High Definition Fiber Tracking Exposes Circuit Diagram for Brain Showing Triarchic Representation, Domain General Control, and Metacognitive Subsystems

Walter Schneider, Sudhir K. Pathak, Jeff S. Phillips, Michael W. Cole

University of Pittsburgh
629 Learning Research and Development Center
3939 O’Hara Street, Pittsburgh, PA 15260, USA
wws@pitt.edu

Abstract
Dramatic advances in the last six months in High Definition Fiber Tracking (HDFT) make it possible to image brain fiber connectivity from source to destination, mapping hundreds of thousands of fiber tracts with sufficient resolution to identify the cable-level circuit diagram of the human brain. Brain activity imaging studies using functional Magnetic Resonance Imaging (fMRI) identify differential activation patterns as a function of task and level of practice. These data show subnetworks which suggest communication of high bandwidth vector associations, scalar priority and control signals, and interactions with control and meta-cognition. The connectivity and activity data support a triarchic cognitive architecture. In this model, processing is the synergistic interaction of three interlinked cognitive computational systems with differential computation roles and evolutionary histories. This model provides a guide to interpretation of the interaction of the major systems levels of the human brain.

Triarchic Human Cognitive Architecture
We propose that the effectiveness of human learning and cognition is a result of the synergistic interaction of three hierarchically organized systems (Figure 1): a representation system, a cognitive control network, and a meta-cognitive system.

The representation system (RS: Figure 2) involves sensory-motor processing in a multilevel hierarchy (e.g., visual features to visual objects). It provides associative information and transformations (e.g., rotations of a face) within modality. The representation system also provides cross links between modalities (e.g., connecting the visual, auditory, tactile, and motor representations of a

“hammer”). Information is coded in many (~10,000) hypercolumn modules, each containing a million neurons with the modules organized in about ~500 functionally identifiable brain areas. Activity within each of these hypercolumns can be viewed as a vector of representation information which is communicated to other representation modules. Each representation module also outputs three report signals providing a scalar coding of priority (importance to be processed), activity (number of units active in module) and value (hedonic value of the information in the module scoring if good or bad). The modules are organized in modality-specific hierarchies with all modalities feeding into an inner loop of cross modal communication (see Schneider1993). The representation system hardware evolved in reptilian brains 300 million years ago and represents most (>90%) of cortical space.

Triarchic Control
Meta –Cognitive Net
Cognitive Control Net
Representation System

Figure 1. Three interrelated classes of computation.

Neuroimaging studies of cognitive control in humans suggest the existence of a domain general Cognitive Control Network (CCN: Figure 3) in dorsal frontal and parietal cortices (e.g., Cole & Schneider, 2007), which monitors and controls the representation systems. Functional neuroimaging studies indicate that this network performs multiple operations, which include: monitoring scalar module reports of priority, activity, and value;
selection of representation modules control module feedback/categorization and output transmission gain, and initiating module learning of associative and scalar coding of information. The brain areas and functions of the putative CCN include dorsolateral prefrontal cortex (DLPFC, sequencing); anterior cingulate cortex / presupplementary motor area (ACC/pSMA, monitoring/decision); inferior frontal junction (IFJ, task loading); anterior insular cortex (AIC, autonomic arousal); dorsal pre-motor cortex (dPMC, response release/sequencing); posterior parietal cortex (PPC, selective attention); and thalamus (Thal, scalar signal routing for priority and attention). The CCN evolved during mammalian evolution 60 million years ago and represents a small (<10% of cortex).

**Figure 2. Representation level. A: major division, B: representation brain areas shown in color control/episodic store in grey/white.**

In addition to the RS and CCN, we propose the existence of a Meta-Cognitive Network (MCN: Figure 4), which monitors the inner loop of the RS, alters cognitive control routines, and creates new control routines. As with the CCN, evidence for these meta-cognitive functions precedes from human functional neuroimaging studies which demonstrate that MCN brain areas are engaged during early stages of task learning (Cole et al., in prep.) and in semantic retrieval and selection (Badre et al., 2005). The brain areas involved are in the anterior prefrontal cortex and middle temporal lobe, which expanded dramatically (3-7x) in the human brain 160,000 years ago and represent a small portion (~1%) of cortex.

**Figure 3. Domain General Cognitive Control Network. A: major divisions, B: brain areas CCN shown in red.**

attending/increasing gain of the output vector of a given module, comparing with the activity of a second module, and--depending on the degree of activity--releasing output from a third module. These comparative operations are not done on symbol codes devoid of meaning (as is typical of symbolic machines), but on vectors incorporating the robustness of the representation experience.

Finally, in our model the MCN is on top of the CCN, monitoring the CCN behavior, and modifying and creating CCN routines (e.g., if a CCN routine is running too long without improved result, switch to a second CCN routine). It is worth noting that in most complex domains the human
brain is the most effective learning/computing engine known, even though it operates on hardware that is $10^{12}$ times slower than current digital devices. Of course our brains operate with high parallel computation.

An important characteristic of the triarchic computing architecture that we propose is how it avoids the homunculus trap by reducing and compacting information at each level of the computing system. How does a single domain general control system not become overloaded by the information from approximately 500 cortical regions, containing 10,000 modules, each with a million neurons?

Figure 4. Meta Cognitive Network. A: major components, B: MCN brain area in tan.

In our model, the CCN performs data ranking via scalar coding of representation system activity to dramatically reduce the complexity of data the CCN works on, relative to the messages in the representation system. For example, in determining which modules should be attended, the CCN cannot look at the representations within the modules, because doing so would involve a great deal of data ($10^4$ modules x $10^3$ long vectors) with no scheme for ranking messages.

We propose that the CCN works on the scalar report signals, and can implement attentive selection by controlling the scalar gain of the module (see Chein & Schneider 2003). Each module of the representation system sends scalars for its activity and priority to the thalamus, and a third scalar for “goodness” (i.e., hedonic value) to the amygdala. The CCN has the comparatively simple task of identifying the module with the highest report value. This can be accomplished by a simple winner-take-all network or peak-finding routine. The CCN is domain-general: it can examine report scalars from all the modules with simple computation (e.g., find the cluster of modules with the highest priority). Search engines like Google use an analogous ranking scheme, a scalar PageRank (0-10). The search engine does not analyze the content of the pages, instead it rank-orders the pages for listing based on the PageRank. This ranking system is domain-general, operating on all types of content.

An example CCN routine, based on classical cognitive psychology experiments, would be a visual search to find words from one of two categories in a display of words (e.g., Fisk & Schneider 1983).

- Goal: search display for the category “animals” or “vehicles.”
- Load working memory with the set of target categories: “animals”, “vehicles.”
  - Load the first item into verbal working memory.
- Search through all visual words that are active in order of priority and determine if the attended item increases the activity of the searched-for category module.
- If activity is increased above criterion, respond.
- If all display items have been searched, change to the next target category held in verbal working memory.
- If all target categories have been exhausted, terminate search.

Note that the CCN is doing this by changing gains and monitoring a scalar activity report. The CCN does not process the content of vectors from the representation system, only the scalar activity report. Processing is slow and serial (e.g., 200 ms per category comparison, based on behavioral data). The observer has minimal awareness of objects that did not match the target category (e.g., cannot recognize a distracter word even if it had been semantically compared 20 times (see Fisk & Schneider 1983)). After the CCN has found a likely target, the MCN could then verify the identity of the target display item and provide a verbal report of it.

The CCN is thus “message-blind”: it is not able to examine the specific content of the message, and instead works with scalar rankings of priority, activity, and goodness. In a second example, let us suppose that the CCN uses this scalar information to satisfy a general goal, such as “find good food.” The CCN could alter processing in the multimodal representation system until gustatory modules have the highest priority, activity, and goodness report scalars. This would enable a search for the broad stimulus category of “food,” with gustatory modules coding the target class of representations.

In contrast to the CCN, which is presumed to be message blind, the MCN can monitor traffic in the inner loop of the representation system, allowing it to “sniff” specific codes and alter processing to evoke or avoid it. In our food example, the MCN might enable search for a specific sensory representation (e.g., if “chocolate covered
almonds” are the target food, then search for the message \(<\text{cca}\>\).

As this example indicates, the MCN performs two major processing functions in our model. The first is meta-control of the CCN. It monitors/alters CCN processing (e.g., it changes CCN routines if current processing has not brought about the goal state). Note that this is different from CCN monitoring of the match process itself (e.g., monitoring the metric of match in a comparison operation in memory). The MCN is assumed to get input from the CCN. It can decide to terminate the routine (e.g., if the search is too long without match), repeat the search, or modify the parameters of the search (e.g., the speed of comparison or criterion for match).

The second role of the MCN is message monitoring/injection. We assume that MCN has a representation inner loop monitoring port in which it can inspect messages traveling on the inner loop of the representation system to identify the presence of a specific representation code. This is analogous to packet analyzers (http://en.wikipedia.org/wiki/Packet_analyzer) or internet sniffers for monitoring internet traffic passing through a node to a specific packet address. To avoid the homunculus trap, the MCN must limit search to a small subset of the codes in the representation system. The hierarchical arrangement of the representation system facilitates this limited search (Figure 2). We assume that the MCN can look only at reduced vectors of modules on the top, inner loop layer of the representation systems, with the limited goal of detecting whether a particular message was present or injecting a specific message in the inner loop. Monitoring only the inner loop versus the full representation system reduces the monitoring to only a small number (~10) of regions. Using the scalar codes, the highest-ranking region (based on priority, activity, or goodness) can be loaded into a single item cache of the vector (like a sniffer having a specific packet address it is watching for).

Note that the MCN does not “comprehend” vectors traveling in the inner loop of the representation system, as the vector is a symbolic code. Rather, the MCN has the ability to detect if a target representation was transmitted on the inner loop of the representation system. This detection process is based on neuronal activity patterns, which act like bits in symbol coding. For example, the MCN could detect whether the \(<\text{cca}\>\) code (“chocolate covered almonds”) appeared on the inner loop, without processing all of the perceptual and semantic features associated with that food. The “meaning” of the vector is assumed to reside in the representation space. Although the MCN and CCN could query the representation space to derive detailed perceptual and semantic feature information, neither the MCN nor the CCN directly code that information.

In this model, the MCN thus alters the CCN, which in turn alters the representation system, to seek a particular code on the inner loop. This interactive search might be analogous to using a search engine to find a particular string of characters—the text query—and rank-ordering them by PageRank. The human performing the search plays the role of the MCN in this analogy. It should be noted that the MCN expanded dramatically in humans relative to other primates (Semendeferi 2001), perhaps enabling complex message-aware sub-routing to reach a specific goal. For example, to get “chocolate covered almonds” \(<\text{cca}\>\) the MCN could put a seek routine in the CCN, put a target code in the monitoring port, and project the target code into the representation system. In pseudo-code terms this process could be summarized:

- Goal: get \(<\text{cca}\>\)
  - Load \(<\text{cca}\>\) into inner loop monitoring port
  - If \(<\text{cca}\>\) activates a recent personal episodic trace, reanimate that event.
    - Identify the location message of that event.
    - If it can go to the location easily, go there.
  - Else-If some location module is activated (e.g., grocery store),
    - If not at location
      - Determine constraints of getting object from there
        (are the conditions met to have money, transportation, route, ability/willingness of the store to sell the object?).
      - Go to that location.
    - Else If at a location of \(<\text{cca}\>\),
      - Activate search for \(<\text{cca}\>\) at that location.
      - Execute routine to acquire “cca.”

### Brain Connectivity Architecture: Expectations / Match to Triarchic Architecture

We have laid out a rich set of information processing assumptions for the triarchic theory. If processing is done on priority scalars, we expect them to be carried on a small number of fibers relative to the module message vectors (e.g., 100 fibers versus 10,000). Recent High Definition Fiber tracking Methods (HDFT) allow MRI based non-invasive diffusion imaging of small fiber bundles from source to destination through crossings (Hagmann et. al., 2006). In the last three months our group has seen dramatic improvement of the precision of such tracking. We now feel we can quantify the connectivity between cortical areas and from cortical to subcortical areas, enabling us to test the predictions described above.

We can use fMRI brain imaging methods to map the CCN (Cole & Schneider 2007), MCN (Cole & Schneider, in preparation), and representation areas [e.g., map visual system (Wandell et. al., 2007)]. Additional work on the nature of the semantic coding in the representation system (Goldberg et. al., 2006, 2007) allows us to specifically identify the inner-loop modules.
Together, these functional imaging methods permit identification of the component parts of the three systems. We believe that these three systems evolved sequentially, so we can utilize the evolutionary development of the brain as an additional guide for localization of their components. For example, we note that the representation system evolved in the reptilian brain and includes major subcortical and limbic brain structures. As mammals, we have the additional presence of the CCN in dorsal neocortical areas. We note that in the evolution of the human brain there was an explosive expansion (3-7x expansion over 160,000 years of evolution) in two areas (anterior PFC area 10, Semendeferi et al. 2001; temporal areas related to language, Rilling et al. 2002). By combining the data based on functional activity and evolution, we can thus tentatively identify the brain locations of the three systems.

Critically, we can then use HDFT to test the connective topology of these systems. HDFT measures the size of the diffusion pipe between areas; thus, if we have a diffusion pipe that is 1mm in diameter, we may estimate a connection of approximately 40,000 axons between regions. The width of these connections varies dramatically between areas. Assuming that higher bandwidth is required to carry vector than scalar information between regions, we view large differences in diffusion pipe width (e.g., 50 times or more) as suggestive of scalar vs. vector bandwidth connections. Although the thalamus is connected to many parts of cortex, those connections are thin, suggesting scalar communication. In contrast, a typical connection from cortex to cortex (e.g., V4 to V2) is orders of magnitude higher. For example, connections from the fusiform face area (FFA) to V4 are 76 times wider than connections between FFA and thalamus (Figure 5). This connectivity pattern is consistent with the transmission of high-bandwidth, feature-rich vectors between representation system modules.

![Figure 5. Thalamus projections to cortex, connecting to most areas with thin fibers. Consistent with interpretation of scalar processing.](image)

There is also widely-varying connectivity between CCN and MCN. One structure that is particularly interesting is the claustrum. It connects to the inner loop representation areas, to the anterior frontal cortex (BA10), and to the temporal language areas. The claustrum has high connectivity to the brain association areas and connects to all of those areas. This suggests that the claustrum could be the inner loop monitoring port between the representation areas and the MCN.

We have provided an overview of a triarchic theory of cognitive processing, in which each system performs different roles and has different input/output patterns which support those functions. The HDFT data provide quantification of connectivity and patterning of connection distributions to test the predictions of the theory and suggest the information roles of the processing systems. The current resolution of HDFT is sufficient to begin to quantify connectivity between brain regions. These connectivity data, in turn, will allow us to test theories of cortical specialization and synergy.

References


