ACKNOWLEDGMENTS
Preparation of this commentary was supported by grants from the Office of Naval Research (ONR# N000140310046) as well as by the Air Force Office of Scientific Research (AFOSR# F49620–03–1–0143).

## Bring ART into the ACT

Stephen Grossberg

*Department of Cognitive and Neural Systems, Boston University, Boston, MA 02215.* **steve@bu.edu** **http://www.cns.bu.edu/Profiles/Grossberg**

**Abstract:** ACT is compared with a particular type of connectionist model that cannot handle symbols and use nonbiological operations which do not learn in real time. This focus continues an unfortunate trend of straw man debates in cognitive science. Adaptive Resonance Theory, or ART-neural models of cognition can handle both symbols and subsymbolic representations, and meet the Newell criteria at least as well as connectionist models.

The authors' use of the nomenclature, "classical connectionist models," falsely suggests that such models satisfy the Newell criteria better than other neural models of cognition. The authors then dichotomize ACT with "classical" connectionism based on its "failure to acknowledge a symbolic level to thought. In contrast, ACT-R includes both symbolic and subsymbolic components" (target article, Abstract). Actually, neural models of cognition such as ART include both types of representation and clarify how they are learned. Moreover, ART was introduced before the "classical" models (Grossberg 1976; 1978a; 1980) and naturally satisfies key Newell criteria. In fact, Figures 2 and 3 of ACT are reminiscent of ART circuits (e.g., Carpenter & Grossberg 1991; Grossberg 1999b). But ART goes further by proposing how laminar neocortical circuits integrate bottom-up, horizontal, and top-down interactions for intelligent computation (Grossberg 1999a; Raizada & Grossberg 2003).

Critiques of classical connectionist models, here called CM (Carnegie Mellon) connectionism, show that many such models cannot exist in the brain (e.g., Grossberg 1988; Grossberg et al. 1997b; Grossberg & Merrill 1996). We claim that ART satisfies many Newell criteria better, with the obvious caveat that no model is as yet a complete neural theory of cognition.

***Flexible behavior.*** ART models are self-organizing neural production systems capable of fast, stable, real-time learning about arbitrarily large, unexpectedly changing environments (Carpenter & Grossberg 1991). These properties suit ART for large-scale technological applications, ranging from control of mobile robots, face recognition, remote sensing, medical diagnosis, and electrocardiogram analysis to tool failure monitoring, chemical analysis, circuit design, protein/DNA analysis, musical analysis, and seismic, sonar, and radar recognition, in both software and VLSI microchips (e.g., Carpenter & Milenova 2000; Carpenter et al. 1999; Granger et al. 2001). The criticism of CM connectionism "that complex, sequentially organized, hierarchical behavior" cannot be modeled also does not apply to ART (e.g., Bradski et al. 1994; Cohen & Grossberg 1986; Grossberg 1978a; Grossberg & Kuperstein 1989; Grossberg & Myers 2000; also see the section on dynamic behavior later in this commentary).

***Real-time performance.*** ART models are manifestly real-time in design, unlike CM connectionist models.

***Adaptive behavior.*** ART provides a rigorous solution of the *stability-plasticity dilemma,* which was my term for *catastrophic forgetting* before that phrase was coined. "Limitations like short-term memory" (target article, sect. 5.3) can be derived from the LTM Invariance Principle, which proposes how working memories are designed to enable their stored event sequences to be stably chunked and remembered (Bradski et al. 1994; Grossberg 1978a; 1978b).

***Vast knowledge base.*** ART can directly access the globally best-matching information in its memory, no matter how much it has learned. It includes additional criteria of value and temporal relevance through its embedding in START models that include cognitive-emotional and adaptive timing circuits in addition to cognitive ART circuits (Grossberg & Merrill 1992; 1996).

***Dynamic behavior.*** "Dealing with dynamic behavior requires a theory of perception and action as well as a theory of cognition" (sect. 2.5). LAMINART models propose how ART principles are incorporated into perceptual neocortical circuits and how high-level cognitive constraints can modulate lower perceptual representations through top-down matching and attention (Grossberg 1999a; Raizada & Grossberg 2003). ART deals with novelty through *complementary* interactions between attentional and orienting systems (Grossberg 1999b; 2000b), the former including corticocortical, and the latter, hippocampal, circuits. Action circuits also obey laws that are *complementary* to those used in perception and cognition (Grossberg 2000b), notably VAM (Vector Associative Map) laws. VAM-based models have simulated identified brain cells and circuits and the actions that they control (e.g., Brown et al. 1999; Bullock et al. 1998; Contreras-Vidal et al. 1997; Fiala et al. 1996; Gancarz & Grossberg 1999; Grossberg et al. 1997), including models of motor skill learning and performance (Bullock et al. 1993a; 1993b; Grossberg & Paine 2000).

***Knowledge integration.*** ART reconciles distributed and symbolic representations using its concept of resonance. Individual features are meaningless, just as pixels in a picture are meaningless. A learned category, or symbol, is sensitive to the global patterning of features but cannot represent the *contents* of the experience, including their conscious qualia, because of the very fact that a category is a compressed, or symbolic, representation. Resonance between these two types of information converts the *pattern* of attended features into a coherent context-sensitive state that is linked to its symbol through feedback. This coherent state, which binds distributed features and symbolic categories, can enter consciousness. ART predicts that *all conscious states are resonant states.* In particular, resonance binds spatially distributed features into a synchronous equilibrium or oscillation. Such synchronous states attracted interest after being reported in neurophysiological experiments. They were predicted in the 1970s when ART was introduced (see Grossberg 1999b). Recent neurophysiological experiments have supported other ART predictions (Engel et al. 2001; Pollen 1999; Raizada & Grossberg 2003). Fuzzy ART learns explicitly decodable Fuzzy IF-THEN rules (Carpenter et al. 1992). Thus ART accommodates symbols and rules, as well as subsymbolic distributed computations.

***Natural language.*** ART has not yet modeled language. Rather, it is filling a gap that ACT-R has left open: "ACT-R lacks any theory of the processes of speech perception or speech production" (sect. 4.5, para. 3). ART is clarifying the *perceptual units* of speech perception, word recognition, working memory, and sequential planning chunks on which the brain builds language (e.g., Boardman et al. 1999; Bradski et al. 1994; Grossberg 1978a; 1978b; 1999b; Grossberg et al. 1997a; Grossberg & Myers 2000; Grossberg & Stone 1986a; 1986b). Such studies suggest that a radical rethinking of psychological space and time is needed to understand language and to accommodate such radical claims as, "Conscious speech is a resonant wave" (cf. Grossberg, 1999b). ACT-R also does not have "mechanisms . . . [of] perceptual recognition, mental imagery, emotion, and motivation" (sect. 4.5). These are all areas where ART has detailed models (e.g., Grossberg 2000a; 2000c). Speech production uses complementary VAM-like mechanisms (Callan et al. 2000; Guenther 1995). After perceptual units in vision became sufficiently clear, rapid progress ensued at all levels of vision (cf. http://www.cns.bu.edu/Profiles/Grossberg). This should also happen for language.

***Development.*** ART has claimed since 1976 that processes of cortical development in the infant are on a continuum with processes of learning in the adult, a prediction increasingly supported recently (e.g., Kandel & O'Dell 1992).

***Evolution.*** "Cognitive plasticity . . . What enables this plasticity in the architecture?" (sect. 5.11). ART clarifies how the ability to



https://doi.org/10.1017/S0140525X03290130 Published online by Cambridge University Press

learn quickly and stably throughout life implies cognitive properties like intention, attention, hypothesis testing, and resonance. Although Bayesian properties emerge from ART circuits, ART deals with novel experiences where no priors are defined.

***Brain.*** CM connectionism is said to be "best," although its main algorithms are biologically unrealizable. ART and VAM are realized in verified brain circuits.

It might be prudent to include more ART in ACT. I also recommend eliminating straw man "debates" that do not reflect the true state of knowledge in cognitive science.

ACKNOWLEDGMENTS

Preparation of the this commentary was supported in part by the Air Force Office of Scientific Research (AFOSR F49620–01–1–0397) and the Office of Naval Research (ONR N00014–01–1–0624).

## Developing a domain-general framework for cognition: What is the best approach?

James L. McClelland[a], David C. Plaut[a], Stephen J. Gotts[b], and Tiago V. Maia[c]

[a]*Department of Psychology and Center for the Neural Basis of Cognition, Carnegie Mellon University, Pittsburgh, PA 15213;* [b]*Department of Psychology and Center for the Neural Basis of Cognition, Carnegie Mellon University, and Laboratory of Neuropsychology, NIMH/NIH, Bethesda, MD 20892;* [c]*Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213.* **jlm@cnbc.cmu.edu http://www.cnbc.cmu.edu/~jlm plaut@cmu.edu http://www.cnbc.cmu.edu/~plaut gotts@nih.gov http://www.cnbc.cmu.edu/~gotts tmaia@cmu.edu http://www.cnbc.cmu.edu/~tmaia**

**Abstract:** We share with Anderson & Lebiere (A&L) (and with Newell before them) the goal of developing a domain-general framework for modeling cognition, and we take seriously the issue of evaluation criteria. We advocate a more focused approach than the one reflected in Newell's criteria, based on analysis of failures as well as successes of models brought into close contact with experimental data. A&L attribute the shortcomings of our parallel-distributed processing framework to a failure to acknowledge a symbolic level of thought. Our framework does acknowledge a symbolic level, contrary to their claim. What we deny is that the symbolic level is the level at which the principles of cognitive processing should be formulated. Models cast at a symbolic level are sometimes useful as high-level approximations of the underlying mechanisms of thought. The adequacy of this approximation will continue to increase as symbolic modelers continue to incorporate principles of parallel distributed processing.

In their target article, Anderson & Lebiere (A&L) present a set of criteria for evaluating models of cognition, and rate both their own ACT-R framework and what they call "classical connectionism" on the criteria. The Parallel Distributed Processing (PDP) approach, first articulated in the two PDP volumes (Rumelhart et al. 1986) appears to be close to the prototype of what they take to be "classical connectionism." While we cannot claim to speak for others, we hope that our position will be at least largely consistent with that of many others who have adopted connectionist/PDP models in their research.

There are three main points that we would like to make.

1. We share with A&L (and with Newell before them) the effort to develop an overall framework for modeling human cognition, based on a set of domain-general principles of broad applicability across a wide range of specific content areas.

2. We take a slightly different approach from the one that Newell advocated, to pursuing the development of our framework. We think it worthwhile to articulate this approach briefly and to comment on how it contrasts with the approach advocated by Newell and apparently endorsed by A&L.

3. We disagree with A&L's statement that classical connectionism denies a symbolic level of thought. What we deny is only the idea that the symbolic level is the level at which the principles of processing and learning should be formulated. We treat symbolic cognition as an emergent phenomenon that can sometimes be approximated by symbolic models, especially those that incorporate the principles of connectionist models.

In what follows, we elaborate these three points, addressing the first one only briefly since this is a point of agreement between A&L and us.

***The search for domain-general principles.*** There is a longstanding tradition within psychological research to search for general principles that can be used to address all aspects of behavior and cognition. With the emergence of computational approaches in the 1950s and 1960s, and with the triumph of the von Neumann architecture as the basis for artificial computing devices, this search could be formulated as an effort to propose what Newell called "a unified architecture for cognition." An architecture consists of a specification of (1) the nature of the building blocks out of which representations and processes are constructed, (2) the fundamental rules by which the processes operate, and (3) an overall organizational plan that allows the system as a whole to operate. Newell's SOAR architecture and A&L's ACT-R architecture are both good examples of architectures of this type. For our part, we have sought primarily to understand (1) the building blocks and (2) the fundamental rules of processing. Less effort has been devoted to the specifics of the overall organizational plan as such, although we do take a position on some of the principles that the organizational plan instantiates. Because the organization is not fully specified as such, we find it more congenial to describe what we are developing as a framework rather than an architecture. But this is a minor matter; the important point is the shared search for general principles of cognition.

We are of course well aware that this search for general principles runs counter to a strong alternative thread that treats distinct domains of cognition as distinct cognitive modules that operate according to domain-specific principles. Such a view has been articulated for language by Chomsky; for vision, by Marr. Fodor and Keil have argued the more general case, and a great deal of work has been done to try to elucidate the specific principles relevant to a wide range of alternative domains. Although we cannot prove that this approach is misguided, we have the perspective that the underlying machinery and the principles by which it operates are fundamentally the same across all different domains of cognition. While this machinery can be tuned and parameterized for domain-specific uses, understanding the broad principles by which it operates will necessarily be of very broad relevance.

***How the search for domain-general principles is carried out.*** If one's goal is to discover the set of domain-general principles that govern all aspects of human cognition, how best is the search for such principles carried out? Our approach begins with the fundamental assumption that it is not possible to know in advance what the right set of principles are. Instead, something like the following discovery procedure is required:

1. Begin by formulating a putative set of principles.
2. Develop models based on these principles and apply them to particular target domains (i.e., bodies of related empirical phenomena).
3. Assess the adequacy of the models so developed and attempt to understand what really underlies both successes and failures of the models.
4. Use the analysis to refine and elaborate the set of principles, and return to step 2.

In practice this appears to be the approach both of Newell and of A&L. Newell and his associates developed a succession of cognitive architectures, as has Anderson; indeed, Newell suggested that his was only really one attempt, and that others should put forward their own efforts. However, Newell argued for broad application of the framework across all domains of cognition, suggesting that an approximate account within each would be satisfactory. In contrast, we advocate a more focused exploration of a few informative target domains, using failures of proposed models to guide further explorations of how the putative set of principles should be elaborated. To illustrate the power of this approach,



https://doi.org/10.1017/S0140525X03290130 Published online by Cambridge University Press