Unified Theories of Cognition: Newell's Vision after 25 Years

Glenn Gunzelmann (glenn.gunzelmann@us.af.mil)
Cognitive Models & Agents Branch, Air Force Research Laboratory
711 HPW/RHAC, 2620 Q St., Building 852
Wright Patterson Air Force Base, OH 85212-6061 USA

Introduction

It has been 25 years since Unified Theories of Cognition was published (Newell, 1990). In it, Newell outlines a vision to inspire generations of cognitive scientists and cognitive modelers; a quest for theories that provide comprehensive accounts of the human mind. As he put it:

> "A single system (mind) produces all aspects of behavior. It is one mind that minds them all. Even if the mind has parts, modules, components, or whatever, they all mesh together to produce behavior… If a theory covers only one part or component, it flirts with trouble from the start. It goes without saying that there are dissociations, interdependencies, impenetrabilities, and modularities… But they don't remove the necessity of a theory that provides the total picture and explains the role of the parts and why they exist." (Newell, 1990, pp. 17-18).

The intervening years have produced a wealth of research progress on many fronts. One important measure is the number of candidate theories that have emerged. In his book, Newell explicitly pointed to the need for multiple unified theories of cognition to drive progress through model comparison. The 1990's and early 2000's saw a large expansion in theories providing broad accounts of human cognitive capacities (see, e.g., Gluck & Pew, 2005). In addition, there have been some attempts to formally compare alternative theories to address their associated strengths and weaknesses (e.g., Gluck & Pew, 2005; Gonzalez, Lebiere, & Warwick, 2009).

In conjunction with the increase in candidate architectures, the scope of these theories also has broadened, leading to more complete theories that have incorporated many critical aspects of human cognition (e.g., Anderson, 2007; Bach, 2008; Laird, 2012). This was a critical component of Newell's vision, addressing the need to account for a wider array of cognitive mechanisms to provide ever more comprehensive and inclusive accounts of the capacities and limitations of human cognition. It is worthwhile to take stock of these achievements.

Despite the evidence of progress and sustained contributions to cognitive science that have emerged from the pursuit of unified theories of cognition, the zeitgeist has evolved. Opinions differ significantly within the community represented at this conference regarding the current state of cognitive architectures, trends in their development, and where they should go in the future (e.g., Kurup, Gunzelmann, Lewis, Salvucci, & Taatgen, 2012).

In the broader cognitive science community, there is also a tendency to focus on phenomena and challenges that play to the strengths of the modeling formalisms that are used for model development. As McClelland (2009) notes, different approaches are often adopted because they are "particularly apt for addressing certain types of cognitive processes and phenomena. Each has its core domains of relative advantage, its strengths and weaknesses, and its zones of contention where there is competition with other approaches" (p. 25). This perspective is not new. It contrasts with Newell's vision, which stood in opposition to his perception of the prevailing trends in the field. Specifically, Newell perceived that cognitive science had become "too focused on specific issues and had lost sight of the big picture needed to understand the human mind." (Anderson & Lebiere, 2003, p. 587).

The International Conference on Cognitive Modeling is the premier venue for cognitive modeling research. Moreover, it emerged from research pursuits directly aligned with Newell's vision. This year's conference provides a unique opportunity to revisit Newell's vision, and look to the future of our community.

Presenters

The presenters in this symposium will focus on Newell's vision for unified theories of cognition, discuss whether and how that vision drives their research, and comment on the extent to which it still defines an appropriate vision for the community. The participants in the symposium have been selected to represent a cross-section of the community, each of whom will provide a unique perspective on the topic.

Glenn Gunzelmann

Current cognitive architectures capture many of the capacities and limitations of human cognition. However, the current state of the art falls well short of Newell's vision for unified theories. Many critical cognitive abilities identified by Newell remain poorly understood (e.g., perception, language, emotions). In addition, the overwhelming majority of computational cognitive models are based on the implicit assumption that the human mind continually operates in an efficient, effective, and goal-directed manner. Our models do not get hungry, fatigued, angry, or distracted. Part of Newell's vision entailed developing theories that make it "further down the list" of phenomena that characterize the human mind (Newell, 1990, p. 16). Unfortunately, too little research in the cognitive modeling community addresses this challenge today. Instead, unified theories are used increasingly to explain isolated phenomena and validate micro-theories. For cognitive architectures to remain relevant in the future of cognitive science, the community must take seriously Newell's vision, and refocus on the challenges of developing a theory that explains the roles for the various components, why they exist, and how they are integrated to create the human mind.

Paul Rosenbloom

Newell's call for integrated approaches to cognition is as relevant as ever, but broad progress over the past 25 years in both the natural and artificial sciences enables, and even demands, we be even more ambitious today when thinking about integration. Can we build single systems that span from the biological band, through the cognitive and rational bands, up to the social band? Can we complete the processing path from perception and attention, through cognition and affect, out to motor control without arbitrary boundaries between these parts? And can integrated approaches inform us about both natural and artificial cognition? I will discuss how an attempt to answer such questions, toward ultimately yielding what could be called a *grand unification*, has driven the development of the Sigma cognitive architecture and system (Rosenbloom, 2013).

Dario Salvucci

Newell's vision for unified theories of cognition has no doubt stood as the centerpiece of cognitive-architecture research since his seminal "20 Questions" paper (Newell, 1973). In this paper, Newell proposed three complementary activities in this effort: the use of "complete processing models," exemplified by production systems; the analysis of complex tasks, beyond those involved in simple psychological paradigms; and the development of "one program for many tasks," a single model that acts in a variety of task domains. Arguably the cognitive-architecture community has focused largely on the first and second activities, while the third activity has received much less attention. I will discuss some recent efforts (e.g., Salvucci, 2013) that aim to extend the capabilities of cognitive architectures in this third direction.

Iris van Rooij

New formal and conceptual tools for theorizing about cognition have developed since Newell voiced his concerns about experimental psychology in his seminal "20 questions" paper, and proposed specific ways of dealing with them. Using these new tools we can cast our theoretical net even wider than Newell perhaps envisioned. For instance, important advances have been made in theorizing about cognition at a level *above* that of mechanism, viz., what Marr (1982) called the 'computational level' (and Anderson (1990) calls the 'rational level'). I will discuss how theorizing at this level may be useful for addressing a challenge that remains to this day: How to make models that can scale beyond specific experimental tasks and explain cognition in its full domain generality?

Marieke van Vugt

As a relative outsider, it struck me that the adoption of Unified Theories of Cognition (UTCs) is relatively low. Moreover, the community faces criticisms of not being falsifiable. Those criticisms are raised especially by modelers who focus on modeling a single essential cognitive operation. Indeed, I share these concerns, and I think it is crucial that the community develops good methodology for comparing and testing models, and their dependence on model parameters. On the other hand, I think UTCs can provide insight not only in what happens during the trials of a single task but also what happens between those trials as a person gets bored, tired, etc. Furthermore, in the domain of neuroscience UTCs have the potential to describe how different parts of the brain collaborate to guide information flows across different task stages. In the age of big data, both simple mathematical models and UTCs are more necessary than ever.

References

Anderson, J. R. (1990). The adaptive character of thought. Psychology Press.

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?*. Oxford: Oxford University Press.

Anderson, J. R., & Lebiere, C. (2003). The Newell test for a theory of cognition. *Behavioral and Brain Sciences, 26(05)*, 587-601.

Bach, J. (2009). *Principles of Synthetic Intelligence. Psi, an architecture of motivated cognition*. Oxford: Oxford University Press.

Gluck, K. A., & Pew, R. W. (Eds.). (2005). *Modeling human behavior with integrated cognitive architectures: Comparison, evaluation, and validation*. New York, NY: Psychology Press.

Lebiere, C., Gonzalez, C., & Warwick, W. (2009). A comparative approach to understanding general intelligence: Predicting cognitive performance in an open-ended dynamic task. In B. Goertzel, P. Hitzler, & M. Hutter (Eds.), *Proceedings of the Second Conference on Artificial General Intelligence* (pp. 103-107). Amsterdam, The Netherlands: Atlantis Press.

Kurup, U., M. D., Gunzelmann, G., Lewis, C., Salvucci, D. & Taatgen, N. (2012, July). Cognitive architectures: State, trends, and roadmap. Symposium presented at the *19th Annual ACT-R Workshop*. Pittsburgh, PA.

Laird, J. (2012). *The Soar cognitive architecture*. Boston, MA: MIT Press.

Marr, D. (1982). *Vision*. Freeman: New York.

McClelland, J.L. (2009). The place of modeling in cognitive science. *Topics in Cognitive Science, 1(1)*, 11-38.

Newell, A. (1990). *Unified Theories of Cognition*. Cambridge, MA: Harvard University Press.

Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In *Visual Information Processing* (pp. 283-308). New York: Academic Press.

Rosenbloom, P. S. (2013). The Sigma cognitive architecture and system. *AISB Quarterly*, *136*, 4-13.

Salvucci, D. D. (2013). Integration and reuse in cognitive skill acquisition. *Cognitive Science, 37*, 829-860.