

## **Essential circuits of cognition: The brain's basic operations, architecture, and representations**

**R. Granger**

University of California Irvine, and Dartmouth College  
Richard.Granger@Gmail.com / www.BrainEngineering.org

### **Introduction**

The goals of artificial intelligence have always been twofold: i) formal explanation of the mechanisms underlying human (and animal) intelligence and ii) construction of powerful intelligent artifacts based on those mechanisms. The latter engineering goal may pragmatically benefit from the former scientific one: extant face recognition systems and automated telephone operators might have been considered the best possible mechanisms were it not for our own abilities. The only reason that we know that these industrial systems can be outperformed is that humans do so.

Biological systems achieve their cognitive capabilities solely through brain mechanisms: the physiological operation of anatomical circuitries. Brain circuits are circuits; that is, they can be understood in computational terms. An explosion of knowledge in neuroscience and related fields is revealing the data crucial for characterizing the layout and properties of these circuits. For purposes of artificial intelligence, this information can be organized into three key topics:

- basic operations: what are the elemental operators carrying out fundamental mental steps?
- architecture: in what organizational control structure are the operators embedded?
- representation: how are memories and knowledge structured, stored, and retrieved?

Not one of these questions is yet answered. But findings from a range of fields including anatomy, physiology, plasticity, pharmacology, and neuroimaging, enable formulation of strict and interlocking constraints on the space of admissible answers. The ensuing review will discuss key findings that provide such constraints, and will point

to publications in the primary literature offering further background on these topics.

### **Basic operations**

A perennial challenge for the fields of artificial intelligence and psychology is the question of distinguishing composite operations from their underlying fundamental constituents. Analyses of the primary basic circuits of the human brain has led to derivation of a specific proposed set of fundamental psychological operators.

The human brain consists of evolutionarily recent forebrain circuit designs (telencephalic circuits) layered on top of preserved ancient (e.g., reptilian) circuits, with the new designs accounting for more than 90% of the volume of the human brain. There are four primary divisions of telencephalic forebrain (cortex, striatal complex, hippocampal formation, amygdala nuclei), and many subdivisions (e.g., anterior vs posterior cortex, five cortical layers, local circuits, striatal components, hippocampal fields CA1, CA3, dentate gyrus, subiculum, ...), each with its own cell types and local circuit design layouts, thus presumably each conferring unique computational properties. A multi-year program of stepwise bottom-up analyses of many of these constituents has yielded extensive observations of the responses of these circuits to hypothesized typical input signals based on physiological activity in vivo. This in turn has led to a principled array of these circuits' proposed algorithmic formulations.

Table 1 presents a compact summary of the findings from these bottom-up analysis efforts, listing primary circuits and derivations of their proposed characteristic algorithms, along with behaviors that have been identified from the primary cortical-subcortical loops in which the circuits are embedded (cortico-striatal, cortico-amygdala, cortico-hippocampal).

Table 1: Summary of derived brain circuit algorithms and initial citations

Circuits studied	Functions derived
thalamocortical core loops	clustering, hierarchical clustering (Rodriguez et al., 2004)
thalamocortical matrix loops	sequences, chaining, hash codes (Rodriguez et al., 2004)
striatal complex / basal ganglia	reinforcement learning, exploratory action selection (Granger 2005)
cortico-striatal loops	automatization; variation; power law (Granger 2005)
amygdala nuclei	filters / “toggles” (Parker et al., in prep)
cortico-amygdala loop	state-dependent storage & retrieval; category broadening (“
hippocampal fields	time dilation / compression (Granger et al. 1996)
cortico-hippocampal loops	spatiotemporal relations; navigation (Kilborn 1996)

Derivation of each of these different brain circuit algorithms has been detailed in prior publications; each has been simplified and subjected to formal treatment, resulting in time and space analyses showing, in each case, linear or better time costs, a crucial scaling characteristic of any algorithm or circuit design that is to be applied to large real data sets. (For an extensive literature detailing these briefly summarized derivations see, e.g., Lynch & Granger 1992; Coultrip & Granger 1994; Granger et al. 1994, 1996; Kilborn et al. 1996; Shimono et al. 2000; Rodriguez et al. 2004; Granger et al. 2005, 2006). The present article will attempt to address the integration of these components into the overall well-specified architecture of the human brain, and their resulting emergent function as that brain architecture is grown to human size.

## Architecture

### 1. Telencephalon

As the individual mechanisms are derived bottom-up from basic brain circuits, they may constitute irreducible fundamental operators, which combine to form composite operators. Their combinations are determined by the larger systems-level architectures within which they are embedded. There is (perhaps

surprisingly) a single large-scale architecture that organizes all telencephalic components. Figure 1 illustrates an outline of this encompassing architecture (see, e.g., Striedter 1997, 2005).

For almost any given region of posterior cortex, there is a corresponding region of anterior cortex (e.g., the frontal eye fields, connected to posterior visual cortical areas), as well as corresponding regions of striatum, pallidum and thalamus, connected in register. These complementary cortical and subcortical regions are connected in a characteristic pattern: reciprocal connections between posterior and anterior cortex, converging anterior and posterior cortical projections to a related region of striatum, which in turn connects (via pallidum and thalamus) back to the same region of anterior cortex. The overall “systems circuit” is by far the largest coherent loop in the mammalian brain, and it is repeated for multiple regions of posterior cortex, with dedicated regions corresponding to individual sensory modalities, as well as non-cortical telencephalic regions including components of hippocampus and amygdala, connected with dedicated regions of striatum and anterior cortex, as shown in Figure 2.

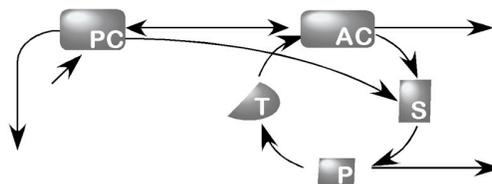


Figure 1. Architectural organization of regions of posterior cortex (PC), anterior cortex (AC), striatum (S), pallidum (P) and thalamus (T). Primary inputs arrive at posterior cortex, which projects cortico-cortically to other posterior regions as well as innervating (in two places) the large loop of anterior cortex, striatum, pallidum, thalamus. Outputs from both pallidum and anterior cortex project to motor control systems.

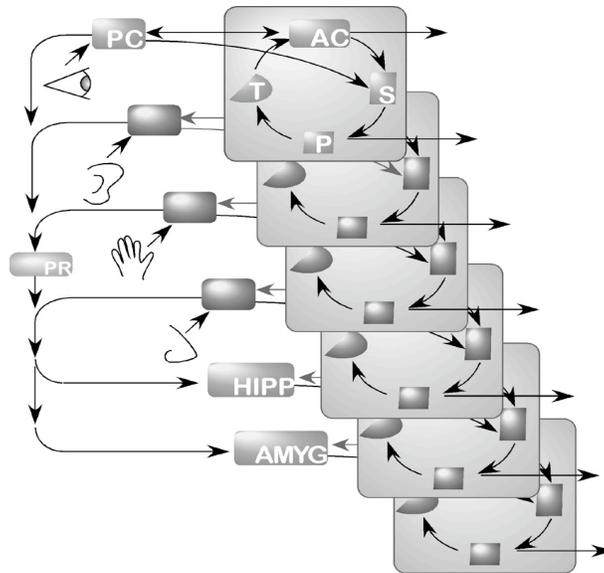


Figure 2. Repetition of telencephalic architecture across multiple cortical, striatal, and limbic regions.

This repeated architectural design alone accounts for the vast majority of the circuitry of the human brain. If this architecture determines the input / output relations among the disparate constituent modules, it may be possible to arrive at initial hypotheses of the architecture's overall operation. The system changes via statistical modifications at the myriad circuit connectors (synapses). Inputs come to generate selectively responsive pathways, which can be thought of as distributed internal representations (despite the baggage that that term may carry) in posterior cortex. Portions of anterior cortex directly produce motor outputs as well as generating inputs to striatum and pallidum, which in turn also control brainstem motor systems. Striatum and pallidum, together constituting most of the basal ganglia, have been widely argued to use dopaminergic signaling to carry out trial and error training, formally characterizable as reinforcement learning (cf. Schultz et al. 1997). Hippocampal regions and cortico-hippocampal loops are sources of much controversy but much of the published literature is concordant with a putative role in representational coding especially during storage, and relational coding among multiple features, as seen in navigation as well as other relational or configural demands.

## 2. Allometry

Additional constraints aid in analysis. In particular, it is noteworthy that the brain retains its internal

structures and connectivity throughout all mammals, but the relative sizes of these structures and pathways change as the brain grows in absolute size. Figure 3 illustrates more or less canonical small (left) and large (right) mammalian brains, highlighting their architectural differences. Interacting posterior and anterior cortical areas, although not shown to scale in the illustration, constitute far and away the largest components of the design, and they are disproportionately (allometrically) further enlarged with evolution, so that the ratio of cortical to subcortical tissue increases with brain size. Relations among cortical areas and between cortical and subcortical areas are also allometrically altered. The three largest relational changes, as shown in the figure, are:

- increase in connection pathways (fasciculi) between anterior and posterior cortex;
- rebalance of pallidal outputs, chiefly projecting to brainstem motor nuclei in small-brained mammals (left) but, in large-brained mammals (right), predominantly projecting back to anterior thalamocortical networks, "closing the loop" between cortex and basal ganglia;
- increase in motor projections from cortex, compensating for reduced pallidal motor outputs.

(See e.g., Nudo & Masterson 1990; Striedter 2005).

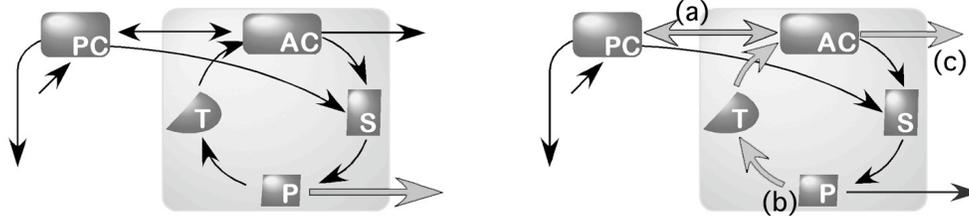


Figure 3. Allometric connectivity changes with brain growth: small-brained mammals (left) vs large-brained mammals (right) exhibit changes in relative size of a) anterior-posterior cortical connectivity, b) striatal output pathways, and c) cortical motor control paths.

To the extent that cortical, thalamic, striatal and pallidal circuitry compute similarly in small and large brains, they must be able to contribute to the range of different configurations in which they find themselves embedded. The basal ganglia's outputs (from pallidum), then, must presumably be intelligible both to motor nuclei and to thalamocortical circuitry. This imposes a notable computational constraint on the possible functions of the cortico-striatal communication system (Granger 2005; 2006).

A final substantial architectural constraint arises from the connections among the cortical regions across these telencephalic "blocks". In particular, the output of a given cortical area becomes input to a subsequent or "downstream" area, and in turn receives projections backwards, with forward and backward projections exhibiting very different patterns of connectivity. The outputs of a given area must be able to be "read" by downstream areas, and their outputs must in turn be intelligible to upstream regions (see, e.g., Lorente de No 1938; Szentagothai 1975; Rockel et al. 1980; White & Peters 1993; Peters & Payne 1993; Valverde 2002).

## Representation

### 1. Regularities

With this set of processing elements, connected as prescribed in the overall telencephalic architecture, we may ask what it is that is being computed. Perceptual inputs arrive at peripheral structures, e.g., retina, certain thalamic nuclei, and even early sensory cortical areas, all of which contain unique and often exotic structures, not replicated elsewhere in the brain, designed to deal with the specific physics of input signals from photons and sound waves to skin touch sensors. Beyond these front end

systems, the internal signaling system of the brain remains unsolved, but again there are many constraints from multiple sources that severely restrict the possibilities.

Among the most prevalent characteristics of brain circuits is their plasticity: they alter the "strengths" of their neuron-to-neuron synaptic connections, subtly rewiring themselves to yield slightly different responses as a function of prior exposure. The apparent "magic" of human cognition arises in large measure from its remarkable ability to learn. It is noteworthy that learning alone does not automatically confer "magic," as witness the many so-called "neural networks" that perform various kinds of learning, achieving intriguing and useful statistical capabilities but little more. To achieve the capabilities of the brain, learning must be embedded in the circuits, architectures and operating rules of the brain.

Two features of internal brain circuits (past the sensory periphery) are particularly notable: i) circuits for different modalities (e.g., vision, audition) are remarkably similar and ii) the majority of circuits receive inputs from multiple modalities. Taken together, these two observations provide yet another constraint on possible mechanism: communication among cortical regions likely consists of a single, shared, cross-modal (or amodal) internal representation language, regardless of the particular information being conveyed.

Another constraint arises from the variation in capabilities arising from the same repeated telencephalic architectural design. Most engineering circuits (e.g., in the CPUs of typical computers) do not have the property of giving rise to wholly new

functions when larger versions of them are constructed. Large computers can have larger memories and address spaces, but they do not intrinsically perform different kinds of functions from their smaller counterparts. Brains, in contrast, somehow accrue new faculties with growth: dogs are capable of cognitive feats unknown to mice, such as their ability to be trained as seeing-eye dogs; chimps can learn complex social interactions; humans of course attain language and reasoning faculties that have not been found in any other animals. Brain architectures must be constituted such that making more of them enables interactions that confer new powers to larger assemblies.

2. Grammars

As specified earlier, individual cortical regions are hypothesized to compute clusters (i.e., similarity-based categories) and sequences (chaining), via different components of their intrinsic circuitry. These two components, thalamocortical core and matrix loops, interact to produce sequences of clusters (see Rodriguez et al. 2004).

The output of one thalamocortical circuit is input to others with identical or near-identical structure; these thus produce sequences of clusters of sequences of clusters ..., effectively nesting the product of one “level” of processing into downstream processing products. Successive nesting creates increasingly deep hierarchical “trees” of sequences of clusters. (Feedback from downstream to upstream regions participates actively in this process; partial activation of a downstream region has the consequence of increasing the probability of response of its potential upstream input constituents, acting in effect like

“expectations” that those inputs will occur.) These cortical mechanisms interact with hippocampal time dilation and contraction, amygdala “toggling” of salient features, and striatal reinforcement learning in cases of relevant feedback. Together the system produces incrementally constructed and selectively reinforced hierarchical representations consisting of nested sequences of clusters (Granger 2006).

Figure 4 is an abstract illustration of successive stages of a representation so constructed. Initial simple input features (e.g., visual spots or edges; auditory frequencies or formants) transduced by front end mechanisms are learned by earliest, specialized stages (denoted in the figure by single letters A, B, etc). Their encoded outputs are input to downstream structures which learn clusters (categories of similar inputs) and sequences of clusters; further downstream regions learn sequences of clusters of sequences of clusters, and so on. Each downstream region, depending on its pattern of connectivity with its inputs, may exhibit a “bias” preferring inputs with particular characteristics; these are genetically programmed and little is yet known of their layout, though work in quantitative neuroanatomy is advancing knowledge in this realm. Prohibitive space would be required to learn all such combinations, but combinatorial explosion is avoided by two primary mechanisms: i) Bias: Of all the possible combinations of features that could occur, only some actually do, and, as just mentioned, some combinations are preferred over others; ii) Competition: With learning, oft-traversed regions become increasingly strengthened and, via lateral inhibition of neighboring regions, become what may be thought of as “specialists” in certain types of inputs, competing to respond.

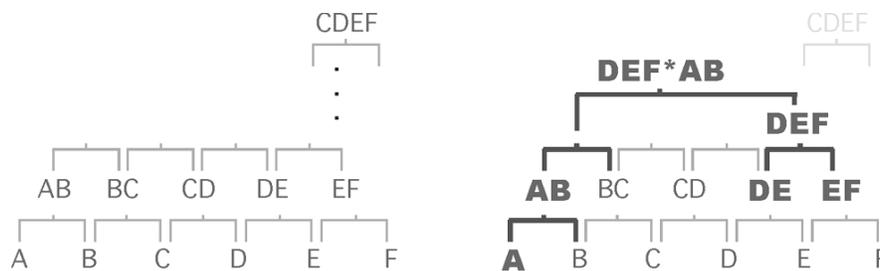


Figure 4. Illustration of hierarchies constructed by telencephalic architecture. Initial features generate successively nested sequences of categories of features (left). Additional exposure eventually (right) selectively strengthens sequences that recur (e.g., AB), weakens those that do not (e.g., CDEF), and constructs new sequences of categories as they occur and recur (e.g., DEF followed by a category that may include any of A-F (denoted here by a \*) followed by AB).

The emergent data structure of the telencephalic system, statistically learned nested sequences of clusters (as illustrated in Figure 4) is a superset of the structures that constitute formal grammars. The nested sequences of clusters are equivalent to ordered sequences of “proto-grammatical” elements such that each element represents either a category (in this case a cluster) or expands to another such element (nesting), just as grammatical rewrite rules establish new relations among grammatical elements.

### 3. Predicted characteristics

This arrangement leads to a series of implied characteristics. Whereas early upstream areas respond to generic features and simple feature assemblies, downstream regions respond with increasing selectivity to only specific assemblies, typically those that occur as patterns within oft-seen stimuli. As a concomitant, further downstream regions should be expected to selectively respond to larger or longer (visual or auditory) patterns. This expectation agrees with experimental findings (see, e.g., Griffiths et al. 1998).

As most visual inputs consist simply of different arrangements of the same sets of primitive input features, it is expected that patterns of brain activation should be extremely similar in response to many different visual inputs, but that the similarity of those brain activation patterns ought to correspond to the similarity of their inputs, that is, activation patterns ought to be more similar for similar inputs, and more different for different inputs. Moreover, if cortical regions are competing to respond to a given input, they should exhibit “category boundaries,” that is, the responses to images within a category (e.g., faces versus houses) should be more similar to each other than the images themselves are. Put differently, even highly different faces are likely to generate very similar cortical response patterns, whereas the similarity between any face and any house (as long as it is not a house that looks like a face!) should be more different than any two faces or any two houses.

These three predictions of the model (distributed representations, similarity of patterns, and category boundaries) turn out to be controversial: depending on the analysis methods, neuroimaging studies have been used to support a number of still-conflicting hypotheses. The present prediction is concordant with some prominent findings, in which distributed, overlapping patterns occur in response to images of, say, faces vs houses; more similar inputs tend to generate more similar responses; and responses to images within perceptual categories are more similar than responses to images across categories (Haxby et al. 2001; Pietrini et al. 2004; Furey et al. 2006; Hanson et al. 2004). These findings are the subject of ongoing study.

### 4. Specializations

The incremental nature of the nested sequences of clusters data structure enables it to grow simply by adding new copies of telencephalic thalamo-cortico-striatal-limbic loops, corresponding to the incremental addition of “rules” acquired by the grammar. As more telencephalic “real estate” is added, the data structures that are constructed correspond to both longer and more abstract sequences, due to iterative nesting. Even regions of telencephalon with identical computational function nonetheless receive inputs from different sources, thus changing the feature combinations on which they operate (but see Galuske et al. 2000; Preuss 1995; 2000).

Current study is focused on possible mechanisms by which successively more complex data structures, corresponding to differential downstream pathways, might capture increasingly complex representational concepts. Table 2 suggests examples of such pathways, meant to highlight potential relations between anatomical pathways and cognitive outgrowths of those pathways. (The table emphasizes mapped pathways in the visual domain, but corresponding pathways, of increasing representational complexity, exist for auditory and somatosensory modalities.)

Table 2: Partial mapping of representational pathways to anatomical pathways

Peripheral sensors ↔ core thalamus (e.g., LGN); superior colliculus [sensory input]
↔ Posterior cortex
↔ Primary visual cortex [features]
Lateral occipital cortex (LOC) ↔ Inferior temporal cortex (IT)
[intra-object (“what”): categorical, structural, configural]
V3a ↔ Medial temporal (MT) ↔ Posterior parietal (PPC)
[inter-object (“where”): spatiotemporal adjustment]
↔ Angular gyrus [cross-modal; associational]
↔ Parahippocampal cortex ↔ hippocampal region [navigation; spatiotemporal]
↔ (fasciculi)
↔ Anterior cortex
↔ Dorsolateral prefrontal cortex (DLPFC)
[expectation, causality, naïve physics, affordance, simulation]
↔ Frontal eye fields [covert and overt selective attention]
↔ Anterior cingulate [self/other; goals] ↔ amygdala nuc. [motivation; state; “toggles”]
↔ areas 44, 45 [social self/other; language?]
↔ Orbitofrontal cortex (OFC) [abstract sequences ↔ plans ↔ motor behavior]
↔ DLPFC; OFC [match/mismatch] ↔ basal ganglia [reinforcement learning]
↔ motor effectors

Inverting these relations, i.e., listing successively emergent functions first and their anatomical regions second, yields an organization of functions that appear to fall naturally into a series of conceptual categories (Table 3):

As mentioned, of the large set of all possible assemblies of features, only a small subset seem to be readily learned by biological organisms; there apparently exist species-specific biases that shape animals’ (including humans’) interpretations of various inputs. For instance, in response to very little data, humans will interpret certain coherent point-source motions as biological motion (e.g., when lights are affixed to the limbs of people moving in an otherwise dark environment); will interpret many distorted inputs as face-like; will interpret many sounds as speech-like, and so on. It is assumed that these biases may arise from

developmental pre-selection (via still-unknown mechanisms) of some cortico-cortical pathways that will selectively respond to particular types of feature assemblies.

Without yet understanding the biological means by which such paths may be pre-shaped, it is possible nonetheless to observe data from multiple sources such as neuroimaging and to use the information to similarly pre-shape the telencephalic simulation artificially, in hopes of studying potential effects of these hypothesized pre-set biases. Work in this vein is in progress, artificially establishing predefined pathways initially for faces and for voices. It is hoped that perhaps by this expedient it will be possible to study emergent specializations in downstream cortical regions after sufficient training on a range of related inputs. Initial work in this direction is being done on linguistic inputs.

Table 3: Inverted partial mapping; anatomical domains to representational domains

Perception (“assembly” of features into objects)
within-object: categorical, structural, configural [V1 – V3; LOC; IT]
between-object: spatial, temporal [MT; PPC]
Cross-modal (relations among objects)
association, successive abstraction [Angular gyrus]
location, navigation, spatiotemporal adjustment [Parahippocampal ctx; hippocampus]
Function (coherent pan-relational sequence categories)
expectation, causation, naïve physics [DLPFC]
utility, affordance, simulation [DLPFC; OFC]
Action (“re-assembly” of abstracted sequences)
sequence reassembly, plans, acts [DLPFC; OFC]
match/mismatch; reinforcement, learning by doing [OFC; PFC; basal ganglia]
Interaction (relations among re-assembled sequences)
selective attention, goals [FEF; ACC]
social interaction, self/other [ACC; areas 44, 45]
symbolic descriptors, language [DLPFC; areas 44, 45]

## 5. Language

In current work on language, far-downstream areas are assumed to come to identify symbolic descriptors (see Tables 2, 3) that are statistically repeated in relevant situations, such as words. The following theoretical question is explored: if further downstream regions arose (e.g., in the evolution of human primates) beyond primary symbolic descriptors (words), and if those downstream regions carried out the same computations as the other telencephalic regions that led up to them, what resulting internal representational structures would be produced?

Figure 5 illustrates structures occurring in response to simple sentences (“John hit Sam”) as input. Construction of sequences (e.g., S11, “John” followed by “hit”), and categories (e.g., C21, “hit” and “kissed,” items that can follow “John”) are combined in successive downstream regions (n+1, n+2, etc.) to create “proto-grammatical fragments” corresponding to internal representations of linguistic structure information.

It is worth noting that the generated structures can be used both a) in the processing of subsequent novel

inputs and b) in the generation of arbitrary new strings that will conform to the rules inherent in the learned internal representational structures. The generative nature of the resulting representations is worth emphasizing, addressing a crucial aspect of linguistic grammars that can otherwise be absent from some purely input-processing or parsing mechanisms. A potentially infinite set of strings can be generated from the internal sequences of clusters, and the strings will be consistent with the internal grammar (see, e.g., Pinker 1999; Hauser et al. 2002; Fitch & Hauser 2004; Pinker & Jackendoff 2005). It is also noteworthy that the grammar does not take the form typically adopted in attempts to formally characterize the syntactic structure of natural languages (such as English). The protogrammatical fragments capture regularities that are empirically seen to suffice for both parsing and generation, and have the structure to account for rule-like behaviors that characterize linguistic behavior. Research is currently in progress to study the formal relations between typical linguistic grammars, and protogrammatical fragments that are emergent from nested sequences of clusters.

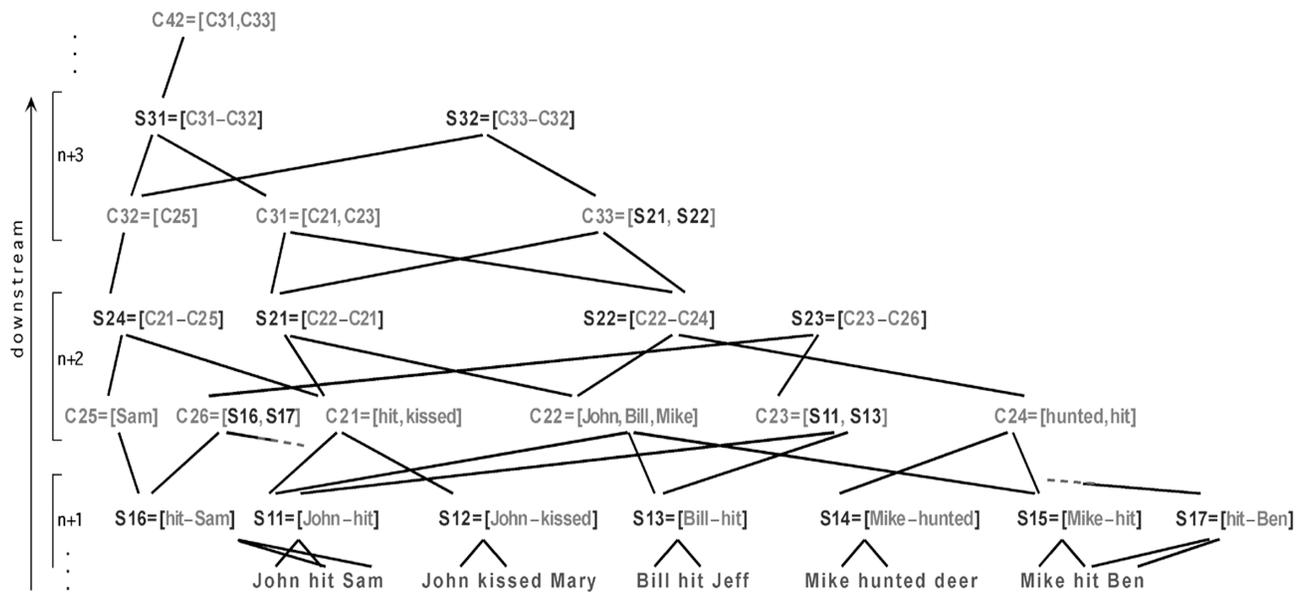


Figure 5. Nested sequences of clusters as sample proto-grammatical fragments educed from input strings.

An additional characteristic of language that challenges researchers is the seeming effortless with which children learn language – readily contrasted even with the comparatively laborious training typically required for adults learning a second language. Some presumably innate tendencies enable children to master complex language structure solely by exposure rather than by intensive schooling. If an innate bias related to sequences of categories of vocal utterances (speech) led (in larger-brained organisms) to a downstream bias for certain sequences of categories of assemblies of speech sounds (words), then this may at least in part account for this much-studied but still elusive nature of innate language capacity.

**Discussion**

The myriad tasks of intelligence (visual and auditory recognition, planning, language, and many more) are, with few exceptions, ill-specified. That is, unlike engineering tasks in which a formal specification precedes solution, these tasks are approached with a sole point of reference: observation of intelligent systems (humans and other animals) that perform them, by means unknown. Many approaches thus refer to attempts at “reverse engineering,” i.e., observing behaviors and attempting to educe their underlying mechanisms. But observation is highly prone to

underspecification, that is, many nontrivially-different mechanisms can give rise to any particular observed behavior, and thus the reverse engineering task is itself ill-specified.

The many related fields of neuroscience (anatomy, physiology, biochemistry, pharmacology, neuroimaging) have generated an unprecedented profusion of new data in the last decade, perhaps earning it its extrinsically-endowed label “the decade of the brain.” That vast trove of data may for the first time enable bottom-up approaches to the understanding of intelligence, a task that is well specified: formal characterization of the behaviors arising from the anatomical structure and physiological operation of biological circuits in their endogenous systems architectures. That task is oddly posed compared to standard artificial intelligence tasks. Whereas AI has traditionally attempted to identify mechanisms that could arrive at a predetermined functional outcome, we attempt to identify what outcomes emerge from a predetermined set of biological mechanisms.

This new task, which might be termed “brain engineering,” will intermittently falter in the face of still-ambiguous biological findings and still-imperfect understanding of brain structures and mechanisms, but in principle, via study and analysis of the circuits of the brain, it has the potential to

derive the actual individual and composite operations of the brain that constitute the “instruction set,” or basic mental operations, from which all complex behavioral and cognitive abilities are constructed. At present this field is in its infancy, struggling to identify emergent functions, resolve ambiguous data, and integrate sometimes disparate and apparently irreconcilable findings. To the extent that it generates coherent hypotheses, and the extent to which those hypotheses find occasional application in behaviorally relevant target domains from perception and navigation to reasoning and language, the field is perhaps making incremental progress towards its goal.

(A final digression may be made in this regard. It has been posited here that added brain regions, judiciously sited, may have given rise to the qualitative leap from simple symbol use in apes to true language, in all its complexity, in humans. The question is thus raised of what additional capabilities, perhaps currently unimagined, would be birthed if further brain regions were added, either by next natural steps of evolution or by the engineering artifice of man. If profoundly useful and transformative linguistic abilities arose almost full-blown via the brain expansion from ape to human, might there be leaps of equal size if brain systems are engineered to the size of human brains – and beyond? Two goals of artificial intelligence were mentioned at the outset of this article: the scientific understanding of the brain and the engineering replication of it, but success in these endeavors may lead to a third goal, that of surpassing human intelligence, possibly creating thinking machines as far beyond our comprehension as we may be beyond apes. This specter has lived in the realm of science fiction, but in our specific understanding of how the powerful abilities of language may have been spliced on to pre-human brain systems, perhaps a glimpsed route to future new capabilities is revealed.)

It should never have been expected that the burdens of artificial intelligence would be light ones. It may well be that the science of the mind will be no less challenging than the far older sciences of life (biology) and of matter (physics), to which it is logically subsequent, and on which it may depend. Fifty years have passed since the inception of the field of artificial intelligence. Those years have seen

much frustration and many impediments, but also have witnessed the successful creation of systems that have made strong initial strides in realms of navigation, language processing, problem-solving, and data mining and analysis, among many others; steps along the path to the unexpectedly monumental goals set by the creators of the field here at Dartmouth a half-century ago. With its burgeoning new tools of neuroscience data and brain engineering, possibly new wonders will be created, and perhaps the coming fifty years will be seen as at least as fruitful, when we duly reconvene for this purpose, at this place, in 2056.

**Acknowledgments:** This work was supported in part by DARPA/ONR awards NBCHC050074 and N00014-05-C-0517.

## References

- Aboitiz F (1993) Further comments on the evolutionary origin of mammalian brain. *Medical Hypotheses* 41: 409-418.
- Braitenberg V, Schüz A (1998) *Cortex: statistics and geometry of neuronal connectivity*, NY: Springer.
- Coultrip R, Granger R (1994) LTP learning rules in sparse networks approximate Bayes classifiers via Parzen's method. *Neural Networks* 7:463-476.
- Coultrip R, Granger R, Lynch G (1992) A cortical model of winner-take-all competition via lateral inhibition. *Neural Networks* 5:47-54.
- Fitch T, Hauser M (2004) Computational constraints on syntactic processing in a nonhuman primate. *Science* 303: 377-380.
- Furey M, Tanskanen T, Beauchamp M, Avikainen S, Uutela K, Hari R, Haxby J (2006). *Proc Nat'l Acad Sci.* 103:1065-1070.
- Galuske RA, Schlote W, Bratzke H, Singer W (2000) Interhemispheric asymmetries of the modular structure in human temporal cortex. *Science* 289:1946-1949.
- Granger R, Whitson J, Larson J, Lynch G (1994) Non-Hebbian properties of long-term potentiation enable high-capacity encoding of temporal sequences. *Proc Nat'l Acad Sci* 91:10104-10108.
- Granger R, Wiebe S, Taketani M, Ambros-Ingerson J, Lynch G (1996) Distinct memory circuits comprising the hippocampal region. *Hippocampus* 6:567-578.
- Granger R. (2005) Brain circuit implementation: High-precision computation from low-precision components. In: *Replacement Parts for the Brain* (T.Berger,Ed) MA: MIT Press.

- Granger R. (2006) Engines of the brain: The computational instruction set of human cognition. *AI Magazine*. (in press).
- Griffiths T, Buchel C, Frackowiak R, Patterson R (1998). Analysis of temporal structure in sound by the human brain. *Nature Neurosci*, 1(5):422-427.
- Hanson S, Matsuka T, Haxby J (2004). Combinatorial codes in ventral temporal lobe for object recognition: Haxby (2001) revisited: is there a "face" area? *NeuroImage* 23: 156-166.
- Hauser M, Chomsky N, Fitch T (2002). The language faculty: What is it, who has it, and how did it evolve? *Science* 298: 1569-1579.
- Haxby J, Gobbini M, Furey M, Ishai A, Schouten J, Pietrini P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293:2425-2430.
- Jones E (1998) A new view of specific and nonspecific thalamocortical connections. *Advances in Neurology* 77:49-71.
- Kilborn K, Granger R, Lynch G (1996) Effects of LTP on response selectivity of simulated cortical neurons. *J Cognitive Neuroscience* 8:338-353.
- Lorente de No R (1938) Cerebral cortex: Architecture, intracortical connections, motor projections. In: *Physiology of the nervous system* (Fulton J, ed), pp 291-340. London: Oxford.
- Lynch G, Granger R. (1992). Variations in synaptic plasticity and types of memory in cortico-hippocampal networks. *J. Cognitive Neurosci.*, 4: 189-199.
- Nudo R, Masterson R. (1990) Descending pathways to the spinal cord, IV: Some factors related to the amount of cortex devoted to the corticospinal tract. *J Comp Neurol* 296:584-97.
- Peters A, Payne B (1993) Numerical relationships between geniculocortical afferents and pyramidal cell modules in cat primary visual cortex. *Cerebral Cortex* 3:69-78.
- Pietrini P, Furey M, Ricciardi E, Gobbini M, Wu W, Cohen L, Guazelli M, Haxby J (2004). Beyond sensory images: object-based representation in the human ventral pathway. *Proc Nat'l Acad Sci*, 101:5658-5663.
- Pinker S (1999). *Words and rules: the ingredients of language*. New York: HarperCollins.
- Pinker S, Jackendoff R (2005). *The faculty of language: what's special about it?* *Cognition*, 95: 201-236.
- Preuss T (1995). Do rats have prefrontal cortex? The Rose-Woolsey-Akert program reconsidered. *J Cognitive Neuroscience*, 7: 1-24.
- Preuss T (2000). What's human about the human brain? In: *The New Cognitive Neurosciences*. M.Gazzaniga (Ed.), Cambridge, MA: MIT Press, pp.1219-1234.
- Rockel A, Hiorns R, Powell TPS (1980) Basic uniformity in structure of the neocortex. *Brain* 103:221-244.
- Rodriguez A, Whitson J, Granger R (2004) Derivation & analysis of basic computational operations of thalamocortical circuits. *J. Cog Neuro*, 16: 856-877.
- Schultz W, Dayan P, Montague P (1997) A neural substrate of prediction & reward. *Science* 275:1593-9.
- Shimono K, Brucher F, Granger R, Lynch G, Taketani M (2000) Origins and distribution of cholinergically induced beta rhythms in hippocampal slices. *Journal of Neuroscience* 20:8462-8473.
- Striedter G (1997) The telencephalon of tetrapods in evolution. *Brain Behavior Evolution*. 49:179-213.
- Striedter G (2005) *Principles of Brain Evolution*. NY: Sinauer.
- Szentagothai J (1975) The 'module-concept' in cerebral cortex architecture. *Brain Research* 95:475-496.
- Valverde F (2002) Structure of cerebral cortex. Intrinsic organization & comparative analysis. *Revista de Neurologia*. 34:758-780.
- White E, Peters A (1993) Cortical modules in the posteromedial barrel subfield (Sml) of the mouse. *J Comparative Neurology* 334:86-96.