Complex, Creative and Adaptive Learning
Review of Stellan Ohlsson’s
Deep Learning: How the Mind Overrides Experience

Jared Freeman
Aptima, Inc.

In the past decade, educational institutions of the U.S. military have increasingly focused their attention on developing adaptive warfighters and adaptable organizations. The immediate catalyst is the challenge in Afghanistan, where troops must simultaneously engage the enemy and engage with the people, two populations that intersect with troubling frequency and deadly effects. In my experience, conversations with military educators begin with the present challenge to the individual soldier or Marine, and often turn to historic battles won (or escaped) through the inventive tactics of legendary armies. Research concerning cognitive technologies likewise grapples with the dynamics of individuals as learners, the fluid behaviors of teams in discourse and coordinated action, and the peculiar viscosity of institutional memory over time. Across these and other domains, the phenomenon of adaptability spans the individual and the organization, the instant and the epoch. As such, adaptability represents a core construct for the development of cognitive technologies for the 21st century - technologies supporting cognition across multiple levels that can help us understand and accelerate the process of learning. This is the same, immense range to which Stellan Ohlsson applies himself and three theories of adaptation—specifically, theories of creativity, learning, and belief conversion—in Deep Learning: How the Mind Overrides Experience (Cambridge University Press, 2011).

Deep Learning is impressive not just for its scope, but for the utility of its solutions and the rigor with which Ohlsson tests them. The author, a Professor of Psychology and Adjunct Professor of Computer Science at the University of Illinois at Chicago, opens by defining the scope of his inquiry as complex systems. These stand in contrast to "clockwork" systems, such as mechanical artifacts and planetary bodies, that can be modeled with such precision that we reliably estimate their state far into the future. Complex systems, rivers and social systems for example, exhibit emergent (rather than cyclic) patterns of behavior across multiple levels (e.g., from molecule to stream, from individual to organization and society), cannot be analyzed separately from their environment, and seem to be governed by rules too complex or mutable to codify. Complex systems, says Ohlsson, are in the majority. They are the systems our minds must comprehend, and our minds themselves are such systems. He takes as his mission the development of theories concerning the processes by which the mind adapts, enabling us to escape the bonds of experience to generate new ideas, develop new abilities, and change our beliefs.

To explain creativity, Ohlsson posits a "redistribution theory." Impasses during problem solving generate negative feedback, which decrements the activation of the failed concept. The activation is redistributed to competing concepts, potentially raising a correct one to working memory. We commonly call this experience insight. This theory of creativity has a number of virtues relative to alternatives, such as blind combination of reusable steps or the miraculous gestalt accounts of restructuring of representations. Among these benefits are that redistribution theory defines a process in testable detail, it does so using validated mechanisms of analytical problem solving (heuristic search and subgoaling), and it explains why creativity is rare: because it is massively contingent on the contents and state of memory and on provision of and response to timely, diagnostic feedback. Ohlsson tests this theory in computational experiments at the level of the individual. He then scales these tests in historical and computational analyses to the level of the organization and society. At the higher levels of complex systems, redistribution theory helps account for several
familiar phenomena: the often-slow pace of creativity over time as the inventor elicits and leverages feedback from studies; the rarity of creativity within bureaucracies, which favor the timely delivery of familiar ideas over the slow and unpredictable delivery of novel ones; and the relatively rapid advances achieved in the scientific community, which systematically generates feedback and rewards innovative responses to it.

The second focus of Ohlsson's book is learning. How, he asks, do people acquire accurate, practical knowledge of what to do when? He offers answers at two levels of granularity. At the highest level, he proposes that successful learners differentially tap at least nine learning modes at each of three stages of learning: getting started, mastery, and optimization. The nine modes, a familiar but admittedly incomplete set, are: learning from verbal instruction, from reasoning over declarative knowledge, from analogy to related tasks, from solved examples, from trial and error, from generalization, from error, by eliminating unnecessary steps, and by optimizing processing based on statistical regularities detected over applications of the mastered skill. The author then turns to a particularly dark and difficult corner of the learning research: how we learn practical knowledge given that most practice is erroneous. He first characterizes a large class of these errors (exemplified by over-regularization in early language learning) as mistaken applications of rules that are overly general, rules that fail to specify important constraints or exceptions (such as irregular verbs). These errors are cured, according to "specialization theory," when the learner interprets negative feedback from errors as violations of expectations or constraints, and generates a specialized form of the general rule for use in similar circumstances. The more general form of the rule is retained, and it is used in situations that are novel or lack characteristics of the specialized rule. This theory leverages the well-tested formalism of productions; thus it is rooted in firm ground. It explains the capability of experts to recognize and respond accurately to specific cases, yet handle common and novel cases reasonably well. Above the level of the individual, organizations have developed business processes (e.g., Capability Maturity Model Integration) that systematically identify the exception conditions underlying failures, and drive the organization to specialize its business rules to handle those cases correctly in the future. Thus, this cognitive theory resonates at larger scales.

Why do people hold erroneous beliefs in the face of conflicting evidence? Ohlsson offers an explanation in chapters concerning conversion of beliefs. The author first catalogs the mechanisms of resistance to conflicting evidence. These include selective attention to evidence, discrediting the source of conflicting evidence, and allowing a conflicting case as special case (perhaps as a function of specialization theory). These mechanisms are ubiquitous and effective; there is little reason to hope that gaps in the intellectual and emotional armor against evidence should give way to belief change. Rather, says Ohlsson, the cause of belief change is collision between two beliefs that are incompatible under some deeper belief. Such conflicts are resolved either by rejecting one belief, or in the most extreme cases, altering the core belief to accommodate them both. The victor in these rare conflicts is the belief that has the greatest utility and confidence, thus it is most likely to be "true" in a practical sense (and possibly an absolute one). This "resubsumption theory" is built upon the very foundation of the edifice it replaces. Our compulsion to understand, to find predictability in the world, leads us to invent theories, or to apply theory from one domain to another domain. The proliferation and repurposing of theories produces a complex belief system ripe with conflicts, but it is only rarely that conflicts are discovered and sufficiently contested to change fundamental beliefs. Resubsumption theory thus explains the rarity of belief change, as well as its occurrence.

In his final chapters, Ohlsson takes initial steps to unify the three theories of redistribution, specialization, and resubsumption. This is challenging in part because the theories themselves vary in type, from a neural model (i.e., redistribution theory), to the mechanical logic of production rules (evident in specialization theory), to a conceptual model of, appropriately, concept change (resubsumption theory). He approaches unification not by positing a foundational theory that encompasses the three, but by enumerating ten well-validated cognitive principles or processes that
support them all. Drawn from the canon of cognitive science, these collectively describe mind as a process of ubiquitous cognitive activity over a structured but unbounded space of representations developed through a layered architecture of constrained capacity, an architecture that develops and competes representations to optimize utility over problems over time. Ohlsson's explication of the principles reassures the reader that a coherent if complex system of processes may account for adaptive behaviors. It does not constitute a parsimonious, unified theory; but then the author does not claim it does so.

Ohlsson's exposition of his theories is impressive not just for its scope but for its rigor. For each of the three theories, above, Ohlsson defines the phenomenon, specifies requirements for a sufficient theory, evaluates standing theories against those criteria, explicates his own theory, tests it in computational studies, and identifies the implications and limitations of that theory over scales of time and social organization. The inclusion of computational studies in this volume is particularly welcome. There is a hot debate in the cognitive sciences concerning the need for and utility of a complexity theory of cognition (c.f., TOPICS in Cognitive Science, v4, n1, Jan 2012). Critics of the complexity theorists argue that the complexity movement has failed to define specific models that explain interesting cognitive phenomena. The current work explores complexity, develops predictions from theory, and tests several of them in enlightening ways. The presentation would have been enriched had the author described more human subjects experiments and computational studies. Ohlsson offers too few for my taste, and they are briefly rendered. Perhaps researchers will interpret this as a come on, as a challenge to test the author's theories and to evaluate them against alternatives on at least the criteria he recommends. In this sense, the book has high utility for cognitive psychologists and cognitive science.

Deep Learning invites other readers to action, as well. Those interested in accelerating the development of more adaptive students, staff, and leaders may find inspiration here for new instructional techniques and technologies. Building on Ohlsson's theories, these practitioners might lift students to creative insight by ensuring they master and apply a useful variety of representations to problems; efficiently develop expert knowledge (i.e., specializations of general productions) by applying practical knowledge under conditions that vary systematically and meaningfully; and transform their beliefs by inventing or applying theories that challenge them. Those who lead institutions may be reminded, by Ohlsson’s emphasis on attending to errors, to seek out organizational failures, not just successes. They can systematically learn from these failures by creating more effective, context-specific solutions as well as better theories of the complex systems that are organizations operating in dynamic environments.

**AUTHOR NOTES**

Jared Freeman is Chief Scientist of Aptima, Inc. His work spans the range from model-driven instructional of individuals and teams, to simulation of workers and work processes within organizations, to analysis of behavioral patterns in big data. He received a Ph.D. in human cognition and learning from Columbia University. He can be reached at freeman@aptima.com.

**REFERENCE**
