Explainable and Reliable AI and Autonomous Adaptive Intelligence: Deep Learning, Adaptive Resonance, and Models of Perception, Emotion, and Action

Stephen Grossberg

Center for Adaptive Systems
Graduate Program in Cognitive and Neural Systems
Departments of Mathematics & Statistics, Psychological & Brain Sciences, and Biomedical Engineering
Boston University

steve@bu.edu
sites.bu.edu/steveg
This lecture is based on the following article:

Frontiers in Neurobotics, June 25, 2020
https://doi.org/10.3389/fnbot.2020.00036 (OPEN ACCESS)

The article summarizes core problems of DEEP LEARNING, such as its untrustworthiness (unexplainable) and unreliability (catastrophic forgetting),

explains how ADAPTIVE RESONANCE overcomes them, indeed overcomes 17 problems of Deep Learning,

and outlines a blueprint for achieving autonomous adaptive intelligence
“Though Deep Learning is the main pillar of current AI techniques and is ubiquitous in basic science and real-world applications, it is also flagged by AI researchers for its black-box problem: it is easy to fool, and it also cannot explain how it makes a prediction or decision”

Deep Learning is NOT TRUSTWORTHY

No life or death decision, such as a medical or financial decision, can confidently be made based upon a Deep Learning prediction
Deep Learning uses the back propagation (BP) algorithm to learn how to predict output vectors in response to input vectors.

BP was based upon perceptron learning principles
Rosenblatt (1958, 1987)

It has a complicated history; cf., Schmidhuber (2020)

Major contributors include:

BP reached its modern form with simulated applications in
Werbos (1974)

It was popularized by
Rumelhart, Hinton, and Williams (1986)
BACK PROPAGATION CIRCUIT

figure reprinted from Carpenter (1989)

Information flows FEEDFORWARD
BACK PROPAGATION CIRCUIT
figure reprinted from Carpenter (1989)

Information flows FEEDFORWARD

Learning is SUPERVISED
An external teacher on each trial
BACK PROPAGATION CIRCUIT
figure reprinted from Carpenter (1989)

Information flows FEEDFORWARD

Learning is SUPERVISED
An external teacher on each trial

Teaching signal is the ERROR or MISMATCH between ACTUAL and TARGET outputs

\[
\delta_k = f'(x_k) (b_k - S_k)
\]

\[
\delta_j = f'(x_j) \sum_k w_{jk} \delta_k
\]

\[
x_k = \sum_j S_j w_{jk} + \theta_k
\]

\[
x_j = \sum_i S_i w_{ij} + \theta_j
\]
Information flows **FEEDFORWARD**

Learning is **SUPERVISED**

An external teacher on each trial

Teaching signal is the **ERROR** or **MISMATCH** between **ACTUAL** and **TARGET** outputs

Teaching signal in level $F_3$ of adaptive weights $w_{ij}$ in level $F_2$ have no network pathway whereby to reach $F_2$

Use **WEIGHT TRANSPORT** of the teaching signal!

**NON-LOCAL!**

**NON-BIOLOGICAL!**
SLOW LEARNING
Adaptive weights change just a little to reduce error on each learning trial

REQUIRES MANY TRIALS (i.e., repetitions of database) to learn, possibly hundreds or thousands of trials
SLOW LEARNING
Adaptive weights change just a little to reduce error on each learning trial

REQUIRES MANY TRIALS (i.e., repetitions of database) to learn, possibly hundreds or thousands of trials

CONTRAST FAST LEARNING
Adaptive weights zero error signals on EACH trial

Cf. learn a face that you see just once, and remember it for a long time
CATASTROPHIC FORGETTING
During any learning trial, an unpredictable part of its learned memory can collapse
Ratcliff (1990), French (1999)
CATASTROPHIC FORGETTING
During any learning trial, an unpredictable part of its learned memory can collapse
Ratcliff (1990), French (1999)

Deep Learning is thus neither RELIABLE nor TRUSTWORTHY
CATASTROPHIC FORGETTING
During any learning trial, an unpredictably part of its learned memory can collapse
Ratcliff (1990), French (1999)

WHY?
All inputs are processed by a shared set of learned weights

It cannot selectively buffer learned weights that are still predictively useful (no attention)

This problem occurs in ANY learning algorithm whose shared weight updates follow the gradient of the error in response to the current batch of data points, while ignoring past batches
MULTIPLE EFFORTS TO FIX BACK PROPAGATION

figure reprinted from Carpenter (1989)

Selectively slow learning
“on the weights important for...supervised learning and reinforcement learning problems...by optimizing...parameters...using Bayes’ rule”
Kirkpatrick et al (2017)

Assumes:
omniscient observer who can discover and alter “important weights”

non-local computations e.g., Bayesian computation

Same problems with evolutionary algorithms
Clune et al (2013)

and diffusion-based neuromodulation
Velez & Clune (2017)
These efforts to overcome catastrophic forgetting created additional conceptual and computational problems. I view them as adding **EPICYCLES** to ameliorate a fundamental flaw in the model. Reminiscent of adding epicycles to correct problems in the Ptolemaic model of the solar system. The Copernican model that we now accept did not require epicycles!
Perhaps this is why Geoffrey Hinton said in Axios (LeVine, 2017) that he is “deeply suspicious of back propagation... I don’t think it’s how the brain works. We clearly don’t need all the labeled data... My view is, throw it all away and start over”
Perhaps this is why Geoffrey Hinton said in AXIOS (LeVine, 2017) that he is “deeply suspicious of back propagation… I don’t think it’s how the brain works. We clearly don’t need all the labeled data… My view is, throw it all away and start over”

We do not have to start over!

These problems were solved in the 1970s and 1980s!
17 PROBLEMS OF BACK PROPAGATION OVERCOME BY ADAPTIVE RESONANCE
Grossberg (1988, Neural Networks, 1, 17-41)

- Real-time (on-line) learning vs. lab-time (off-line) learning
- Learning in nonstationary unexpected world vs. in stationary controlled world
- Self-organized unsupervised or supervised learning vs. supervised learning
- Dynamically self-stabilize learning to arbitrarily many inputs vs. catastrophic forgetting
- Maintain plasticity forever vs. externally shut off learning when database gets too large
- Effective learning of arbitrary databases vs. statistical restrictions on learnable data
- Learn internal expectations vs. impose external cost functions
- Actively focus attention to selectively learn critical features vs. passive weight change
- Closing vs. opening the feedback loop between fast signaling and slower learning
- Top-down priming and selective processing vs. activation of all memory resources
- Match learning vs. mismatch learning: Avoiding the noise catastrophe
- Fast and slow learning vs. only slow learning: Avoiding the oscillation catastrophe
- Learning guided by hypothesis testing and memory search vs. passive weight change
- Direct access to globally best match vs. local minima
- Asynchronous learning vs. fixed duration learning: A cost of unstable slow learning
- Autonomous vigilance control vs. unchanging sensitivity during learning
- General-purpose self-organizing production system vs. passive adaptive filter
Perhaps this is why Geoffrey Hinton said in Axios (LeVine, 2017) that he is

“deeply suspicious of back propagation… I don’t think it’s how the brain works. We clearly don’t need all the labeled data… My view is, throw it all away and start over”

We do not have to start over!

These problems were solved in the 1970s and 1980s!
17 PROBLEMS OF BACK PROPAGATION OVERCOME BY ADAPTIVE RESONANCE
Grossberg (1988, Neural Networks, 1, 17-41)

• Real-time (on-line) learning vs. lab-time (off-line) learning
• Learning in nonstationary unexpected world vs. in stationary controlled world
• Self-organized unsupervised or supervised learning vs. supervised learning
• Dynamically self-stabilize learning to arbitrarily many inputs vs. catastrophic forgetting
• Maintain plasticity forever vs. externally shut off learning when database gets too large
• Effective learning of arbitrary databases vs. statistical restrictions on learnable data
• Learn internal expectations vs. impose external cost functions
• Actively focus attention to selectively learn critical features vs. passive weight change
• Closing vs. opening the feedback loop between fast signaling and slower learning
• Top-down priming and selective processing vs. activation of all memory resources
• Match learning vs. mismatch learning: Avoiding the noise catastrophe
• Fast and slow learning vs. only slow learning: Avoiding the oscillation catastrophe
• Learning guided by hypothesis testing and memory search vs. passive weight change
• Direct access to globally best match vs. local minima
• Asynchronous learning vs. fixed duration learning: A cost of unstable slow learning
• Autonomous vigilance control vs. unchanging sensitivity during learning
• General-purpose self-organizing production system vs. passive adaptive filter
**SLOW LEARNING**
Adaptive weights change just a little to reduce error on each learning trial

**REQUIRES MANY TRIALS**
(i.e., repetitions of database) to learn, possibly hundreds or thousands of trials

**CONTRAST FAST LEARNING**
Adaptive weights zero error signals on EACH trial

Cf. learn a face that you see just once, and remember it for a long time

**VERSUS…**
17 PROBLEMS OF BACK PROPAGATION OVERCOME BY ADAPTIVE RESONANCE
Grossberg (1988, *Neural Networks*, 1, 17-41)

- Real-time (on-line) learning vs. lab-time (off-line) learning
- Learning in nonstationary unexpected world vs. in stationary controlled world
- Self-organized unsupervised or supervised learning vs. supervised learning
- Dynamically self-stabilize learning to arbitrarily many inputs vs. catastrophic forgetting
- Maintain plasticity forever vs. externally shut off learning when database gets too large
- Effective learning of arbitrary databases vs. statistical restrictions on learnable data
- Learn internal expectations vs. impose external cost functions
- Actively focus attention to selectively learn critical features vs. passive weight change
- Closing vs. opening the feedback loop between fast signaling and slower learning
- Top-down priming and selective processing vs. activation of all memory resources
- Match learning vs. mismatch learning: Avoiding the noise catastrophe
- Fast and slow learning vs. only slow learning: Avoiding the oscillation catastrophe
- Learning guided by hypothesis testing and memory search vs. passive weight change
- Direct access to globally best match vs. local minima
- Asynchronous learning vs. fixed duration learning: A cost of unstable slow learning
- Autonomous vigilance control vs. unchanging sensitivity during learning
- General-purpose self-organizing production system vs. passive adaptive filter
ART can learn to classify an entire database using fast learning on a single learning trial

Carpenter and Grossberg (1987, 1988)
ART OVERCOMES ALL 17 PROBLEMS OF BP

without EPICYCLES!

Moreover...
All the core ART predictions have been supported by subsequent psychological and neurobiological data.

ART is a principled biological and technological THEORY.

ART has explained data from hundreds of experiments.

ART has made scores of predictions that have subsequently received experimental support.
All the core ART predictions have been supported by subsequent psychological and neurobiological data.

ART is a principled biological and technological THEORY.

ART has explained data from hundreds of experiments.

ART has made scores of predictions that have subsequently received experimental support.

Why is ART so successful?
ART CAN BE DERIVED FROM A THOUGHT EXPERIMENT ABOUT A UNIVERSAL PROBLEM IN ERROR CORRECTION
Grossberg (1980, Psychological Review, 87, 1-51)

The thought experiment asks the question:

How can a coding error be corrected if no individual cell knows that one has occurred?

“The importance of this issue becomes clear when we realize that erroneous cues can accidentally be incorporated into a code when our interactions with the environment are simple and will only become evident when our environmental expectations become more demanding.

Even if our code perfectly matched a given environment, we would certainly make errors as the environment itself fluctuates”

AUTONOMOUS LOCAL LEARNING IN A CHANGING WORLD
ART CAN BE DERIVED FROM A THOUGHT EXPERIMENT ABOUT A UNIVERSAL PROBLEM IN ERROR CORRECTION

AUTONOMOUS LOCAL LEARNING IN A CHANGING WORLD

A purely logical inquiry into error correction is translated at every step of the thought experiment into processes learning autonomously in real time with only locally computed quantities

The thought experiment uses familiar environmental facts about how we learn as its hypotheses ART circuits naturally emerge

ART circuits may thus, in some form, be embodied in all future autonomous adaptive intelligent devices, whether biological or artificial

ART has, probably for this reason, already been used in many large-scale engineering and technological applications
EARLY ARTMAP BENCHMARK STUDIES

Database benchmark:
MACHINE LEARNING (90-95% correct)
ARTMAP (100% correct on a training set an order of magnitude smaller)

Database benchmarks:
BACKPROPAGATION (10,000 – 20,000 training epochs)
ARTMAP (1-5 epochs)

Medical database:
STATISTICAL METHOD (60% correct)
ARTMAP (96% correct)

Letter recognition database:
GENETIC ALGORITHM (82% correct)
ARTMAP (96% correct)

Used in applications where other algorithms fail
e.g. Boeing CAD Group Technology
Part design reuse and inventory compression
Need fast learning and stable memory to learn and search a huge
(16 million 1 million dimensional vectors) and continually growing
non-stationary parts inventory
ART WORKS!
Large-scale applications in engineering and technology

Boeing parts design retrieval (used to design Boeing 777)
satellite remote sensing
radar identification
robot sensory-motor control and navigation
machine vision
3D object and face recognition
Macintosh operating system software
automatic target recognition
ECG wave recognition
protein secondary structure identification
character classification
musical analysis
air quality monitoring and weather prediction
medical imaging and database analysis
multi-sensor chemical analysis
strength prediction for concrete mixes
signature verification
decision making and intelligent agents
machine condition monitoring and failure forecasting
chemical analysis
electromagnetic and digital circuit design
ART WORKS!
Large-scale applications in engineering and technology

techlab.bu.edu

Boeing parts design retrieval (used to design Boeing 777)
satellite remote sensing
radar identification
robot sensory-motor control and navigation
machine vision
3D object and face recognition
Macintosh operating system software
automatic target recognition
ECG wave recognition
protein secondary structure identification
character classification
musical analysis
air quality monitoring and weather prediction
medical imaging and database analysis
multi-sensor chemical analysis
strength prediction for concrete mixes
signature verification
decision making and intelligent agents
machine condition monitoring and failure forecasting
chemical analysis
electromagnetic and digital circuit design
**AI Expert system – 1 year**

- Field identification of natural regions
- Derivation of ad hoc rules for each region, by expert geographers
- Correct 80,000 of 250,000 site labels

**230m (site–level) scale**

**ARTMAP system – 1 day**

- Rapid, automatic, no natural regions or rules
- Confidence map
- **30m (pixel–level) scale**: can see roads
- Equal accuracy at test sites
Multimodal integration of information from many sources to learn a knowledge structure:

- CONSISTENT
- STABLE
- ROBUST
- LEARNED ONLINE
- SELF-ORGANIZED

Boston testbed

Carpenter et al. (2004)
CONSISTENT KNOWLEDGE FROM INCONSISTENT DATA
Automatically learns and stably stores one-to-many mappings

PROBLEM: Integrate multiple sources into a coherent knowledge structure

Solution 1:
HUMAN MAPPING EXPERT:
Slow, expensive, possibly unavailable

Solution 2:
Distributed ARTMAP MODEL:
Fast, automatic, easy to deploy
NO PRIOR RULES OR DOMAIN KNOWLEDGE

Self-organizing expert system
SELF-ORGANIZES a HIERARCHY of COGNITIVE RULES

Distributed predictions across test set pixels

RULE DISCOVERY

Confidence in each rule = 100%, except where noted

CONSISTENT MAPS, LABELED BY LEVEL

Boston testbed

ocean
beach
park
ice
road
river
water
open space
residential
industrial
built-up
natural
man-made
ART WORKS!
Large-scale applications in engineering and technology

Some more recent work about ART:

Special issue of *Neural Networks* in December, 2019:


BP is a feedforward adaptive filter

ART is more than a feedforward adaptive filter

ART IS AN EXPLAINABLE SELF-ORGANIZING PRODUCTION SYSTEM IN A NON-STATIONARY WORLD
BP is a feedforward adaptive filter
ART is more than a feedforward adaptive filter
ART IS AN EXPLAINABLE SELF-ORGANIZING PRODUCTION SYSTEM IN A NON-STATIONARY WORLD

It is SELF-ORGANIZING because it can autonomously carry out arbitrary combinations of unsupervised or supervised learning trials with the world as its only teacher
BP is a feedforward adaptive filter

ART is more than a feedforward adaptive filter

ART IS AN EXPLAINABLE SELF-ORGANIZING PRODUCTION SYSTEM IN A NON-STATIONARY WORLD

It is a PRODUCTION SYSTEM because it uses HYPOTHESIS TESTING to discover and learn RULES via a top-down matching process that focuses attention on CRITICAL FEATURE PATTERNS that predict behavioral success while suppressing irrelevant features
BP is a feedforward adaptive filter
ART is more than a feedforward adaptive filter

ART IS AN EXPLAINABLE SELF-ORGANIZING PRODUCTION SYSTEM IN A NON-STATIONARY WORLD

It is EXPLAINABLE using both its activities, or short term memory (STM) traces and adaptive weights, or long term memory (LTM) traces:

Observing the STM TRACES in a critical feature pattern explain what recognition categories code and what features predict goal-oriented actions

The LTM TRACES in fuzzy ARTMAP translate into fuzzy IF-THEN rules that code what features, in what numerical ranges, control predictions
ART MECHANISMS THAT DEFINE IT AS AN EXPLAINABLE SELF-ORGANIZING PRODUCTION SYSTEM include:

*Bottom-up* adaptive filter (feedforward neural network) is supplemented by *top-down learned expectations* and *two types of recurrent inhibitory feedback interactions* that help to choose recognition categories and critical feature patterns.

Top-down expectations use the *ART MATCHING RULE* to learn how to **FOCUS ATTENTION** on **CRITICAL FEATURES** that control predictive success.
ART MATCHING RULE for OBJECT ATTENTION stabilizes learning (avoids catastrophic forgetting)

Top-down, modulatory on-center, off-surround network

Carpenter and Grossberg (1987, CVGIP) and many later articles

SG/BI/2020
ART MATCHING RULE for OBJECT ATTENTION

stabilizes learning (avoids catastrophic forgetting)

Top-down, modulatory on-center, off-surround network

Carpenter and Grossberg (1987, CVGIP) and many later articles
LAMINAR CORTICAL CIRCUIT FOR OBJECT ATTENTION
Grossberg (1999, Spatial Vision)

Attention acts via a TOP-DOWN MODULATORY ON-CENTER OFF-SURROUND NETWORK
LAMINAR CORTICAL CIRCUIT FOR OBJECT ATTENTION

Grossberg (1999, Spatial Vision)

Attention acts via a TOP-DOWN MODULATORY ON-CENTER OFF-SURROUND NETWORK

INTRAcortical feedback from groupings

INTERcortical attention

FOLDED FEEDBACK
Illustrates **NEW PARADIGMS** for brain computing

**INDEPENDENT MODULES**
Computer Metaphor

**COMPLEMENTARY COMPUTING**
What is the nature of brain specialization?

**LAMINAR COMPUTING**
Why are all neocortical circuits organized in layers? How do laminar circuits give rise to biological intelligence?
ATTENDED FEATURE CLUSTERS REACTIVATE BOTTOM-UP PATHWAYS

ACTIVATED CATEGORIES REACTIVATE THEIR TOP-DOWN PATHWAYS

FEATURE-CATEGORY RESONANCE SYNCHRONIZES, AMPLIFIES, PROLONGS SYSTEM RESPONSE

RESONANCE TRIGGERS LEARNING IN BOTTOM-UP AND TOP-DOWN ADAPTIVE WEIGHTS: **ADAPTIVE RESONANCE**!
“ALL CONSCIOUS STATES ARE RESONANT STATES”
Grossberg (1980)

Surface-shroud resonances support conscious seeing of visual qualia

Feature-category resonances support conscious recognition of visual objects and scenes

Stream-shroud resonances support conscious hearing of auditory qualia

Spectral-pitch-and-timbre resonances support conscious recognition of sources in auditory streams

Item-list resonances support conscious recognition of speech and language

Cognitive-emotional resonances support conscious feelings and recognition of them
“ALL CONSCIOUS STATES ARE RESONANT STATES”

Surface-shroud resonances support conscious seeing of visual qualia

Feature-category resonances support conscious recognition of visual objects and scenes

Stream-shroud resonances support conscious hearing of auditory qualia

Spectral-pitch-and-timbre resonances support conscious recognition of sources in auditory streams

Item-list resonances support conscious recognition of speech and language

Cognitive-emotional resonances support conscious feelings and recognition of them
SUPPORT FOR ART PREDICTIONS

ATTENTION HAS AN ON-CENTER OFF-SURROUND

Bullier, Jupe, James, and Girard, 1996
Caputo and Guerra, 1998
Downing, 1988
Mounts, 2000
Reynolds, Chelazzi, and Desimone, 1999
Smith, Singh, and Greenlee, 2000
Somers, Dale, Seiffert, and Tootell, 1999
Sillito, Jones, Gerstein, and West, 1994
Steinman, Steinman, and Lehmkuhne, 1995
Vanduffell, Tootell, and Orban, 2000

“BIASED COMPETITION”

Desimone, 1998
Kastner and Ungerleider, 2001
SUPPORT FOR ART PREDICTIONS

ATTENTION CAN FACILITATE MATCHED BOTTOM-UP SIGNALS

Hupe, James, Girard, and Bullier, 1997
Luck, Chellazi, Hillyard, and Desimone, 1997
Roelfsema, Lamme, and Spekreijse, 1998
Sillito, Jones, Gerstein, and West, 1994
and many more…

INCONSISTENT WITH MODELS WHERE TOP-DOWN MATCH IS SUPPRESSIVE

Mumford, 1992
Rao and Ballard, 1999
Bayesian Explaining Away
ART IS EXPLAINABLE (TRUSTWORTHY)

STM: critical feature patterns determine attentional focus that controls information processing

LTM: critical feature patterns determine adaptive weights learned by the BU adaptive filter and TD expectation

Later: fuzzy ARTMAP learns fuzzy IF-THEN rules
Outlier features not in critical feature patterns are suppressed

Only predictive features are processed and coded
BP is a feedforward adaptive filter

ART is more than a feedforward adaptive filter

ART IS AN EXPLAINABLE SELF-ORGANIZING PRODUCTION SYSTEM IN A NON-STATIONARY WORLD

It is a PRODUCTION SYSTEM because it uses HYPOTHESIS TESTING to discover and learn RULES via a top-down matching process that focuses attention on CRITICAL FEATURE PATTERNS that predict behavioral success while suppressing irrelevant features
ART 1 MODEL

ATTENTIONAL SYSTEM

STM F₂

ORIENTING SYSTEM

Reset and Search

Matching criterion: vigilance parameter

Carpenter and Grossberg (1987)

Nonspecific inhibitory gain control

INPUT

SG/BI/2020
ART HYPOTHESIS TESTING AND LEARNING CYCLE

VIGILANCE
How big a mismatch causes reset?

Mismatch
Reset:
Novelty-
Sensitive
Arousal
Burst

Choose category, or symbolic representation

Test hypothesis: ART matching rule

Choose another category

ART HYPOTHESIS TESTING AND LEARNING CYCLE

Choose category, or symbolic representation

Test hypothesis: ART matching rule

Mismatch
Reset:
Novelty-
Sensitive
Arousal
Burst

Choose another category

ART HYPOTHESIS TESTING AND LEARNING CYCLE

Choose category, or symbolic representation

Test hypothesis: ART matching rule

Mismatch
Reset:
Novelty-
Sensitive
Arousal
Burst

Choose another category
COGNITIVE LEARNING AND MEMORY CONSOLIDATION CYCLE

A dynamic cycle of RESONANCE and RESET

As categories are learned, search automatically disengages
Modulatory novelty potentials subside as this type of memory consolidation ends
Direct access to globally best-matching category

Mathematical proof in: Carpenter & Grossberg, CVGIP, 1987

Many supportive psychological and neurobiological data

Explains how we can quickly recognize familiar objects even if, as we get older, we store enormous numbers of memories.
ERP SUPPORT FOR HYPOTHESIS TESTING CYCLE

Event-Related Potentials: Human Scalp Potentials

ART predicted correlated sequences of P120-N200-P300

Event Related Potentials during oddball learning

P120 - mismatch; N200 - arousal/novelty; P300 - STM reset

Confirmed in: Banquet and Grossberg, 1987
Cells in inferotemporal cortex are actively reset during working memory tasks.

There is an “active matching process that was reset between trials.”

Miller, Li, Desimone, 1991
COMPLEMENTARY COMPUTING IN ART
Complementary Match/Mismatch Event-Related Potentials

PN AND N200 ARE COMPLEMENTARY WAVES

MATCH

PN
Top-Down
Conditionable
Specific
Match

MISMATCH

N200
Bottom-Up
Unconditionable
Nonspecific
Mismatch
Illustrates NEW PARADIGMS for brain computing

INDEPENDENT MODULES
Computer Metaphor

COMPLEMENTARY COMPUTING
What is the nature of brain specialization?

LAMINAR COMPUTING
Why are all neocortical circuits organized in layers?
How do laminar circuits give rise to biological intelligence?
COMPLEMENTARY COMPUTING

New principles of UNCERTAINTY and COMPLEMENTARITY clarify why

Multiple parallel processing streams exist in the brain

 Lots of specialization!

Van Essen et al
WHAT ARE COMPLEMENTARY PROPERTIES?

Anallogies:
Key fits lock, puzzles pieces fit together

Computing one set of properties at a processing stage prevents that stage from computing a complementary set of properties

Complementary parallel processing streams are BALANCED against one another

INTERACTIONS between streams overcomes their complementary weaknesses and support intelligent and creative behaviors
### SOME COMPLEMENTARY PROCESSES

<table>
<thead>
<tr>
<th>Visual Boundary</th>
<th>Visual Surface</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interbob Stream V1-V4</td>
<td>Blob Stream V1-V4</td>
</tr>
<tr>
<td>Visual Boundary</td>
<td>Visual Motion</td>
</tr>
<tr>
<td>Interbob Stream V1-V4</td>
<td>Magno Stream V1-MT</td>
</tr>
<tr>
<td>WHAT Steam</td>
<td>WHERE Stream</td>
</tr>
<tr>
<td>Perception &amp; Recognition</td>
<td>Space &amp; Action</td>
</tr>
<tr>
<td>Inferotemporal and Prefrontal areas</td>
<td>Parietal and Prefrontal areas</td>
</tr>
<tr>
<td>Object Tracking</td>
<td>Optic Flow Navigation</td>
</tr>
<tr>
<td>MT Interbands and MSTv</td>
<td>MT Bands and MSTd</td>
</tr>
<tr>
<td>Motor Target Position</td>
<td>Volitional Speed</td>
</tr>
<tr>
<td>Motor and Parietal Cortex</td>
<td>Basal Ganglia</td>
</tr>
</tbody>
</table>
BP AND DEEP LEARNING DO NOT HAVE

STM activation patterns
STM critical feature patterns
ATTENTION

ANY FAST INFORMATION PROCESSING
LTM top-down learned expectations
HYPOTHESIS TESTING
using interacting STM and LTM traces

No NEURAL ARCHITECTURE
e.g., Complementary Computing
CATASTROPHIC FORGETTING EXAMPLES
Carpenter & Grossberg (1987)

You do not need a large database to show catastrophic forgetting if the ART MATCHING RULE does not hold.

Learning lists of JUST FOUR INPUT VECTORS A, B, C, and D can exhibit catastrophic forgetting if they are repeated cyclically in the order:

ABCAD   ABCAD   ABCAD   ABCAD

and are related to each other in the following way:
CODE INSTABILITY INPUT SEQUENCES

\[
D \subset C \subset A \\
B \subset A \\
B \cap C = \emptyset \\
|D| < |B| < |C|
\]

where \(|E|\) is the number of features in the set \(E\).

Any set of input vectors that satisfy the above conditions will lead to unstable coding if they are periodically presented in the order

\[ABCAD\]

and the top-down ART Matching Rule is shut off.
### STABLE AND UNSTABLE LEARNING

#### (a) UNSTABLE CODING

<table>
<thead>
<tr>
<th>Trial</th>
<th>BU</th>
<th>RES</th>
<th>NODE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td>RES</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td>RES</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td>RES</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td>RES</td>
</tr>
<tr>
<td>9</td>
<td>*</td>
<td></td>
<td>RES</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td>RES</td>
</tr>
<tr>
<td>11</td>
<td>*</td>
<td></td>
<td>RES</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td>RES</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td>RES</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td>RES</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td>RES</td>
</tr>
</tbody>
</table>

#### (b) STABLE CODING

<table>
<thead>
<tr>
<th>Trial</th>
<th>BU</th>
<th>RES</th>
<th>NODE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td>RES</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td>RES</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td>RES</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td>RES</td>
</tr>
<tr>
<td>9</td>
<td>*</td>
<td></td>
<td>RES</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td>RES</td>
</tr>
<tr>
<td>11</td>
<td>*</td>
<td></td>
<td>RES</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td>RES</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td>RES</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td>RES</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td>RES</td>
</tr>
</tbody>
</table>

**Learning finished on trial 2
DIRECT ACCESS!**
ART HYPOTHESIS TESTING AND LEARNING CYCLE

Choose category, or symbolic representation

Mismatch Reset:
Novelty-Sensitive Arousal Burst

Test hypothesis: ART matching rule

VIGILANCE
How big a mismatch causes reset?

Choose another category
VIGILANCE determines what features are learned in the CRITICAL FEATURE PATTERN.

It clarifies how our brains learn CONCRETE knowledge for some tasks and ABSTRACT knowledge for others.

High Vigilance – Narrow Categories; CONCRETE
Mom’s face

Low Vigilance – Broad Categories; ABSTRACT
A face

Critical feature patterns are explainable at every level of vigilance!
CLASSIFYING ALPHABET AT TWO VIGILANCE LEVELS

Vigilance .5

Fewer, and more abstract, categories

Vigilance .8

More, and more concrete, categories
VIGILANCE DATA IN INFEROTEMPORAL CORTEX

RECEPTIVE FIELD SELECTIVITY MUST BE LEARNED

Some cells respond selectively to particular views of particular faces
Other cells respond to broader features of an animal’s environment

Desimone, Gross, Perrett, …

EASY vs. DIFFICULT DISCRIMINATIONS: VIGILANCE!

“In the difficult condition the animals adopted a stricter internal criterion for discriminating matching from non-matching stimuli... The animal’s internal representations of the stimuli were better separated ... increased effort appeared to cause enhancement of the responses and sharpened selectivity for attended stimuli…”

Spitzer, Desimone, and Moran, 1988
VIGILANCE CONTROL

\[ \rho \left| I \right| - \left| X \right| \leq 0 \quad \rho \leq \frac{X}{|I|} \]

resonate and learn

\[ \rho \left| I \right| - \left| X \right| > 0 \quad \rho > \frac{X}{|I|} \]

reset and search

\[ \rho \text{ is a sensitivity or gain parameter} \]
VIGILANCE CONTROL

\[ \rho |I| - |X| \leq 0 \quad \rho \leq \frac{|X|}{|I|} \quad \text{resonate and learn} \]

\[ \rho |I| - |X| > 0 \quad \rho > \frac{|X|}{|I|} \quad \text{reset and search} \]

\[ \rho I \]

\( \rho \) is a sensitivity or gain parameter

How to change vigilance based on predictive success?
FROM UNSUPERVISED
TO SUPERVISED ART MODELS
Extend UNSUPERVISED ART to SUPERVISED or UNSUPERVISED ARTMAP

**FUZZY ARTMAP**

![Diagram of Fuzzy ARTMAP](image)

MATCH TRACKING realizes Minimax Learning Principle:
Vigilance increases to just above the match ratio of prototype / exemplar, thereby triggering search
**LEARN MANY-TO-ONE and ONE-TO-MANY MAPS**

<table>
<thead>
<tr>
<th>Many-to-One</th>
<th>One-to-Many</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Compression, Naming</strong></td>
<td><strong>Expert Knowledge</strong></td>
</tr>
<tr>
<td>((a_1,b))</td>
<td>((a,b_1))</td>
</tr>
<tr>
<td>((a_2,b))</td>
<td>((a,b_2))</td>
</tr>
<tr>
<td>((a_3,b))</td>
<td>((a,b_3))</td>
</tr>
<tr>
<td>((a_4,b))</td>
<td>((a,b_4))</td>
</tr>
</tbody>
</table>

- **Fruit**: Orange, Grapes, Apple, Banana
- **Animal**: Dog
- **Mammal**: Dalmatian
- **Pet**: Dog
- **Fireman’s Mascot**: “Rover”
MANY-TO-ONE MAP
Two Stages of Compression

VISUAL CATEGORIES

AUDITORY CATEGORIES

SEE

HEAR

“AY”
MANY-TO-ONE MAP

Two Stages of Compression

Medical Database Prediction

Symptoms test treatments

Length of stay in hospital
ONE-TO-MANY MAP
Expert Knowledge

Visual categories

Auditory categories

See

Hear

Recognize dog

Search

Mismatch

Reset

Expect “dog”

“Rover”
MINIMAX LEARNING PRINCIPLE

How to conjointly

minimize predictive error

and

maximize generalization

using error feedback

in an incremental fast learning context

in response to nonstationary data?
MATCH TRACKING realizes MINIMAX LEARNING PRINCIPLE

Given a predictive error, vigilance increases just enough to trigger search and thus sacrifices the minimum generalization to correct the error.

...and enables expert knowledge to be incrementally learned.
Are ART mechanisms like vigilance control realized within LAMINAR cortical and thalamic circuits?

YES!

**SMART model**

**Synchronous Matching ART**

Grossberg and Versace, 2008

**MAIN QUESTIONS:**

How are multiple levels of brain organization spikes
local field potentials
inter-areal synchronous oscillations
spike-timing dependent plasticity

coordinated to
regulate stable category learning and attention
during cognitive information processing via
laminar cortical circuits
specific and nonspecific thalamic nuclei?
Illustrates NEW PARADIGMS for brain computing

INDEPENDENT MODULES
Computer Metaphor

COMPLEMENTARY COMPUTING
What is the nature of brain specialization?

LAMINAR COMPUTING
Why are all neocortical circuits organized in layers?
How do laminar circuits give rise to biological intelligence?
THE MODEL FUNCTIONALLY EXPLAINS LOTs OF ANATOMICAL DATA

<table>
<thead>
<tr>
<th>Connections</th>
<th>Type</th>
<th>Functional interpretation</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>thalamic core A → 4 A</td>
<td>D</td>
<td>Primary thalamic relay cells drive layer 4.</td>
<td>Blasdel and Lund (1983)</td>
</tr>
<tr>
<td>thalamic core A → 6(^l) A</td>
<td>D</td>
<td>Primary thalamic relay cells prime layer 4 via the 6 → 4 modulatory circuit.</td>
<td>Blasdel and Lund (1983) for LGN → 6; Callaway (1998) LGN input to 6 is weak and Layer 5 projections to 6 [Note 1]</td>
</tr>
<tr>
<td>thalamic core A → RE A</td>
<td>D</td>
<td>Recurrent inhibition to primary and secondary thalamic relay cells.</td>
<td>Sherman and Guillery (2001); Jones (2002)</td>
</tr>
<tr>
<td>RE A → thalamic core A</td>
<td>I</td>
<td>Off-surround to primary and secondary thalamic relay cells, synchronization of thalamic relay cells.</td>
<td>Cox et al. (1997); Pinault and Deschenes (1998); Sherman and Guillery (2001)</td>
</tr>
<tr>
<td>RE A (B) → RE B(A)</td>
<td>GJ</td>
<td>Synchronize RE and thalamic relay cells.</td>
<td>Landisman et al. (2002)</td>
</tr>
<tr>
<td>RE A → nonspecific thalamic A</td>
<td>I</td>
<td>Inhibition of nonspecific thalamic cells, participates in the reset mechanism.</td>
<td>Kolmac and Mitrofanis (1997); Van der Werf et al. (2002)</td>
</tr>
<tr>
<td>nonspecific thalamic A → 5 A</td>
<td>M</td>
<td>To 5 through apical dendrites in 1, participates in the reset mechanism.</td>
<td>Van der Werf et al. (2002)</td>
</tr>
<tr>
<td>4 A → 2/3 A</td>
<td>D</td>
<td>Feedforward driving output from 4 to 2/3.</td>
<td>Fitzpatrick et al. (1985); Callaway and Wiser (1996)</td>
</tr>
<tr>
<td>2/3 A → 4 B</td>
<td>D</td>
<td>Feedforward output from Area A to Area B.</td>
<td>Van Essen et al. (1986)</td>
</tr>
<tr>
<td>2/3 A → 6(^l) B</td>
<td>D</td>
<td>Feedforward output from Area A to Area B.</td>
<td>Van Essen et al. (1986)</td>
</tr>
<tr>
<td>2/3 A → 5 A</td>
<td>D</td>
<td>Conveys layer 2/3 output to layer 5.</td>
<td>Callaway and Wiser (1996)</td>
</tr>
<tr>
<td>2/3 A → 6(^l) A</td>
<td>D</td>
<td>Conveys layer 2/3 output to layer 6(^l).</td>
<td>Callaway (1998)</td>
</tr>
</tbody>
</table>
### THE MODEL FUNCTIONALLY EXPLAINS LOTS OF ANATOMICAL DATA

<table>
<thead>
<tr>
<th>Connections</th>
<th>Type</th>
<th>Functional interpretation</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>5A → thalamic core B</td>
<td>D</td>
<td>Feedforward connections from Area A to Area B through secondary thalamic relay neurons.</td>
<td>Rockland (1999); Sherman and Guillery (2001)</td>
</tr>
<tr>
<td>5A → 6I A</td>
<td>D</td>
<td>Delivers feedback to the 6 → 4 circuit from higher cortical areas, sensed at the apical dendrites of 5 branching in 1.</td>
<td>Callaway (1998); Callaway and Wiser (1996), class B&lt;sup&gt;®&lt;/sup&gt; cells [Note 2]</td>
</tr>
<tr>
<td>6II A → thalamic Core A</td>
<td>M</td>
<td>On-center to primary thalamic relay cells.</td>
<td>Sillito et al. (1994); Callaway (1998);</td>
</tr>
<tr>
<td>6II A → RE A</td>
<td>D</td>
<td>Off-surround to primary thalamic relay cells mediated by thalamic RE.</td>
<td>Guillery and Harting (2003); Sherman and Guillery (2001)</td>
</tr>
<tr>
<td>6II B → 2/3, 2/3 inh., 5 A</td>
<td>M</td>
<td>Intercortical feedback from 6II area B to 1 area A, where it synapses on 2/3 excitatory and inhibitory neurons, as well as 5 apical dendrites branching in 1</td>
<td>Rockland and Virga (1989); Rockland (1994); Salin and Bullier (1995)</td>
</tr>
</tbody>
</table>

**Abbreviations:** inh. = inhibitory neurons; RE = reticular nucleus; A = primary (thalamic, cortical) loop; B = secondary (thalamic, cortical) loop; D = driving excitatory connections; M = modulatory connections; I = inhibitory connections; GJ = gap junctions; int. = inhibitory interneuron. [Note 1]: Callaway (1998) subdivides Layer 6 neurons in 3 classes: Class I: provide feedback to 4C, receive input from LGN, and project back to LGN; Class IIa: dendrites in 6, axons from 2/3, project back to 2/3 with modulatory connections; Class IIb: dendrites in 5, project exclusively to deep layers (5 & 6) and claustrum. In the model, these populations are clustered in 2 classes, layer 6I and 6II, which provide feedback to thalamic relay cells and layer 4, respectively. [Note 2]: Callaway (1998) subdivides Layer 5 neurons in 3 classes: Class A: dendrites in 5, axons from 2/3, project back to 2/3 with modulatory connections; Class B: dendrites in 5, axons from 2/3, project laterally to 5 and PULVINAR; Class C: dendrites in 1, project to superior colliculus. In the model, these differences are ignored, and it is assumed that the model layer 5 neuron receives input from 2/3 (Classes A and B), as well modulatory input from the nonspecific thalamic nuclei (Class C, apical dendrites in layer 1), and provide output to 6I and second-order thalamic nuclei. The inner, recurrent loop with 2/3 has also been ignored.
BRAIN OSCILLATIONS DURING MATCH/MISMATCH

(a) TD CORTICOTHALAMIC FEEDBACK increases SYNCHRONY Sillito et al., 1994

(b) MATCH
Increases γ oscillations

(c) MISMATCH
increases θ, β oscillations

---

SG/BI/2020

88
(a) TD CORTICOTHALAMIC FEEDBACK increases SYNCHRONY  Sillito et al., 1994

(b) MATCH
Increases $\gamma$ oscillations

(c) MISMATCH
Increases $\theta$, $\beta$ oscillations

V1  Buffalo et al., 2011
van Kerkoerle et al., 2014

V4  Buschman and Miller, 2009

Hippocampus  Berke et al., 2009
Acetylcholine (Ach) regulation by NONSPECIFIC THALAMIC NUCLEI via NUCLEUS BASALIS OF MEYNERT reduces AHP in layer 5

Ach thereby facilitates RESET (compare ART VIGILANCE control)

HIGH Vigilance ~ Sharp Code
LOW Vigilance ~ Coarse Code

CHOLINERGIC DENSITY AXONS IN V1 AND HOMOLOGS
Gu (2003)

SG/BI/2020

Gu (2003)
A dynamic cycle of

RESONANCE

and

RESET

As categories are learned, search automatically disengages

Modulatory novelty potentials subside as

this type of memory consolidation ends

Direct access to globally best-matching category

Mathematical proof in: Carpenter & Grossberg, CVGIP, 1987

Many supportive psychological and neurobiological data

Explains how we can quickly recognize familiar objects even if, as we get older, we store enormous numbers of memories
Catastrophic forgetting occurs if top-down expectations fail.

What goes wrong if the ORIENTING SYSTEM fails? AMNESIA OCCURS!

DYNAMIC PHASE OF MEMORY CONSOLIDATION

While input exemplar still drives memory search

before direct access occurs

An emergent property of the entire circuit
A FORMAL AMNESIC SYNDROME

Due to damaged medial temporal brain structures – Hippocampus

ORIENTING SYSTEM!

1. Unlimited anterograde amnesia
   Cannot search for new categories

2. Limited retrograde amnesia
   Direct access

3. Failure of consolidation
   Squire & Cohen, 1994

4. Defective novelty reactions
   Perseveration
   O’Keefe & Nadel, 1978

5. Memory consolidation and novelty detection
   Mediated by same structures
   Zola-Morgan & Squire, 1990
5. **Normal priming**
   - Baddeley & Warrington (1970)
   - Mattis & Kovner (1984)

6. **Learning of first item dominates**
   - Gray (1982)

7. **Impaired ability to attend to relevant dimensions of stimuli**
   - Butters & Cermak (1975); Pribram (1986)
Fornix transection “impairs ability to change an established habit...impaired learning when one habit is to be formed in one set of circumstances and a different habit in a different set of circumstances that is similar to the first…”

Gaffan (1985)

Memory consolidation and novelty detection mediated by same neural structures

Zola-Morgan & Squire (1990)

Reduction in novelty-related hippocampal potentials as learning proceeds in rats

Deadwyler, West, & Lynch (1979)
Deadwyler, West, & Robinson (1981)
VIGILANCE control during MEDIAL TEMPORAL AMNESIA

Knowlton and Squire (1993) assume that two memory systems are needed to explain their amnesia data.

Nosofsky and Zaki (1998) showed that a single exemplar model with a low SENSITIVITY parameter can quantitatively fit their data.

Low SENSITIVITY plays a role similar to low VIGILANCE in ART.

His exemplar model implicitly needs BU and TD feedback between exemplars and a category PROTOTYPE.
VIGILANCE DATA IN INFEROTEMPORAL CORTEX

RECEPTIVE FIELD SELECTIVITY MUST BE LEARNED

Some cells respond selectively to particular views of particular faces
Other cells respond to broader features of an animal’s environment

Desimone, Gross, Perrett, ...

EASY vs. DIFFICULT DISCRIMINATIONS: VIGILANCE!

“In the difficult condition the animals adopted a stricter internal criterion for discriminating matching from non-matching stimuli... The animal’s internal representations of the stimuli were better separated ... increased effort appeared to cause enhancement of the responses and sharpened selectivity for attended stimuli...”

Spitzer, Desimone, and Moran, 1988
Predictive ART, or pART, architecture macrocircuit
How prefrontal cortex learns to control all higher-order intelligence
Grossberg (2018; see sites.bu.edu/steveg)

Working memory, learned plans, prediction, optimized action

Reinforcement learning, emotion, motivation, adaptively-timed learning,

Visual perception, category learning, object attention
EACH BRAIN REGION IN NATURE AND IN pART CARRIES OUT A DIFFERENT FUNCTION

CONTRAST THE HOMOGENEOUS ORGANIZATION OF A TYPICAL DEEP LEARNING NETWORK
EXPLAINABLE VISUAL PERCEPTS

The functional units of BIOLOGICAL VISION are

COMPLETED DEPTH-SELECTIVE BOUNDARIES

which gate

FILLING-IN OF DEPTH-SELECTIVE SURFACES

BOUNDARIES and SURFACES are COMPUTATIONALLY COMPLEMENTARY
EXPLAINABLE VISUAL PERCEPTS

The functional units of BIOLOGICAL VISION are

**COMPLETED DEPTH-SELECTIVE BOUNDARIES**

which gate

**FILLING-IN OF DEPTH-SELECTIVE SURFACES**

**BOUNDARIES and SURFACES are COMPUTATIONALLY COMPLEMENTARY**
Spatially abutting and colinear Boundary Contours (BC) and Feature Contours (FC) can trigger depth-selective filling-in of the FC color in the Filling-In-Domain that is surrounded by its BC.
Boundaries and surfaces are explainable by observing their depth-selective spatially distributed representations.
EXPLAINABLE VISUAL PERCEPTS

The functional units of BIOLOGICAL VISION are

COMPLETED DEPTH-SELECTIVE BOUNDARIES

which gate

FILLING-IN OF DEPTH-SELECTIVE SURFACES

BOUNDARIES and SURFACES are COMPUTATIONALLY COMPLEMENTARY
Illustrates **NEW PARADIGMS** for brain computing

**INDEPENDENT MODULES**
Computer Metaphor

**COMPLEMENTARY COMPUTING**
What is the nature of brain specialization?

**LAMINAR COMPUTING**
Why are all neocortical circuits organized in layers? How do laminar circuits give rise to biological intelligence?
SOME COMPLEMENTARY PROCESSES

Visual Boundary
Interbob Stream V1-V4

Visual Surface
Blob Stream V1-V4

Visual Boundary
Interbob Stream V1-V4

Visual Motion
Magno Stream V1-MT

WHAT Steam
Perception & Recognition
Inferotemporal and
Prefrontal areas

WHERE Stream
Space & Action
Parietal and
Prefrontal areas

Object Tracking
MT Interbands and MSTv

Optic Flow Navigation
MT Bands and MSTd

Motor Target Position
Motor and Parietal Cortex

Volitional Speed
Basal Ganglia
Predictive ART, or pART, architecture macrocircuit
How prefrontal cortex learns to control all higher-order intelligence
Grossberg (2018)

Working memory, learned plans, prediction, optimized action

Reinforcement learning, emotion, motivation, adaptively-timed learning,

Visual perception, category learning, object attention
WHAT IS A VISUAL BOUNDARY OR GROUPING?

- Illusory contour
- Texture pop-out
- 3D shape from texture
- Figure-ground separation
VISUAL BOUNDARY AND SURFACE COMPUTATIONS ARE COMPLEMENTARY

BOUNDARY COMPLETION

- oriented inward
- insensitive to direction of contrast

SURFACE FILLING-IN

- unoriented outward
- sensitive to direction of contrast

Grossberg (1984)

Neon color spreading
BOUNDARY AND SURFACE CORTICAL STREAMS

WHAT

Inferotemporal Areas

V4

V3

V2 Interstripe

V2 Thick

Interblob

V1 Interblob

V1 Blob

V2 Thin

LGN Parvo

WHERE

Parietal Areas

MT

V2 4B

LGN Magno

Retina

DeYoe and van Essen (1988)

Boundaries

interblob stream

Surfaces

blob stream
VISUAL BOUNDARY AND SURFACE COMPUTATIONS ARE COMPLEMENTARY

BOUNDARY COMPLETION

SURFACE FILLING-IN

 oriented inward insensitive to direction-of-contrast

unoriented outward sensitive to direction-of-contrast

Neon color spreading

Grossberg (1984)

What about
SEEING vs. KNOWING

SEEING
an object

Epstein, Gregory, Helmholtz, Kanizsa, Kellman, Michotte,…

KNOWING
what it is

SEEING
Ehrenstein Figure

SEEING
vs.

RECOGNIZING
Offset Grating

See

Recognize

Do not see

Recognize

Some boundaries are invisible, or amodal

SG/BI/2020
ALL BOUNDARIES ARE INVISIBLE!
within the Boundary Stream
Grossberg (1984)

WHY? To recognize object boundaries in front of textured backgrounds
VISUAL BOUNDARY AND SURFACE COMPUTATIONS ARE COMPLEMENTARY

BOUNDARY COMPLETION

SURFACE FILLING-IN

All Boundaries Are Invisible!

Neon color spreading

oriented inward
insensitive to
direction-of-contrast

unoriented outward
sensitive to
direction-of-contrast
IF BOUNDARIES ARE INVISIBLE, HOW DO WE SEE?

*Filling-In* of Surface Color

Boundaries define the compartments within which lightness and color spread

Ehrenstein (1941)

Varin (1971)

Neon color spreading
Craik-O’ Brien-Cornsweet Effect

Boundary completion defines filling-in compartments

Filling-in determines what we see in each compartment

Grossberg (1984)
Todorović (1987)
VISUAL BOUNDARY AND SURFACE COMPUTATIONS ARE COMPLEMENTARY

All Boundaries Are Invisible!

BOUNDARY COMPLETION

oriented inward
insensitive to direction-of-contrast

SURFACE FILLING-IN

unoriented outward
sensitive to direction-of-contrast

Neon color spreading

Filling-in of Visible Color and Lightness

All Boundaries Are Invisible!
PREDICTIONS

Grossberg (1984)

ALL BOUNDARIES ARE INVISIBLE
in the interblob stream

VISIBLE QUALIA ARE SURFACE PERCEPTS
in the blob stream
PREDICTIONS

ALL CONSCIOUS STATES ARE RESONANT STATES

Grossberg (1976)

VISIBLE QUALIA ARE SURFACE PERCEPTS

Grossberg (1984)
PREDICTIONS

ALL CONSCIOUS STATES ARE RESONANT STATES
Grossberg (1976)

VISIBLE QUALIA ARE SURFACE PERCEPTS
Grossberg (1984)

WHAT KIND OF RESONANCE SUPPORTS
VISIBLE SURFACE QUALIA?
How do we see?!
PREDICTIONS

ALL CONSCIOUS STATES ARE RESONANT STATES
Grossberg (1976)

VISIBLE QUALIA ARE SURFACE PERCEPTS
Grossberg (1984)

WHAT KIND OF RESONANCE SUPPORTS VISIBLE SURFACE QUALIA?
How do we see?!

A SURFACE-SHROUD RESONANCE
Grossberg (2009+)
“ALL CONSCIOUS STATES ARE RESONANT STATES”

Surface-shroud resonances support conscious seeing of visual qualia EXPLAINABLE!

Feature-category resonances support conscious recognition of visual objects and scenes

Stream-shroud resonances support conscious hearing of auditory qualia

Spectral-pitch-and-timbre resonances support conscious recognition of sources in auditory streams

Item-list resonances support conscious recognition of speech and language

Cognitive-emotional resonances support conscious feelings and recognition of them
“ALL CONSCIOUS STATES ARE RESONANT STATES”

Surface-shroud resonances support conscious seeing of visual qualia

Feature-category resonances support conscious recognition of visual objects and scenes

Stream-shroud resonances support conscious hearing of auditory qualia

Spectral-pitch-and-timbre resonances support conscious recognition of sources in auditory streams

Item-list resonances support conscious recognition of speech and language

Cognitive-emotional resonances support conscious feelings and recognition of them

SEEING

KNOWING

SG/BI/2020
WHAT IS A SURFACE-SHROUD RESONANCE?
WHAT IS AN ATTENTIONAL SHROUD?

Surface-fitting spatial attention

ATTENTIONAL SHROUD!

marks the object-hood of the as-yet-undefined object category

Tyler and Kontsevich (1995)
used shrouds to study perceptual transparency

Cf. Cavanagh, Pylyshyn, Yantis,…

PREDICTION:
Shrouds enable learning of invariant object categories
Not explained in this talk, but see:

Fazl, Grossberg, and Mingolla (2009)
Cao, Grossberg, and Mingolla (2011)
Grossberg, Markowitz, and Cao (2011)
Foley, Grossberg, and Mingolla (2012)
Chang, Grossberg, and Cao (2014)
BOTTOM-UP SPATIAL ATTENTIONAL COMPETITION

Spatial Attention

Competition

More luminous

Less luminous

Perceptual Surfaces
SURFACE-SHROUD RESONANCE

Spatial Attention

Competition

Perceptual Surfaces

Neurophysiology: Reynolds and Desimone (2003)
SURFACE-SHROUD RESONANCE

An active SURFACE-SHROUD RESONANCE means that sustained SPATIAL ATTENTION IS FOCUSED ON THE OBJECT SURFACE
WHAT KINDS OF RESONANCES SUPPORT KNOWING VS. SEEING?

What Stream | Where Stream
---|---
IT | PPC
V2/4

**KNOWING**
- Feature-Category Resonance

**SEEING**
- Surface-Shroud Resonance

Synchronous linkage between resonances enables us to **KNOW** what the object is as we **SEE** it.

Many data support this prediction; e.g.:
WHAT KINDS OF RESONANCES SUPPORT KNOWING VS. SEEING?

KNOWING
Feature-Category
Resonance

SEEING
Surface-Shroud
Resonance

VISUAL AGNOSIA: reaching without knowing
Patient DF Goodale et al, 1991
WHAT KINDS OF RESONANCES SUPPORT KNOWING VS. SEEING?

**KNOWING**
- Feature-Category Resonance

**SEEING**
- Surface-Shroud Resonance

**VISUAL AGNOSIA:** reaching without knowing

**Patient DF** Goodale et al, 1991

Attention and Intention both parietal cortical functions

Andersen, Essick, and Siegel, 1985; Gnadt and Andersen, 1988; Snyder, Batista, and Andersen, 1997, 1998
“ALL CONSCIOUS STATES ARE RESONANT STATES”

Surface-shroud resonances support conscious seeing of visual qualia

Feature-category resonances support conscious recognition of visual objects and scenes

Stream-shroud resonances support conscious hearing of auditory qualia

Spectral-pitch-and-timbre resonances support conscious recognition of sources in auditory streams

Item-list resonances support conscious recognition of speech and language

Cognitive-emotional resonances support conscious feelings and recognition of them
UNIFYING THREE BASIC BEHAVIORAL COMPETENCES

1. Pay attention quickly to salient events, both positive and negative
   The fast motivated attention pathway includes the AMYGDALA

   However, a rapid attention shift to focus on a salient event could cause a premature response to that event

   This problem is eliminated by the second and third competences

2. Adaptively time and maintain motivated ATTENTION on a salient event until the response is executed
   This ability involves the HIPPOCAMPUS, notably its dentate-CA3 region

3. Adaptively time and execute an appropriate RESPONSE to the salient event
   This ability involves the CEREBELLUM

SG/BI/2020
Cognitive-Emotional-Motor (CogEM) model
...and its further developments...
Cognitive-Emotional-Motor (CogEM) model

Need converging cue and incentive inputs to fire

CS → SENSORY CORTEX → ORBITOFrontal CORTEX → AMYGDALA → DRIVE

Conditioned Reinforcer Learning

Incentive Motivational Learning
Adapted from Barbas (1995)
Cognitive-Emotional-Motor (CogEM) model

Need converging cue and incentive inputs to fire

CS → SENSORY CORTEX → ORBITOFRONTAL CORTEX

Conditioned Reinforcer Learning

Amygdala

Motivated attention closes the cognitive-emotional feedback loop, focuses on relevant cues, and causes blocking of irrelevant cues

Incentive Motivational Learning

Drive

SG/BI/2020
COGNITIVE-EMOTIONAL RESONANCE

Basis of “core consciousness” and “the feeling of what happens”
Damasio (1999) derives heuristic version of CogEM model from his clinical data

![Diagram showing the brain regions involved in cognitive-emotional resonance.]

- SENSORY CORTEX
- PREFRONTAL CORTEX
- AMYGDALA
- DRIVE

**COGNITIVE-EMOTIONAL RESONANCE**

Basis of “core consciousness” and “the feeling of what happens”

Damasio (1999) derives heuristic version of CogEM model from his clinical data

**Diagram:**

- SENSORY CORTEX
- PREFRONTAL CORTEX
- AMYGDALA
- DRIVE

**Questions:**

- How is this resonance maintained long enough to become conscious?!
HIPPOCAMPUS CAN SUSTAIN A COGNITIVE-EMOTIONAL RESONANCE that can support “the feeling of what happens” and knowing what event caused that feeling.
HIPPOCAMPUS CAN SUSTAIN A COGNITIVE-EMOTIONAL RESONANCE
that can support “the feeling of what happens”
and knowing what event caused that feeling.
A representation can be explainable without potentially being conscious

e.g., MOTOR REPRESENTATIONS

Their matching and learning laws are not resonant!
### SOME COMPLEMENTARY PROCESSES

<table>
<thead>
<tr>
<th>Visual Boundary</th>
<th>Visual Surface</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interbob Stream V1-V4</td>
<td>Blob Stream V1-V4</td>
</tr>
<tr>
<td>Visual Boundary</td>
<td>Visual Motion</td>
</tr>
<tr>
<td>Interbob Stream V1-V4</td>
<td>Magno Stream V1-MT</td>
</tr>
<tr>
<td><strong>WHAT Steam</strong></td>
<td><strong>WHERE Stream</strong></td>
</tr>
<tr>
<td>Perception &amp; Recognition</td>
<td>Space &amp; Action</td>
</tr>
<tr>
<td>Inferotemporal and Prefrontal areas</td>
<td>Parietal and Prefrontal areas</td>
</tr>
<tr>
<td>Object Tracking</td>
<td>Optic Flow Navigation</td>
</tr>
<tr>
<td>MT Interbands and MSTv</td>
<td>MT Bands and MSTd</td>
</tr>
<tr>
<td>Motor Target Position</td>
<td>Volitional Speed</td>
</tr>
<tr>
<td>Motor and Parietal Cortex</td>
<td>Basal Ganglia</td>
</tr>
</tbody>
</table>
WHAT and WHERE/HOW LEARNING and MATCHING are COMPLEMENTARY

Spatially-invariant object learning and recognition
Fast learning without catastrophic forgetting

Spatially-variant reaching and movement
Continually update sensory-motor maps and gains

WHAT

MATCHING

LEARNING

WHERE/HOW

MATCHING

EXCITATORY

INHIBITORY

MATCH

MISMATCH

ART

VAM
SENSORY EXPECTATION vs MOTOR EXPECTATION

ART MATCH

Category

Features

Match Amplifies Match Learning

VAM MATCH

Target Position

Difference Vector

Present Position

Match Suppresses Mismatch Learning

Target Position

Difference Vector

Present Position

Match Suppresses Mismatch Learning

SG/BI/2020 146
VECTOR INTEGRATION TO ENDPOINT (VITE): EXPLAINABLE!

Bullock & Grossberg (1988)

Target Position Vector $T$

Present Position Vector $P$

Difference Vector $V$

GO signal $G$

Outflow Movement Speed $[D]^+G$

are all EXPLAINABLE!
VITE HAS BEEN MEASURED PHYSIOLOGICALLY

e.g., DIFFERENCE VECTOR

Georgopoulos et al., (1982)
THE THREE S’S OF MOVEMENT CONTROL

SYNERGY
Defining T determines the muscle groups that will contract during the movement

SYNCHRONY
When G turns on, all muscle groups for which $D \neq 0$ contract by variable amounts in equal time. Because G multiplies $D$, it does not change the direction in which P moves to acquire T: Straight line movement

SPEED
$P$ integrates $D$ at rate $G$ until $P = T$. Increasing (decreasing) $G$ makes the movement faster (slower)
VITE HAS BEEN UNLUMPED TO REVEAL FINER STRUCTURE

Bullock, Cisek, & Grossberg (1998)
ALL THESE BIOLOGICAL MODELS OF PERCEPTION, COGNITION, EMOTION, AND ACTION ARE EXPLAINABLE

Perceptual and cognitive processes use ART-like excitatory matching and match-based learning to create self-stabilizing attentive and conscious representations of objects and events that embody increasing expertise about the world.

Complementary spatial and motor processes use inhibitory matching and mismatch-based learning to continually update spatial and motor representations to compensate for bodily changes throughout life.

Together they provide a self-stabilizing perceptual and cognitive front end for conscious awareness and knowledge acquisition, which can intelligently manipulate more labile spatial and motor processes that enable our changing bodies to act effectively on a changing world.