Proceedings of the 1989 Winter Simulation Conference
E.A. MacNair, K.J. Musselman, P. Heidelberger (eds.)

EXPERIMENTAL DESIGN ISSUES FOR LARGE SCALE SIMULATION MODELS

Chair
Russell R.Barton
Dept. of Operations Research and Industrial Engineering
Cornell University
Ithaca, NY 14850

Panelists
George S. Fishman
Malvin H. Kalos
W. David Kelton
Jack P.C. Kleijnen

INTRODUCTION

Experiment design issues are important at each stage in a simulation model's life-cycle, yet we have found that supercomputer users are poorly supported in this area. This panel will address experimental design techniques that can be used to make more efficient use of supercomputing resources and researcher resources. Problems associated specifically with large sale simulation models will be discussed by example.

AUTHOR'S BIOGRAPHY

Russell R. Barton was born in Buffalo, New York, on August 11, 1951. He received the B.S. degree in Electrical Engineering from Princeton University in 1973. He received the M.S. degree in Operations Research from Cornell University in 1975, and Ph.D. in 1978. His thesis described and analyzed failure rate regression models, with applications to exoffender behavior. From 1975 to 1978, he was an operations research consultant, working on contracts for several government agencies. In 1978, Dr. Barton accepted a position as member of technical staff at the RCA David Sarnoff Research Center. There he provided statistical and operations research consulting. He was co-creator and co-instructor for two continuing education video courses on the design of experiments. During 1985-1987 he developed and implemented optimization methods for simulation models of electron optics and microwave circuits. Since 1987, Dr. Barton has been a visiting Associate Professor in the School of Operations Research at Cornell University, and Laboratory Director for the Cornell Computational Optimization Project. His current research interests are in the development and analysis of optimization methods for simulation models used in engineering design.

POSITION STATEMENT

George S. Fishman
Department of Operations Research
University of North Carolina at Chapel Hill
Chapel Hill, NC 27514

INTRODUCTION

For present purposes, let us regard experimental design in simulation in the broadest possible terms. At one extreme, this approach allows us to contemplate a range of alternative sampling plans each aimed at producing an estimate of a single quantity of interest. Alternative methods of generating samples and conventional variance reducing techniques are principal considerations in this area. At the other extreme, we consider how to layout a sequence of simulation experiments, of different but related systems, each of which produces an estimate of a quantity of interest for the ultimate objective of making inferences about how these quantities are ordered. For example, which of several simulated systems performs best? Here emphasis often focuses on identifying the order in which experiments should be performed to provide the richest possible information base at the end of each experiment to decide whether or not the remaining level of uncertainty warrants continued experimentation. If the systems being compared differ only in the values assigned to critical parameters, then sequence of experiments may be replaced by a single one which incorporates either a form of infinitesimal perturbation analysis or importance sampling, provided that conditions exist for either or both these techniques to apply.

The record of success in helping analysts address these issues for simulation experiments executed on sequential computers has been modest. This experience accords with the general experience of consulting statisticians with regard to guiding experimental design in other areas of study. They and the simulation expert are rarely called in early enough to influence the fundamental course that experimentation takes and are rarely prepared to learn all the essential details of the problem at hand so as to exploit the inherent special structure of the problem to improve statistical efficiency.

When we turn to the super computing environment, the availability of options such as vectorization and distributed processing seem to offer new tools that can be turned to advantage in experimental design. No doubt, this is true. However, these features also create new global problems that simulation experts need to address before they can expect their advice on design to carry any weight. First, there is the issue of random number generation. Today most super computers use a traditional multiplicative congruential generator designed to work in a sequential, not a parallel, environment. While it is true that these generators have been modified to acknowledge certain features of parallelism (e.g., Durst 1987 and Percus and Kalos 1989), the fact remains that little agreement exists among simulation experts as to how to advise clients with regard to choosing a random number generator that exploits concurrency, ensures reproducibility, guarantees long period sequences, and exhibits the desired theoretical distributional features.

A second issue concerns the development of parallel algorithms for randomly generating samples from a variety of distributions or sample combinatorial objects. Little has appeared on this topic in the published literature, yet the potential for exploiting concurrency appears high. For example, for a network of m arcs, and n nodes randomly generating a spanning tree takes $O(mn^3)$ time on a sequential computer and relies heavily on search (Kulkarni 1988). Being able to reduce this time with some type of parallel search procedure does not seem remote.

One area in which simulation experts can provide constructive guidance is in devising a sampling plan that distributes the execution time of independent replications over a bank of connected processors. Bhavsar and Isaac (1987) and Heidelberger (1988) analyze the affects of alternative allocation schemes and demonstrates that care must be taken to avoid inducing bias in the resulting estimate.

In summary, I believe that the best role that the simulation methodologist can play at this moment in time is to acquaint herself or himself with the problems that users of simulation encounter in a concurrent computing environment, digest these problems and to try to develop broadly based methodologies that will solve classes of problem that regularly arise in this new computing environment.

REFERENCES

Bhavsar, V.C. and J.R. Isaac (1987). Design and analysis of parallel Monte Carlo algorithms, SIAM J. Sci. Stat. Comp., 8, s73-s95.

Durst, M. (1987). Multiprocessing pseudo random number generators, Lawrence Livermore National Laboratory, Livermore, California.

Heidelberger, P. (1988). Discrete event simulations and parallel processing: statistical properties, SIAM J. Stat. Comput., 9, 1114-1132.

Kulkarni, V.K. (1988). Generating random combinatorial objects, Technical Report No. UNC/OR/TR-88/10, Department of Operations Research, University of North Carolina at Chapel Hill, to appear in J. of Algorithms.

Percus, O.E. and M.H. Kalos (1987). Random number generators for MIMD parallel processors, J. Parallel and Dist. Comput., 6, 477-497.

AUTHOR'S AUTOBIOGRAPHY

George S. Fishman is professor and chairman of the Department of Operations Research at the University of North Carolina at Chapel Hill. His principal interest is the development of statistical methodology applicable to the analysis of output from discrete event digital simulation models. He is the autohor of Concepts and Methods in Discrete Event Digital Simulation puslished by Wiley in 1973 and of Principles of Discrete Event Simulation published by Wiley in 1978. He is a frequent contributor to the operations research and statistical literature on this topic. At present, he is working on variance reducing methods for network reliability estimation and on the influence of concurrent processing on simulation program structure. Professor Fishman has been simulation departmental editor for Management Science and is a member of the Operations Research Society of America, the Institute of Management Science, and the American Statistical Association.

George S. Fishman
Department of Operations Research
University of North Carolina
at Chapel Hill
Chapel Hill, NC 27514

POSITION STATEMENT

Malvin H. Kalos
Cornell Theory Center
Cornell University
Ithaca, NY 14853-5201

For the future in which more science and engineering depend on large scale computing, we will need to improve some of the basic tools of the trade. Better languages are urgently needed, languages that are more expressive of a wide variety of algorithms, are more robust against errors, and which are expressive of parallelism, or better yet in which parallelism can be found without having to be specified. Another need will be powerful methods of debugging, capable of treating very large programs on highly parallel machines. It is likely that many future parallel machines will not always be deterministic and styles and techniques of coping with this problem will have to evolve.

AUTHOR'S BIOGRAPHY

Malvin H. Kalos is Director of the Center for Theory and Simulation in Science and Engineering at the Cornell University National Supercomputer Facility. His professional interests include the theory and practice of Monte Carlo methods;

statistical, many-body and nuclear physics; radiation transport theory; scientific programming and computer architecture.

Malvin H. Kalos
Director
Center for Theory and Simulation in Science and Engineering
265 Olin Hall
Cornell University
Ithaca, NY 14853-5201
(607) 255-9398
kalos@tcgould.tn.cornell.edu

POSITION STATEMENT

W. David Kelton
Department of Operations and Management Science
Carlson School of Management
University of Minnesota
Minneapolis, Minnesota 55455

INTRODUCTION

My experience in using supercomputers and other advanced computing equipment has been primarily in the realm of conducting experiments to evaluate competing methodologies for (1) statistical analysis of simulation output, and (2) design of simulation experiments. The realm is primarily that of discrete-event, dynamic, stochastic models. Specifically, my work on these types of machines has been on a four-processor Cray-2 at the Minnesota Supercomputer Center (and supported through the Minnesota Supercomputer Institute), and vicariously through one of my students, Murali Shanker, on an eight-node hypercube that Intel Corporation has made available to us.

Our use for such facilities stems from the fact that, in conducting experiments to evaluate a methodology, it is possible (even desirable) to use vast quantities of computing resources, especially time (as opposed to memory). To make things concrete, take as an example two alternate methods for constructing a confidence interval for an unknown process parameter. In evaluating and comparing these methods, one would simulate comparatively simple models with known parameter values, apply both techniques to the same simulation output data, and then see which method appears to yield confidence intervals that contain this value with a frequency that is the most consistent with the nominal level of the interval. Clearly, the more replications we have of this interval, the sharper our comparison can be of the competing methods, and the better our conclusions and recommendations. In many settings, such as steady-state or long-run simulation models, each individual replication involves perhaps several long simulation runs, and computing-time requirements can easily become prohibitive except on the most powerful machines.

The capabilities of these types of machines (and, not insignificantly, the operating systems' and compilers' capabilities) currently present both opportunities and perceived needs that we have experienced, as briefly mentioned next.

OPPORTUNITIES

The raw speed of supercomputers makes them a natural platform for these kinds of studies, and it was this that drew us initially to them. Machines such as the Cray-2 achieve this primarily by relying on the programmer to write code consistent with vectorization capabilities. While this speed is certainly an advantage, the vast memory of machines such as the MSC Cray-2 (a half-billion 64-bit words) can be exploited to enable vectorization to take place (discussed below).

On the other hand, distributed-processing machines such as the Intel Hypercube present opportunities of a different sort. Multiple replicates can be assigned to separate nodes for simultaneous processing, an obvious advantage. In stochastic simulation, the comparatively recent availability of random-number generators with extremely long periods (and subsegments attainable without generating through all the intermediate values) provide a vehicle to ensure independence of the replicates in progress simultaneously on the different nodes.

NEEDS

While the above opportunities are real and have been used to considerable advantage, there remain significant possibilities for further improvement. Some of these are:

o Better automatic methods of recognizing potential vectorization opportunities. The speed advantages of vectorization can be significant. However, writing vectorizable code can consume a lot of researcher time (a commodity with some value, one would hope) and may involve thinking in ways that are not particularly natural. One specific example is that subprogram invocations within deeply nested loops may prevent vectorization. Random-number generators are traditionally written as subprograms that may well be referenced several layers down in a nested loop structure. We have dealt with this on the Cray-2 by taking advantage of its huge memory to pregenerate the required random numbers and store them for use down deep in the loops. While we were able to recognize this, it would have been more convenient from our point of view if the compiler had been able to recognize this opportunity itself. Thus, we see a need to continue the advances in code-optimization capabilities of compilers.

o High-level language support. From the point of view of simulation users, it would be desirable for high-level simulation languages to take better advantage of vectorization opportunities. There are limits, though, to the extent to which this may be possible due to the inherent serial nature of many simulation models themselves.

o Enhanced capabilities of the hardware and operating systems of distributed-processing machines. In particular, distributed simulation models can, in the case of complex real-world models, involve heavy use of message-passing buffers in message-passing schemes that can become overloaded, hanging the machine. Capabilities for heavier message volume would expand the realm of models for which distributed processing is possible.

SUMMARY

The potential for vectorizing supercomputers and distributed-processing machines in simulation methodological research is being partially realized at this time, but there appear to be a number of opportunities for improvement of capabilities. It is hoped that hardware and software designers will take into account these kinds of needs, since the benefits would doubtless spill over into other areas as well.

ACKNOWLEDGEMENT

Bob Crain of Wolverine Software participated in a helpful conversation with the author on this topic.

AUTHOR'S BIOGRAPHY

W. David Kelton is an Associate Professor in the Operations and Management Science Department at the Carlson School of Management, University of Minnesota, in Minneapolis. He received a B.A. in Mathematics from the University of Wisconsin-Madison, an M.S. in Mathematics from Ohio University, and M.S. and Ph.D. degrees in Industrial Engineering from the University of Wisconsin-Madison. His research interests are in simulation methodology, stochastic modeling and estimation, and quality control. He is a member of ORSA, TIMS, and ASA. He served as Program Chair for the 1987 Winter Simulation Conference.

W. David Kelton
Department of Operations and Management Science
Carlson School of Management
University of Minnesota
Minneapolis, Minnesota 55455
(612) 624-8503
dkelton@vx.acss.umn.edu
dkelton@umnacvx.bitnet

POSITION STATEMENT

Jack P.C. Kleijnen
Department of Information Systems and Auditing
Katholieke Universiteit Brabant (KUB), Tilburg
5000 LE Tilburg, Netherlands.

Simulation is a technique applied in many areas, because of its flexibility, simplicity and realism. By definition, simulation involves experimenting (namely with the model of the real system); hence simulation requires STATISTICAL DESIGN AND ANALYSIS. We concentrate on strategic issues, namely, which variants of the simulation model should be run (i.e., which combinations of parameter values are input), and how can the resulting output be analyzed? Such issues arise in both random and deterministic simulations (whereas tactical issues - such as runlength and confidence intervals - arise only in random simulation). See Kleijnen (1987).

The methodological issues indicated above, concern problems arising in practice under such names as MODEL VALIDATION, WHAT-IF ANALYSIS, GOAL SEEKING, and OPTIMIZATION. The relevant statistical techniques are experimental design and regression analysis. Tentative prior knowledge leads to a TENTATIVE REGRESSION MODEL that specifies which input factors may be important, which interactions (among these factors) may be important, which scaling seems appropriate (e.g. a logarithmic scale may seem best), etc.

Based on this tentative regression model a DESIGN is selected that specifies which combinations of inputs is to be simulated. Classic statistical theory of experimental design gives designs that are 'efficient' and 'effective'. EFFICIENCY means that the number of factor combinations (or 'simulation runs') is minimal so that computer time is minimal, and yet the estimates of the factor effects are 'accurate'; accuracy is measured by the variance of the estimated effects. EFFECTIVENESS means that - given a small number of runs - interactions among factors can still be estimated. (The one-factor-at-a-time method cannot estimate interactions; moreover, accuracy of estimated main or first-order effects is low.) Also see Box et al. (1978).

Simulation models with HUNDREDS OF INPUTS need experimental designs not discussed in the textbooks on statistics. We have experience with group screening techniques, which aggregate individual inputs when searching for important inputs. Recently sequential bifurcation has been further developed for random simulation models with interacting inputs. See Bettonvil (1989).

The resulting output of the simulation experiment needs to be analyzed. This is done through REGRESSION (META)MODELS. For deterministic simulation models it suffices to use Ordinary Least Squares; for random simulation models Generalized Least Squares is needed. The regression analysis of standard textbooks needs some minor adjustments to account for the special characteristics of simulation. See Kleijnen (1987).

APPLICATIONS of the approach outlined above, are getting numerous. For demonstration purposes, a simple - but realistic case study is presented concerning a Flexible Manufacturing System (FMS). Input to the FMS simulation is the 'machine mix', that is, the number of machines of type i with i. = 1,...,4.

Intuitively selected combinations of these four inputs give inferior results when compared to a fractional factorial design. The throughput predicted by the simulation is analyzed through two different regression models. These models are validated. A regression model in only two inputs but including their interaction, gives valid predictions and sound explanations. See Kleijnen and Standridge (1988).

Another application concerns a decision support system (DSS) for production planning, developed for a Dutch company. To evaluate this DSS, a simulation model is built. The DSS has 15 control variables that are to be optimized. The effects of these 15 variables are investigated, using a sequence of fractional factorial designs. Originally, 34 response variables were distinguished. These 34 variables, however, can be reduced to one criterion variable, namely productive machine hours, that is to be maximized, and one commercial variable measuring lead times, that must satisfy a certain side-condition. For this optimization problem the Steepest Ascent technique is applied to the experimental design outcomes. The resulting Response Surface Methodology is developed theoretically. In practice a number of complications arise. See Kleijnen (1988).

A final case study concerns a set of deterministic ecological simulation models that require sensitivity analysis to support the Dutch government's decision making. First results are given by Rotmans and Vriese (1989) who treat a model for the 'greenhouse' effect; additional results will be reported at the conference.

Other approaches not covered in this presentation but referenced in Kleijnen (1987, pp. 241-242), are: piecewise linear, spline, and inverse polynomial metamodels, spectral analysis, and perturbation analysis.

REFERENCES

Bettonvil, B.: Sequential Bifurcation: The Design of a Factor Screening Method, Research Memorandum FEW 381, Katholieke Universiteit Brabant, 1989.

Box, G.E.P., W.G. Hunter, and J.S. Hunter: Statistics for Experimenters, John Wiley & Sons, Inc., New York, 1978.

Kleijnen, J.P.C.: Statistical Tools for Simulation Practitioners, Marcel Dekker, Inc., New York, 1987.

Kleijnen, J.P.C.: Simulation and Optimization in Production Planning: a Case Study, Research Memorandum FEW 308, Katholieke Universiteit Brabant, 1988.

Kleijnen, J.P.C. and C.R. Standridge: Experimental Design and Regression Analysis in Simulation: an FMS Case Study, European Journal of Operational Research, 33, 1988, pp. 257-261.

AUTHOR'S BIOGRAPHY

Jack P.C. Kleijnen received his master's and doctor's degrees in business administration from the "Katholieke Universiteit Brabant" in Tilburg (Netherlands), in 1964 and 1971, respectively. He was a research associate at this university from 1965 until 1980; in 1980 he was promoted to professor of simulation and information systems. He also spent several years in the USA: UCLA (1967/1968), Duke University (Summer 1968, Summer 1969), IBM Research San Jose (1974), Indiana University (Summer 1979), IBM Research Yorktown (Summer 1981), Pritsker & Associates (Summer 1984), Rutgers University (Summer 1988).

He received a number of fellowships and awards, both nationally and internationally. He published approximately 100 articles in international journals on simulation, statistics, operations research, computer science, etc.

His first book, "Statistical techniques in simulation" (Dekker, N.Y., 1974/1975), received a Lanchester-Prize Honorable-Mention from the Operations Research Society of America, and was translated into Russian. He is now finishing his fifth book.

He lectured at conferences in many European countries and the USA, consulted a number of organizations, and is a member of several scientific committees, editorial boards, computer committees, and many professional organizations. His research interests are in simulation, mathematical statistics, and management information systems.

Jack P.C. Kleijnen
Department of Information Systems and Auditing
Katholieke Universiteit Brabant (KUB), Tilburg
5000 LE Tilburg, Netherlands.
E-mail: t435klei@htikub5
Fax: 013-663019)