologia—Biol Forum 1992, 85, 405–418.
69. Ruppin, E.; Schwartz, E.L.;Yeshurun, Y. Examining the volume efficiency of the cortical architecture in a multi-processor network model. Biol Cybern 1993, 70, 89–94.
70. Van Essen, D.C. A tension-based theory of morphogenesis and compact wiring in the cntral nervous system. Nature 1997, 385, 313–319.
71. Changizi, M.A. The economy of the shape of limbed animals. Biol Cybern 2001, 84, 23–29.
72. Changizi, M.A.; Shimojo, S. Parcellation and area-area connectivity as a function of neocortex size. Brain Behav Evol, in press.
73. Chklovskii, D.B.; Koulakov, A.A. A wire length minimization approach to ocular dominance patterns in mammalian visual cortex. Physica A 2000, 284, 318–334.

74. Murray, C.D. A relationship between circumference and weight in trees and its bearing on branching angles. J Gen Physiol 1927, 10, 725–729.
75. Kamiya, A.; Togawa, T. Optimal branching structure of the vascular tree. Bull Math Biophys 1972, 34, 431–438.
76. Zamir, M. Optimality principles in arterial branching. J Theor Biol 1976, 62, 227–251.
77. Roy, A.G.; Woldenberg, M.J. A generalization of the optimal models of arterial branching. Bull Math Bio 1982, 44, 349–360.
78. Woldenberg, M.J.; Horsfield, K. Finding the optimal lengths for three branches at a junction. J Theor Biol 1983, 104, 301–318.
79. Woldenberg, M.J.; Horsfield, K. Relation of branching angles to optimality for four cost principles. J Theor Biol 1986, 122, 187–204.
80. Schreiner, W.; Neumann, M.; Neumann, F.; Roedler, S.M.; End, A.; Buxbaum, P.; Müller, M.R.; Spieckermann, P. The branching angles in computer-generated optimized models of arterial trees. J Gen Physiol 1994, 103, 975–989.
81. Schreiner, W.; Neumann, F.; Neumann, M.; End, A.; Müller, M.R. Structural quantification and bifurcation symmetry in arterial tree models generated by constrained constructive optimization. J Theor Biol 1996, 180, 161–174.
82. Changizi, M.A.; Cherniak, C. Modeling the large-scale geometry of human coronary arteries. Can J Physiol Pharmacol 2000, 78, 603–611.
83. Bern, M.; Graham, R. The shortest-network problem. Sci Am 1989, 260, 84–89.
84. Sherwani, N. Algorithms for VLSI physical design automation. Kluwer Academic: Boston, 1995.
85. Changizi, M.A. Universal scaling laws for hierarchical complexity in languages, organisms, behaviors and other combinatorial systems. J Theor Biol 2001, 211, 277–295.
86. Wolfram, S.A. New Kind of Science; Wolfram Media; Champaigne, IL, 2002.
87. Changizi, M.A. Mathematica's first academic monograph. A review of Stephen Wolfram A New Kind of Science. Complexity 2003, 8(2), 63–65.
88. Fentress, J.C.; Stilwell, F.P. Grammar of a movement sequence in inbred mice. Nature 1973, 244, 52–53.
89. Fentress, J.C. Ethological models of hierarchy and patterning of species-species behavior. In: Handbook of Behavioral Neurobiology; Satinoff, E.; Teitelbaum, P., Eds.; Plenum: New York, 1983, pp 185–234.
90. Slater, P.J.B. Describing sequences of behavior. In: Perspectives in Ethology; Bateson, P.P.G.; Klopfer, P.H., Eds.; Plenum Press: New York, 1973, pp 131–153.
91. Dawkins, R.; Dawkins, M. Hierarchical organization and postural facilitation: Rules for grooming in flies. Anim Behav 1976, 24, 739–755.
92. Douglas, J.M.; Tweed, R.L. Analysing the patterning of a sequence of discrete behavioural events. Anim Behav 1979, 27, 1236–1252.
93. Rodger, R.S.; Rosebrugh, R.D. Computing the grammar for sequences of behavoural acts. Anim Behav 1979, 27, 737–749.
94. Gallistel, C.R. The Organization of Action: A New Synthesis; Lawrence Erlbaum Associates: Hillsdale, 1980.
95. Lefebvre, L. Grooming in crickets: Timing and hierarchical organization. Anim Behav 1981, 29, 973–984.
96. Schleidt, W.M.; Yakalis, G.; Donnelly, M.; McGarry, J. A proposal for a standard ethogram, exemplified by an ethogram of the bluebreasted quail (Coturnix chinensis). J Comp Ethol 1984, 64, 193–220.
97. Berkinblit, M.B.; Feldman, A.G.; Fukson, O.I. Adaptability of innate motor patterns and motor control mechanisms. Behav Brain Sci 1986, 9, 585–638.
98. Greenfield, P.M. Language, tools and brain: The ontogeny and phylogeny of hierarchically organized sequential behavior. Behav Brain Sci 1991, 14, 531–595.
99. Allott, R. The motor theory of language: Origin and function. In: Language Origin: A Multidisciplinary Approach; Wind, J.; Bichakjian, B.H.; Nocentini, A.; Chiarelli, B., Eds.; Kluwer: Dordrecht, 1992.
100. Bizzi, E.; Mussa-Ivaldi, F.A. Neural basis of motor control and its cognitive implications. Trends Cogn Sci 1998, 2, 97–102.
101. Tower, D.B.; Elliott, K.A.C. Activity of acetylcholine system in cerebral cortex of various unanesthetized mammals. Am J Physiol 1952, 168, 747–759.
102. Tower, D.B. Structural and functional organization of mammalian cerebral cortex: The correlation of neurone density with brain size. J Comp Neurol 1954, 101, 9–52.
103. Passingham, R.E. Anatomical differences between the neocortex of man and other primates. Brain Behav Evol 1973, 7, 337–359.
104. Shultz, J.R.; Wang, S.S.-H. How the cortex got its folds: Selection constraints due to preservation of cross-brain conduction time. Proceedings of the Neuroscience Conference, 2001.
105. Murray, C.D. The physiological principle of minimum work. I. The vascular system and the cost of blood volume. Proc Natl Acad Sci USA 1926, 12, 207–214.
106. Murray, C.D. The physiological principle of minimum work applied to the angle of branching of arteries. J Gen Physiol 1926, 9, 835–841.
107. Sherman, T.F. On connecting large vessels to small. J Gen. Physiol 1981, 78, 431–453.
108. Chklovskii, D.M.; Stepanyants, A. Power-law for axon diameters at branch point. BMC Neuroscience 2003, 4, 18.
109. Frahm, H.D.; Stephan, H.; Stephan, M. Comparison of brain structure volumes in Insectivora and Primates. I. Neocortex. J Hirnforschung 1982, 23, 375–389.
110. Allman, J.M. Evolving Brains. Scientific American Library: New York, 1999.
111. Changizi, M.A. Scaling the brain and its connections. In: Evolution of Nervous Systems; Kaas, J.H., Ed.; Elsevier: New York, 2005.
112. Kaas, J.H. Sensory representations in mammals. In: Function and Formation of Neural Systems; Stent, G.S., Ed.; Dahlem Konferenzen: Berlin, 1977, pp 65–80.
113. Kaas, J.H. Why does the brain have so many visual areas? J Cogn Neurosci 1989, 1, 121–135.
114. Kaas, J.H. The evolution of isocortex. Brain Behav Evol 1995, 46, 187–196.
115. Kaas, J.H. Topographic maps are fundamental to sensory processing. Brain Res Bull 1997, 44, 107–112.
116. Braitenberg, V. Cortical architectonics, general and areal. In: Architectonics of the Cerebral Cortex; Brazier, M.; Petsche, H., Eds.; Raven: New York, 1978.
117. Barlow, HB. Why have multiple cortical areas? Vis Res 1986, 26, 81–90.
118. Ringo, J.L.; Doty, R.W.; Demeter, S.; Simard, P.Y. Time is of the essence: A conjecture that hemispheric specialization arises from interhemispheric conduction delay. Cereb Cortex 1994, 4, 331–343.
119. Braitenberg, V. Brain size and number of neurons: An exercise in synthetic neuroanatomy. J Comput Neurosci 2001, 10, 71–77.
120. Bollabás, B. Random Graphs; Academic Press: London, 1985.
121. Amaral, L.A.N.; Scala, A.; Barthélémy, M.; Stanley, H.E. Classes of small-world networks. Proc Natl Acad Sci USA 2000, 97, 11149–11152.
122. Albert, R.; Jeong, H.; Barabási, A.-L. Diameter of the world-wide web. Nature 1999, 401, 130–131.
123. Watts, D.J.; Strogatz, S.H. Collective dynamics of 'small-world' networks. Nature 1998, 393, 440–442.
124. Bilke, S.; Perterson, C. Topological properties of citation and metabolic networks. Phys Rev E 2001, 64, 036106, 1–5.
125. Newman, M.E.J. The structure of scientific collaboration networks. Proc Natl Acad Sci USA 2001, 98, 404–409.
126. Barabási, A.L.; Jeong, H.; Néda, Z.; Ravasz, E.; Schubert, A.; Vicsek, T. Evolution of the social network of scientific collaborations. Physica A 2002, 311, 590–614.
127. Ferrer, R.; Janssen, C.; Solé, R.V. Topology of technology graphs: small world patterns in electronic circuits. Phys Rev E 2001, 64, 46–119.
128. Changizi, M.A. Vagueness, rationality and undecidability: A theory of why there is vagueness. Synthese 1999, 120, 345–374.
129. Adami, C. Introduction to Artificial Life; Springer-Verlag: New York, 1998.

130. Lenski, R.E.; Ofria, C.; Collier T.C.; Adami, C. Genome complexity, robustness and genetic interactions in digital organisms. Nature 1999, 400, 661–664.
131. Ofria, C.; Adami, C.; Collier, T.C.; Hsu, G.K. Evolution of differentiated expression patterns in digital organisms. Lect Notes Artific Intell 1999, 1674, 129–138.
132. Hof, P.R.; Glezer, I.I.; Condé, F.; Flagg, R.A.; Rubin, M.B.; Nimchinsky, E.A.; Weisenhorn, D.M.V. Cellular distribution of the calcium-binding proteins parvalbumin, calbindin, and calretinin in the neocortex of mammals: phylogenetic and developmental patterns. J Chem. Neuroanat 1999, 16, 77–116.
133. Nimchinsky, E.A.; Gilissen, E.; Allman, J.M.; Perl, D.P.; Erwin, J.M. A neuronal morphologic type unique to humans and great apes. Proc Natl Acad Sci USA 1999, 96, 5268–5273.
134. Changizi, M.A. The relationship between number of muscles, behavioral repertoire, and encephalization in mammals. J Theor Biol 2003, 220, 157–168.
135. Marc, R.E.; Murry, R.F.; Basinger, S.F. Pattern recognition of amino acid signatures in retinal neurons. J Neurosci 1995, 15, 5106–5129.
136. Marc, R.E.; Murry, R.F.; Fisher, S.K.; Linberg, K.A.; Lewis, G.P.; Kalloniatis, M. Amino acid signatures in the normal cat retina. Investig Ophthalm Visual Sci 1998, 39, 1685–1693.
137. Kalloniatis, M.; Marc, R.E.; Murry, R.F. Amino acid signatures in the primate retina. J Neurosci 1996, 16, 6807–6829.
138. Manger, P.; Sum, M.; Szymanski, M.; Ridgway, S.; Krubitzer, L. Modular subdivisions of Dolphin insular cortex: Does evolutionary history repeat itself? J Cogn Neurosci 1998, 10, 153–166.
139. Changizi, M.A.; Shimojo, S. Character complexity and redundancy in writing systems over human history. Proc Roy Soc Lond B 2005, 272, 267–275.
140. Abeles, M. Corticonics: Neural Circuits of the Cerebral Cortex. Cambridge University Press: Cambridge, 1991.
141. Horowitz, P.; Hill, W. The Art of Electronics; Cambridge University Press: Cambridge, 1989.
142. Kaas, J.H. The organization of neocortex in mammals: implications for theories of brain function. Annu Rev Psychol 1987, 38, 129–151.
143. Krubitzer, L. The organization of neocortex in mammals: Are species differences really so different? Trends Neurosci 1995, 18, 408–417.
144. Northcutt, R.G.; Kaas, J.H. The emergence and evolution of mammalian neocortex. Trends Neurosci 1995, 18, 373–379.
145. Krubitzer, L.; Manger, P.; Pettigrew, J.; Calford, M. Organization of somatosensory cortex in monotremes: In search of a prototypical plan. J Comp Neurol 1995, 351, 261–306.
146. Beck, P.D.; Pospichal, M.W.; Kaas, J.H. Topography, architecture, and connections of somatosensory cortex in opossums: Evidence for five somatosensory areas. J Comp Neurol 1996, 366, 109–133.
147. Krubitzer, L.; Künzle, H.; Kaas, J. Organization of sensory cortex in a madagascan insectivore, the tenrec (Echinops telfairi). J Comp Neurol 1997, 379, 399–414.
148. Lyon, D.C.; Jain, N.; Kaas, J.H. Cortical connections of striate and extrastriate visual areas in tree shrews. J Comp Neurol 1998, 401, 109–128.
149. Catania, K.C. Cortical organization in shrews: Evidence from five species. J Comp Neurol 1999, 410, 55–72.
150. Krubitzer, L.; Huffman, K.J. Arealization of the neocortex in mammals: Genetic and epigenetic contributions to the phenotype. Brain Behav Evol 2000, 55, 322–355.
151. Manger, P.R.; Kiper, D.; Masiello, I.; Murillo, L.; Tettoni, L.; Hunyadi, Z.; Innocenti, G.M. The representation of the visual field in three extrastriate areas of the ferret (Mustela putorius) and the relationship of retinotopy and field boundaries to callosal connectivity. Cereb Cortex 2002, 12, 423–437.
152. Coogan, T.A.; Burkhalter, A. Hierarchical organization of areas in rat visual cortex. J Neurosci 1993, 13, 3749–3772.
153. Young, M.P. The organization of neural systems in the primate cerebral cortex. Proc R Soc Lond B 1993, 252, 13–18.
154. Scannell, J.W.; Young, M.P. The connectional organization of neural systems in the cat cerebral cortex. Curr Biol 1993, 3, 191–200.
155. Hackett, T.A.; Stepniewska, I.; Kaas, J.H. Subdivisions of auditory cortex and ipsilateral cortical connections of the parabelt auditory cortex in macaque monkeys. J Comp Neurol 1998, 394, 475–495.
156. Kaas, J.H.; Hackett, T.A. Subdivisions of auditory cortex and processing streams in primates. Proc Natl Acad Sci USA 2000, 97, 11793–11799.