CHAPTER

12

Cognitive Architectures for Human Factors in Aviation

Kevin Gluck
Air Force Research Laboratory

INTRODUCTION, MOTIVATION, AND ORGANIZATION

There has been a great deal of investment and resultant progress in the development and evaluation of, improvements to, and comparisons of cognitive architectures over the last several decades. Not all—however, certainly the majority—of that work has taken place since the publication of Weiner and Nagel's (1988) first volume on *Human Factors in Aviation,* so it is understandable both that there was no mention of computational cognitive modeling or cognitive architectures in that first edition, and also that the editors of the second edition are interested in expanding coverage of the text to include this relevant development in the scientific study of human performance and learning.

The overarching interest and motivation for the existence of the aviation human factors community is improving the operational safety of current and future aviation systems. The people serving in the roles of pilots, navigators, maintainers, controllers, or other user-operator positions in all aviation-related sociotechnical systems are both enabled and constrained by their cognitive architectures. By improving our understanding of the human cognitive architecture, we improve our understanding of an important component of the larger system of systems in which those people are performing. Through better understanding of aviation systems, to include the human components of those systems, we can improve their overall performance standards and safety levels. This is why the aviation human factors community should care about basic and applied research on the human cognitive architecture.

The purpose of this chapter is to provide an introduction to cognitive architectures that will be useful to anyone who wants to understand what a cognitive architecture is and who wants some pointers regarding what to read and consider in order to use an architecture effectively in their research. Given the broad range of educational and professional backgrounds among the scientists, engineers, acquisition professionals, instructors, and students that one might reasonably consider to be in the target readership for a volume such as this, I have positioned the chapter as a relatively high-level introduction to the topic. The first section of the chapter provides definitions and characteristics of cognitive architectures to help the reader develop an understanding of what cognitive architectures are and what they are intended to be. The second section identifies recently published key reference materials for those interested in a deeper investigation of the topic than is possible in this single chapter. The third section of the chapter describes current efforts by my colleagues and me at the Air Force Research Laboratory to improve on an existing cognitive architecture. The fourth and final section summarizes challenges and vectors for those interested in evolving cognitive architectures from research programs to applied products that are useful in aviation and elsewhere.

WHAT ARE COGNITIVE ARCHITECTURES?

The concept of *cognitive architecture* and its formal study in humans has its scientific origins in calls for an information-processing psychology and its use in the development of computational process models as means for improving our understanding of human cognition (Simon, 1956; Newell, Shaw, & Simon, 1958, Newell & Simon, 1972). At their core all attempts to implement computational theories of the human cognitive architecture are inherently consistent with the theoretical and methodological position first expressed in Newell, Shaw, and Simon's (1958) *Psychological Review* paper that, ". . . an explanation of an observed behavior of the organism is provided by a program of primitive information processes that generates this behavior." Thus it was that the research agenda was set in motion for cognitive architects, who demonstrate an improving understanding of the mind with computational information processing models that replicate human behavior in simulation.

Anderson (2007) traces the introduction of the actual term *cognitive architecture* into the lexicon of cognitive science, crediting Allen Newell as its progenitor. Anderson's (1983) and Newell's (1990) seminal books provided what may still be the best-known and most widely cited proposals for specific cognitive architectures and their important role as engines of theoretical integration and unification within cognitive science.

Much has been written about cognitive architectures since their inception as a scientific endeavor, and naturally some of that writing has involved attempts at (more or less) concise definitions of what they are. Here is a sampling of such definitions, organized chronologically:

> Anderson (1983)—" . . . a theory of the basic principles of operation built into the cognitive system."
>
> Pylyshyn (1984)—" . . . those basic information-processing mechanisms of the system for which a nonrepresentational or nonsemantic account is sufficient."
>
> Newell (1990)—". . . the fixed structure that forms the framework for the immediate processes of cognitive performance and learning."
>
> VanLehn (1991)—"In general, the architecture of a computing system leaves out details about the implementation of the system and includes only a description of its basic operations and capabilities. An architecture for the mind would describe the way memory and attention operate but it would not describe how they are implemented biologically."
>
> Byrne (2003)—"A cognitive architecture is a broad theory of human cognition based on a wide selection of human experimental data, and implemented as a running computer simulation program."

> Sun (2004)—". . . a cognitive architecture is the overall essential structure and process of a domain-generic computational cognitive model, used for a broad, multiple-level, multiple-domain analysis of cognition and behavior."
>
> Kieras (2007)—"Cognitive architectures are the current form of the traditional computer metaphor in which human cognition is analyzed in information-processing terms."
>
> Anderson (2007)—"A cognitive architecture is a specification of the structure of the brain at a level of abstraction that explains how it achieves the function of the mind."
>
> Laird (2008)—"Cognitive architectures must embody strong hypotheses about the building blocks of cognition that are shared by all tasks, and how different types of knowledge are learned, encoded, and used, making a cognitive architecture a software implementation of a general theory of intelligence."

The details of these definitions vary considerably with respect to how they position cognitive architectures in classic and ongoing debates regarding levels of abstraction, metaphors for mind, and the complicated intersections of neuro-bio-cogno-silico methods and theories. For instance, Anderson (1983), Pylyshyn, Newell, and Sun all associated cognitive architectures with the invariant "basic principles," "mechanisms," and "fixed" or "essential" structure. Byrne, Sun, and Laird approach it from the other side of that coin, explicitly emphasizing domain generality in their definitions, while others left this characteristic implicit. VanLehn stated that cognitive architectures leave out details regarding biological implementation, whereas Anderson (2007) associated cognitive architectures with abstract specifications of the structure of the brain. Note the interesting historical contrast with Anderson's 1983 definition, which does not mention brain structure. Of course, that earlier definition was written well before the Decade of the Brain and more than 20 years before he explicitly adopted the goal of answering Newell's question, "How can the human mind occur in the physical universe?" Finally, both Byrne and Laird included implementation in software as a feature of cognitive architectures, while there is no mention of this feature in the other definitions.

Despite the coverage provided in the set of definitions here, there are a couple of characteristics of cognitive architectures that are missing from that set, and that I think are important and worth emphasizing. One characteristic is that cognitive architectures that are being developed and used in ongoing research efforts are *evolving*. This often is not clear to those outside the cognitive architecture community. Although there are some core architectural features that are similar or identical, the Soar described in Newell's (1990) *Unified Theories of Cognition* is different than the Soar architecture that exists today. Similarly, the ACT theory described in Anderson's (1976) *Language, Memory, and Thought* is different than the ACT-R architecture that exists today. Indeed, both of these architectures' websites reference architectural evolution. The Soar Web site[1] states, "It has been in use since 1983, evolving through many different versions to where it is now Soar, Version 8.6." The ACT-R Web site[2] states, "As the research continues, ACT-R evolves ever closer into a system which can perform the full range of human cognitive tasks: capturing in great detail the way we perceive, think about, and act on the world." Of course, this characteristic holds not just for ACT-R and Soar but also for any cognitive architecture that is being used in an ongoing research program. Their evolutionary nature is a consequence of the intention to continually expand the functional and explanatory breadth and depth of the architectures.

None of the architectures that exist today is considered complete. They all have weaknesses, deficiencies, and idiosyncrasies that serve as pointers toward the next evolutionary adaptation. Rather than being cast aside they are modified and extended. In this manner, cognitive architectures serve as a formal instantiation of progress in cognitive science (Cooper, 2007; Lakatos, 1970; Newell, 1990).

A second characteristic of cognitive architectures that is not represented in the set of definitions above is that their development is often motivated by an interest in *application*. That is, an objective of cognitive architects is often that the architectures have some applied utility. Note the careful inclusion of the adverb "often" here, to reflect the fact that this characteristic is not universally present, or at least there is not explicit evidence of an interest in

[1] (http://sitemaker.umich.edu/soar/home)

[2] (http://act-r.psy.cmu.edu/)

application in the publications and other available materials on every existing architecture. However, the evidence for this is clear in some cases. For example, Anderson (1976) actually ends his book introducing the ACT theory with a statement of the importance of application for his research program by saying, "I would like to conclude this chapter with a remark about one of the ultimate goals I have set for my research efforts . . . that is, that it produce a theory capable of practical applications" (p. 535). Newell's (1990) position was that, "Applications are an important part of the frontier of any theory. . . . A unified theory of cognition is the key to successful applied cognitive science" (p. 498). A third example is found in the EPIC architecture, for which some of the earliest publications (Kieras & Meyer, 1997; Kieras, Wood, & Meyer, 1997) make it clear that applications in system design served an important motivational role in its creation. The fact that this motivation persists is clear on the EPIC Web site[3] which states that EPIC is, ". . . for constructing models of human-system interaction that are accurate and detailed enough to be useful for practical design purposes."

Regardless of which definition and combination of characteristics one prefers as a description of cognitive architectures, there has been enough progress in the creation of cognitive architectures over the course of the last half century, and especially in the last two decades, that lately there have been published a number of reports, articles, comparisons, and collections on this topic. These publications are the focus of the next section, for the benefit of any readers who are interested in more detailed and complete coverage of the state of the science and practice of cognitive architectures than could possibly be achieved in a single chapter on the topic.

RELEVANT RECENT PUBLICATIONS

Review Reports

There have been three noteworthy review reports that focused on, or at least included the topic of, cognitive architectures. In all

---

[3] http://www.eecs.umich.edu/~kieras/epic.html

three cases the reports were requested and/or funded by defense-related organizations. This is no surprise, given that financial support for the development of and improvements to cognitive architectures has originated primarily (though not exclusively) from the military services.

The first of these reviews was a report by Pew and Mavor (1998), published as a book summarizing the conclusions of their Panel on Modeling Human Behavior and Command Decision Making: Representations for Military Simulations. The panel was organized by the U.S. National Research Council, at the request of an organization then known as the U.S. Defense Modeling and Simulation Office (DMSO). Pew and Mavor used the term "integrative architecture," rather than cognitive architecture, in order to accommodate the fact that complete processing systems of this sort include functionalities beyond exclusively the cognitive. Their report reviewed the state of theory and modeling across a variety of important aspects of human information processing from the individual to military unit levels of analysis, with chapters on attention, multitasking, learning, memory, decision making, situation awareness, planning, and behavior moderators (e.g., physiological stressors, intelligence, personality, emotions). Most directly relevant to this chapter is Pew and Mavor's Chapter 3, in which they describe and compare 10 cognitive architectures across these dimensions: purpose and use, underlying assumptions, architecture and functionality, operation, current implementation, support environment, extent of validation, and applicability for military simulations. Their general conclusions were that substantial progress has been made in the formal computational modeling of human behavior, no single one of the architectures provides all that is needed to address the range of the military's interests and requirements in human modeling and simulation, and significant ongoing investment in this area is warranted and advised. Pew and Mavor also made recommendations regarding challenges and future directions. I return to the topic of recommendations later in the chapter, so will not elaborate here.

Ritter, Shadbolt, Elliman, Young, Gobet, and Baxter (2003) responded to and extended the Pew and Mavor (1998) report. Their supplementary review identified several additional architectures either overlooked by or unavailable to the Pew and Mavor report and also emphasized some important challenges and issues that Pew and Mavor did not, such as integration (both across architectures and between architectures and simulation environments) and usability. It is interesting to note, at least to those who follow such things as sources of research support, that preparation of the Ritter et al. report was supported by a combination of the U.K. Defence Evaluation and Research Agency (DERA), the Australia Defence Science and Technology Organization (DSTO), and the U.S. Office of Naval Research (ONR). It was eventually published through the Human Systems Information Analysis Center at Wright-Patterson Air Force Base, making this the most ecumenically supported publication on cognitive architectures and related technologies to date.

Another National Research Council report was published recently which includes material on cognitive architectures (Zacharias, MacMillan, & van Hemel, 2008). This review was requested by the U.S. Air Force Research Laboratory's Human Effectiveness Directorate, with a focus on the state of the science and practice in societal modeling. To their credit, the assembled experts on the NRC review panel recognized the relevance of individual and organizational modeling approaches, given AFRL's interests, and expanded the scope of their study accordingly. Thus, the Zacharias et al. report encompasses individual, organizational, *and* societal (IOS) modeling and simulation accomplishments, challenges, and prospects, oriented around the military's application interests. They propose a four-part framework for IOS models: (1) micro, (2) meso, (3) macro, and (4) integrated, multilevel models. Cognitive architectures explicitly fall into their micro-level modeling category, along with affective models and cognitive-affective architectures, expert systems, and decision theory and game theory. Within the material on cognitive architectures, Zacharias et al. provide brief descriptions of 12 architectures and discuss current trends and persistent issues.

Articles

There are two "must read" articles for those in the aviation human factors community interested in knowing more about cognitive architecture. Conveniently, these articles were written with the

Human-Computer Interaction (HCI) and the Human Factors and Ergonomics communities in mind. The first is Byrne's (2003) article in the Jacko and Sears (2003) *Human-Computer Interaction Handbook.* The paper begins with background information on the characteristics, strengths, and limitations of cognitive architectures and their relevance to HCI, then explains the relationship of cognitive architectures to three systems and theories that have had an impact in HCI research, including the Model Human Processor (MHP; Card, Moran, & Newell, 1983), Cognitive Complexity Theory (CCT; Bovair, Kieras, & Polson, 1990; Kieras & Polson, 1985), and Collaborative Activation-based Production System (CAPS; Just & Carpenter, 1992). Byrne then describes four cognitive architectures that were under active, ongoing development and application within HCI at the time he wrote that article. Those four architectures were LICAI/CoLiDeS (Kitajima & Polson, 1997; Kitajima, Blackmon, & Polson, 2000), Soar (Newell, 1990), EPIC (Kieras & Meyer, 1997), and ACT-R/PM (Byrne & Anderson, 1998). The article ends with some prospective comments on issues, challenges, and the future of cognitive architectures.

A second helpful article for those interested in a historical perspective on developments of and in cognitive architectures is Gray's (2008) article in the 50th anniversary special issue of the journal *Human Factors.* Gray emphasized the use of cognitive architectures in the context of cognitive engineering, making the point that accomplishing this involves acknowledging and taking advantage of the tight coupling between mind and environment. Indeed, the article is actually subtitled "Choreographing the Dance of Mental Operations with the Task Environment." An interesting contribution of Gray's article is his proposed cognitive architecture taxonomy. The taxonomy is intended to help organize and explain the relationships among the 50 or so architectures that exist in the world today. The top-level branch in Gray's taxonomy divides the space of existing cognitive architectures into "Architectures for Developing Cognitive Science Theory" and "Hybrid Architectures for Cognitive Engineering." This is a quite reasonable initial discriminator among the architectures available today, and some of them do fall cleanly into one branch and not the other. Yet we find the situation gets a bit more complicated another level into the taxonomy, where some of the theory-oriented architectures are "Occasionally Used for Cognitive Engineering" and where some of the engineering architectures include "Cognitive Theories of Control of Cognitive Systems." This reflects the fact that many of these architectures exist today at the productive, exciting, bidirectional interplay of basic and applied research.

Comparisons

The Air Force Research Laboratory and NASA both sponsored research efforts in model and architecture comparison within the last decade. These are of relevance to this chapter not only because they involved using and improving on cognitive architectures, but also because both projects focused on task contexts, such as air traffic control and piloting, that are central to the interests of those working in aviation human factors.

AFRL's AMBR Model Comparison Project. The Gluck and Pew (2005) book on modeling human behavior describes the human data, models, and lessons learned from the U.S. Air Force Research Laboratory's Agent-based Modeling and Behavior Representation (AMBR) Model Comparison. This research effort involved four modeling teams using different architecture-based modeling systems (ACT-R, COGNET/iGEN, DCOG, and EASE), a moderator team (BBN Technologies), and several related architecture-based model development, evaluation, and validation efforts over a period of more than four years. The processes and performance levels of computational cognitive process models are compared to each other and to human operators performing the identical tasks. The tasks are variations on a simplified en route air traffic control hand-off task and emphasize multitasking, workload, and category learning. The book is divided into three sections. The first section of the book is background material, including: an overview of the effort, followed by a description of the method and results from the human experiments, the rationale for the choice of tasks, a detailed description of the task software and its dynamics, the human operator requirements, how the software was set up to allow seamless introduction of either a human operator or a computational process model that simulates the human operator, and the way in which the models were connected into the simulation. The second section of the book includes a separate chapter for each of the participating architectures (ACT-R, COGNET, DCOG, and EASE) and the models that were developed with those architectures. At the end of each of these chapters the authors answered a set of summary questions about their models. The last third of the book presents a discussion of the practical and scientific considerations that arise in the course of attempting this kind of model development and validation effort. It starts with a discussion of how the architectures and models were similar and different and how they performed the target tasks as compared with human data. Included are comments on how the results of the models' performances were related to and derived from the architectures and assumptions that went into the models. The last three chapters are of general interest to those working in the area of cognitive modeling, including a chapter that relates the AMBR models of category learning to other models of category learning in the contemporary psychological literature (Love, 2005), a chapter on a variety of important issues associated with the validation of computational process models (Campbell and Bolton, 2005), and the final chapter, which includes reflections on the results of the project and a proposed research agenda to carry the field of human behavior modeling forward (Pew, Gluck, & Deutsch, 2005).

NASA's HPM Project. NASA recently sponsored a Human Performance Modeling (HPM) project within its Aviation Safety and Security Program. The results are published in Foyle and Hooey (2008). There are many high-level similarities between this research project and the AFRL research described previously, primary among them being that the NASA project also involved quantitative and qualitative comparisons among models, in this case developed with four cognitive architectures (ACT-R, IMPRINT/ACT-R, Air MIDAS, and D-OMAR) and a model of situation awareness called A-SA. There also are a variety of differences between the AFRL and NASA efforts. For instance, the tasks used in the NASA HPM project were piloting tasks (navigating around airport taxiways, approach and landing with synthetic vision systems), whereas AFRL's AMBR project used a simplified en route air traffic control task. Another difference is that the NASA project was more applied in orientation. They list as goals of the project: (1) investigating and informing specific solutions to actual aviation safety problems, and (2) exploring methods for integrating human performance modeling into the design process in aviation. The first section of the Foyle and Hooey text includes a chapter that introduces the NASA HPM project, a background chapter on human performance modeling in aviation, and a chapter that describes the simulators and human subjects studies used in their project. The second section of the book provides details regarding the participating architectures, with a full chapter on each. The third section includes a cross-model comparison chapter, a "virtual roundtable" chapter in which the model developers all respond to each of 13 questions, ranging from the more general (e.g., "Why model?") to the relatively specific (e.g., "In terms of supporting the aviation research and development community, what issues and research questions are HPMs best able to address? What issues are HPMs not yet capable of addressing? What will it take to address those issues?"), and a final chapter that includes comments on the achievements of the NASA HPM research project and ongoing challenges for the science and application of human performance models.

Collections

Polk and Seifert's (2002) book, titled *Cognitive Modeling,* is a collection of previously published journal articles pulled together by the editors into a single reference book. Part I of the book covers modeling architectures, which are divided into *symbolic* and *neural network* categories. The "symbolic architectures" include Construction-Integration, ACT, Soar, EPIC, and CAPS. The "neural network architectures" (also sometimes referred to as "approaches" or "paradigms" in their preface) include backpropagation networks, recurrent networks, Hopfield networks, Adaptive Resonance Theory, and Optimality Theory). Part II of the book is a collection of papers on specific use cases of these architectures and approaches, and Part III includes articles on issues and controversies in cognitive modeling.

Gray's (2007a) book on *Integrated Models of Cognitive Systems* is also an edited collection of writings, but it is new material, rather than previously published articles. It is the first book in Oxford University Press' new series on Cognitive Models and Architectures. Gray (2007b) draws a distinction between *single-focus models of cognitive functions* (e.g., visual attention, categorization, situation awareness, working memory) and *integrated models of cognitive systems* (e.g., ACT-R, CLARION, EPIC, Polyscheme, Soar) and emphasizes the complementary and preferably congenial relationships that should exist among people working in those two areas. Gray identifies, and contributors to the book elaborate on, three theory "types" that must be involved in any integrative cognitive modeling effort. Type 1 theories involve an implementation of the control mechanisms among functional components of the system. If one is going to propose a cognitive architecture that integrates multiple functional components into a full, end-to-end, simulation system, it is necessary to specify in detail the relationships among those functional components. How, if at all, do the visual, motor, and knowledge components of the architecture interact? Which contents, if any, of the various functional components are accessible by the other components? What are the roles of the different components in learning new knowledge and skill, in adapting based on performance feedback, and in prioritizing across different possible courses of action? Within Gray's tripartite typology, the implementation details that address these sorts of questions are Type 1 theories of the cognitive architecture. Type 2 theories are implementations of the internal processes within the functional components. What are the representations and processes that are produced and used by vision? By audition? What kinds of memories are available to the system, and how are they created, maintained, strengthened, and/or lost? What are the representations and processes that enable the motor system to take action in the environment? Type 3 theories are implementations of knowledge that grounds the control mechanisms and functional components within a particular context. All by themselves, the theories that are the functional components of the architecture and the theory of how they interrelate can not produce performance in any particular context. There must be at least some knowledge in the system in order to get it moving, and the Type 3 theory addresses the structure and content of that knowledge.

The set of articles, reviews, comparisons, and collections described in this section provide a thorough overview of what has been done and where things stand in computational cognitive modeling, and especially in the development of cognitive architectures. It is an active, vibrant, ongoing research area. The next section of the chapter describes the ways in which my colleagues and I at the Air Force Research Laboratory are working to advance the theory and application of cognitive architectures in ways that are relevant to aviation and other complex contexts.

IMPROVING HUMAN PERFORMANCE AND LEARNING MODELS FOR WARFIGHTER READINESS

The role of the Air Force Research Laboratory (AFRL), like the other service laboratories, is to conduct the basic and applied research and advanced technology development necessary to create future technology options for the Department of Defense. At the Warfighter Readiness Research Division of AFRL's Human Effectiveness Directorate we have a research program focused on mathematical and computational cognitive process modeling for replicating, understanding, and predicting human performance and learning. This research will lead to new technology options in the form of human-level synthetic teammates, simulation-based cognitive readiness analysis tools, and performance tracking and prediction algorithms. Creating a future in which these objectives become realities requires tightly coupled, multidisciplinary, collaborative interaction among scientists and engineers dedicated to overcoming the myriad challenges standing between the reality of the present and our vision for the future.

The Performance and Learning Models (PALM) team was formed to pursue this agenda. Our research approach is organized around a set of methodological strategies with associated benefits. First, we are using and improving on the ACT-R (Adaptive Control of

Thought—Rational) cognitive architecture (Anderson et al., 2004; Anderson, 2007) because it provides important, well-validated theoretical constraints on the models we develop, facilitates model reuse among members of the ACT-R research community, and serves the integrating, unifying role intended for cognitive architectures. Second, we use the architecture, or equations and algorithms inspired by it, to produce precise, quantitative forecasts about the latencies and probabilities of human performance and learning in order to facilitate eventual transition to applications that require such capabilities. Third, we develop models in both abstract, simplified laboratory tasks and in more realistic, complex synthetic task environments in order to bridge the gap between the laboratory and the real world. Fourth, we compare the predictions of our models to human data, in order to evaluate the necessity and sufficiency of the computational mechanisms and parameters that are driving those predictions and in order to evaluate the validity of the models. We are pursuing this research strategy in several lines of research, which are briefly described next.

We have one research line that is entirely mathematical modeling and does not involve a computational simulation component. Progress to date involves an extension and (we think) improvement to the general performance equation proposed by Anderson and Schunn (2000) that allows us to make performance predictions or prescribe the timing and frequency of training, both in aviation-related and other domains (Jastrzembski, Gluck, & Gunzelmann, 2006; Jastrzembski, Portrey, Schreiber, & Gluck, submitted). On the computational modeling side we have research underway in all of the following areas: (1) natural language communication in knowledge-rich, time-pressured team performance environments similar to those encountered in real-world situations, such as unmanned air vehicle reconnaissance missions (Ball, 2008; Ball, Heiberg, & Silber, 2007); (2) a neurofunctional and architectural view of how spatial competence is realized in the brain and the mind (Gunzelmann & Lyon, 2008; Lyon, Gunzelmann, & Gluck, 2008) and how spatial cognition interacts with vision and language to produce situated action (Douglass, 2007; Douglass & Anderson, 2008); (3) implementing new architectural mechanisms and processes that allow us to replicate the effects of sleepiness on the cognitive system, in order to predict what the precise effects of sleep deprivation or long-term sleep restriction will be in a given performance context (Gunzelmann, Gluck, Kershner, Van Dongen, & Dinges, 2007; Gunzelmann, Gross, Gluck, & Dinges (2009)); (4) the interactive dynamics of cognitive coordination for development of a synthetic teammate (Myers, Cooke, Ball, Heiberg, Gluck, & Robinson, submitted); (5) the creation of a metacomputing software infrastructure for faster, broader, and deeper progress in computational cognitive modeling (Gluck & Harris, 2008; Harris & Gluck, 2008; http://www.mindmodeling.org); and (6) a new initiative at the intersection of cognitive modeling and ultra-large-scale software engineering and systems simulation that will create new methods and capabilities that enable the development, exploration, understanding, and validation of computational cognitive process models and software agents (whether in standard ACT-R, some modified version of ACT-R, or some other formalism) on an unprecedented scale.

These ambitious lines of research were carefully selected on the basis of scientific merit and relevance to the U.S. Air Force mission, aviation-related and otherwise. They represent the range of basic and applied research efforts we chose to pursue with the resources made available to date. It is easy to make the case for the relevance of these lines of research to civilian aviation contexts, as well, where capabilities such as natural language communication, spatial reasoning, and vision all are required, where performance and learning take place in complex, time-pressured, dynamic situations, where better performance often requires that people work effectively as teammates, and where stressors like sleepiness may lead to undesirable or even catastrophic degradations in performance. Even this ambitious range of carefully considered and relevant research lines is only a small sampling of the possible investments that could and should be made in improving on the state of the science and practice in cognitive architectures. In the final section of this chapter I review and elaborate on opinions regarding significant scientific and technical gaps in cognitive architectures and recommendations for future investments in this area.

CONSIDERATIONS, CHALLENGES, AND RECOMMENDATIONS FOR THE FUTURE OF COGNITIVE ARCHITECTURE RESEARCH

Many of the published reports and comparisons described earlier included material in which the authors recorded their advice regarding recommendations for future research. In some cases, such as the Pew and Mavor (1998) and Zacharias et al. (2008) National Research Council reports, generating such recommendations was explicitly a motivation for preparing the report in the first place. In all cases, the recommendations are helpful and worthy of attention, especially by those who are in positions of influence or authority at organizations that control basic and applied research and development investments, such as the FAA, for instance. Cognitive architectures are ambitious, incomplete, research programs in progress, with an emphasis on *in progress.* For all of the impressive progress that already has occurred, there is important work yet to be done. Given that the space of possible work that could be done is infinitely large, it seems advisable to review what some of those with experience in cognitive architectures have said about what should be done, both from the perspective of improving on their completeness as unifying theories of human cognition and also from the perspective of improving on their application potential.

Pew and Mavor (1998) summarize their conclusions and recommendations in the form of a research and development framework, with near-, mid-, and far-term objectives, with investment toward objectives at all three timeframes advised to begin immediately and run concurrently. Their framework, also explicitly described as a research program plan, addresses a variety of key issues and challenges faced by the human behavior modeling community, military or otherwise. The Pew and Mavor program plan for the development of models of human behavior is:

- Collect and disseminate human performance data
- Create accreditation procedures (including verification and validation)
- Support sustained model development in focused domains
- Develop task analysis and structure

- Establish model purposes
- Support focused modeling efforts
- Promote interoperability
- Employ substantial resources
- Support theory development and basic research in relevant areas

Ritter et al. (2003) propose 22 specific project-level research activities as important ways to contribute to the science and technology of architecture-based human performance modeling. I point the interested reader to the Ritter et al. report for details regarding each of the 22 projects, choosing here to mention the three higher-level issues which the full set of projects is intended to address. These issues, which Ritter et al. clearly propose as important ongoing or future directions for research in this area, are:

- More Complete Performance (extending the ability/ functionality of cognitive architectures)
- Integration (within architectural modules and also between architectures and external simulations of task contexts)
- Usability

Byrne (2003) describes major ongoing challenges for cognitive architectures in the context of HCI. Although his focus is on challenges and limitations, Byrne's description of these seems written with an eye toward suggested research directions. Indeed, he even elaborates on a few existing research efforts addressing some of the limitations. Byrne mentions:

- Subjectivity (e.g., preference, boredom, aesthetics, fun)
- Social interaction
- The knowledge engineering bottleneck
- Usability, especially for larger-scale models
- Interfacing architectures with simulation environments

The final chapter of Gluck and Pew (2005) describes challenges for and guidance to those who may be interested in conducting a model comparison, followed by a list of improvements that are needed in the theory and practice of computational human modeling

(Pew, Gluck, & Deutsch, 2005). That list of needed improvements included:

- Robustness
- Integrative behavior
- Validation
- Establishing the necessity of architectural and model characteristics
- Inspectability and interpretability
- Cost-effectiveness

The final chapter of Foyle and Hooey (2008) begins with an explanation of the synergistic manner in which they used a combination of human-in-the-loop simulations and human performance model runs to advance their objectives (Hooey & Foyle, 2008). They then describe key modeling advances that were achieved in the context of the NASA project, including modeling the human-environment interaction, visual attention, situation awareness, and human error. They end the chapter by considering important challenges for modeling complex aviation tasks:

- Selecting an architecture (matching architectural strengths to intended application)
- Developing models (knowledge engineering; strategic variability)
- Interpreting model output (Transparency of tasks, procedures, architectures, and models)
- Verification, Validation, and Credibility (matching method, to model, to intended use).

Zacharias et al (2008) listed 10 suggested future directions for research and development with cognitive architectures. Their report was requested by the Air Force Research Laboratory and was written primarily with the military behavioral modeling and simulation communities as the intended audience, but their suggestions regarding research directions are just as relevant when considering the current challenges and prospects for using cognitive architectures to explain and predict human cognition and performance in non-military contexts, such as commercial aviation. Zacharias et al.'s 10 suggested directions are:

- Facilitate architecture development (standardization, interoperability, IDEs)
- Facilitate architecture instantiation (domain ontologies and data repositories)
- Facilitate knowledge base development (address knowledge engineering bottleneck)
- Enhance model explanation capabilities (inspection and visualization tools)
- Address the brittleness problem (larger knowledge bases and learning for robustness)
- Enhance realism (embodiment, emotion, personality, and cultural factors)
- Validation (develop common methods, metrics, and test suites)
- Explore new modeling formalisms (e.g., chaos theory, genetic algorithms)
- Models of groups and teams (via abstraction to group/team-level processes)
- Context and task models (formal estimates of generalizability of architecture-based cognitive models via task and context taxonomies)

This collection of suggested research directions for cognitive architectures is not exhaustive, but certainly is comprehensive. There is plenty of scientific and technical justification for pursuing any one or more of these and other avenues of research in cognitive architecture. In closing this section I will elaborate on a practical issue that must be addressed through additional scientific and technical progress: cost. At Wayne Gray's Integrated Models of Cognitive Systems workshop in 2005, I polled the attendees regarding their estimates of person-years and dollars required to implement and validate an architecture-based model that interacts with a simulation of a moderately complex task. "Moderately complex" in this case would be something like the simplified air traffic control simulation used in AFRL's AMBR project (Gluck & Pew, 2005). The average of the estimates offered by the 13 respondents (all people with some firsthand experience with the development of such models) was 3.4 years and $400,000. That still strikes me as a reasonable estimate today. If it is, in fact, a reasonable estimate, then it is an indication of how far we have left to go as a research community.

That sort of time and money for a single model is prohibitive and clearly stands in the way of transitioning the use of cognitive architectures to the aviation community or any other application context. Cognitive architectures will never have the kind of applied impact that many of us would like as long as model development is measured in a timescale of years and a budget of hundreds of thousands of dollars. Cognitive architectures are still very much *research* programs. They are research programs with a great deal of potential for revolutionizing the way we evaluate and improve on our system designs and training programs, but that revolution has not yet been realized. Perhaps progress toward those revolutionary new capabilities can be measured, at least partly, in terms of the money and time required to develop and validate models with cognitive architectures.

SUMMARY AND CONCLUSION

This chapter has been an introduction to cognitive architectures, intentionally written at a relatively high level of description and review in the hope that the chapter would be an approachable and useful reference for people in the aviation human factors community. As mentioned in the beginning of the chapter, the topic of human cognitive architecture is central to the interests of researchers and practitioners in aviation human factors. This is true whether or not they actually think of themselves as working on or within cognitive architecture theory or application. It is in the theory and application of the human cognitive architecture that an understanding of component processes, phenomena, and stressors, such as sensing and perception, information processing, situation awareness, group and team dynamics, and fatigue (to draw some important examples from the contents of other chapters in this volume), must come together in an integrated, generative, action-producing system. No analysis, representation, or simulation of a planned or operational aviation system is complete without the inclusion of the human component. Cognitive architectures are both theoretical claims about the fundamental nature, strengths, and weaknesses of that human component, and also modeling and simulation research tools for formally exploring the performance implications of changes in next generation system design or training regimen. Progress to date in the development of cognitive architectures has been impressive, yet scientific gaps, technical challenges, and practical issues remain. The research and application contexts in aviation human factors are fertile ground for continuing the evolution of cognitive architectures from promising research programs to useful products for improving the operational safety of current and future aviation systems.

ACKNOWLEDGMENTS

The views expressed in this chapter are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. government. Many thanks to Don Lyon, Bob Patterson, Ed Salas, and Dan Maurino for their helpful suggestions regarding how to improve on a previous version of this chapter.

References

Anderson, J. R. (1983). The architecture of cognition. Cambridge, MA: Harvard University Press.

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

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111*, 1036–1060.

Anderson, J. R., & Schunn, C. D. (2000). Implications of the ACT-R learning theory: No magic bullets. In R. Glaser (Ed.)*Advances in instructional psychology: Educational design and cognitive science: Vol. 5* . Mahwah, NJ: Erlbaum.

Ball, J. (2008). A naturalistic, functional approach to modeling language comprehension. *Paper presented at the AAAI fall symposium on naturally inspired AI*. Arlington, VA.

Ball, J., Heiberg, A., & Silber, R. (2007). Toward a large-scale model of language comprehension in ACT-R 6. *Proceedings of the 8th international conference on cognitive modeling*.

Byrne, M. D. (2003). Cognitive architecture. In J. Jacko & A. Sears (Eds.), *The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications* (pp. 97–117). Mahwah, NJ: Lawrence Erlbaum.

Cooper, R. P. (2007). The role of falsification in the development of cognitive architectures: Insights from a Lakatosian analysis. *Cognitive Science*, *31*(3), 509–533.

Douglass, S. A. (2007). *A computational model of situated action* Unpublished doctoral dissertation. Carnegie Mellon University.

Douglass, S. A., & Anderson, J. R. (2008). A model of language processing and spatial reasoning using skill acquisition to situate action. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th annual conference of the cognitive science society* (pp. 2281–2286). Austin, TX: Cognitive Science Society.

Foyle, D. C., & Hooey, B. L. (Eds.), (2008). *Human performance modeling in aviation*. Boca Raton, FL: CRC Press.

Gluck, K. A., & Pew, R. W. (Eds.), (2005). *Modeling human behavior with integrated cognitive architectures: Comparison, evaluation, and validation*. Mahwah, NJ: Lawrence Erlbaum Associates.

Gluck, K. A., & Harris, J. (2008). MindModeling@Home [Abstract]. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th annual conference of the cognitive science society*, p. 1422. Austin, TX: Cognitive Science Society.

Gray, W. D. (Ed.). (2007a). *Integrated Models of Cognitive Systems*. New York: Oxford University Press.

Gray, W. D. (2007b). Composition and control of integrated cognitive systems. In W. D. Gray (Ed.), *Integrated models of cognitive systems* (pp. 3–12). New York: Oxford University Press.

Gray, W. D. (2008). Cognitive architectures: Choreographing the dance of mental operations with the task environment. *Human Factors*, *50*(3), 497–505.

Gunzelmann, G., Gluck, K. A., Kershner, J., Van Dongen, H. P. A., & Dinges, D. F. (2007). Understanding decrements in knowledge access resulting from increased fatigue. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th annual meeting of the cognitive science society* (pp. 329–334). Austin, TX: Cognitive Science Society.

Gunzelmann, G., & Lyon, D. R. (2008). Mechanisms of human spatial competence. In T. Barkowsky, C. Freksa, M. Knauff, B. Krieg-Bruckner, & B. Nebel (Eds.), *Lecture notes in artificial intelligence #4387* (pp. 288–307). Berlin, Germany: Springer-Verlag.

Gunzelmann, G., Gross, J. B., Gluck, K. A., & Dinges, D. F. (2009). Sleep deprivation and sustained attention performance: Integrating mathematical and cognitive modeling. *Cognitive Science, 33*(5), 880–910.

Harris, J., & Gluck, K. A. (2008, July). MindModeling@Home. *Poster presented at the 30th annual conference of the cognitive science society*. Washington, DC.

Hooey, B. L., & Foyle, D. C. (2008). Advancing the state of the art of human performance models to improve aviation safety. In D. C. Foyle & B. L. Hooey (Eds.), *Human performance modeling in aviation* (pp. 321–349). Boca Raton, FL: CRC Press.

Jacko, J. A., & Sears, A. (Eds.), (2003). *The human-computer interaction handbook: Fundamentals, evolving technologies, and emerging applications*. Mahwah, NJ: Lawrence Erlbaum Associates.

Jastrzembski, T. S., Gluck, K. A., & Gunzelmann, G. (2006). Knowledge tracing and prediction of future trainee performance. In: *The proceedings of the interservice/industry training, simulation, and education conference*. Orlando, FL: National Training Systems Association (pp. 1498–1508).

Kieras, D. (2007). Control of cognition. In W. Gray (Ed.), *Integrated models of cognitive systems* (pp. 327–355). New York: Oxford University Press.

Kieras, D., & Meyer, D. E. (1997). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human-Computer Interaction*, *12*, 391–438.

Kieras, D. E., Wood, S. D., & Meyer, D. E. (1997). Predictive engineering models based on the EPIC architecture for a multimodal high-performance human-computer interaction task. *ACM Transactions on Computer-Human Interaction*, *4*, 230–275.

Laird, J. E. (2008). Extending the Soar cognitive architecture. In *Proceedings of the 1st artificial general intelligence conference*, Memphis, TN.

Lakatos, I. (1970). Falsification and the methodoligy of sceintific research program. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the Growth of Knowledge* (pp. 91–196). Cambridge, UK: Cambridge University Press.

Lyon, D., Gunzelmann, G., & Gluck, K. A. (2008). A computational model of spatial visualization capacity. *Cognitive Psychology*, *57*(2), 122–152.

Newell, A. (1990). Unified theories of cognition. Cambridge, MA: Harvard University Press.

Newell, A., Shaw, J. C., & Simon, H. A. (1958). Elements of a theory of human problem solving. *Psychological Review*, *65*, 151–166.

Newell, A., & Simon, H. A. (1972). Human problem solving. Englewood Cliffs, NH: Prentice-Hall.

Pew, R. W., Gluck, K. A., & Deutsch, S. (2005). Accomplishments, challenges, and future directions for human behavior representation. In K. A. Gluck & R. W. Pew (Eds.), *Modeling human behavior with integrated cognitive architectures: Comparison, evaluation, and validation* (pp. 397–414). Mahwah, NJ: Lawrence Erlbaum Associates.

Pew, R. W., & Mavor, A. S. (Eds.), (1998). *Modeling human and organizational behavior: Applications to military simulations*. Washington, DC: National Academy Press.

Polk, T. A., & Siefert, C. M. (2002). Cognitive Modeling. Cambridge, MA: MIT Press.

Pylyshyn, Z. W. (1984). Computation and Cognition: Towards a Foundation for Cognitive Science. Cambridge, MA: MIT Press.

Ritter, F. E., Shadbolt, N. R., Elliman, D., Young, R. M., Gobet, F., & Baxter, G. D. (2003). *Techniques for modeling human performance in synthetic environments: A supplementary review (HSIAC-SOAR-2003–01)*. Wright-Patterson Air Force Base, OH: Human Systems Information Analysis Center.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*, 129–138.

Sun, R. (2004). Desiderata for cognitive architectures. *Philosophical Psychology*, *17*(3), 341–373.

VanLehn, K. (Ed.). (1991). *Architectures for Intelligence*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Wiener, E. L., & Nagel, D. C. (Eds.), (1988). *Human factors in aviation*. New York: Academic Press.

Zacharias, G. L., MacMillan, J., & Van Hemel, S. B. (Eds.), (2008). *Behavioral modeling and simulation: From individuals to societies*. Washington, DC: National Academies Press.