The Virtual Driver: Integrating Task Planning and Cognitive Simulation with Human Movement Models

Omer Tsimhoni and Matthew P. Reed
University of Michigan Transportation Research Institute

Reprinted From: Military Vehicles, 2007
(SP-2110)
The Virtual Driver: Integrating Task Planning and Cognitive Simulation with Human Movement Models

Omer Tsimhoni and Matthew P. Reed
University of Michigan Transportation Research Institute

ABSTRACT

Digital human modeling has traditionally focused on the physical aspects of humans and the environments in which they operate. As the field moves towards modeling dynamic and more complex tasks, cognitive and perceptual aspects of the human's performance need to be considered. Cognitive modeling of complex tasks such as driving has commonly avoided the complexity of physical simulation of the human, distilling motor performance to motion execution times. To create a more powerful and flexible approach to the modeling of human/machine interaction, we have integrated a physical architecture of human motion (the Human Motion Simulation Ergonomics Framework—HUMOSIM) with a computational cognitive architecture (the Queueing network model human processor—QN-MHP). The new system combines the features of the two separate architectures and provides new capabilities that emerge from their integration.

INTRODUCTION

Software representations of the human body as a kinematic linkage are widely used in the design and assessment of products and workplaces. These human figure models provide scaling to represent the range of human body size and, to a lesser extent, body shape. Recent research in this area has focused on improvements in posture prediction, motion prediction, and high-level control.

The current use paradigm for digital human models in ergonomics assumes that the software user will simulate a series of tasks that are known in advance. For instance, an evaluation of a vehicle interior using a human figure model could include simulations of the driver reaching to each of the controls and looking at targets inside and outside of the vehicle. The analyst would determine if the targets are reachable by each of a family of manikins differing in size and assess the visibility of displays. In contrast, the utility of models of perceptual and cognitive behavior lies primarily in their ability to produce emergent behavior by responding in a human-like way to sets of stimuli according to “knowledge” (represented variously under different modeling paradigms). Recent advances in models of both physical and cognitive human performance now provide the opportunity for integration of these tools to perform more complete and meaningful simulations.

In a discussion of issues and technologies involving the development and integration of human performance models, McDaniel (2000) argued that large-scale simulations are making satisfactory progress in all areas except human performance. His call for a substantial effort to integrate various computational models also applies to the case of integrating computational cognitive and physical models. McDaniel’s conclusion is motivated by a pragmatic goal of better modeling capabilities. It is not different, however, from the approach driven by purely academic motivation more than 30 years ago.

In 1973, Allen Newell published a book chapter entitled “You Can't Play 20 Questions with Nature and Win” (Newell, 1973), in which he advocated the development of unified theories of cognition and made theoretical unification of micro-models and theoretical constructs an immediate and principal goal. The need for integration of cognitive models with physical models follows a similar rationale. Namely, phenomena that involve both physical and cognitive aspects can be modeled best with models that combine those aspects. There needs to be a technical integration of such existing models and a theoretical unification of the theories that underlie them.

The need for integration between cognitive models and physical models has been addressed for various purposes and to various levels of detail. For example, CSDT—the crew station design tool (Brett, Bzostek, and Jia, 2005)—integrates OpenInventor (a 3D graphics environment) with IMPRINT (a task network modeling tool) while allowing designers to visualize the results using Jack (an anthropometric human figure modeling tool) and optimize their choices of controls and displays and their position in the cockpit.

Zhang (2003) connected a cognitive-performance modeling environment with Safework, a human figure model to conduct a preliminary feasibility study. The Integrated Performance Modeling Environment (IPME), which uses a discrete-event simulation paradigm to simulate human task selection and execution, was linked to Safework. The human figure model in the 3D graphics environment simulated motions commanded by the IPME and Safework.

The opportunities for productive integration between physical and cognitive models have improved due to progress in both areas. Considerable progress on the high-level control of figure models has been made over
the last several years. Badler et al. (2005) presented a
vision for a comprehensive system of control of avatars
implemented as the Human Model Testbed. The
approach builds on other developments including
Parameterized Action Representation (Balder et al.
1999) and recent progress in obstacle avoidance (Zhao
et al. 2005). This work has found both direct and
indirect application in the developments reported in the
current paper. Commercial software developers are
also addressing the need for integrated, high-level
control of human simulations. Raschke et al. (2005)
described the Task Simulation Builder, an approach to
task programming in the Jack™ human modeling system
that incorporates aspects of the Parameterized Action
Representation. Recently, researchers in the Human
Motion Simulation Laboratory have developed a new
approach to motion simulation and analysis, the
HUMOSIM Ergonomics Framework (Reed et al. 2006).

The long-term objective of the current project is to create
an integrated suite of computational models that
produces quantitatively realistic simulations of vehicle
occupants performing multiple tasks, including driving,
remote piloting, and other communication and control
tasks. We plan to conduct these simulations using
networks of commercial, academic, and military software
that provide an effective and efficient development and
deployment environment. The outcomes of such
simulations provide accurate assessment of task timing,
physical and mental workload, safety, and performance,
primarily—but not limited to—the context of driving.

We describe here, as a first step towards that goal, the
integration of a physical architecture of human motion
(the HUMOSIM Ergonomics Framework) with a
computational cognitive architecture (the Queueing
network model human processor—QN-MHP) to create a
new, integrative approach to the modeling of
human/machine interaction. The new system combines
the features of the two separate architectures and
provides new capabilities that emerge from their
integration.

METHODS

HUMOSIM Framework

The Human Motion Simulation Laboratory at the
University of Michigan has developed a suite of
algorithms that improve the realism of the postures and
motions produced by these figures. Based on analysis
of thousands of human motions captured in the
laboratory, modular implementations of these algorithms
are organized into a framework that simulates the
coordinated motion of the human body performing gaze,
reach, and object transfer tasks in a standing or seated
environment (Reed et al. 2006). The HUMOSIM
framework is designed to be applicable to any figure
model used for ergonomic analysis, and the algorithms
have been demonstrated in Jack, Safework/Delmia, and
3DSSPP. The current work uses a reference
implementation of the Framework in the commercial
Jack™ human modeling software.

The HUMOSIM Framework provides control of human
motions at the level of task primitives such as reach,
grasp, touch, and gaze. The whole-body coordination
necessary perform these tasks is produced by the
Framework. In typical usage, the software user
analyzes tasks by issuing commands to simulate
particular actions (look at a control, reach to the
control). Under the current integration effort, these
commands are issued by a computational model of
human cognition, rather than specified a priori by the
software user.

Fig. 1. Illustration of the reference implementation of the
HUMOSIM framework in Jack™

QN-MHP

Queueing Network-Model Human Processor (QN-MHP)
(Liu, Feyen, and Tsimhoni, 2006, Wu and Liu, 2006) is a
computational architecture that integrates two
complementary approaches to cognitive modeling: the
queueing network approach and the symbolic approach
(exemplified by the MHP/GOMS family of models,
ACT-R, EPIC, and SOAR). Queueing networks are
particularly suited for modeling parallel activities and
complex structures. Symbolic models have particular
strength in generating a person's actions in specific task
situations. By integrating the two approaches, QN-MHP
offers an architecture for mathematical modeling and
real-time generation of concurrent activities in a truly
concurrent manner. QN-MHP expands the three discrete
serial stages of MHP of perceptual, cognitive, and motor
processing into three continuous-transmission
subnetworks of servers, each performing distinct
psychological functions specified with a GOMS-style
language. Multitask performance emerges as the
behavior of multiple streams of information flowing
through a network, with no need to devise complex, task
specific procedures to either interleave production rules
into a serial program or for an executive process to
interactively control task processes.
The queuing network modeling approach has been successfully used to generate human behavior in real time, including simple and choice reaction time, psychological refractory period, visual menu search, driver performance, and transcription typing (Liu, Feyen & Tsimhoni, 2006; Lim & Liu, 2004, Wu & Liu, 2004a, 2004b). Fig. 2 shows the structure layout of the QN-MHP.

**Integration Approach**

Two modes of integration between QN-MHP and HUMOSIM have been implemented. First, as a feasibility test, an open loop was implemented. In the open loop integration mode, QN-MHP generates a list of motor actions and eye movement requests. The HUMOSIM Framework reads this list at a later time and produces a visualization of the list using Jack. The HOMOSIM framework does not return any information to the QN model.

This basic level of integration is useful in demonstrating the cognitive model’s actions and is used as a visual tool for verification and validation of both models and of the communication protocol between them.

The second, more elaborate integration mode involves continuous communication between the two model with two simulation programs and environments (ProModel and Jack) connected via socket-based (TCP/IP) client server communication. In the closed loop integration mode, the QN model sends motor action requests to the HUMOSIM framework and monitors task completion, providing the QN with a feedback loop analogous to human somatosensory feedbacks. The QN is able to monitor where the hands are, for example, what they are currently trying to achieve (e.g., reaching for a control) and how far they’ve progressed toward their goals. This information allows the QN to schedule tasks that require the same physical resource (gaze or a hand) and allows the QN to respond to changes in task timing that result from physical aspects of the interaction. For example, if a control is moved farther away, the reach task will take longer to execute and the QN will respond appropriately, possibly by delaying the next task.

**RESULTS**

The integration builds upon the transfer of information between the two architectures in two phases: the task analysis phase and the runtime phase. Fig. 3 lists the required steps for the integrated model.

**STEP 1: TASK ANALYSIS**

In the task analysis phase, the physical ergonomics and cognitive factors are analyzed by each of the respective systems. Analyzed physical information such as the relative position of a display and the expected hand movement time to a cluster of controls is transferred to the cognitive model. Analyzed cognitive information such as the set of possible hand movements and their sequences are then transferred back to the physical model prior to running the simulation. The cognitive task analysis involves a detailed walk through the possible actions that the modeled human will take to achieve any of a set of possible goals. The analysis uses an NGOMSL-style “task description” roughly following the GOMS method of analysis of human performance (Card, Moran, and Newell, 1983) by identifying Goals, Objects, Methods, and Selection rules.

**STEP 2: NAMING**

Common names are assigned in both programs to objects of interest. After this step is completed, only the physical model deals with the actual position of the objects relative to the modeled person. The cognitive model refers to the objects by their abstract name and need not identify their absolute or relative position. Any information about objects required by the cognitive
model (e.g., can the human see an object?) can be obtained through processes that simulate a human’s information acquisition, e.g., try to look at an object and report if it was possible to obtain an unobstructed line of sight.

Fig. 4. Example of matching physical attributes to abstract symbols.

STEP 3: GENERATING THE MODEL SETTING

A model setting is defined as a specific combination of modifiable parameters. The parameters are categorized into the cognitive properties of the human (e.g., perceptual and cognitive response time as correlated to age, preferred strategies and driving style), the physical attributes of the human (e.g., height, body link dimensions, motor speed, weight of armored clothing), and attributes of the environments (e.g., difficulty of the primary task, positions of controls, traffic density). The modeler may choose to simulate a single setting (e.g., use one set of parameters of a typical operator in a normal environment), or to simulate in succession several settings that consist of various combinations of parametric changes (e.g., several driver heights and control positions, or ).

STEP 4: PRE-RUN

In the pre-run step, the physical model pre-runs the actions that are required for all the set of specific parameters as defined in the current scenario and stores those actions in its memory. In addition, it reports back to the cognitive model the completion time and physical effort for each of the targets. For example, it reports the reach time and effort for a control and the glance time and effort for a display.

STEP 5: INITIALIZATION

Before the programs begin simulating the scenario, the physical model sends the times and efforts to the cognitive model, which, in turn, stores them as preliminary estimates into its long term memory. Work on biases in the estimation of reach and gaze times and efforts will later complement this information by assigning estimated biases and rounding to the data stored.

STEP 6: RUNTIME

In the runtime phase, information is exchanged in real time between the models via a TCP/IP connection. The cognitive model initiates movements and notifies the physical model to execute them. The physical model executes and returns the completion time and current status, at which time the cognitive simulation continues to the next steps. The concept of limited resources, which is simulated as limited capacities in some of the nodes of the Queueing Network, limits the number of some simultaneous movements requested from the physical model while allowing others. The exact prediction of time to target and preferred path, along with an estimation of the exerted physical effort is communicated to the cognitive model in simulation time and guides the choice of concurrent and future actions.

Simulation clock and Timing

The integrated models use a synchronized clock that progresses in discrete predefined time steps (e.g., 50 ms) in simulation time but can move faster or slower than real time, depending on computational requirements by the models. This approach accommodates the execution of computational intensive processes while allowing the integrated model to move faster than real time otherwise. For visualization purposes, the simulation can be recorded and replayed in real-time steps. Additionally, the model can be forced not to exceed real-time so that visualization is possible except when delays occur due to high computational demand. As computer power continues to increase in the next few years, we expect that our model will be able to work in real time, even for difficult computations.

INTEGRATION EXAMPLE

As an example of the potential of the integration, the system was used to model vehicle drivers who interacted with an in-vehicle display system while driving on a road with several degrees of steering difficulty based on road geometry. The modeled virtual driver initiated glances and hand movements to drive a vehicle on a simulated road in the UMTRI driving simulator. At the same time, it adapted those motions to perform in-vehicle tasks based on the perception of its driving performance, its anthropometric data, and the positions of the in-vehicle display and controls. Table 1 lists the modeling steps as described above and provides specific examples to the actions taken by each model and the communication between the models.
Table 1. Example of modeling steps for the integrated model

<table>
<thead>
<tr>
<th>Step</th>
<th>QN</th>
<th>Jack</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task Analysis</td>
<td>Steer the vehicle Retrieve position from Navigation system (press button, read map…)</td>
<td>Survey the absolute and relative position of knobs and displays.</td>
</tr>
<tr>
<td>Naming</td>
<td>RoadFar, RoadNear Knob1, NavDisplay</td>
<td>Get positions of named objects</td>
</tr>
<tr>
<td>Model Settings</td>
<td>Young driver, difficult road, no traffic</td>
<td>95% male Cab layout 1</td>
</tr>
<tr>
<td>Pre-Run</td>
<td></td>
<td>Pre-run of the scenario. Estimate time and effort</td>
</tr>
<tr>
<td>Initialization</td>
<td>Enter time and effort estimates to LTM INITIALIZE 0.0 50 ms</td>
<td>GazeTargetList ReachTargetList</td>
</tr>
<tr>
<td>Runtime</td>
<td>Steer .5 rad, Reach Knob1 2 s</td>
<td>Status 12.05 …</td>
</tr>
</tbody>
</table>

**FURTHER SYSTEM INTEGRATION**

The integration between the cognitive and physical model is the first step towards an integration of a more extensive set of modules, using interfaces that in some cases have already been established. In addition to the two models described above, we plan to integrate a model of the driving environment that is based on the commercial DriveSafety driving simulator at UMTRI. A two-way integration of the QN-MHP with the UMTRI driving simulator has previously been demonstrated (Tsimhoni & Liu, 2003). The purpose of the integration with the UMTRI driving simulator in the current effort is threefold. First, the UMTRI driving simulator provides a straightforward method of defining realistic road models and traffic scenarios in a way that is consistent with our design of human factors simulator experiments. Second, it provides real-time input of vehicle dynamics as the models are running. Third it serves as a valuable visualization tool for the actions taken by the human figure driver.

An additional piece of the integrated model is a higher level model of task analysis. Although all the cognitive tasks can be modeled directly with the QN model, it might be useful to add a task network analysis tool that models the higher level decision making and strategies. Fig. 5 shows an abstract sketch of the integration of QN-MHP as the cognitive model with the HUMOSIM framework as the physical model and the UMTRI driving simulator as the vehicle/traffic model, and a task analysis model. The task analysis model could include consideration of multiple operators and systems whose performance may affect the focal simulation. The cognitive/physical linkage developed in the current work could be used to simulate multiple interacting humans functioning in a task environment. Under the current implementation, multiple Jack figures controlled by unique instances of the QN-MHP could interact with a representation of the physical world through the motor actions mediated by the HUMOSIM Ergonomics Framework.

**DISCUSSION AND CONCLUSION**

Recent advancements in the modeling of both cognitive and physical aspects of human performance have made possible the initial steps toward integrated models. Even the preliminary work reported here has highlighted some advantages of the combined approach. In particular, a visual examination of the driving motions in Jack produced by instructions from the QN-MHP provides a valuable first-level check of the plausibility of the model performance. Any unrealistic movements that are observed may be indicative of problems with the simulation.

Looking forward, the preliminary integration highlights the advantages of representing and manipulating the cognitive, perceptual, and physical elements of the simulation in the domain that provides the most efficient interface. Current approaches to modeling physical task elements (e.g., reaching to a control) within cognitive simulations usually model the activity with an expected duration, possibly with a stochastic aspect, and potentially as an algebraic function of other parameters such as body size. In contrast, representing the task in the physical domain allows the effects of many task parameters that have complex, nonlinear effects on task duration and difficulty to be modeled more directly. Taking the example of a reach to a control, the physical simulation will start with drivers of various body sizes appropriately placed with respect to the control, so that people of different sizes have different reach distances as a function of vehicle interior geometry. Changing the location of the control will affect the reach distances for all simulated operators appropriately. Simulating the encumbrance produced by heavy clothing will be more accurately simulated when the result is dependent on the joint angles required for different operators. The
consequences of display location for gaze direction, and the consequent variation in the ability to detect peripheral stimuli, are also more readily and more accurately modeled in a three-dimensional environment than using simple functions within a cognitive simulation.

The physical simulation also benefits from having the sequence of task elements emerge from the task requirements, and their interaction with the environmental scenarios, rather than being scripted in advance. Creating a physical interaction script for a complex scenario is time-consuming and error prone in the absence of a model of how a human would sequence the elements to achieve goals.

The current work uses implementations of the QN-MHP and HUMOSIM Framework in ProModel and Jack environments, respectively, but both models are defined independently of any particular implementation environment. Consequently, the current effort can be viewed more as the development of a modeling methodology than of a single tool. The modular architecture is intended to allow components to be swapped out as needed to achieve particular applied objectives.

The work presented here bears superficial similarities to so-called “artificial intelligence” or AI engines used to control characters in video games used for entertainment and training. Without discounting the substantial capabilities of the rule-based methods used to guide these characters, using a validated cognitive architecture such as the QN-MHP along with data-based motion models enables more complex behaviors to be simulated with greater confidence in the fidelity of the results. The models used for the current work are based on an understanding of the underlying mechanisms that produce the behavior, rather than a direct mapping between stimuli and responses.

The preliminary work presented in this paper is focused on driving, but the methodology could be applied more generally to any task scenario that can be simulated in cognitive and physical model components. The HUMOSIM Framework provides algorithms for standing task elements analogous to the seated motions (reach, grasp, gaze) as well as walking and acyclic stepping. Consequently, it would be feasible to simulate multiple humans collaborating to accomplish a task that involves moving around in a three-dimensional environment, interacting with physical components and each other.

The current approach does not place a high emphasis on real-time or human-in-the-loop operation. All of the current application scenarios envisioned for the model are best served by a simulation-time approach that allows the model to run faster or slower than real-time as necessary to accomplish the objectives. Nonetheless, real-time or human-in-the-loop operation is feasible if the subsystems can be made to operate fast enough. The current performance relative to real time suggests that this is an issue addressable at the hardware level (using faster computers) rather than one requiring significant software improvements.

ACKNOWLEDGMENTS

The model integration work reported in this paper was sponsored by the Automotive Research Center at the University of Michigan. Research in the Human Motion Simulation Laboratory is supported by the partners of the Human Motion Simulation program at the University of Michigan. The current HUMOSIM partners are DaimlerChrysler, Ford, General Motors, International Truck and Engine, the United States Postal Service, and the U.S. Army Research and Development Engineering Command (RDECOM). TRW, Johnson Controls, and Lockheed Martin have also supported the program. UGS is a Technology Partner of the HUMOSIM program.

REFERENCES


**CONTACT**

Omer Tsimhoni, Ph.D.
e-mail: omert@umich.edu
Matthew P. Reed, Ph.D.
http://mreed.umtri.umich.edu/

The University of Michigan
Transportation Research Institute
2901 Baxter Road
Ann Arbor, MI 48109-2150, U.S.A.