The ACT-R/PM Project

Michael D. Byrne

Department of Psychology
Rice University, MS-25
Houston, TX 77005
byrne@acm.org

Abstract

Simulating human agents requires going beyond merely simulating the cognitive abilities of a human. It requires an analysis of human perceptual-motor abilities, the task being performed, and the artifact being used to perform the task. ACT-R/PM is a system designed to foster such analyses.

1. Overview

Cognitive Science is a relatively young discipline which is beginning to realize some of the initial promise. Part of that promise is the advancement of well-specified theories of cognition that are comprehensive enough to be a unifying theoretical force (e.g. Newell, 1990, Anderson, 1993). However, the promise of such theories goes beyond purely theoretical concerns. Advanced computational theories of cognition also have great potential for applications, particularly in the domain of human-computer interaction (HCI). For example, computational cognitive theories can provide the baseline model for intelligent tutoring systems, or serve as "virtual users" in the evaluation of an interface. In order to do this, however, the scope of traditional theories of cognition have to be expanded to include concerns about human perceptual-motor capabilities, the tasks people are asked to do (or learn), and the artifacts with which people must interact to do those tasks.

ACT-R/PM is a computational cognitive model designed to help address such questions, with the ultimate goal of making predictions about human performance. However, it is important not to neglect the other constraints on interactive behavior. The range of tasks to which ACT-R/PM has been applied is considerable, but the range of artifacts somewhat more restricted, due in part to the tasks considered but also as a function of software integration issues. There are several quite successful applications of ACT-R/PM, but many issues and challenges remain.

2. Analysis of Interactive Behavior: The ETA Triad

In this framework, interactive behavior is seen as being a joint function of three things: Embodied cognition, the Task, and the Artifact (after Gray & Altmann, in press). As Gray and Altmann describe, traditional disciplines have generally considered these pairwise rather than as a triad. Computer scientists have traditionally considered the design of artifacts to support particular tasks, but often ignored constraints imposed by the capabilities and limitations of the user. Conversely, the experimental psychology community has typically considered the user, but often with artificial tasks or in context that minimize or eliminate the role of the artifact. Ethnographic analysis typically considers the context of artifacts and the tasks, but often overlooks issues rooted in the capabilities and limitations of the human element. Simulating human agents alone is not enough; it is necessary to also consider the environment in terms of tasks and artifacts as well. Computational cognitive modeling forces the analyst to consider all three at once. A modeling system such as Soar or EPIC provides a description of the capabilities and limitations of the user, but contains no task or artifact. However, for a simulation model to run, it must be given both a task to perform and a complete and detailed description of the artifact being used. Verbal theories and abstract AI architectures can exist in a task- and artifact-free vacuum, but to successfully simulate human agents, all three must be considered.

2.1 Embodied Cognition

The model of "embodied cognition" in the triad is obviously the one most related to traditional AI. In models of highly interactive tasks, there are a number of features and capabilities that are clearly required in order to successfully capture performance. Goal-driven behavior and reactivity are critical, and production systems are well-suited for both. The ability to model the effects of time pressure and workload are also key. And, of course, to be truly useful, the simulation models must also be as human-like as possible. Training to an unattainable standard of performance or interface analysis based on an overly artificial model are unlikely to be of much benefit.

Further, cognitive science and AI have been criticized for ignoring not only the tasks and artifacts used by people in accomplishing tasks, but of ignoring perceptual-motor capabilities as well. AI domains like reasoning and planning have traditionally been quite removed from considerations of perceptual-motor capabilities, though this has been less true of late. While a failure to consider perceptual-motor capabilities and limitations may suffice for abstract tasks such as chess, such neglect will not serve in high-performance applications such as air traffic control and in-car navigation systems. As computer systems become increasingly embedded and mobile, the demands they place on our perceptual-motor systems are likely to become increasingly central in understanding interactive behavior. Thus, we need theories and applications that pay more than mere lip service to these issues.

2.2 The Task

The next component to be considered is the Task. Issues in determining the true task to be analyzed are overlooked with surprising frequency. For example, recent studies of WWW behavior—certainly a "hot" topic— (e.g. Nielsen, 1997; Tauscher & Greenberg, 1997) have failed to consider whether the tasks they ask users to do are typical of the tasks users actually use the WWW to accomplish in their normal use of the Web. Optimizing interfaces or training regimens for tasks for which they are not actually being employed is a waste of time and effort. Methods like protocol analysis, contextual inquiry, and ethnographic analysis can be invaluable in understanding the actual tasks in which users are engaged. A thorough analysis of the goal structure/decomposition, such as GOMS, entailed by the actual task plays a critical role in informing the model-building process.

A second important issue is the way by which success in performing a task is measured. Is it time, user satisfaction, or some other metric? In high-performance systems, time and errors are likely to be the most central measures with things like user preference and satisfaction less critical (though still not completely unimportant).

2.3 The Artifact

In general, the HCI community has been much better at understanding and augmenting the Artifact. The artifact determines which operators the user can apply to reach their goals and often plays a central role in maintaining state information for the task. The artifact is the component that is most subject to design—it is often much easier to redesign the device than change the underlying task or change the cognitive, perceptual, or motor characteristics of the user, though there are exceptions. For example, it can be nearly impossible to modify a physical artifact for a space mission during flight.

The design of the artifact is typically fraught with tradeoffs, such as the tradeoff between the goal of making information available to the user and limitations of screen space. In fact, one of the central potential uses of performance analysis such as computational modeling is to help evaluate such tradeoffs.

One of the important pieces of this framework is fidelity to true artifacts. In computing systems, the artifact in the analysis is more often than not a piece of software. In that spirit, one goal of researchers in computational modeling and HCI is the use of the same software both by users and by the computational cognitive models. This can require solving non-trivial software integration problems.

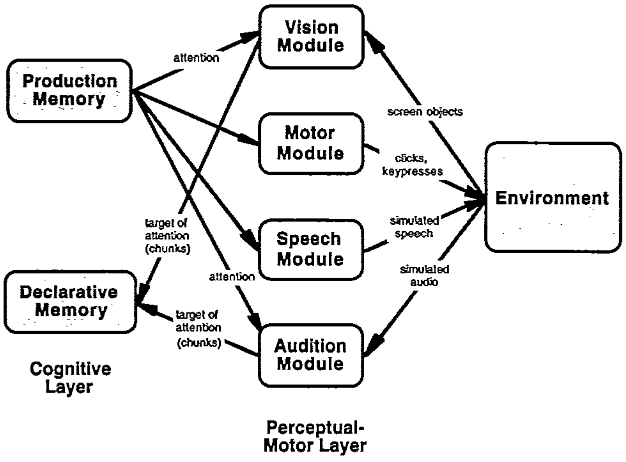

Figure 1. ACT-R/PM System

3. The ACT-R/PM Project

ACT-R/PM is a computational cognitive architecture based on ACT-R (Anderson & Lebiere, 1998), an activation-based production system with many neural-network-like properties. While production systems are not particularly popular among current AI researchers, the ACT-R hybrid approach has been quite successful at modeling many aspects of human cognition (Anderson & Lebiere, 1988), and thus seems appropriate for building human-like agents.

The ACT-R production system forms the "cognitive layer" of the system. Interaction between the cognitive system and a device (simulated or "real" in the sense that it is the same software used by human subjects) is mediated by the perceptual-motor layer. The perceptual-motor system contains an attention-based visual system, an attention based auditory system, and motor and speech systems based on those found in EPIC (Kieras & Meyer, 1996). These modules are not all necessarily detailed or complete models of human capabilities and limitations; many details are abstracted out. However, they are all based on empirical human performance data, which again is deemed critical to the success of constructing human-like agents. These modules are actively being refined as the approximations break down, though the refinement process is decidedly conservative.

ACT-R/PM can form the basis of the theory of "embodied cognition" in the ETA triad given its roots in the human performance literature. However, this literature is not as complete in many respects as one would want it to be. Psychologists have not always asked the right kinds of questions to address issues involved in simulating human agents. ACT-R/PM represents a selective synthesis of the current literature along with numerous guesses and approximations that have up to this point served well. However, as the breadth of application widens, further limitations in our knowledge about the human cognitive-perceptual-motor system will undoubtedly surface. These provide both opportunity and challenge. The opportunity lies in the empirical work these limitations will inspire; the challenges are what to do in terms of system design while these questions are being answered.

The range of tasks to which ACT-R/PM has been applied is substantial, from very simple laboratory experiments on dual-task interference (Byrne & Anderson, 1998) to complex command-and-control like tasks (Schoelles & Gray, 2000; Lee, 2000). Taking on a broad range of tasks provides evidence for the soundness of the underlying theory; it is possible to model large and complex tasks without having to sacrifice veridicality in the details for smaller tasks in more tightly-controlled situations.

Typical GUIs are the primary domain of artifacts to which ACT-R/PM has been applied. This is an excellent starting domain for several reasons. First, it is relatively straightforward to work with this domain. Issues with 3D vision can mostly not be addressed, and the set of actions available to the agent is limited, which constrains the motor domain as well. Second, there is no shortage of interest in this domain, both theoretically and practically. The domain is both rich enough to be interesting and simple enough to be tractable.

ACT-R/PM "understands" simple GUI widgets (labels, buttons) implemented in the Macintosh Common Lisp environment (MCL), and extension to more complex visual environments and dialogs is relatively straightforward. Work is underway at Carnegie Mellon to integrate this functionality with Allegro Common Lisp for the Windows platform, which would also be a win. In principle, it should be possible to link ACT-R/PM to nearly any software interface; in practice this can actually be quite difficult, particularly when access to the underlying code driving the interface cannot be modified. Recent work on this kind of integration problem has been done elsewhere (e.g. Ritter, et al., in press; St. Amant, et al., in press), but ACT-R/PM lags somewhat behind these efforts.

In fact, there are a great many challenges that remain. There are places where the integration of the cognitive layer and the perceptual-motor layer is not entirely smooth, and many unanswered questions about how to manage that integration. Many of these questions revolve around vision: cognitive control of attention vs. exogenous control, representation of the visual scene in a language that is sensible for a production system, making those memories consistent with ACT-R's activation-based system, enforcing visual guidance constraints on aimed movements, and so on.

Furthermore, building a large ACT-R/PM model for a task like air traffic control is not a trivial undertaking by any means. There has been some, though clearly not enough, modularization of the models constructed in ACT-R/PM such that pieces of one simulation can easily be adapted for use in another. The people constructing such models are academic researchers, typically ACT-R specialists, who hand-code knowledge with diligence and patience.

Finally, ACT-R/PM is not unlimited in scope. ACT-R//PM has little to say about numerous things that clearly impact human performance on task to which agents are applied, such as fatigue. ACT-R/PM can be used to model execution time, learning, and error rates in human performance, but has little to say about personal preferences or user satisfaction (though one might be inclined to presume that in the long run, users will prefer interfaces which allow them to achieve their goals with a minimum of time and effort).

Despite these limitations, ACT-R/PM as a project has so far been quite successful. As mentioned, the range of tasks to which it has been applied is substantial, and this range is being expanded by researchers at multiple institutions. Admittedly, most of the projects for which ACT-R/PM has been utilized are more along the lines of basic psychological research than direct application or advances in AI technology, but certainly the project is relevant to both of these concerns. One of the advantages of working in the context of ACT-R is that there is an active research community working on a variety of issues and the underlying cognitive theory is revised to maximize the breadth and veridicality of the theory.

4. References

Anderson, J. R. and C. Lebiére, Eds. (1998). *Atomic components of thought.* Hillsdale, NJ, Erlbaum.

Byrne, M. D., & Anderson, J. R. (1998). Perception and Action. In J. R. Anderson & C. Lebiere (Eds.) *The Atomic Components of Thought* (pp. 167–200). Mahwah, NJ: Lawrence Erlbaum.

Lee, F. J. (2000). Does learning of a complex task have to be complex? A study in learning decomposition. Ph.D. Thesis, Carnegie Mellon University. Pittsburgh, PA.

Nielsen, J. (1997). How Users Read on the Web. http://www.useit.com/alertbox/9710a.html

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

Ritter, F. E., Baxter, G. D., Jones, G., & Young, R. M. (in press). Supporting cognitive models as users. *ACM Transactions on Computer-Human Interaction.*

Schoelles, M. J., & Gray, W. D. (2000). Argus Prime: Modeling emergent microstrategies in a complex, simulated task environment. Paper presented at the 2000 International Conference on Cognitive Modeling.

St. Amant, R., Riedl, M. O., & Zettlemoyer, L. S. (in press). A practical perception substrate for cognitive modeling in HCI. *International Journal of Human-Computer Studies.*

Tauscher, L., & Greenberg, S. (1997). Revisitation patterns in World Wide Web navigation. In *Human Factors in Computing Systems: Proceedings of CHI 97* (pp.399–406). New York: ACM Press.