

1997-11

# Representations Need Self-Organizing Top-Down Expectations to Fit a Changing World

---

<https://hdl.handle.net/2144/2131>

*"Downloaded from OpenBU. Boston University's institutional repository."*

BBS COMMENTARY  
on Representation is Representation of Similarities  
by Shimon Edelman

**REPRESENTATIONS NEED SELF-ORGANIZING  
TOP-DOWN EXPECTATIONS TO FIT A CHANGING WORLD**

Commentator: Stephen Grossberg

**November 1997**

**Technical Report CAS/CNS-97-019**

Permission to copy without fee all or part of this material is granted provided that: 1. The copies are not made or distributed for direct commercial advantage; 2. the report title, author, document number, and release date appear, and notice is given that copying is by permission of the BOSTON UNIVERSITY CENTER FOR ADAPTIVE SYSTEMS AND DEPARTMENT OF COGNITIVE AND NEURAL SYSTEMS. To copy otherwise, or to republish, requires a fee and / or special permission.

Copyright © 1997

Boston University Center for Adaptive Systems and  
Department of Cognitive and Neural Systems  
677 Beacon Street  
Boston, MA 02215

BBS COMMENTARY

on Representation is Representation of Similarities

by Shimon Edelman

**REPRESENTATIONS NEED SELF-ORGANIZING TOP-DOWN EXPECTATIONS  
TO FIT A CHANGING WORLD**

Commentator: Stephen Grossberg<sup>1</sup>

Department of Cognitive and Neural Systems

Boston University

677 Beacon Street

Boston, MA 02215

steve@cns.bu.edu

Technical Report CAS/CNS-TR-97-019

Boston, MA: Boston University

Word count: 1000

---

<sup>1</sup> Supported in part by the Defense Advanced Research Projects Agency and the Office of Naval Research (ONR N00014-92-J-1309, ONR N00014-95-1-0409, and ONR N00014-95-1-0657)

### Abstract

The author's model "Chorus embodies an attempt to find out how far a mostly bottom-up approach to representation can be taken" (p. 22). Models which embody both bottom-up and top-down learning have stronger computational properties and explain more data about representation than feedforward models.

### Text

Adaptive Resonance Theory, or ART, models self-organize "second-order isomorphisms" using either unsupervised learning, supervised learning, or mixtures of both. This self-organizing capability is needed to learn in the real world. Regularization networks are not self-organizing in this sense. They cannot do fast stable learning in complex changing environments. These properties depend upon learned top-down expectations, matching of bottom-up data with these expectations, and mismatch-driven search for new representations (Carpenter and Grossberg, 1991; Grossberg, 1980, 1987). These mechanisms allow ART to automatically "ignore those directions...that are irrelevant to the identity of the stimulus" (p. 13) by focusing attention upon critical features while suppressing irrelevant features. This ART matching rule has been supported by many psychophysical and neurobiological data (e.g., Grossberg, 1995; Grossberg and Merrill, 1996). ART matching also allows a dynamical control of attentive vigilance, through a process of "match tracking", which automatically controls how general learned representations become to match world statistics (Carpenter and Grossberg, 1991). Other models in which bottom-up and top-down processes are employed do not yet have these properties; e.g., Back Propagation and the Helmholtz Machine.

The author criticizes winner-take-all decisions because they violate the "principle of least commitment" (p. 20), but such decisions can quantitatively simulate categorical perception data; e.g., Grossberg et al. (1997a). ART systems such as masking fields (Cohen and Grossberg, 1986), ART-EMAP (Carpenter and Ross, 1995), Distributed ARTMAP (Carpenter, 1996), and Gaussian ARTMAP (Williamson, 1996) also show how distributed codes may improve recognition, and how the distribution reflects data uncertainty. Gaussian ARTMAP, in particular, is a self-organizing RBF production system.

Self-organizing view-invariant 3-D object categories fuse view-specific categories in ARTMAP systems (e.g., Bradski and Grossberg), as in the IT data reviewed in Section 7.2. The 3-D categories occur in the Map Field, wherein outputs from multiple categories, whether of different letter fonts or different object views, are adaptively fused.

Edelman's Measurements and Dimensionality Reduction stages (p. 11) are typically called Vision and Learned Recognition stages. Although ART top-down matching occurs within the vision system, even as peripherally as the LGN (Gove et al, 1995; Grossberg et al., 1997b), vision uses different principles and circuits than the recognition system. Edelman describes measurement as "a convolution with a number of filters, followed by the application of a nonlinearity" (p. 11), including light source compensation (p. 11) and figure-ground separation (p. 22). Cortical models of visual perception, called FACADE models, suggest additional mechanisms (e.g., Arrington, 1994; Chey *et al.*, 1997; Francis and Grossberg, 1996; Gove *et al.*, 1995; Grossberg and Todorovic, 1988; Grossberg, 1994, 1997; Grossberg *et al.*, 1997b). For example, parallel processing streams for boundary representation (interblob stream) and surface representation (blob stream) compute complementary computational properties. Feedback between these streams assures their mutual consistency and initiates figure-ground pop-out. Diffusive filling-in completes surface representations from signals that discount the illuminant.

Edelman summarizes a sensible approach to representation, but one that is limited by its feedforward character. ART models self-organize stable representations that achieve second-order isomorphism to arbitrarily large and changing environments, but only by using learned top-down expectations, attention, and memory search. FACADE models have clarified many data about vision, but only by introducing new concepts about how complementary streams of boundary, surface, and motion processes achieve mutual consistency and coherence using other types of feedback. A major intellectual watershed separates feedforward models from self-organizing feedforward/feedback models. This watershed needs to be crossed to deeply understand how humans autonomously form representations of the real world.

## References

- Arrington, K.F. (1994). The temporal dynamics of brightness filling-in. *Vision Research*, **34**, 3371-3387.
- Bradski, G. and Grossberg, S. (1995). Fast-learning VIEWNET architectures for recognizing 3-D objects from multiple 2-D views. *Neural Networks*, **8**, 1053-1080.
- Carpenter, G.A. (1996). Distributed activation, search, and learning by ART and ARTMAP neural networks. *Proceedings of the International Conference on Neural Networks: Plenary, Panel and Special Sessions*. Piscataway, NJ: IEEE, pp. 244-249.
- Carpenter, G.A. and Grossberg, S. (1991). Pattern recognition by self-organizing neural networks. Cambridge, MA: MIT Press.
- Carpenter, G.A. and Ross, W.D. (1995). ART-EMAP: A neural network architecture of object recognition by evidence accumulation. *IEEE Transactions on Neural Networks*, **6**, 805-818.
- Chey, J., Grossberg, S., and Mingolla, E. (1997). Neural dynamics of motion grouping: From aperture ambiguity to object speed and direction. *Journal of the Optical Society of America A*, **14**, 2570-2594.
- Cohen, M.A. and Grossberg, S. (1986). Neural dynamics of speech and language coding: developmental programs, perceptual grouping, and competition for short term memory. *Human Neurobiology*, **5**, 1-22.
- Francis, G. and Grossberg, S. (1996). Cortical dynamics of form and motion integration: Persistence, apparent motion, and illusory contours. *Vision Research*, **36**, 149-173.
- Gove, A., Grossberg, S., and Mingolla, E. (1995). Brightness perception, illusory contours, and corticogeniculate feedback. *Visual Neuroscience*, **12**, 1027-1052.
- Grossberg, S. (1980). How does a brain build a cognitive code? *Psychological Review*, **87**, 1-51.
- Grossberg, S. (1987). The adaptive brain, Vol. I and II. Amsterdam: Elsevier/North-Holland.
- Grossberg, S. (1994). 3-D vision and figure-ground separation by visual cortex. *Perception and Psychophysics*, **55**, 48-120.
- Grossberg, S. (1995). The attentive brain. *American Scientist*, **83**, 438-449.
- Grossberg, S. (1997). Cortical dynamics of 3-D figure-ground perception of 2-D pictures. *Psychological Review*, **104**, 618-658.
- Grossberg, S., Boardman, I., and Cohen, M.A. (1997a). Neural dynamics of variable-rate speech categorization. *Journal of Experimental Psychology: Human Perception and Performance*, **23**, 481-503.
- Grossberg, S. and Merrill, J.W.L. (1996). The hippocampus and cerebellum in adaptively timed learning, recognition, and movement. *Journal of Cognitive Neuroscience*, **8**, 257-277.
- Grossberg, S., Mingolla, E., and Ross, W.D. (1997b). Visual brain and visual perception: How does the cortex do perceptual grouping? *Trends in Neurosciences*, **20**, 106-111.

Grossberg, S. and Todorovic, D. (1988) Neural dynamics of 1-D and 2-D brightness perception. *Perception and Psychophysics*, **43**, 723-742.

Williamson, J.R. (1996). Gaussian ARTMAP: A neural network for fast incremental learning of noisy multidimensional maps. *Neural Networks*, **9**, 881-897.