Integrating Cognitive Architectures with External Environments: Approaches and Contributions to Validation

Glenn Gunzelmann
(glenn.gunzelmann@mesa.afmc.af.mil)
Air Force Research Laboratory
6030 South Kent St., Mesa, AZ  85212

Art Pope (apope@setcorp.com)
SET Corporation
1005 North Glebe Rd., Arlington, VA  22201

Keywords: Computational Model; Virtual Environment; Interaction; Validation.

Introduction
The history of cognitive architectures and computational modeling is rooted in the kinds of tasks and environments that are typical of experimental psychology. For the last 10 years or more, cognitive models have been able to interact directly with small-scale tasks of this sort with relative ease. In contrast, enabling computational models to interact in more complex environments requires more substantial effort. Models that have been developed successfully to operate in such environments have demonstrated the potential for computational cognitive models to provide detailed explanations of human behavior in naturalistic tasks. To further extend the theoretical breadth and cognitive plausibility of computational cognitive architectures, however, they need to be able to interact easily with a wide variety of external environments, both real and simulated. This capability would enable a wider variety of testing of existing architectures, but would also allow a different class of problems to be explored.

All of the panelists in this symposium have experience with developing computational models and linking them to an external environment. They relate some of the issues involved in such efforts, and also discuss their approaches for developing a general infrastructure for integrating computational models with complex real and virtual environments. Tools providing this capability would greatly decrease the time and effort required to develop models that interact within a variety of virtual and robotic environments. They would reduce the need to focus on the technical issues of how to manage the interaction of the model within the virtual environment, thereby allowing modelers to concentrate their efforts on understanding and integrating the cognitive mechanisms underlying human performance. Finally, these efforts promise to facilitate validation efforts as well, by providing more naturalistic and realistic contexts for demonstrating the capabilities of the models.

Moderator: Glenn Gunzelmann

Potential applications of computational cognitive models include acting as synthetic teammates or adversaries in virtual environments and training simulations, as well as serving as intelligent tutors in such environments to provide time-sensitive, individualized feedback and support to individuals as they acquire new skills and knowledge. In all these roles, computational models need the ability to reason in humanlike ways about the tasks that are being performed. In complex real and virtual environments, this entails the ability to represent and process visuospatial information in a psychologically valid manner. This capability requires an infrastructure that provides models with information about the environment in a form that is appropriate for human-level encoding and reasoning about the visuospatial aspects of the space. In addition, models must be able to generate actions within the task environment in the same way as humans, so that those actions are appropriately constrained as well.

Ideally, this infrastructure would be a seamless and nearly invisible part of the modeling activity. Unfortunately, virtual simulation environments are designed to maximize processor efficiency rather than to facilitate integration with cognitive architectures and models. This presents a challenge – even a barrier – to cognitive modelers interested in developing models that interact in such environments.

Art Pope

We describe a framework for linking cognitive models with virtual environment simulations of the sort that are typically used in military training applications. The framework, called Castle, is designed to accommodate a variety of different cognitive models and simulations. It supplies a cognitive model with sensory information obtained from a simulation, and enacts cognitive model motor and other commands in the simulation, while imposing human-like limitations on sensory information processing and control.

The framework is based on a formal analysis of requirements, including surveys of candidate simulations, cognitive architectures, and human performance data. It includes its own graphics and physics engines so that, where necessary, it can supplement an attached simulation in order
to present cognitive models with complete and consistent phenomena. It is based on portable software and distributed processing middleware, allowing the framework to operate on and be accessed from many common computing platforms and languages. Lastly, it includes logging, replication and monitoring mechanisms to facilitate experimentation.

**Robert Wray**

We are currently exploring how to standardize and simplify the functional perception problem in virtual environments. Because a virtual environment (VE) must already unambiguously identify individual objects and their positions in space, many challenges for realizing a complete, functional model of visual perception can be ignored or abstracted. Instead, representations of the ground-truth visual scene can be provided to models, transformed and/or annotated in a manner consistent with neuroscience and psychology.

We have prototyped a VE-model interface that defines 1) a standard, common scene format, which unifies the similar but distinct data representations used to render visual scenes in virtual environments and 2) transformations of the common format for flexible, psychologically-realistic input at three different levels of precision. The interface includes functions to support visual imagery and spatial reasoning as well as visual perception. It builds on existing middleware that simplifies syntactic integration of simulation and model components and directly supports the modeling process via logging, playback, and situation reconstruction.

The existing prototype is being evaluated with a Soar-based model. However, the interface and supporting tools are designed to support a range of models and modeling architectures.

**Bradley J. Best**

Development of a general purpose software interface between cognitive models and 3D virtual environments (VEs) requires innovation in theoretically grounded cognitive abstractions and interfaces, as well as a thoughtful application of software architecture and engineering principles. We are working to embed these cognitive abstractions in a middleware layer that offers "plug-and-play" operation for arbitrary pairings of cognitive architectures and simulation environments, while supporting potential future extensions to robotic environments. This requires identifying the common ground between cognitive models and virtual environments both in terms of spatial representations and in terms of supported spatial processing within the cognitive architectures. In some cases, however, cognitive architectures may lack sufficient processing capabilities to interact in even a minimal way, and thus augmentation of spatial processing must also be addressed.

Our middleware design involves four key components: 1) a processing layer that abstracts the spatial layout of an entity's surroundings in a general VE agnostic way, 2) a processing layer that constrains the information available by considering the limitations of human cognition, 3) a method for mapping entities in various VEs onto a common ontology, and 4) a spatial reasoning module to enable cognitive architectures without such capacities to function in spatial environments. The goal of developing this interface is to advance research exploring interactions of cognitive models with VEs. In addition, this interface will also be useful for applied research that requires embedding agents based on cognitive architectures within VEs.

**J. Gregory Trafton**

Most computational cognitive architectures are extremely limited in what they can “see” and “hear.” We have been working on creating a fully embodied agent that can see, hear, and communicate with others in a cognitively plausible manner. Specifically, we have connected ACT-R’s visual and auditory modules to a set of state-of-the-art vision and auditory AI systems. Our full system, ACT-R/E (ACT-R/Embodied) allows us to see and hear objects and actions in the external world as well as move around in the environment using cognitively plausible spatial representations. Allowing ACT-R/E to see and act upon the world allows us to bring cognitive plausibility to a wide range of interactive, communicative, and environmental tasks and situations.

Connecting ACT-R to the environment through sensors has not been a trivial job. First, many of the AI sensor systems are not perfect (though they are quite good). The imperfections in the sensor systems can lead to problems in the modeling itself (e.g., if a sound or object is perceived but not actually there). Second, the three dimensional world provides a different set of problems than a simple 2D world. Third, how important is cognitive plausibility at the perceptual or action level (i.e., how long it takes for the AI systems to execute)? In order to solve some of these issues, we use multiple simulators at different levels of fidelity.

Our approach allows advantages from many different perspectives. From a cognitive science point of view, we can explore issues of embodiment, representation, and modeling. From a robotics point of view, we are able to facilitate interaction between people and robots because our ACT-R/E system uses similar representations, strategies, and knowledge that people do. From an AI point of view, we are striving toward building a robotic system with human-level intelligence.