2003

Smart experimental designs provide military decision-makers with new insights from agent-based simulations

Lucas, Thomas. W.

FEATURED PROJECT

SMART EXPERIMENTAL DESIGNS PROVIDE MILITARY DECISION-MAKERS WITH NEW INSIGHTS FROM AGENT-BASED SIMULATIONS

Associate Professor Thomas W. Lucas, Department of Operations Research
Professor Susan M. Sanchez, Department of Operations Research and the Graduate School of Business and Public Policy

The 6th International Project Albert Workshop, described elsewhere in this newsletter, is part of an ongoing effort that seeks to exploit the advances in computing power and new technologies in order to "provide quantitative answers…to important questions facing military decision-makers" (Brandstein, 1999). In particular, in his former position as Chief Scientist of the U.S. Marine Corps, Dr. Brandstein was frustrated with legacy models because he felt they were unable to support analysis needs in the rapidly evolving global environment. Areas of particular concern were, and continue to be, our inabilities to adequately deal with the chaos inherent in military engagements, the human dimensions of warfare (e.g., leadership, courage, trust, unit cohesiveness), and adversaries who adapt their behavior based on perceptions of

> "Imagination is more important than knowledge. Knowledge is limited. Imagination encircles the world."
> **Albert Einstein**

--continued on page 21

About the INVESTIGATORS

Thomas W. Lucas is an Associate Professor in the Department of Operations Research (OR). Dr. Lucas received a B.S. in Industrial Engineering and Operations Research from Cornell University, an M.S. in Statistics from Michigan State University, and a Ph.D. in Statistics from the University of California at Riverside. He joined the NPS faculty in 1998, and has been teaching courses in statistics and combat modeling. He was recognized by the OR Department for his Outstanding Instructional Performance in 1999.

Thomas W. Lucas

Dr. Lucas is a member of the American Statistical Association, the Military Applications Society of the Institute for Operations Research and the Management Sciences (INFORMS), and the Military Operations Research Society (MORS). His primary research interests are combat analysis, design of simulation experiments, and robust Bayesian statistics. Previously, he worked as a statistician at RAND and as a systems engineer at Hughes Aircraft Company.

Susan M. Sanchez is a Professor and Associate Chair of Instruction in the Department of Operations Research. She also holds a joint appointment in the Graduate School of Business and Public Policy. Dr. Sanchez received a B.S. in Industrial and Operations Engineering from the University of Michigan, and her M.S. and Ph.D. in Operations Research from Cornell University. Dr. Sanchez came to NPS as a Senior Postdoctoral Associate under a National Research Council fellowship in 1999, and joined the NPS faculty in 2000. She teaches courses in statistics, operations research, and simulation analysis.

Susan M. Sanchez

Dr. Sanchez is a member of the Institute for Operations Research and the Management Sciences (INFORMS), the American Statistical Association, and the American Society for Quality. She is currently President of the INFORMS College on Simulation, and is also President of the INFORMS Forum on Women in Operations Research and Management Science. She serves as Simulation Area Editor for the *INFORMS Journal on Computing* and as an Associate Editor for *Naval Research Logistics.* Her research interests include the design and analysis of simulation experiments, selection procedures, and active learning.

SMART EXPERIMENTAL DESIGNS, *continued from page 20*

our strategies and tactics.

Project Albert focuses on *Operational Synthesis*—that is, the process of combining the information gleaned from a family of diverse analytical tools to provide the most compelling analyses. The majority of Project Albert's efforts have involved the building of relatively simple models, along with data farming and visualization environments in which they can be explored. These models by design are fast-running, flexible, and easy to use. They contain only the essence of a given question or scenario and utilize only that detail absolutely necessary to capture the relevant aspects.

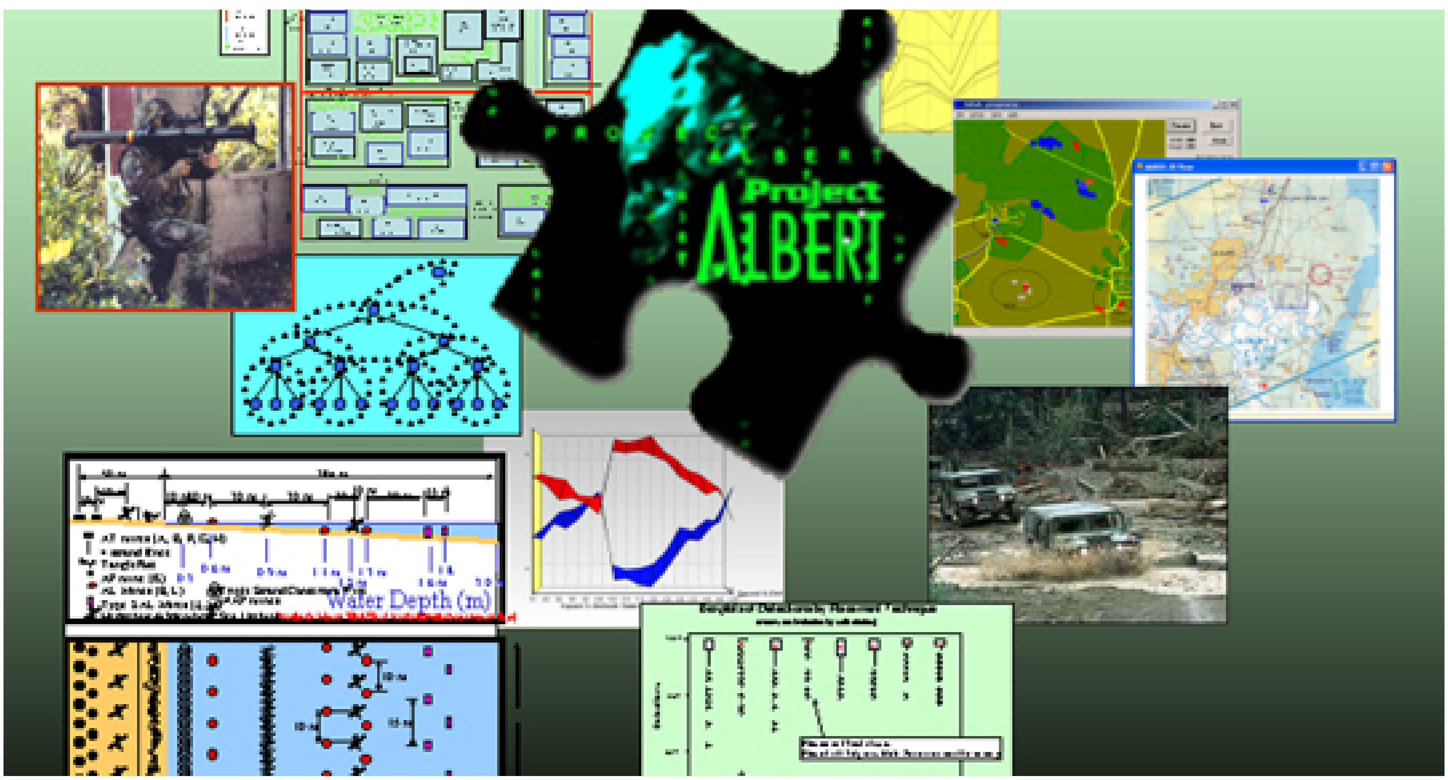
In order to fully evaluate all of the combinations of a model containing only 100 factors, each with only two settings, 2^{100} (about 10^{30}) runs of the model are necessary. Is this feasible? Former Air Force Major General Jasper Welch succinctly summarized the analyst's dilemma by the phrase "10^{30} is forever." Using a computer that can evaluate a model run in a nanosecond, an analyst who started making runs at the dawn of the universe would just be finishing his runs -- hence it would have taken him or her "forever" to explore the model.

To date, several modeling platforms have been developed by a diverse set of researchers under the Project Albert umbrella. Most of these are *agent-based simulations*. While the definition varies, we use this term to mean a simulation composed of agents, objects, or entities that make decisions (where to go, whom to shoot at, etc.) autonomously. These agents are aware of, and interact with, their local environment through relatively simple internal decision rules. The rules determine an agent's "personality" traits, such as their drive to move toward or away from a destination, and alive or injured friendly (or enemy) agents. Additionally, group characteristics can be defined which affect group behavior—such as the difference in forces required for an agent in a unit to want to advance toward an enemy. An agent's physical characteristics include their ability to sense, communicate, and engage with other agents.

Motivation

While Project Albert's distillations are quite simple by traditional Department of Defense (DoD) simulation standards, they nonetheless contain many variables that an analyst might desire to explore. Thus, a key thrust of the project is to utilize supercomputing to "farm" or run the models many times—millions of computation experiments are

--continued on page 57



FEATURED PROJECT

SMART EXPERIMENTAL DESIGNS, *continued from page 21*

often made on a given scenario. While millions of runs seem like a lot, the number of runs required to comprehensively explore even the simplest distillation can be astronomically large.

Most of our models have more than 100 factors, many of which are continuous or can take on a large number of discrete values. Our analyses are often further complicated by the uncertainty corresponding to many (if not most) of the factors. Therefore, even with super computers and "simple" models, we typically cannot use brute force searches on more than about 5-10 factors at a time. Moore's Law suggests that we will be able to extend this only by about two factors (through an increase of two orders of magnitude in processing power) each decade. Thus, if we want computational experiments that look broadly across these models, we need better designs. Our research objective is to develop search strategies that give DoD analysts flexibility in fitting models when exploring high-dimensional computer simulations in situations in which there is considerable *a priori* uncertainty about the shapes of the response surfaces. To this end, we are working with several NPS students to develop new search algorithms and assess their performance (analytically and empirically) over a broad set of models and scenarios.

In one sense, this need to examine many factors is an old problem. Situations we have chosen to explore via experimental designs have always been complicated. However, practical limitations (for physical experiments) and computational limitations (for simulation experiments) have forced decision-makers to focus on only a handful of factors at a time—those deemed the most important. We have found that this narrow window into a system's behavior can give rise to misleading results. Important factors or interactions may be ignored, or the results may be highly sensitive to a model input that was set arbitrarily. If the exploration of the model's behavior begins broadly, this reduces the likelihood of inap-

--continued on page 58

--continued from page 56

ability and Testability Group within Boeing Commercial Airplane Company to participate in presentations/seminars on software and systems reliability. The intention of the seminars is to bring the latest research and theory before Boeing's designers.

OPERATIONS RESEARCH

T.C. Barkdoll, **D.P. Gaver**, K.D. Glazebrook, **P.A. Jacobs**, and S. Posadas, "Suppression of Enemy Air Defenses (SEAD) as an Information Duel," *Naval Research Logistics*, No. 49, 2002.

THE MODELING AND SIMULATION INSTITUTE

Prof M. Zyda has been nominated for the 2003 World Technology Award for Information Technology - Software and the MOVES Institute has also been nominated in the "corporate category." The winners of these prestigious will be announced at the 2003 World Technology Awards and World Technology Summit in San Francisco, California on 24-25 June 2003.

PROMOTION AND AWARD OF TENURE FOR ACADEMIC YEAR 2003

Award of Tenure

Professor Leonard Ferrari
Department of Electrical and Computer Engineering

Professor Thomas Housel
Department of Information Science

Associate Professor Bret Michael
Department of Computer Science

Promotion to Professor

Professor Douglas Fouts
Department of Electrical and Computer Engineering

Professor David Jenn
Department of Electrical and Computer Engineering

Professor James Suchan
Graduate School of Business and Public Policy

Promotion to Research Associate Professor

Research Associate Professor Susan Hutchins
Department of Information Science

Research Associate Professor Haf Jonsson
Center for Interdisciplinary Remotely Piloted Aircraft Studies

Research Associate Professor Douglas Miller
Department of Meteorology

Research Associate Professor Robin Tokmakian
Department of Oceanography

CONGRATULATIONS!

SMART EXPERIMENTAL DESIGNS, *continued from page 57*

propriate results and substantially improves on our ability to provide rapid responses to new questions as they arise.

Methodological Approach

An extensive body of literature on designing experiments exists—most with roots in agriculture and laboratory experiments. That is, they were developed for situations with a relatively small number of experimental units (e.g., plots of land, patients, widgets) on which experiments could be conducted. Consequently, there are not many readily available tools for high-dimensional computer explorations where we can precisely control all of the factors and potentially take millions of runs. Furthermore, most of the existing designs also assume many of the following: linear effects, sparse effects, negligible higher order interactions, homogeneous normal errors, and a single measure of performance. Experience suggests that these are risky assumptions to make with models of combat, so we need alternative methods of selecting the best set of experiments from the vast ensemble of possibilities.

Clearly, the appropriate design depends on both the type of information needed and the nature of the model's response surfaces. In general, for exploring distillations, we want designs that can look at a large number of factors, isolate interactions, identify non-linearities (such as diminishing rates of impact, synergistic or redundant effects), and find thresholds where responses change dramatically. To accomplish this, we have developed some new experimental designs, and devising adaptive strategies that combine these with other well-known designs. In particular, we are looking at search strategies that use adaptive mixtures of full-factorial (or grid), fractional-factorial, group screening, random perturbations, Latin hypercube, and frequency-based designs. Further details are available in Sanchez et al. (2002) and Lucas et al. (2002).

Our findings to date show that the adaptive search strategies can greatly enhance an analyst's ability to explore agent-based models efficiently and effectively. It is not surprising that the appropriateness of the design depends critically on the shape of the model's surface and the number of samples feasible. There is no one-fits-all design, but we have developed some guidelines that depend on the total computational budget and the number of factors involved. They also reflect knowledge of the nature of the response surface, if this is available from previous experiments or subject-matter experts. As examples,

- For relatively smooth surfaces, fractional factorial designs are an efficient means of looking at a dozen or so factors.
- For high-dimensional surfaces with sparse effects, group screening designs work well.
- When large samples are feasible (hundreds of thousands or millions), regular Latin hypercube designs work very well, particularly on highly nonlinear surfaces.
- For high-dimensional searches of highly nonlinear surfaces when only a few hundred or a few thousand runs can be taken, special near-orthogonal Latin hypercube designs are more efficient.
- Frequency-based designs also work well on highly nonlinear surfaces when moderate or large samples are feasible, even in the presence of substantial error. Furthermore, they allow for a natural multi-resolution search.

Our current work focuses on combining these designs into an adaptive sequential framework. This is richer and substantially more powerful than any single one-stage design. The one-stage designs often used in practice correspond to categorizing all factors into two classes: those evaluated (typically at a common level of resolution) and those ignored.

Applications and Assessments

This section summarizes some of what we have learned by a series of empirical explorations on a variety of models and scenarios.

Brown (2000) examined how the personalities of leaders and subordinates can affect Blue agents' ability to reach a goal in a simulated urban environment. In his scenario, he found that losses are reduced for a local commander who has a strong propensity to mass his forces while maneuvering away from the enemy, and who assigns a relative degree of importance to the mission of reaching the objective without letting this objective dominate his actions. He identified an interesting interaction between friction (modeled as inhibiting the subordinates' ability to listen to their local commander) and the bond (modeled as their desire to stay with their local commander). Even if the subordinate agents cannot hear, comprehend, or otherwise act on the local commander's orders, their losses are reduced if they stay with him.

Of course, we cannot tell without additional data involving real people, perhaps under real combat conditions, whether these insights extend to real combat. Nonetheless, there are some interesting insights gleaned regarding the effectiveness of potential designs. Specifically, a 2^{5-1} fractional factorial design was almost as informative, in terms of variance explained, as a 5^5 full factorial design, despite requiring less than one percent

--continued on page 59

SMART EXPERIMENTAL DESIGNS, *continued from page 58*

as many runs. Similarly, Wan (2002) used both a full factorial design (with 174,000 runs) and a Latin hypercube design (with only 4,800 runs) to examine the effects of human factors in a small unit infantry engagement. He found that almost all of the information extracted from the full factorial design could be more efficiently obtained by the Latin hypercube.

We have found Latin hypercubes (McKay et al., 1979) particularly valuable. Ordinary Latin hypercubes have received widespread use because they are extremely flexible and easy to generate. However, if the number of runs is moderate relative to the number of factors, this inhibits the analyst's ability to obtain precise estimates of some parameters. To rectify this, Cioppa (2002) developed an algorithm that generates "nearly orthogonal" Latin hypercubes. These designs also have excellent space-filling properties, which make them amenable to fitting non-parametric surfaces. They have been used to study peace-enforcement operations and guerrilla combat, and are currently being considered for use in important Army transformational studies.

The output of ordinary and near-orthogonal Latin hypercube designs can be analyzed by a host of analytical techniques. Some examples include:

- Vinyard and Lucas (2002) used linear modeling and intensity plots to identify regions of non-monotonicity and explore the mitigating effects of making some variables stochastic elements in the infamous 18 dimensional Dewar combat model.
- Pee (2002) applied neural networks, in combination with visualization techniques, while assessing the impact of information systems and procedures on battle outcomes. He found that the Blue force can ensure a positive outcome if it can sufficiently control two of its process latencies—regardless of the values of the other nine factors examined.
- Ipekci (2002) used classification trees, multiple additive regression trees (MART), Bayes nets, and Trellis plots in a 22 dimensional exploration of a simulation of a guerrilla infiltration attack he experienced as a platoon commander. The results of his sequential analysis indicate that the outcome of an infiltration scenario was dominated by Red agent parameters. This suggests that when combating guerrillas, we might best use our resources to restrict the terrorists' abilities to mass, move, train, and acquire materiel.
- Wu (2002) explored the use of a frequency-based approach to designing experiments for terminating simulations. He applied these to multiple performance measures for a peace-enforcement scenario. His results show once again that the set of factors classified as important depends on the construct of the performance measure, but that frequency-based designs can be a natural and efficient way of administering the experiments. Wu also applied human factors principles in creating visual and auditory displays of the results.

Future Directions

This ongoing project will continue to advance both the theory and application of high-dimensional simulation exploration in several ways. The experiments described above have used designs generated locally and then exported for batch processing at the supercomputing facilities in Maui, Hawaii or Woodbridge, Virginia. We are in the process of converting our algorithms so they are readily available to all Project Albert researchers before the next Project Albert workshop in September 2003. Second, we are enhancing the sequential performance of our designs. Our goal is to provide enough guidance that the procedure can be used as a decision-support tool for an analyst with little experience in experimental design. Finally, we continue to apply our approaches to a broad spectrum of application areas, such as Marine logistics, Army force design, and Navy littoral area analysis.

It is worth noting that our approach is not restricted to agent-based simulations. These designs can be applied to any computer model with many factors—deterministic or stochastic—where the analyst is interested in gaining insight into how the performance varies across a wide range of possibilities.

Additional Reading

Brandstein, A., 1999. "Operational Synthesis: Applying Science to Military Science," *PHALANX*, **32**, December.

Cioppa, T.M., 2002. "Efficient Nearly Orthogonal and Space-filling Experimental Designs for High-dimensional Complex Models," Ph.D. Dissertation, Operations Research Department, Naval Postgraduate School, Monterey, CA.

Ipekci, A., 2002. "How Agent Based Models Can Be Utilized To Explore and Exploit Non-linearity and Intangibles Inherent in Guerrilla Warfare," Masters Thesis, Operations Research Department, Naval Postgraduate School, Monterey, CA.

Lucas, T.W., S.M. Sanchez, L. Brown, and W. Vinyard, 2002. "Better Designs for High-dimensional Explorations of Distillations," *Maneuver Warfare Science 2002*, eds. G. Horne and S. Johnson, USMC Project Albert, Quantico, VA, 17-45.

McKay, M.D., R.J. Beckman, and W.J. Conover, 1979. "A

--continued on page 63

NAVAL POSTGRADUATE SCHOOL CHAIR PROFESSORSHIPS

School of International Graduate Studies

Chair of Strategic Planning
CDR David Spain, USN
Intelligence Chair
CDR Steve Recca, USN
Transportation Security Chair
Dr. James L. Fobes

Graduate School of Operational and Information Sciences

Navy Warfare Development Command Chair for Warfare Innovation
CAPT Jeff Kline, USN
Chair of Manpower Modeling
Dr. Sam Buttrey
Chair in Applied Systems Analysis
CDR Matthew Boensel, USN
Special Operations Chair
COL Joseph Tyner, USA

Graduate School of Engineering and Applied Sciences

Space Systems Academic Chair
Charles Racoosin
Navy Tactical Exploitation of National Capabilities (TENCAP) Chair
Dr. Alan Ross
Naval Space Technology Program Chair
RADM Tom Betterton, USN (Ret.)
Michael J. Smith Space Systems Chair
Dr. Chirold Epp
Lawrence Livermore National Laboratory Chair
Dr. William Kruer
Measurement and Signature Intelligence (MASINT) Chair Professor
Col David Trask, USAF (Ret.)
Chair of Arctic Marine Science
Dr. Max Coon
National Reconnaissance Office Chair Professor
CDR David Kretzmann, USN (Ret.)
Engineering Acoustics Chair
Joe Rice
Lockheed Martin Missiles and Space Chair Professorship
Salim Abbas
NAVSEA Chair of Total Ship Systems Engineering
Professor Chuck Calvano
NAVSEA Chair of Total Ship Systems Engineering in Combat Systems
Senior Lecturer James Harney
Battlespace Environments Chair
Vacant

Graduate School of Business and Public Policy

Conrad Chair of Financial Management
VADM Thomas Hughes, USN (Ret.)
Chair of Logistics
RADM Donald Eaton, USN (Ret.)
Admiral Boorda Chair of Management and Analysis
RADM David Bill, USN (Ret.)
RADM George F. Wagner Chair in Public Management
Dr. Lawrence Jones
Chair of Acquisition Management
RADM Jim Green, USN (Ret.)

Wayne E. Meyer Institute of Systems Engineering

Chair of Mine Warfare
RADM John Pearson, USN (Ret.)
Chair of Expeditionary Warfare
Dr. Phil Depoy
Chair of Undersea Warfare
VADM Roger F. Bacon, USN (Ret.)

Cebrowski Institute for Information Innovation and Superiority

National Security Agency Cryptologic Innovation Chair
Ms. Joanne Kim

SMART EXPERIMENTAL DESIGNS, *continued from page 63*

Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code." *Technometrics*, **21**, 239-245.

Pee, E., 2002. "An Exploratory Analysis on the Effects of Information Superiority on Battle Outcomes," Masters Thesis, Operations Research Department, Naval Postgraduate School, Monterey, CA.

Sanchez, S.M and T.W. Lucas, 2002. "Exploring the World of Agent-based Simulations: Simple Models, Complex Analyses," *Proceedings of the Winter Simulation Conference*, eds. J.L. Snowdon, J. Charnes,, C.-H. Chen, and E. and Yucesan, Institute of Electrical and Electronic Engineers: Piscataway, NJ.

Vinyard, W. and T.W. Lucas, 2002. "Exploring Combat Models for Non-monotonicities and Remedies," *PHALANX,* **35**, March.

Wan, S., 2002. "An Exploratory Analysis on the Effects of Human Factors on Combat Outcomes," Masters Thesis, Operations Research Department, Naval Postgraduate School, Monterey, CA.

Wu, H.-F., 2002. "Spectral Analysis and Sonification of Simulation Data Generated in a Frequency Domain Experiment," Masters Thesis, Operations Research Department, Naval Postgraduate School, Monterey, CA.