SIMULATION MODELING AND METHODOLOGY *

Robert E. Shannon
The University of Alabama in Huntsville
Huntsville, Alabama

Introduction

Simulation is one of the most powerful analysis tools available to those responsible for the design and/or operation of complex processes or systems. It is heavily based upon computer science, mathematics, probability theory and statistics; yet the process of simulation modeling and experimentation remains very much an intuitive art. Simulation is a very general and somewhat ill-defined subject. For the purpose of this paper, we will define simulation as, "the process of designing a computerized model of a system (or process) and conducting experiments with this model for the purpose either of understanding the behavior of the system and/or of evaluating various strategies for the operation of the system." Thus we will understand the process of simulation to include both the construction of the model and the analytical use of the model for studying a problem.

Even though simulation is considered a brute force approach or court of last resort by those with extensive mathematical training, numerous surveys have shown that it is the most widely used technique for operations research or management science type studies. Simulation is used when one or more of the following conditions exist:

1. A complete mathematical formulation of the problem does not exist or analytical methods of solving the mathematical model have not yet been developed. Many waiting line (queueing) models are in this category.

2. Analytical methods are available, but the simplifying assumptions required for their application negate much of the true environment of the problem.

3. Analytical methods are theoretically available but the mathematical procedures are so complex and arduous that simulation provides a simpler method of solution.

4. It is desired to observe a simulated history of the process over a period of time in addition to estimating certain parameters.

5. Simulation may be the only possibility because of the difficulty in conducting experiments in their actual environment, e.g., studies of space vehicles in interplanetary flight.

6. Time compression may be required for systems or processes with long time frames. Simulation affords complete control over time, since a phenomena may be speeded up or slowed down at will. Analysis of urban problems is in this category.

The Simulation Process

There are a number of excellent introductory books dealing with the methodology of simulation [1-7]. Assuming that a simulation is to be used to investigate the properties of a real system, Figure 1 provides one view of the simulation process [7], and the following stages may be distinguished.

1. System (or process) Definition - Determining the boundaries, restrictions and measures of effectiveness to be used in defining and studying the system (or process).

2. Model Formulation - Reduction or abstraction of the real system to a block or logical flow diagram.

3. Data Preparation - Identification of the data needed by the model and their reduction to an appropriate form.

4. Model Translation - Description of the model in an appropriate language acceptable to the computer to be used.

5. Validation - Determining that any inferences drawn from the model about the real system will be correct to some acceptable level of confidence.

6. Strategic Planning - Designing an experiment that will yield the desired information.

* This paper is a distillation of material appearing in Shannon, Robert E., System Simulation: The Art and Science, Prentice-Hall, Inc., Englewood, New Jersey, 1975.

7. Tactical Planning - Determining how each of the test runs specified in the experimental design is to be executed.

8. Experimentation - Execution of the simulation to generate the desired data and to perform sensitivity analyses.

9. Interpretation - Drawing inferences from the data generated by the simulation.

10. Implementation - Putting the model results to use.

11. Documentation - Recording the project activities and results as well as documenting the model and its use.

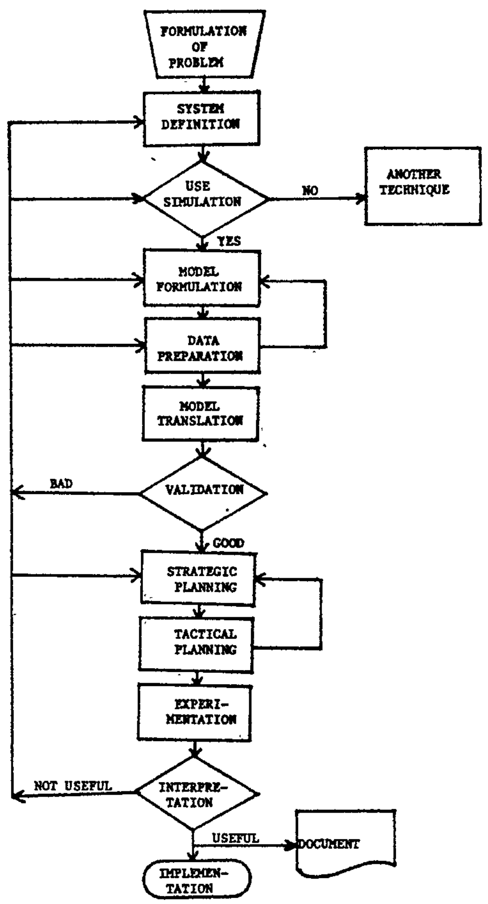

Figure 1. Simulation Process

Problem and Model Definition

To find an acceptable or optimal solution to a problem, one must first know what the problem is. This analysis begins with the specification of the system of interest, the environment in which it exists and operates, the specification of goals of the system and the purpose of the study.

In the process of studying a system or process and converting the resulting knowledge into a mathematical model, we must accomplish four tasks:

1. Specification of the purpose of the model.

2. Specification of the components to be included in the model.

3. Specification of the parameters and variables associated with the components.

4. Specification of the functional relationships among the components, parameters, and variables.

Simulation experiments are conducted for a wide variety of purposes, some of which are:

1. Evaluation - determining how good a proposed system design performs in an absolute sense when evaluated against specific criteria.

2. Comparison - comparing competitive systems designed to carry out a specified function, or comparing several proposed operating policies or procedures.

3. Prediction - estimating the performance of the system under some projected set of conditions.

4. Sensitivity analysis - determining which of many factors are the most significant in affecting overall system performance.

5. Optimization - determining exactly which combination of factor levels will produce the best overall response of the system.

6. Functional relations - establishing the nature of the relationships among one or more significant factors and the system's response.

This list is not exhaustive and merely suggests the most common goals or purposes; the explicit purpose of the model has significant implications for the whole model building and experimentation process. For example, if the model's goal is to evaluate a proposed or existing system in an absolute sense, this imposes a heavy burden upon the accuracy of the model and demands a high degree of isomorphism. On the other hand, if the goal is the relative comparison of two or more systems or operating procedures, the model may be valid in a relative sense even though the absolute magnitude of responses varies widely from that which would be encountered in the real world.

After we have specified (at least tentatively) the specific goal or purpose for which the model is to be constructed, we begin to identify the pertinent components. This process entails itemizing all the components of the system that contribute to the effectiveness or ineffectiveness of its operation. Once a complete list of the components of a system is specified, we next determine whether each component should be included in our model. This is easier said than done, since at this stage of model development it is not always clear whether a component is significant to the overall goal. One of the key questions to be answered is whether a particular component is to be considered part of the model or part of the environment.

One pertinent consideration in deciding which components are to be included and which excluded is the question of how many variables are to be included in the model. In general, we have little difficulty in deciding upon the endogenous or output variables. If we have done a good job in specifying the goals or purposes of the study, the required output variables are fairly obvious. The real difficulty arises in determining which input and status variables produce the effects observed and which can be manipulated to produce the effects desired. We are also faced with conflicting objectives: on one hand, we try to make the model as simple as possible for ease of understanding, ease of formulation, and computational efficiency; on the other hand, we try to make the model as accurate as possible. Consequently, we need to simplify reality but only to the point where there is no significant loss of accuracy.

Once we have decided which components and variables we shall include in our model, we must then determine the functional relationships among them and the values of the parameters to be used. Again, formidable problems confront us. First, it may be difficult (if not impossible) to quantify or measure certain variables that are important to the behavior of the system. Second, the relationships between components and variables may not be clear. Third, the data and information we need may not be available, or perhaps not exist in the form we need. Thus, decisions regarding the data to be used and their validity, form, and goodness of fit to theoretical distributions and past performance, are all critical to the success of the simulation experiment, and far from being academic points.

Having specified the goals and objectives of the study and defined the system, we next reduce the real system to a logical flow diagram or static model. We wish to construct a model of the real system that neither oversimplifies the system to the point where the model becomes trivial (or worse, misleading) nor carries so much detail that it becomes clumsy and prohibitively expensive. The danger is that the model may tend to be too detailed and include elements which contribute little or nothing to the understanding of the problem.

The tendency is nearly always to simulate too much detail rather than too little. Thus, one should always design the model around the questions to be answered rather than imitate the real system exactly. Pareto's law says that in every group or collection there exists a vital few and a trivial many. Nothing really significant happens unless it happens to the vital few. The tendency among systems analysts has too often been to transfer all the detailed difficulties in the real situation into the model, hoping that the computer would solve their problems. This approach is unsatisfactory not only because of the increased difficulty of programming the model and the additional cost of longer experimental runs, but also because the truly significant aspects and relationships may get lost in all the trivial details. Therefore, the model must include only those aspects of the system relevant to the study objectives.

Programming Languages

Early effort in a simulation study is concerned with defining the system to be modeled and describing it in terms of logic flow diagrams and functional relationships. But eventually one is faced with the problem of describing the model in a language acceptable to the computer to be used. Unfortunately, so many general and special purpose programming languages have been developed over the years that it is nearly impossible to decide which language best fits or is even a near best fit to any particular application. There were over 170 programming languages in use in the United States alone in 1972 and new ones are being developed every day [8].

Many writers find it convenient to classify simulation models into two major categories: 1) continuous change models or 2) discrete change models. Continuous change models use fixed increment time advance mechanisms and are appropriate when the analyst considers the system he is studying as consisting of a continuous flow of information or items counted in the aggregate rather than as individual items. In discrete change models, the analyst is interested in what happens to individual items in the system. Most discrete change models, therefore, utilize the next event type of timekeeping. Some problems are clearly described best by one type or the other, whereas either type might be used for other problems. With a few exceptions like GASP IV, simulation languages are restricted to either continuous or discrete change models.

In the most general sense, there are three computer techniques available for simulation; digital, analog and hybrid. One possible classification scheme is depicted in Figure 2. There are several versions and dialects of many of these languages and therefore only generic or family names have been used instead of listing all the various versions.

Since a number of papers dealing with specific languages are being presented at this conference, we will defer any further discussion of languages. Chapter 3 of reference 7 provides further insight and guidance in selecting an appropriate language.

Validation

Validation is the process of bringing to an acceptable level the user's confidence that any inference about a system derived from the simulation is correct. It is impossible to prove that any simulator is a correct or "true" model of the real system. Fortunately, we are seldom concerned with proving the "truth" of a model. Instead, we are mostly concerned with validating the insights we have gained or will gain from the simulation. Thus, it is the operational utility of the model and not the truth of its structure that usually

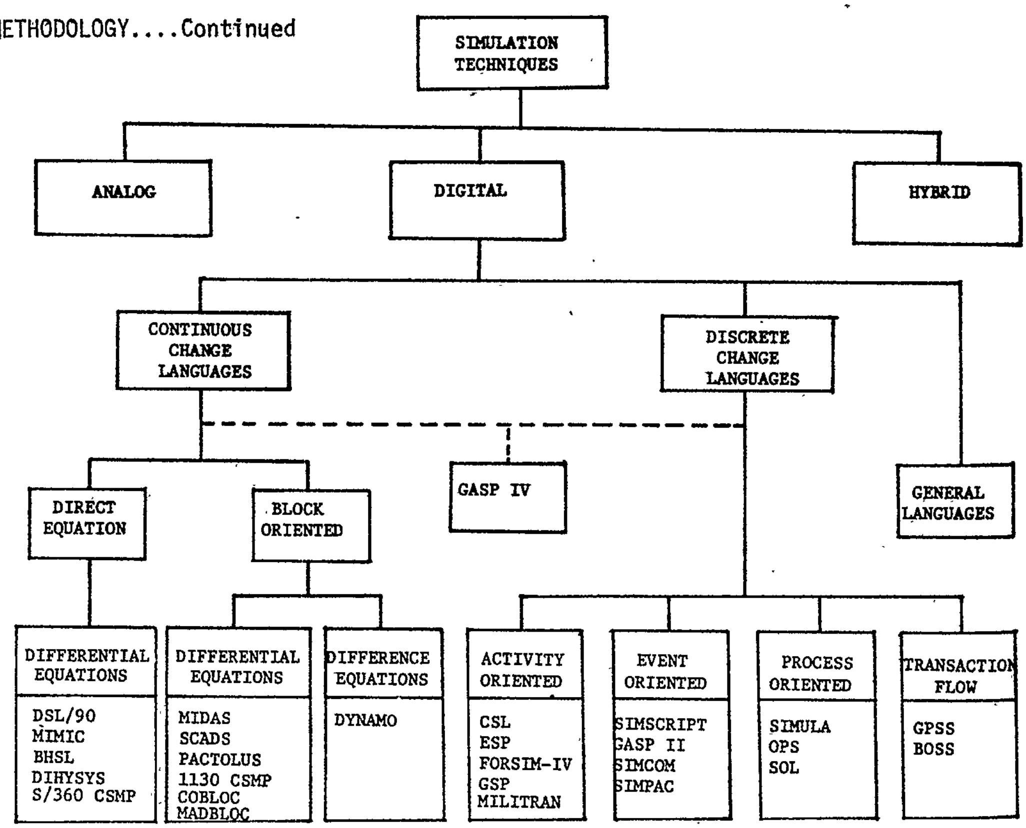

Figure 2. Programming Languages

concerns us.

Validation of the model is extremely important, because simulators look real and both modelers and users find them easy to believe. Unfortunately, simulators often conceal their assumptions from the casual observer and sometimes even from the modeler. Therefore, if validation and evaluation are not carried out carefully and thoroughly, erroneous results may be accepted with disastrous consequences.

There is no such thing as the "test" for validity. Rather, the experimenter must conduct a series of tests throughout the process of developing the model in order to build up his confidence. Three tests may be used to validate a model. First, we must ascertain that the model has face validity. For example, is it possible for the model to give absurd answers if parameters are carried to extreme values? We must also ask if the results of the model appear to be reasonable.

The second method of validation is the testing of assumptions and the third is the testing of input-output transformations. These latter two may entail the use of statistical tests of means and variances, analysis of variance, regression, factor analysis, spectral analysis, auto-correlation, chi-square, and nonparametric tests. Since each of these statistical tests make assumptions about the underlying process, the use of each raises questions of validity. Some statistical tests require fewer assumptions than others, but in general the power of the test decreases as the assumptions are relaxed.

Fishman and Kiviat [9] divide the evaluation of simulations into three categories: (1) verification--insuring that the model behaves the way an experimenter intends; (2) validation--testing the agreement between the vehavior of the model and that of the real system; and (3) problem analysis--the drawing of statisitcally significant inferences from the data generated by the computer simulation. Often a large number of actions are necessary to carry out this evaluation, ranging from testing the model stage-by-stage on a desk calculator before assembling the stages into a machine program, to conducting field experiments. However, the testing itself suffers from the standard problems of empirical research: (1) small samples owing to high cost of data; (2) data that are too aggregated; and (3) data of questionable validity.

The question of validation is thus two-faced: determining that the model behaves in the same fashion as the real life system; validating that the inferences drawn from the experiments with the model are valid or correct. In concept, both these points resolve themselves to the standard decision problem of balancing the cost of each action against the value of the increased information and the consequences of erroneous conclusions.

Validation and analysis of a simulation study is a continuous process that begins from the start of the study. Confidence is built into the model as the study proceeds. It is not something done solely at the end. The greatest possible validity is achieved by:

1. Using common sense and logic throughout the study.

2. Taking maximum advantage of the knowledge and insight of those most familiar with the system under study.

3. Empirically testing by the use of appropriate statistical techniques all of the assumptions, hypothesis, etc. that possibly can be tested.

4. Paying close attention to details, checking and rechecking each step of the model building process.

5. Assuring that the model performs the way it was intended by using test cases etc. during the debugging phase.

6. Comparing the input-output transformation of the model and the real world system (wherever possible).

7. Running field tests or peripheral research where feasible.

8. Performing sensitivity analysis on input variables, parameters, etc.

9. Checking carefully the predictions of the model and actual results achieved with the real world system.

Strategic Planning

We have defined simulation as being experimentation via a model to gain information about a real world system. It then follows that we must concern ourselves with the strategic planning of how to design an experiment that will yield the desired information. The design of experiments is a topic whose relevance to simulation has long been acknowledged but rarely applied in practice. The design of a computer simulation experiment is essentially a plan for purchasing a quantity of information which may be acquired at varying prices depending upon the manner in which the data are obtained. Since the first publication in 1935 of R. A. Fisher's book, The Design of Experiments, a great number of books and papers on experimental design have appeared and the use of designed experiments has found wide spread application. The purpose of using these designs is twofold: (1) they are economical in terms of reducing the number of experimental trials required and, (2) they provide a structure for the investigator's learning process. The running of a simulation experiment is the process of exercising or running the model so as to observe and analyze the resulting information to obtain the desired answers. The experimental design identifies a particular approach for gathering the information needed to allow valid inferences to be drawn.

The development of experimental design techniques which are most suitable for computer simulation experiments has been virtually ignored, probably due to the fact that most analysts are unaware of the fact that conventional techniques are often not completely suitable for simulation. Previous surveys of the state of the art are given by Naylor, Burdick and Sasser [9], and Hunter and Naylor [10]. Odom and Shannon [11] present a series of nomographs which can be useful to the analyst in making the design trade offs required to keep the experimentation costs within available resources.

Depending upon the specific purpose of the experimenter, there are several different types of analysis which may be required. Among the more common are:

1. Comparison of means and variances of alternatives.

2. Determining the importance or effect of different variables and their limitations.

3. Searching for the optimal values of a set of variables.

Designs to accomplish the first type of analysis are generally called single-factor experiments and are fairly straightforward, with the major concerns of the experimenter being such matters as sample size, starting conditions, and the presence or absence of autocorrelation. The second type of analysis is one toward which most textbooks on design and analysis of experiments are directed. These designs primarily utilize analysis of variance and regression techniques for the interpretation of the results. The third type of analysis usually requires search techniques of experimentation.

Tactical Planning

In general, tactical planning involves questions of efficiency and deals with the determination of how each of the test runs specified in the experimental design is to be executed. Primarily, tactical planning is concerned with the resolution of two problem areas: (1) starting conditions, as they affect reaching equilibrium, and (2) the need to reduce the variance of the answer as far as possible while minimizing the required sample sizes.

The first problem (i.e., starting conditions and their effect upon reaching equilibrium) arises from the artificial nature of model operation. Unlike the real world the model represents, a simulation model operates only periodically. That is, the experimenter starts the model, obtains his observations, and shuts it down until the next run. Each time a run is started, it may take a certain period of time for the model to reach equilibrium conditions representative of the real world system operations. Thus, the initial period of operation of the model is distorted owing to the initial start up conditions. The solution is to (1) exclude data for some initial period from consideration, and (2) choose starting conditions that reduce the time required to reach equilibrium. Reasonable starting conditions can reduce but not eliminate the time required for the simulation model to approach equilibrium conditions. Therefore it is still necessary to determine when measurement should begin.

The second part of the tactical planning problem deals with the necessity to estimate the precision of experimental results and the confidence attributable to the conclusions or inferences drawn. This immediately brings us face-to-face with such areas as variability, sample size, and replication. In any experiment, we try to obtain as much information as possible from a limited amount of experimentation. Several techniques for reducing the variance of response have been proposed (mostly in connection with survey sampling procedures), which can significantly reduce the required sample size and number of replications of the experiments. The use of very large sample sizes can overwhelm virtually all the tactical problems of simulation but usually at a great cost in computer and analysis time. The more complex is the simulation model, the more important is good tactical planning before running the experiments.

Experimentation and Sensitivity Analysis

Ultimately, after all development and planning, we run the model to obtain the desired information. At this stage, we begin to find the flaws and oversights in our planning, and to retrace our steps until we achieve our originally established goals.

Sensitivity analysis is one of the most important concepts in simulation modeling. By this we mean determining the sensitivity of our final answers to the values of the parameters used. Sensitivity analysis usually consists in systematically varying the values of the parameters over some range of interest and observing the effect upon the response of the model. In almost any simulation model, many of the set variables are based upon highly questionable data. In many cases, their values may have been determined solely upon the best guess of experienced personnel or very cursory analysis of minimal data. It is therefore extremely important to determine the degree of sensitivity of the results to the values used. If the answer changes greatly with slight variations in the values of some of these parameters, this may provide the motivation and justification for expenditure of more time and money to obtain more accurate estimates. On the other hand, if the results do not change over wide fluctuations in the values of the parameter, no further effort is needed or justified.

Simulation is ideally suited for sensitivity analysis because of the experimenter's degree of control. Unlike experimentation with real world systems, the simulation modeler has absolute control over his model and can vary one parameter at a time if need be, observing the results upon the behavior of the model.

Documentation

The last two elements that must be included in any simulation project are implementation and documentation. No simulation project can be considered successfully completed until it has been accepted, understood, and used. Management scientists' greatest failure has been in gaining acceptance and use of their labors. Rubenstein [12] found one of the greatest causes of failure in operations research and management sciences projects to be the user's inadequate understanding of results, and thus a lack of implementation.

Documentation is closely linked to implementation. Careful and complete documentation of the development and operation of the model can greatly increase its useful life and chances of successful implementation. Good documentation facilitates modification and ensures that the model can be used even if the services of the original developers are no longer available. In addition, careful documentation can help the modeler to learn from his mistakes and perhaps provide a source of subprograms that can be reused in future projects.

Concluding Remarks

The use of simulation has become very extensive in every field of science and technology. As might be expected with such wide-spread use, the state-of-the-art of simulation methodology is fairly advanced. A recent survey of the current state-of-the-art is presented by Shannon [13]. Despite the high level of activity in simulation and developments in mathematics, statistics and computer science, simulation remains almost as much an art as a science. The research reported at conferences such as the present one will help reduce the need for the art and place simulation on a firmer scientific basis.

References

1. Fishman, G.S., Concepts and Methods in Discrete Event Digital Simulation, John Wiley and Sons, Inc., New York, (1973).

2. Emshoff, J.R. and R.L. Sisson, Design and Use of Computer Simulation Models, The MacMillan Co., New York, (1970).

3. Gordon, G., Systems Simulation, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, (1969).

4. Meir, R.C., W.T. Newell and H.L. Pazer, Simulation in Business and Economics, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, (1969).

5. Mirham, G.A., Simulation: Statistical Foundations and Methodology, Academic Press, New York, (1972).

6. Mize, J.H. and J.G.Cox, Essentials of Simulation, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, (1968).

7. Shannon, R.E., Systems Simulation: The Art and Science, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, (1975).

8. Sammet, J.E., "Programming Languages: History and Future," Communications of the ACM, Vol.15, No. 7, (July 1972).

9. Naylor, J.H., D.S. Burdick, and W.E. Sasser, "Computer Simulation Experiments with Economic Systems: The Problem of Experimental Design," Journal of the American Statistical Association, Vol. 62, pp. 1315-1337, (December 1967).

10. Hunter, J.S., and T.H. Naylor, "Experimental Designs for Computer Simulation Experiments," Management Science, Vol. 16, No. 7, (March 1970).

11. Odom, P.R. and R.E. Shannon, "Nomographs for Computer Simulation," Industrial Engineering, Vol. 5, No. 11, (November 1973).

12. Rubenstein, A.H., et al., "Some Organization Factors Related to the Effectiveness of Management Science Groups," Management Science, Vol. 13, No. 8, (April 1967).

13. Shannon, Robert E., "Simulation: A Survey with Research Suggestions," AIIE Transactions, Vol. 7, No. 3, pp. 289-301, (September 1975).