A FRAMEWORK FOR SIMULATING HUMAN COGNITIVE BEHAVIOR AND MOVEMENT WHEN PREDICTING IMPACTS OF CATASTROPHIC EVENTS

Mary Court
Jennifer Pittman

School of Industrial Engineering
University of Oklahoma
Norman, OK 73019, U.S.A.

Christos Alexopoulos
David Goldsman
Seong-Hee Kim
Margaret Loper
Amy Pritchett

School of ISyE
Georgia Institute of Technology
Atlanta, GA 30332, U.S.A.

Jorge Haddock

Lally School of Mgmt. and Technology
Rensselaer Polytechnic Institute
Troy, NY 12180, U.S.A.

ABSTRACT

Our nation has seen an increased need to train its civil authorities and emergency personnel under life-threatening scenarios where human life and critical infrastructure are assumed to be at risk. This training is typically obtained or re-enforced via (human) performance-based tests. At issue is the ability to accurately simulate the scenarios without exposing personnel or human test subjects to injury. In addition, these performance-based tests carry a large monetary cost, and certain scenarios are so complicated, catastrophic or rare that any performance-based test is unrealistic. Our paper outlines the research that must be conducted to develop a framework for modeling and analyzing risk-assessment and decision making when evacuating large populations. The research is aimed at extending an existing construct for simulating passenger and crew behavior during aircraft evacuations, to larger populations, and relies upon rare-event simulation methods, parallel-and-distributed simulation and agent-based simulation.

1 INTRODUCTION

Training of civil authorities, rescue and medical personnel is typically conducted or re-enforced via (human) performance-based tests. The challenge is to provide a virtual environment that accurately simulates the emergency scenarios without exposing the personnel or human test subjects to injury (e.g., toxic or dangerous environments). Current trends in information technology (IT) research are to provide rescue personnel with improved intelligence for increased situational awareness and team-building skills, or to replace authority figures with robotics or sensor technologies (Krane 2001, Sycara and Lewis 2000). While these advances have been commendable, authorities have not been provided with the means to analyze the impact large populations have on disaster mitigation, recovery and evacuation plans. As a result, our country's national, regional, state and local planning methods rely on static urban evacuation plans. Thus, authorities lack the capability of evaluating the plans in terms of their ability to accommodate population growth, urban growth and aging populations. In addition, since the plans are static in nature, there is no convenient means to conduct designs of experiments for determining the factors that influence population survivability. The situations in which such dynamic analysis and planning is needed are vast: logistical support for earthquake victims, highway infrastructure and vehicle movement before hurricane arrival, national airspace congestion during inclement or hostile weather conditions, alternative transportation logistics in response to the possibility of terrorist attack (e.g., near a nuclear power facility), and civilian population movement during and after wartime hostilities. Clearly, the need extends well beyond our nation's borders and includes the ability to respond to world events.

To address the gap in current IT capabilities for evacuation planning and analysis, this paper provides an outline of the research that must be conducted to develop the infrastructure and databases to model the physical, cognitive and emotional (e.g., bonding) characteristics of human populations so that authorities will be able identify, train and plan for the factors that influence population survivability. The approach is to extend an existing construct for simulating passenger and crew behavior during aircraft evacuations, to larger populations, and relies upon research for running and analyzing such emergency systems efficiently — including rare-event simulation methods, fast-time simulation using parallel and distributed simulation and agent-based simulation. The goal is to use this research to establish a framework for a system capable of providing authorities with virtual environments for: (i) testing and

training in realistic, operational environments (involving threats to human lives) with no actual risk to human subjects, (ii) evaluating conceptual designs of various plans and the robustness of those plans as populations grow and age, and (iii) optimizing emergency procedures and personnel requirements. In particular, the system must:

- Analyze various population movements under several types of rare events (e.g., earthquakes) that require no changes to its source code. This will give various end-users (DOT, FEMA, National Oceanic and Atmospheric Administration, airline carriers, etc.) the capability of analyzing movement without requiring them to be versed in a specific programming language (or the model's logic).
- Run in real time. This will support virtual training and give authorities the ability to analyze the procedures they follow, and the impact those procedures have on evacuation time and population survival rates.
- Be able to conduct simulations of both non-emergency and emergency movements. This supports the issue of model validation.
- Consider relationships among populations (e.g., spouses, parents and children) and the ability to age populations.
- Consider the impact authority/rescue personnel have on populations. This feature will allow "population management" to be explored, such as determining the optimal number of rescue personnel and the optimal locations of medical personnel.
- Offer dynamic behavior as opposed to behavior that is fixed at the time of model execution. That is, the prototype system must allow the behavioral characteristics of the populations to change over time. For instance, when humans are exposed to (or spend more time in) a life-threatening environment, they become more willing to take risks (e.g., jumping out of windows to escape fire when all exits are blocked).
- Take into account dynamic and toxic environments and consider the physical and psychological effects those environments have on human behavior.
- Support simulation output analysis, designs of experiments, and sensitivity analysis.
- Provide animation of population movement, to support model validation and training.
- Operate on PC platforms, perhaps linked as a network. This will ensure that the model is portable.
- Have a high degree of user friendliness — so that the model can be used by a wide set of end-users.

2 RESEARCH APPROACH

To achieve the objectives established for the system, research must be conducted to extend an existing construct for evacuation modeling (Court 1999) to larger human populations. This involves updating the construct with the advances made in the areas of modeling of large-scale systems, information technology development, simulation modeling and statistical output analysis. This equates to performing basic research in running and analyzing such systems efficiently — in other words, rare-event simulation methods, variance reduction schemes, fast-time simulation using parallel and distributed simulation, agent-based simulation, and IT design and evaluation.

2.1 The Existing Construct

In 1997, software capable of simulating passenger movement during the evacuation of an MD-80 aircraft was released to the Federal Aviation Administration (FAA). The software is based on an actor-centered, object-oriented (O-O) framework developed by Court and Marcus (1997) and was commissioned by the FAA to transition the performance-based testing of aircraft certification into a simulation-based test. While other evacuation models existed, for instance the Gourary Associates model (Gourary 1993, 1994), AIREVAC (Aviation Research Corporation 1994), and air-EXODUS (Galea, Owen, and Lawrence 1997), all failed to meet the FAA's criteria for developing simulated egress models (Court and Marcus 1997).

The actor-centered approach used in the software is based on the Burns and Morgeson (1988) construct for simulating systems involving endogenous decision-making that draws upon an O-O paradigm. Their work proposes describing the system in terms of a suite of actor classes (object classes) whose endogenous decisions impact the performance and behavior of the system. They suggest a model wherein all actors, including pseudo-actors (environment), follow an actor-centered description. Each actor class requires data structures (assets, attributes and vulnerabilities) and methods (cognitive and physical capabilities), where:

- Assets are marked characteristics and attributes are descriptive characteristics. The actor's own assets and attributes comprise the actual state, while the perceived state is the actor's perception of its surroundings (the environment and other actors). The state of the actor is the combination of the data structures of the actual and perceived state.
- State data is the input to the cognitive inference engine of the actor.
- The actor can physically move (transfer) or change (transform).
- Transformation takes place by modifying the actor's assets/attributes.
- Vulnerabilities represent degradation to the actor's capabilities via the reduction or destruction of the actor's assets.

- An action space for cognitive capabilities and activities describes the decision set and state of each actor.
- A cognitive event (decision) is capable of (i) delaying decisions, (ii) invoking physical activity, and (iii) changing the action space.
- By delaying decisions and changing the action space, an actor then has the ability to "change its mind" (non-monotonic reasoning).

The actor-centered description proposed by Burns and Morgeson (1988) is not a pure O-O paradigm, since knowledge bases (production rules and heuristics) and inference engines are utilized for achieving each actor's cognitive activity. This approach would equate to developing data, knowledge, and method structures for each actor (passenger) and pseudo-actor (environment); but adopting this approach in 1997 would have inhibited the objective of supporting real-time simulation on a stand-alone PC platform, as established by the FAA (Marcus 1994a, 1994b). In response to this problem, an alternate framework (Court 1999) was used to incorporate the actor-centered description for passengers and the crew: a pure O-O paradigm that avoids the use of knowledge bases and inference engines. In this approach, a new set of class objects was defined that contains (i) the data describing the environment that the actors need to make decisions, and (ii) the applicable rules (functions) for the actors when making the decisions. In essence, the actors obtain their data and functions by copying these new objects, or parts of the new objects, into their own class definition when making decisions.

The class objects defined for the framework are the Coordinate, Navigate, Advancement, Path, Responsibility, Panic and Block objects. These objects replace the knowledge bases and inference engines of Burns and Morgeson (1998), while the Passenger and Crew objects generate the actor instances representing the passengers and crew, respectively. The environment (pseudo-actors) is represented by the Fire, Smoke, and Toxicity objects. The Synchronizer object is used to advance the simulation clock, keep track of the state of the environment, and invoke the ordered movement of actors and pseudo-actors. An overview of the construct is described as follows:

- The Coordinate object maintains the aircraft-cabin architecture and generates, through its coordinate functions, a map of the aircraft cabin for the Passenger, Crew, Fire, Smoke, and Toxicity objects. The map contains information on all of the exit, row, and aisle locations.
- The navigate functions of the Navigate object allow the actors and pseudo-actors the capability of choosing headings (directions) for movement.
- The Path object is called upon to generate possible paths for the actors (Passenger and Crew ob-

jects) and pseudo-actors (Fire, Smoke, and Toxicity objects) based on their positions, headings, and cognitive abilities to access the environment.
- The functions of the Advancement object are used to move actors and pseudo-actors to their requested positions.
- Data from the Block object is used to keep actors and pseudo-actors from moving into inaccessible positions. The Block data types consist of architectural (seats, walls, etc.), human (passenger and crew), and environmental (fire, smoke, and toxicity) obstacles.
- The Panic object influences the actor's ability to reason and react. The Panic object is used to degrade the actor's assets/attributes and thus, represents vulnerability. Degradation is realized by the inability to call on other objects (send messages or copy), such as the Advancement object.
- The Responsibility object is used to bind objects together. This is one of the vehicles used to establish a psychological profile for each actor, as well as a means to distinguish flight attendants and crew members from passengers. A flight attendant is expected to assist and direct passengers during an evacuation. In this paradigm, a flight-attendant actor has access to the internal data of other objects. The amount of internal data sharing depends on the relationship type and the amount of responsibility an actor has toward another actor (e.g., mother and child).

How 'capable' the actor is at using the Coordinate, Navigate, Path and Advancement objects depends on its physical and cognitive capability objects. For example, the possible paths a passenger can construct, and the number of times a new path is generated, depends not only on how often the Path object is called upon, but is a function of the actor's actual and perceived states. Thus, path generation is a function of the type of evacuation being performed (certification or accident reconstruction) and the actor's (i) immediate environment (fire, smoke and toxicity levels), (ii) ability to access its current path and blockages, and (iii) time spent in hazardous environments.

The construct supports pre-defined biological hierarchies, but allows distinction between objects within the same biological class. For example, although males have many similar physical characteristics, they may not have the same physical capabilities or cognitive reasoning abilities. Thus, a distinction based on technical knowledge can be made between a male passenger and a male flight attendant. That distinction is incorporated through the ability to copy the Coordinate, Navigate, and Advancement objects. The flight attendant is expected to have knowledge of the aircraft's configuration and, therefore, has more access to the functions and data of the aforementioned objects than

does an average passenger. As another example, consider a passenger traveling alone versus one traveling with an infant. The parent is bound to the child and therefore, would be expected to ensure that the child is evacuated safely. In this construct, when the mother or father actor is generated, a copy of the bonding function from the Responsibility object is copied into the actors. The parent actor is now tied to the child actor, and the child's state is input to the cognitive objects of the parent actor.

2.2 Expanding the Construct

While the construct supports the development of aircraft-cabin egress models by accounting for (i) a varied passenger and crew profile, (ii) a wide variety of aircraft cabin configurations, and (iii) the capability of simulating various hazardous environmental conditions, it is clearly applicable to all types of population movement. However, the construct must now be expanded in order to efficiently and accurately simulate the movements of large populations and the various evacuation scenarios.

2.2.1 Research in Cognitive Performance Analysis and Agent-Based Simulations

Court's work in 1999 represents the first effort to transition simulation modeling for evacuation planning from high-level interactions of catastrophic event scheduling (Farahmand 1997) to interactions of human movement and behavior. That is, algorithms for supporting evacuation planning via simulation prior to 1999 focused on the catastrophic event, and not on the human 'victims' (actors) that are impacted by the event. Advances in information technology have shifted towards human interaction and behavior, but this research is aimed at the rescue personnel/authority figures, not the victims. A void now exists in evacuation planning and emergency preparedness: research must be aimed at predicting the impact catastrophic events have on all actors (human population and authority figures); and human cognitive, physical and emotional behavior must be modeled to predict the impact such behavior has on evacuation planning and population survivability.

We can draw on several studies when undertaking cognitive performance analysis. One set comes from the general field of cognitive engineering, which studies human cognitive behavior in naturalistic settings — including settings in which time-pressure and safety concerns can determine cognitive performance (e.g., see Klein et al. 1993 on Naturalistic Decision Making, or Hutchins 1995). For example, Chen and Pritchett (2001) study human responses in real-time simulations to emergency situations, including pilot responses to time-critical alerts, in-flight system faults, and emergencies requiring re-planning of a flight to effect an immediate landing. In addition, Ockerman and Pritchett (2000) consider methods of encouraging the de-velopment of "good" procedures that establish effective — but not inflexible or brittle — processes.

The other set of studies comes from agent-based simulations of socio-technical system behavior with human behavior represented as cognitive agents (Pritchett, Lee, and Goldsman 2001). Following the widely recognized need for an ecological approach to human cognition, in which cognitive activities are seen less as a response of isolated, internal constructs and more as a response to the broader physical and social environment (e.g., Vicente 1999, Hutchins 1995, Hollnagel 1993), these simulations have focused on explicitly including a representation of the environment and the structure it imposes on human cognitive activity. This model of the environment includes both physical objects and those social and process elements that enable and constrain the human agents.

Of particular interest is examining the range of cognitive activities likely within the diverse population. Following the general Cognitive Control Model (CoCoM) proposed by Hollnagel (1993), four general types of behavior are possible, ranging from Scattered (corresponding to panicked and uncontrolled), to Opportunistic (corresponding to behavior responding to available options in the environment), to Tactical and Strategic (corresponding to more purposeful behavior capable of considering goals and limitations, and planning for the near- and far-future). Different types of behavior are expected out of different agents depending on the resources (e.g., training and experience, fore-planning) and demands (e.g., concern for own safety and the safety of others, lack of time, lack of information). This framework provides a general structure that can be incorporated into the agents used here to provide a simple representation of the variety of behaviors expected in emergency situations from the general populace, from trained workers on the ground, and from planners and managers of the tactical and strategic situation.

The expansion of the construct will also require (i) the implementation of advances in information technology to ensure that the various scenarios will run in real time, and (ii) basic research in simulation and statistical analysis to address the issues involved when simulating rare events (e.g., the ability to run designs of experiments on such systems).

2.2.2 Information Technology to Expand the Simulation Construct to Support Real-Time Execution and Virtual Environments

The construct must be expanded to allow large-scale, discrete-event/continuous simulations to support (i) real-time execution and (ii) realistic virtual environments (animations) for validation and training. But how will the resulting large-scale, discrete-event/continuous simulations of the various evacuation and emergency scenarios be implemented? Unfortunately, there are several problems associ-

ated with the implementation of sequential discrete-event simulations. The first problem to address is the execution time — the larger the physical system being simulated, the longer it takes to execute. Consider a simulation of the commercial air traffic in the United States, where the simulation models the behavior of each plane and each airport. Given that the simulation resides on one machine, it may take enormous amounts of time to run, since only one action is executed at any one time. In this case the simulation user will have to wait long periods of time to receive results from the computation. Further, virtual environments need the simulation to run fast enough to make the computer-generated world "look and feel" like the real system. Again, this may not be possible when the computation is hosted on a single machine.

The second problem to address is interoperability, i.e., the simulation's ability to interoperate with other simulations. In commercial air space, there are numerous types of aircraft in operation. In a sequential simulation of the air traffic, these aircraft may be modeled the same way but have different attributes to distinguish between types. Suppose, however, that different manufacturers have developed simulations (different models) for different types of aircraft. In order to use these simulations in the air traffic simulation, they would have to be ported to a single machine. This may be a costly effort. Further, if one simulation is used to model the entire air space, there may not be adequate resources to represent objects at the desired level of detail. This would force the simulation to represent the behavior of the physical system in a more abstract way than is needed.

A solution to these problems is to distribute the simulation across multiple processors. In other words, divide the simulation into smaller models (e.g., each simulation equals one airport) and execute the smaller models concurrently across multiple processors. Multiprocessor systems may themselves be divided into two categories: parallel processors and distributed systems (Fujimoto 2000). Parallel processors are characterized as tightly coupled homogeneous processors that share the same main memory space. Distributed systems are a loosely coupled collection of processors that communicate by message passing. If the computation is executed on five processors concurrently, the execution time can be reduced by up to a factor of five. Also, distributing the computation across multiple processors allows simulations from different developers to be used together in a single simulation environment. This is particularly useful in virtual environments where multiple simulations are needed to represent the diverse set of objects in the computation. Reduced execution time and interoperability are two of the primary benefits of parallel and distributed (P&D) simulation.

One of the most fundamental aspects of simulation programs is the concept of time. In P&D simulations, it is important to coordinate the advancement of time and ensure that a simulation's clock and its view of global time are managed consistently (Loper 2002). Otherwise, events may not be correctly ordered before they are processed by the simulation — resulting in temporal anomalies. Time management mechanisms are used by P&D simulations to ensure that the temporal aspects of the system are correctly reproduced. The importance of correctly ordering events depends on the simulation application. In some simulations, impossible event orderings can cause a simulator to fail; in others, non-causal event orderings may be acceptable because they are not perceptible to human participants. The result is that two types of time management can be used in these simulations: one that imposes strict ordering of events and one that imposes no ordering of events.

While the major, yet basic, research effort must focus on developing theory and practice to establish "how much" time management is needed for evacuation and emergency scenarios, we feel that this effort will also have applications for conducting other large-scale, discrete-event/continuous simulation in real time.

2.2.3 Statistical Analysis Research to Support Simulation Output Analysis for Rare Events

Simulation has proven to be a reasonable and feasible means to undertake the analysis of novel emergency and evacuation strategies; simulation is also well suited for the study of the safety impact of new preparedness measures, evacuation methods, National Air Space (NAS) configurations and procedures, etc. Further, the use of simulation will enable the ability to make quick, statistically valid inferences. However, these inferences will be challenging to generate for several reasons:

- The volume of data recorded by the simulations can be overwhelming; it can include, at frequent intervals, detailed records of the inner-workings of the agents (e.g., human entities and resources), descriptions of interactions between agents, and records or statistics of the emergent system behavior as a whole.
- During emergency and safety analyses, many of the system metrics of interest are rare (e.g., nuclear power plant incidents); classic design of experiment (DOE) and Monte Carlo methods require a prohibitive number of simulation runs to generate accurate and precise assessments of these events.
- Because the behavior of interest is typically an emergent system behavior, little a priori knowledge of effects may be known before the simulation runs, reducing the efficacy of a priori DOE methods for reducing the number of runs.

To address these issues, statistical methods for analyzing the data for rare events and for efficiently conducting experiments that compare different solution strategies for

various interesting system configurations must be developed. These methods can be applied during run-time as well as for a priori DOE. That is, these methods can create a closed-loop controller of the simulation, reducing the number of simulation runs required in response to the evolution of the statistical analysis during the data collection phase. Ultimately, these methods will be implemented into the simulation architecture as a controller of simulation configurations during runs, and as a controller of the number of runs used to make statistical inference — thus providing automatic mechanisms to drive simulations so as to make valid statistical inferences with a parsimonious number of runs.

Specific efforts will be needed to expand upon current techniques for (1) rare-event analysis — how to efficiently generate and analyze realizations of extremely rare events — and (2) ranking and selection (R & S) of competing configurations — how to determine which of several competing solution strategies is the "best". The following subsections give details on the research approach required to build upon current statistical techniques (which are usually limited by restrictive assumptions that will need to be relaxed), and how we anticipate implementing these methods as automatic mechanisms within the simulation architecture.

2.2.3.1 Rare-Event Analysis

An important task is the design of statistical methods and simulation tools capable of analyzing the potential for rare, catastrophic events. Since existing methods for statistical analysis of rare events currently apply to simple queuing systems and medium-size reliability systems, we must develop new methods that can accommodate the complex, highly interactive characteristics of the behavior of populations during emergencies (e.g., after an earthquake) or the behavior of large-scale distributed systems like the NAS. (Note that this task is tightly coupled with the R & S procedures to be described next.)

The basic approach to emergency and safety analysis relies on data provided by simulations. For example, we observe the NAS as it evolves in a simulation over a finite time interval to estimate the probability p that a rare event (e.g., an "accident") will occur during that time interval. We model the NAS as a multivariate stochastic process X, whose components denote the state of its entities (e.g., location of aircraft, active communications, etc.). The standard Monte Carlo method generates a large number of independent realizations (sample paths) of X and estimates p by the proportion of realizations that result in an accident. Unfortunately, as the event probability p approaches zero, the number of samples required to derive a precise estimate becomes prohibitive. The importance sampling (IS) technique attempts to address this issue by artificially making the occurrence of the rare event more frequent (Alexopoulos and Shultes 1998, 2001). Its ultimate objective is the computation of an estimator of p with smaller mean squared error and bounded relative error. Technically

speaking, IS aims at selecting an alternative probability distribution (change of measure) that reduces the computational cost, subject to a desired accuracy. During the last decade, a variety of IS techniques addressing the estimation of performance measures of highly reliable Markovian systems with repairs have used heuristic, but effective, approaches. Under some conditions that are often hard to verify, these failure biasing (FB) methods yield highly precise estimators. However, they appear to perform poorly in systems with a substantial amount of component redundancy, such as is common in the NAS.

To address the erratic behavior of FB methods in large-scale reliability models, Alexopoulos and Shultes (2001) proposed a new set of techniques — balanced likelihood ratio (BLR) methods — that use adaptive importance sampling transition probabilities aimed at canceling terms of the induced likelihood ratio. (One can think of the likelihood ratio as the product of ratios of transition probabilities within X associated with individual events.) These methods also yield estimators with bounded relative error and large efficiency improvements in systems with more than 100 components. An additional advantage of BLR methods over the FB methods is their robustness in systems with dependent component failures (e.g., failure or error propagation through the NAS).

To carry out effective rare-event analysis, we must achieve two goals. The first is the design of BLR-like methods appropriate for the efficient estimation of rare-event metrics during system design. A link between the reliability models that form the current application domain of the FB and BLR methods and the models under investigation can be established when one considers events taking the system towards a catastrophe as "failure" events, and corrective events as "repair" events. The second goal is the incorporation of the rare-event analysis methods into the simulation code, in parallel with the R & S methods described next. With this in mind, developing the interface to "drive" the simulation equates to:

- creating appropriate initial conditions,
- modeling the data as the appropriate stochastic process X,
- automatically simulating conditions or configurations that are more likely to generate events of interest,
- configuring the simulation to these "event-likely" conditions,
- eliminating conditions found that do not incur events of interest, and
- ending the simulation when certain criteria of statistical accuracy and precision are met.

2.2.3.2 Ranking and Selection (R & S)

Ultimately, the goal of our simulation methodology is to compare and contrast (or, in more formal statistical terms,

rank and select) between different evacuation strategies. Traditionally, post hoc analysis is coupled with a priori DOE, where the DOE serves to dictate the fixed or random effects each configuration is subjected to, as well as the number of runs within each configuration.

Standard methods for a priori DOE face several difficulties in our context. First, we often have little foreknowledge of emergent system behavior resulting from changes in agent behavior or system configuration — beyond that provided by extrapolation from related configurations, expert opinions and anecdotal evidence. As a result, we tend to fit the number of runs to the time available, risking both (i) realization of too little data (or data distributed poorly across configurations or conditions) for statistically valid inferences, and (ii) collection of excessive data from too many simulation runs. Second, common DOE (including many R & S) techniques have restrictive distribution and independence assumptions for the data that do not hold for the output of many simulations. Therefore, we note the need for a number of advances in R & S research:

1. One must develop R & S procedures suitable for the simulation of complex emergency systems. These procedures must not be limited to restrictive distributional and independence assumptions (see, for example, Bechhofer, Santner, and Goldsman 1995; Goldsman et al. 2002; and Kim and Nelson 2001, 2004).

2. One must then instantiate appropriate selection procedures into analysis tools linked with our simulation architecture. The idea is to develop the capability to drive the number and initialization of the simulation runs for statistically valid conclusions about the inferred intermediate metrics — all in a closed-loop manner. For instance, a typical selection procedure dictates that we take a first stage of observations from the competing designs; the procedure then calculates how many additional simulation runs must be taken to determine the "best" alternative with a high probability of correct selection. Ideally, the R & S procedures will drive the simulator and automatically command the simulation runs required to collect the remaining required data.

3. It would also be of great interest to extend R & S procedures to enable comparison of a larger number of emergency scenario configurations. That is, current R & S procedures are typically used to compare only a small number of designs. See, e.g., Nelson et al. (2001), Boesel, Nelson, and Kim (2003), and Pichitlamken and Nelson (2001) for promising approaches on conducting selection in the presence of many competitors.

Our general comments have been intended to support safety analysis for emergency systems using agent-based simulation. However, they also have broad applicability. For example, other data sets describing evacuation behavior may be available that can be analyzed or data-mined using the same methods as described here. Likewise, these mechanisms and statistical methods may also be applicable to other large-scale, distributed systems.

3 CONCLUSIONS AND FUTURE RESEARCH

While the paper has focused on the expansion of the construct in terms of its efficiency to handle large-scale simulations, designs of experiments and the ability to use agent-based simulations for cognitive analysis, other research must also be conducted. The framework must provide for a population that can grow and age — so as to ensure that evacuation planning and emergency preparedness is capable of identifying the factors that influence survivability for large populations across several emergency scenarios. One can only imagine the ability to safely evacuate states, such as Florida, where large retirement communities already exist.

Integral to the validation of the system will be obtaining and developing data sets to be used for establishing baseline performance. Several data sets and case studies exist in the literature that describe in detail a catastrophic event, the movement of the population as it responded to authority figures, the impact of the event on infrastructure, etc. For example, several articles and data sets have been released on hurricane Floyd (Sattler 2000, Covan et al. 2000, Whitehead et al. 2000) and a wealth of information on natural disasters can be found on the FEMA website, <http://www.fema.gov> (e.g., general disaster statistics, major disaster declaration totals by state from 1972-2000, and maps of disaster declarations since 1964). In addition to FEMA, several other centers exist to assist in obtaining case studies (e.g., the Natural Hazards Research and Applications Information Center at the University of Colorado, the National Oceanic and Atmospheric Administration National Weather Service, and the National Weather Service's Storm Prediction Center).

REFERENCES

Alexopoulos, C. and B. C. Shultes. 1998. The balanced likelihood ratio method for estimating performance measures of highly reliable systems. In *Proceedings of the 1998 Winter Simulation Conference*, ed. D. J. Medeiros, E. F. Watson, J. S. Carson, and M. S. Manivannan, 1479-1486, Piscataway, New Jersey: IEEE.

Alexopoulos, C. and B. C. Shultes. 2001. Estimating reliability measures for highly-dependable Markovian systems using balanced likelihood ratios. *IEEE Transactions on Reliability* 50(3): 265-280.

Aviation Research Corporation. 1994. Development of ARCEVAC: An Emergency Evacuation Simulation Model - Beta Version 1.0. Transport Canada Technical

Report: rjs-92-122, Transportation Development Center, Montreal, Canada.

Bechhofer, R. E., T. J. Santner, and D. Goldsman. 1995. *Design and Analysis of Experiments for Statistical Selection, Screening and Multiple Comparisons.* New York: John Wiley and Sons.

Boesel, J., B. L. Nelson, and S.-H. Kim. 2003. Using ranking and selection to 'clean up' after simulation optimization. *Operations Research* 51: 814-825.

Burns, J. R. and J. D. Morgeson. 1988. An object-oriented world-view for intelligent, discrete, next-event simulation. *Management Science* 34 (12): 1425-1440.

Chen, T. L. and A. R. Pritchett. 2001. Cockpit decision aids for emergency flight planning. *Journal of Aircraft* 38 (5): 935-943.

Court, M. 1999. Commercial aircraft-cabin egress: The current state of simulation model development and the need for future research. *Simulation* 73 (4): 218-231.

Court, M. and J. Marcus, 1997. Use of object-oriented programming to simulate human behavior in emergency evacuation of an aircraft's passenger cabin. In *Proceedings of the 1997 International AGARD Conference on Aircraft Fire and Safety*, 341-346.

Covan, E., M. Rosenkeotter, C. Richards, and A. Lane. 2000. The impact of hurricane Floyd on the elderly residing in four southeastern North Carolina counties. Conference on Social, Physical, and Economic Impacts on Groups and Individuals, East Carolina University. May 24-26, 2000.

Farahmand, K. 1997. Application of simulation modeling to emergency population evacuation. In *Proceedings of the 1997 Winter Simulation Conference*, ed. S. Andradóttir, K. J. Healy, D. H. Withers, and B. L. Nelson, 1181-1188, Piscataway, New Jersey: IEEE.

Fujimoto, R. 2000. *Parallel and Distributed Simulation Systems.* New York: John Wiley and Sons, Inc.

Galea, E. R., M. Owen, and P. Lawrence. 1997. The role of evacuation modelling in the development of safer air travel. In *Proceedings of the International AGARD Conference on Aircraft Fire Safety*, 1-12.

Goldsman, D., S.-H. Kim, W. S. Marshall, and B. L. Nelson. 2002. Ranking and selection for steady-state simulation: Procedures and perspectives. *INFORMS Journal on Computing* 14: 2-19.

Gourary, B. S. 1993. Simulation of evacuation from a single-aisle airplane: A comprehensive report. Gourary Associates Inc., Montclair, New Jersey.

Gourary, B. S. 1994. PC-based simulation of the evacuation of passengers from a transport airplane. Gourary Associates Inc., Montclair, New Jersey.

Hollnagel, E. 1993. *Human Reliability Analysis: Context and Control*. London: Academic Press.

Hutchins, E. 1995. *Cognition in the Wild*. Cambridge: MIT Press. (Excerpt: Chapter 7 "Learning in Context", pp. 287-316.)

Kim, S.-H. and B. L. Nelson. 2001. A fully sequential procedure for indifference-zone selection in simulation. *ACM TOMACS* 11: 251-273.

Kim, S.-H. and B. L. Nelson. 2004. Selecting the best system, Chapter 18 in *Elsevier Handbooks in Operations Research and Management Science: Simulation* (eds. S. G. Henderson and B. L. Nelson), Elsevier, forthcoming.

Klein, G. A., J. Orasanu, R. Calderwood, and C. E. Zsambok (eds.) 1993. *Decision Making in Action: Models and Methods*. Norwood, New Jersey: Ablex.

Krane, J. 2001. Military tests software agents for quick intelligence, *Washington Times*, p. 10, October 1, 2001.

Loper, M. L. 2002. *Approximate Time and Temporal Uncertainty in Parallel and Distributed Simulation*, Ph.D. Dissertation, College of Computing, Georgia Institute of Technology, Atlanta, Georgia.

Marcus, J. 1994a. A review of computer evacuation models and their data needs. Technical Report: DOT/FAA/AM-94/11, Dept. of Transportation/Federal Aviation Administration, Washington, DC.

Marcus, J. 1994b. Technical Meeting: FAA Evacuation Model Parameter Sensitivity Study, School of Industrial Engineering, University of Oklahoma, Norman, Oklahoma, November 10, 1994.

Nelson, B. L., J. Swann, D. Goldsman, and W. Song. 2001. Simple procedures for selecting the best simulated system when the number of alternatives is large. *Operations Research* 49: 950-963.

Ockerman, J. J. and A. R. Pritchett. 2000. A review and re-appraisal of task guidance: Aiding workers in procedure following. *International Journal of Cognitive Ergonomics* 4 (3): 191-212.

Pichitlamken, J. and B. L. Nelson. 2001. Selection-of-the-best procedures for optimization via simulation. In *Proceedings of the 2001 Winter Simulation Conference*, ed. B. A. Peters, J. S. Smith, D. J. Medeiros, and M. W. Rohrer, 401-407, Piscataway, New Jersey: IEEE.

Pritchett, A. R., S. M. Lee, and D. Goldsman. 2001. Hybrid-system simulation for National Airspace System safety analysis. *Journal of Aircraft* 38 (5): 835-841.

Sattler, D. 2000. Repeated exposure to hurricanes and willingness to evacuate: Implications of the hurricane Floyd experience. Conference on Social, Physical, and Economic Impacts on Groups and Individuals, East Carolina University. May 24-26, 2000.

Sycara, K. and M. Lewis. 2000. Integrating agents into human teams. In *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting*, 413-417.

Vicente, K. 1999. *Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work.* Mahwah, New Jersey: Lawrence Erlbaum.

Whitehead, J., B. Edwards, M. Van Willigen, J. Mailo, and K. Wilson. 2000. Hurricane evacuation behavior: A preliminary comparison of Bonnie, Dennis, and Floyd. Conference on Social, Physical, and Economic Im-

pacts on Groups and Individuals, East Carolina University. May 24-26, 2000.

AUTHOR BIOGRAPHIES

CHRISTOS ALEXOPOULOS is an Associate Professor in the School of Industrial and Systems Engineering at the Georgia Institute of Technology. He received his Ph.D. in Operations Research from the University of North Carolina at Chapel Hill. His research interests are in the areas of applied probability, statistics, and optimization of stochastic systems. He is an active participant in the WSC, having been Proceedings Co-Editor in 1995. He is the Simulation Department Editor of *IIE Transactions*. His e-mail address is <christos@isye.gatech.edu>.

MARY COURT is an Associate Professor in the School of Industrial Engineering at the University of Oklahoma. She received her BSIE from Rensselaer Polytechnic Institute in May 1984, her M.S. in Operations Management from Rensselaer Polytechnic Institute in May 1989, and her Ph.D. in Decision Sciences and Engineering Systems from Rensselaer Polytechnic Institute in August, 1993. Her research interests lie in simulation output analysis and electronic education. Her research has been funded by the Federal Aviation Administration, the United States Air Force and the Department of Transportation. She is a member of IIE. Her email address is <mcourt@ou.edu>.

DAVID GOLDSMAN is a Professor in the School of ISyE at Georgia Tech. His research interests include simulation output analysis and ranking and selection. He was the Program Chair for the 1995 Winter Simulation Conference and currently serves on the Board of the WSC. His e-mail address is <sman@isye.gatech.edu>.

JORGE HADDOCK is Interim Associate Dean and Professor of Management Science at Rensselaer Polytechnic Institute's Lally School of Management and Technology. He also holds an affiliated position in Industrial Engineering and Operations Research in RPI's Department of Decision Sciences and Engineering Systems. He received his M.S. in Industrial and Management Engineering from RPI, and his Ph.D. in Industrial Engineering from Purdue University. His current research interests involve the design and implementation of effective information technology, production and service systems, as well as their effects on corporate culture. His e-mail address is <hadock@rpi.edu>, and his web page can be found at <http://www.rpi.edu/~haddock/>.

SEONG-HEE KIM is an Assistant Professor in the School of Industrial and Systems Engineering at the Georgia Institute of Technology. Her research interests include simulation output analysis, ranking and selection, and qual-

ity control. She received her Ph.D. in 2001 from Northwestern University's Department of Industrial Engineering and Management Sciences. Her e-mail address is <skim@isye.gatech.edu>, and her website is at <http://www.isye.gatech.edu/~skim/>.

MARGARET LOPER is a Senior Research Scientist at the Georgia Tech Research Institute (GTRI). She earned a B.S. in Electrical Engineering from Clemson University in 1985, an M.S. in Computer Engineering from the University of Central Florida in 1991, and a Ph.D. in Computer Science from the Georgia Tech in 2002. Her research interests include synchronization algorithms, temporal uncertainty, parallel and distributed systems, and theoretical aspects of modeling and simulation. Her e-mail is <margaret.loper@gtri.gatech.edu>.

JENNIFER PITTMAN is a doctoral student at the University of Oklahoma. She received her B.S and M.S. in the accelerated IT program of the School of Industrial Engineering with minors in computer science and psychology. She is an NSF Graduate Fellow and her doctoral research focuses on developing computer simulated environments for analyzing human movement under life threatening conditions. Her email address is <jpittman@ou.edu>.

AMY PRITCHETT is an Associate Professor in the School of Industrial and Systems Engineering at Georgia Tech. She earned her S.B., S.M., and Sci.D. in Aeronautics and Astronautics from the Massachusetts Institute of Technology in 1992, 1994 and 1997, respectively. Her research involves cockpit design, including alerting systems and advanced decision aids. Her e-mail address is <amyp@isye.gatech.edu>, and her website is <http://www.isye.gatech.edu/~amyp/>.