A Standard Model of Mind:
AAAI Technical Report FS-17-05

Towards A Unified Theory of Mind and Brain

Stephen Grossberg
Center for Adaptive Systems, 677 Beacon Street, Boston University, Boston, MA 02215
publications17@aaai.org

Abstract

A unified theory of how brains give rise to minds has been getting steadily developed for 60 years. This theory has explained data from thousands of interdisciplinary experiments, and scores of its predictions have been confirmed. The theory includes new computational paradigms called Complementary Computing and Laminar Computing that clarify the global organization of brain dynamics and behavior. It uses a small number of basic equations and modules to form modal architectures that model different modalities of intelligence. One of its models, Adaptive Resonance Theory, or ART, is currently the most advanced cognitive and neural theory of how advanced brains incrementally and stably learn to attend, recognize, and predict objects and events in a changing world. All the basic ART predictions have been confirmed by psychological and neural data. These results provide a firm foundation for further development of a Standard Theory of the Mind.

Unified Theory of Mind and Brain

Major design principles, mechanisms, and architectures have been discovered and developed during the past 60 years as part of an emerging unified theory of biological intelligence, notably how brain mechanisms give rise to mental functions as emergent properties. As of this writing, thousands of psychological and neurobiological experiments have been explained and predicted in a unified way, including data about perception, cognition, cognitive-emotional dynamics, and action in both normal individuals and clinical patients. Key articles may be downloaded at htty://cns.bu.edu/~steve. These results include a systematic analysis of what is happening in individual brains when they consciously see, hear, feel, or know something, and how we can experience integrated moments of seeing, hearing, feeling, and knowing (Grossberg, 2017).

Brain paradigms: Complementary Computing and Laminar Computing

The possibility of this synthesis is predicated upon the discovery that advanced brains embody novel computational paradigms in order to achieve biological intelligence: Complementary Computing and Laminar Computing. Complementary Computing. Grossberg (2000) describes how the brain is organized into complementary parallel processing streams whose interactions generate biologically intelligent behaviors. A single cortical processing stream can individually compute some properties well, but cannot, by itself, process other computationally complementary properties. Pairs of complementary cortical processing streams interact to generate emergent properties that overcome their complementary deficiencies to compute complete information with which to represent or control some faculty of intelligent behavior. Thus, although brain anatomy embodies a great deal of functional specialization, there are no "independent modules" in advanced brains. Complementary Computing clarifies how different brain regions can achieve a great deal of specialization without being independent modules.

The WHAT and WHERE Cortical Streams are Complementary

For example, the category learning, attention, recognition, and prediction circuits of the ventral, or What, cortical processing stream for perception and cognition are computationally complementary to those of the dorsal, or Where and How, cortical processing steam for spatial representation and action (Grossberg, 2000, 2013, 2017). One reason for What-Where complementarity is that the What stream learns object recognition categories that are substantially invariant under changes in an object's view, size, and position. These invariant object categories enable our brains to recognize valued objects without experiencing a combinatorial explosion. They cannot, however, locate and act upon a desired object in space. Where stream spatial and motor representations can locate objects and trigger actions towards them, but cannot recognize them. What stream dynamics embody many aspects of declarative learning and memory, whereas Where stream dynamics realize procedural learning and memory. By interacting together, the

What and Where streams can recognize valued objects and direct appropriate goal-oriented actions towards them.

Adaptive Resonance Theory

Abundant psychological and neurobiological data have confirmed all of my foundational predictions concerning how perceptual/cognitive processes in the What stream use excitatory matching and match-based learning to create self-stabilizing categorical representations of objects and events, notably recognition categories that can be learned quickly without experiencing catastrophic forgetting during subsequent learning. In other words, this recognition learning process solves the so-called *stability-plasticity dilemma*. These processes enable increasing expertise, and an ever-expanding sense of self, to emerge throughout life. Excitatory matching by object attention is embodied by the ART Matching Rule. This type of attentional circuit enables us to prime our expectations to anticipate objects and events before they occur, and to focus attention upon expected objects and events when they do occur. Good enough matches between expected and actual events trigger resonant states that can support learning of new recognition categories and refinement of old ones, while also triggering conscious recognition of the critical feature patterns that are attended as part of these percepts. Excitatory matching also controls reset of the attentional focus when bottom-up inputs significantly mismatch currently active top-down expectations. Cycles of resonance and reset underlie much of the brain's perceptual and cognitive dynamics.

These matching and learning laws have been articulated as part of Adaptive Resonance Theory, or ART, which has been progressively developed since it was first reported in 1976. ART is a cognitive and neural theory of how the brain autonomously learns to attend, recognize, and predict objects and events in a changing world. ART is currently the most highly developed cognitive and neural theory available, with the broadest explanatory and predictive range. Central to ART's predictive power is its ability to carry out fast, incremental, and stable unsupervised and supervised learning in response to a changing world. ART specifies mechanistic links between processes of consciousness, learning, expectation, attention, resonance, and synchrony (the CLEARS processes) during both unsupervised and supervised learning. I have predicted that all brains that can solve the stability-plasticity dilemma do so using these predicted links between CLEARS processes. Indeed, my 40-year old prediction that "all conscious states are resonant states" is consistent with all the data that I know, and has helped to explain many data about consciousness, as will be briefly noted below.

ART hereby contributes to functional and mechanistic explanations of such diverse topics as 3D vision and figure-ground perception in natural scenes; optic-flow based navigation in natural scenes towards goals around obstacles and spatial navigation in the dark; invariant object and scenic gist learning, recognition, and search; prototype, surface, and boundary attention; gamma and beta oscillations during cognitive dynamics; learning of entorhinal grid cells and hippocampal place cells, including the use of homologous spatial and temporal mechanisms in the medial entorhinal-hippocampal system for spatial navigation and the lateral stream for adaptively timed cognitive-emotional learning; breakdowns in attentive vigilance during autism, medial temporal amnesia, and Alzheimer's disease; social cognitive abilities such as the learning of joint attention and the use of tools from a teacher, despite the different coordinate systems of the teacher and learner; a unified circuit design for all item-order-rank working memories that enable stable learning of recognition categories, plans, and expectations for the representation and control of sequences of linguistic, spatial, and motor information; conscious speech percepts that are influenced by future context; auditory streaming in noise during source segregation; and speaker normalization that enables language learning from adults after a critical period of babbled sounds by a child; cognitive-emotional dynamics that direct motivated attention towards valued goals; and adaptive sensory-motor control circuits, such as those that coordinate predictive smooth pursuit and saccadic eye movements, and coordinate looking and reaching movements. Brain regions that are functionally described include visual and auditory neocortex; specific and nonspecific thalamic nuclei; inferotemporal, parietal, prefrontal, entorhinal, hippocampal, parahippocampal, perirhinal, and motor cortices; frontal eye fields; supplementary eye fields; amygdala; basal ganglia: cerebellum; superior colliculus; and reticular formation.

These results include many important particulars. For example, all working memories obey an LTM Invariance Principle that enables them to temporarily store sequences of events that can be stably chunked and remembered. This fact implies that all linguistic, spatial, and motor working memories are variants of the same network design. Properties like bounded rationality readily follow from this shared working memory design.

ART also clarifies how so-called attentional bottlenecks may arise from basic constraints on how object learning is regulated and dynamically stabilized by interactions that use spatial and object attention.

ART does not, however, describe many spatial and motor processes and their behaviors, which employ different matching and learning laws. ART is thus not "a theory of everything".

Vector Associative Maps for Spatial Representation and Action

Complementary spatial/motor processes in the Where stream often use inhibitory matching and mismatch-based learning to continually update spatial maps and sensory-motor gains in the trajectory formation processes that control our changing bodies throughout life. Inhibitory matching via a Difference Vector occurs between a representation of where we want to move—a Target Position Vector—and where we are now—a Present Position Vector. When we arrive at where we want to be, the match—viz. the Difference Vector—equals zero. Inhibitory matching by this Vector Associative Map, or VAM, Matching Rule thus cannot solve the stability-plasticity dilemma. That is why spatial and motor representations, notably procedural memories, cannot support conscious qualia. Instead, they experience catastrophic forgetting as they learn how to accurately control our changing bodies throughout life.

Together these complementary processes create a self-stabilizing perceptual/cognitive front end in the What stream for learning about the world, gaining increasing expertise along the way, and becoming conscious of it, while it intelligently commands more labile spatial/motor processes in the Where stream that control our changing bodies. Such a complementary synthesis offers a view towards developing a Standard Model of the Mind.

Homologous Laminar Cortical Circuits for All Biological Intelligence

The second computational paradigm is called Laminar Computing. Laminar Computing further contributes to a Standard Model of the Mind since it describes how the cerebral cortex is organized into layered circuits whose specializations support all higher-order biological intelligence, including but not restricted to cognition. Indeed, the laminar circuits of cerebral cortex seem to realize a revolutionary computational synthesis of the best properties of feedforward and feedback processing, digital and analog processing, and data-driven bottom-up processing and hypothesis-driven top-down processing (Grossberg, 2007, 2013). For example, ART mechanisms have, to the present, been naturally embodied in laminar cortical models of vision, speech, and cognition, specifically the 3D LAMINART model of 3D vision and figure-ground separation, the cARTWORD model of speech perception, and the LIST PARSE model of cognitive working memory and chunking. Each model uses variations of the same canonical laminar cortical circuitry, thereby clarifying how specialized resonances use similar types of circuits to support different conscious experiences, and providing an existence proof that other kinds of biological intelligence can also be modeled using variations of this canonical laminar circuit.

Why a Unified Theory is Possible: Shared Equations, Modules, and Architectures

There is another fundamental reason why it is possible for human scientists to discover a unified mind-brain theory that links brain mechanisms and psychological functions, and to demonstrate how similar organization principles and mechanisms, suitably specialized, can support conscious qualia across modalities.

This reason is that a small number of equations suffice to model all modalities. These include equations for short-term memory, or STM; medium-term memory, or MTM; and long-term memory, or LTM, that I first published in PNAS in 1968. See Grossberg (2013) for a review. These equations are used to define a somewhat larger number of modules, or microcircuits, that are used in multiple modalities where they can carry out different functions within each modality. These modules include shunting on-center off-surround networks, gated dipole opponent processing networks, associative learning networks, spectral adaptively timed learning networks, trajectory formation networks, and the like. Each module and its specializations exhibits a rich, but not universal, set of useful computational properties.

For example, shunting on-center off-surround networks can carry out properties like contrast normalization, including discounting the illuminant; contrast enhancement, noise suppression, and winner-take-all choice from multiple parallel alternatives; short-term memory and working memory storage of event sequences; attentive matching of bottom-up input patterns and top-down learned expectations; and synchronous oscillations and traveling waves.

These equations and modules are specialized and assembled into modal architectures, where “modal” stands for different modalities of biological intelligence, including architectures for vision, audition, cognition, cognitive-emotional interactions, and sensory-motor control.

An integrated self is possible because it builds on a shared set of equations and modules within modal architectures that can interact seamlessly together.

Modal architectures are *general-purpose*, in that they can process any kind of inputs to that modality, whether from the external world or from other modal architectures. They are also *self-organizing*, in that they can autonomously develop and learn in response to these inputs. Modal architectures are thus less general than the von Neumann architecture that provides the mathematical foundation of modern computers, but much more general than a traditional AI algorithm. ART networks form part of several different modal architectures, including modal architectures that enable seeing, hearing, feeling, and knowing.

All Conscious States are Resonant States

ART has predicted that "all conscious states are resonant states", but the converse is not true. I know no data that disconfirm this 40 year old prediction. ART resonances clarify questions such as the following, which have been raised by distinguished philosophers (cf., Grossberg, 2017):

What kind of "event" occurs in the brain during a conscious experience that is anything more than just a "whir of information-processing"? What happens when conscious mental states "light up" and directly appear to the subject? ART explains that, over and above "just" information processing, our brains sometimes go into a context-sensitive *resonant state* that can involve multiple brain regions. Abundant experimental evidence supports the ART prediction that "all conscious states are resonant states". Not all brain dynamics are "resonant", and thus consciousness is not just a "whir of information-processing".

Second, when does a resonant state embody a conscious experience? And how do different resonant states support different kinds of conscious qualia? The other side of the coin is equally important: When does a resonant state fail to embody a conscious experience? ART explains (Grossberg, 2017) how various evolutionary challenges that advanced brains face in order to adapt to changing environments in real time have been met with particular conscious states, which form part of larger adaptive behavioral capabilities. ART explains that humans are not conscious just to Platonically contemplate the beauty of the world. Humans are conscious in order to enable them to better adapt to the world's changing demands. To illustrate these claims, ART has proposed how the brain generates resonances that support particular conscious experiences of seeing, hearing, feeling, and knowing. In so doing, it suggests how resonances for conscious seeing help to ensure effective reaching, resonances for conscious hearing help to ensure effective speaking, and resonances for conscious feeling help to ensure effective goal-directed action.

The Varieties of Brain Resonances and the Conscious Experiences That They Support

Towards this end, ART has explained six different types of neural representations of conscious qualia, and has provided enough theoretical background and data explanations based on these representations to illustrate their explanatory and predictive power. The theory's predictions also suggest multiple kinds of interdisciplinary experiments to deepen our mechanistic understanding of the brain mechanisms for generating conscious resonances.

For example, *surface-shroud resonances* are predicted to support conscious percepts of visual qualia. *Feature-category resonances* are predicted to support conscious recognition of visual objects and scenes. Both kinds of resonances may synchronize during conscious seeing and recognition. *Stream-shroud resonances* are predicted to support conscious percepts of auditory qualia. *Spectral-pitch-and-timbre resonances* are predicted to support conscious recognition of sources in auditory streams. Stream-shroud and spectral-pitch-and-timbre resonances may synchronize during conscious hearing and recognition of auditory streams. *Item-list resonances* are predicted to support recognition of speech and language. They may synchronize with stream-shroud and spectral-pitch-and-timbre resonances during conscious hearing of speech and language, and build upon the selection of auditory sources by spectral-pitch-and-timbre resonances in order to recognize the acoustical signals that are grouped together within these streams. *Cognitive-emotional resonances* are predicted to support conscious percepts of feelings, as well as recognition of the source of these feelings. Cognitive-emotional resonances can also synchronize with resonances that support conscious qualia and knowledge about them. All of these resonances have distinct anatomical substrates.

Why Does Resonance Trigger Consciousness?

Detailed analyses of psychological and neurobiological data by ART clarify why resonance is necessary for consciousness. For example, in order to fully compute visual boundaries and surfaces whereby to see the world, the brain computes three pairs of complementary computational properties of boundaries and surfaces, along with the three hierarchical resolutions of uncertainty that require multiple processing stages to overcome. There is thus a great deal of uncertainty in the early stages of visual processing by the brain. Only after all three hierarchical resolutions of uncertainty are complete, and after boundaries are completed and surfaces filled-in, has the brain constructed a contextually informative and temporally stable enough representation of scenic objects on which to base adaptive behaviors.

If this is indeed the case, then why do not the earlier stages undermine behavior? The proposed answer is that *brain resonance, and with it conscious awareness, is triggered at the processing stage that represents boundary and surface representations, after they are complete and stable enough to control visually-based behaviors like attentive looking and reaching*. Consciousness supplies an "extra degree of freedom" to mark these behaviorally predictive representations. ART also explains that object attention obeys what is called the ART Matching Rule, which can operate both top-down and bottom-up to select the information at other processing stages that is consistent with the triggering resonance and to suppress all inconsistent information.

Towards a Standard Model of Mind

The above summary suggests some of the fundamental design principles, mechanisms, and architectures that advanced brains use to create a mind. The summary also notes that these principles, mechanisms, and architectures have been embodied in self-organizing neural network models, which have been used to explain and predict an unrivalled range of mind and brain data. These discoveries thus provide a good foundation for developing a Standard Model of Mind.

All of the models described above have also been applied to large-scale problems in engineering and technology where self-organizing, adaptive, autonomous agents are desired, including mobile agents. A partial list of tech transfers can be found at the CNS Technology Lab website: http://techlab.bu.edu/resources/articles/C5.

References

Grossberg, S. (2000). The complementary brain: Unifying brain dynamics and modularity. *Trends in Cognitive Sciences*, *4*, 233-246.

Grossberg, S. (2007). Consciousness CLEARS the mind. *Neural Networks*, 20, 1040-1053.

Grossberg, S. (2013). Adaptive Resonance Theory: How a brain learns to consciously attend, learn, and recognize a changing world. *Neural Networks, 37*, 1-47.

Grossberg, S. (2017). Towards solving the Hard Problem of Consciousness: The varieties of brain resonances and the conscious experiences that they support. *Neural Networks*, 87, 38-95.