DOCUMENT RESUME

ED 050 765 LI 002 825

AUTHOR	Hawthorne, G. B., Jr.
TITLE	Digital Simulation and Modelling.
INSTITUTION	Mitre Corp., Bedford, Mass.
REPORT NO	MTP-79
PUB DATE	Jun 68
NOTE	23p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	Computer Programs, Computers, Design, *Digital Computers, Evaluation Methods, *Models, Prediction, *Simulation, *Systems Analysis, Techniques
IDENTIFIERS	PEP, *Prepare Educational Planners

ABSTRACT

A basically tutorial point of view is taken in this general discussion. The author examines the basic concepts and principles of simulation and modelling and the application of digital computers to these tasks. Examples of existing simulations, a discussion of the applicability and feasibility of simulation studies, a review of simulation techniques, lists of the advantages and dangers in simulating and a classification of digital computer simulations are also presented. While the emphasis is on the use of digital computers, many of the principles examined hold for simulation done on analogue computers. No attempt is made, however, to compare the uses of the two types of machines for simulation studies. (Author/JN)

OPERATION PEP/EXECUTIVE INFORMATION SYSTEMS

MTP-79

DIGITAL SIMULATION AND MODELLING

G. B. Hawthorne, Jr.

The MITRE Corporation/Bedford, Massachusetts/June 1968

OPERATION PEP/EXECUTIVE INFORMATION SYSTEMS MTP-79

DIGITAL SIMULATION AND MODELLING

by

G. B. Hawthorne, Jr.

June 1968

This paper has been especially adapted for distribution to the California Educational Administrators participating in the "Executive Information Systems" Unit of Instruction as part of the instructional program of OPERATION PEP (Prepare Educational Planners).

THE MITRE CORPORATION
BEDFORD, MASSACHUSETTS

ABSTRACT

A basically tutorial point of view is taken in the following material, the idea being to introduce some basic concepts underlying the simulation technique in general and to indicate some of the ways in which digital computers can be helpful in this process. While the emphasis is on the use of digital computers, many of the principles examined below hold for simulations done on analog computers. No attempt is made, however, to compare the uses of the two types of machine for simulation studies.

This paper draws from a wide variety of opinion, both as expressed in the diverse literature on the subject and as communicated personally to the author. For this reason, no individual credits or specific references are given in the paper itself; however, a short bibliography is given at the end.

ACKNOWLEDGMENT

The work reported in this paper was originally conducted under the sponsorship of the Electronic Systems Division of the U.S. Air Force Systems Command. The original version of this document was deposited with the Defense Documentation Center (formerly ASTIA) as ESD-TDR-64-106.

CONTENTS

DIGITAL SIMULATION AND MODELLING*

DEFINITION AND BASIC IDEAS

If one were to ask an engineer, a computer programmer, an experimental psychologist, and an operations researcher to define the word "simulation," four quite different definitions might well result. So varied are the techniques and so broad is the spectrum of applications for the collection of things which might, in toto, be considered the "field of simulation" that one is hard put to define this "field." Nevertheless, the basic unifying concept, neither new nor complex, is simply that one may construct a more or less faithful representation of some real object or process and then experiment with the representation rather than with the real thing. Two ideas are essential:

(1) Similarity

The representation is in some sense "like" the real thing.

(2) Nonidentity

The representation is not the real thing.

As a consequence of (1), the simulation may be used to predict performance in the real world, and these predictions will be valid to the same extent that the simulation is a faithful copy of reality. As a consequence of (2), the "imitation" frequently costs less, or can be constructed and tested more easily than the thing imitated.

*The original version of this paper was published by The MITRE Corporation in March 1964. It subsequently appeared in Datamation, 10 (10), pp. 25 to 29 (October 1964).

At this point, it is desirable to sharpen these ideas by making a clear distinction between "simulation" and "model." We shall use the noun "model" to refer to our analog, or representation of reality, and the noun "simulation" (act of simulating) to refer to the construction and use of this model for testing, prediction, and design. As an example, consider the testing of a new airfoil design by means of a scale model in a wind tunnel. The model is, perhaps, a piece of wood or metal which, although reduced in size from an actual aircraft wing, is similar in shape, surface smoothness, and other characteristics. Undoubtedly, the list of characteristics in which the model is not similar to the real thing is a more extensive one and might include structural stiffness and strength, relative center of gravity, ratio of area to volume, and, of some importance, cost and ease of manipulation. The simulation, in this case, would consist of the construction of the model, its subsequent testing in the wind tunnel, and the drawing of conclusions about the behavior of a full-size wing. Clearly, "a model" is not synonymous with "a simulation," although "modelling" is an essential part of "simulating."

Models, and hence the simulations involving them, may be classified according to the degree to which elements of the model correspond to elements of the real-world object or process. If there are no human or random factors in the real object or process, the closest approach to reality (and lowest level of abstraction) is obtained by constructing a "scale model," i.e., a reduced-size operating model. By "size" we do not necessarily mean physical dimension, but include such things as complexity, extent of detail, and number of elements. Such "scaling" often introduces problems. For example, one cannot necessarily evaluate the performance of a 20-channel multiplex communication system by studying a 10-channel system loaded to the same percentage of capacity.

Somewhat further removed from the real world, and hence easier to construct and manipulate, are what are commonly called simulation models. These may no longer resemble their real-world counterparts in appearance; being at a higher level of abstraction, their similarity is more logical than physical. The closeness of a model to the real world has been described in terms of a spectrum, or scale of abstraction, with the real world at one end and the mathematical model, whose resemblance is merely symbolic, at the other end. As one moves along this scale toward the mathematical model, the generality of the model increases; i. e., the results obtained from exercising it apply to larger and larger classes of real-world objects. At the same time, its level of validity, when referred to a specific application, as well as its level of detail, decrease.

The discerning reader may see an apparent inconsistency in the foregoing remarks concerning the place of simulation models and mathematical models along the scale of complexity, based upon the fact that simulation models cannot be constructed without first defining the mathematical relationships which exist. Indeed, a simulation model may be viewed as a "mechanization" of a set of mathematical relationships. Why does this set of relationships not constitute a mathematical model, and thus imply that simulation models and mathematical models are in the same region of the scale? The answer is, of course, that it does. The problem, however, is that the "mathematical model" defined by the set of mathematical relationships underlying the simulation is an intractable one. If it were not, there would be no point in going to a simulation. Thus, when we say that the end of the scale corresponding to the greatest generality and the least detail is occupied by mathematical models, we really mean tractable mathematical models. In this sense, such models are "simpler" than those underlying a simulation. Finally, all of these statements assume that one desires to model a fairly complex real-world process; if the real world is simple, then so will be both its simulation model and its mathematical model.

The difficulty of making precise classifications of models is illustrated by the foregoing. A better grasp of the underlying ideas is gained by considering briefly what one actually does in constructing a simulation model of a complex process. Assuming that the process is so complex that one cannot easily formulate a tractable mathematical model from a study of the process as a whole, one seeks to divide the process into a set of interrelated subprocesses. The major requirement on this decomposition process is that it must result in well-defined subdivisions or process elements whose relationships are well understood and for which one can construct tractable mathematical models. In essence, the mathematical models are made tractable by considering the problem in smaller pieces. There still remains the problem of putting the pieces together to form the overall model. It is expected that this overall model will be difficult to handle analytically; nevertheless, it is needed as a basis for the simulation model.

Putting the pieces together involves writing additional mathematical and/or logical expressions describing the relationships between the various subdivisions. For example, the outputs of certain process elements may be used as inputs to certain other elements, or, by means of closed loops, as inputs to themselves. This mathematical description of the "boundary conditions" for each element is then added to the mathematical descriptions of the elements to form the overall mathematical model.

A model constructed in the manner just described generally consists of a large set of mathematical equations, each of which is capable of "mechanization" on a computer. In many cases, a digital computer provides a convenient facility for this purpose. When the set of mathematical relationships is reduced to a computer program, the computer is able simultaneously to take all of them into account, and thus to simulate the operation of the entire process. It is true that the computer's function still could be carried out by humans using desk calculators, reference books, summary sheets, and following some complex master schedule of operations, but the cost and time involved make this impractical.

In considering the simulation spectrum which runs from complete reality to complete abstraction, one is struck by the inherent tradeoff involved in locating a simulation somewhere in this range. This tradeoff is, of course, between realistic detail and cost: as the model moves away from reality, it becomes simpler and hence easier to construct and manipulate, while a move toward realism usually involves added complexity and cost. The two sides of the simulation coin, then, are:

(1) By ignoring relatively unimportant details, it becomes practical to construct and exercise a simulation model at much less cost and time than that required to exercise the real world.

(2) By being a less-than-complete copy of reality, the simulation becomes susceptible to falsely predicting the real world.

There is, in general, no well-defined set of rules by which one can fix the amount of realistic detail in the model: this choice being a partially subjective one depending on judgment and experience. Indeed, the simulation technique itself may be used to fix this choice, by revealing problem parameters which do not significantly affect results.

EXAMPLES OF SIMULATIONS

Consider the air-traffic-control problem, involving various aircraft, their positions, altitudes, and speeds, rules of procedure for dealing with conflicting airspace requirements, landing and take-off priorities, and airport approach congestion. Several simulation studies have been made in this area, utilizing digital computers to keep track of the aircraft, their stacking at congested airports, and delays associated with deviations from scheduled arrival times.

Another simulation example is the processing of various items by collections of machines and workers in production plant "job shops." One problem is to plan the operations so that various pieces of work proceed through the shop in reasonable time with minimum use of overtime, subcontracting, or extra workers. Digital computer simulation of the job shop allows consideration of the physical arrangement of equipment, the decision and priority rules involved in routing the work, the processing times for various machines and workers, temporary storage requirements for partly-finished lots of material, and similar items. Output data from such a simulation might include percentage utilization of various machines and facilities, wasted time (for either machines or workers) due to poor scheduling, overloads on particular machines or workers, and average processing time per item. Other things, such as inventory levels for raw material, stocks of finished goods on hand, and the effect of rush orders can be taken into account.

Here again, in these two examples, the large amount of bookkeeping involved makes the digital computer an extremely useful way of doing the simulation.

APPLICABILITY AND FEASIBILITY OF SIMULATION STUDIES

The primary function of a simulation is to predict. This function can be employed in a number of ways. In evaluation, the predicted behavior or some real-world system (either existent or contemplated) is used to assess the desirability or "goodness" of that system. This notion can be extended to comparison of real-world processes by making the evaluation a relative one. A further extension, consisting of alternating "cut and try" or "propose and evaluate" cycles, results in an iterative design application. For optimization, a process not inherent in simulation itself, the iterative design process would be carried out under some plan which ensures convergence to the optimum value of some parameter (say, by exhaustively trying all possible values) or to within some specified "distance from this optimum value.

All of the foregoing uses of simulation apply, of course, to the study of systems, including static, dynamic, automated, and man/machine. System simulation is indicated:

(1) If the system is too complex to be reduced to a tractable mathematical formulation, such as a small set of equations.

(2) When analytical or mathematical techniques do not exist for the problems involved (such as certain differential equations which are nonlinear or whose coefficients and/or boundary conditions are time-varying).

(3) When it is impossible to experiment with the real-world system (such as astronomical systems) or to observe the system in its natural environment.

(4) When, in systems containing random elements, probability distributions are desired but unobtainable experimentally.

(5) When it is desirable to check the results of an analytical study of the system without building and/or testing the actual system.

(6) When, because of complexity, realistic field tests of the system are not practical. This might arise, for example, in a system having multiple inputs which are difficult to physically reproduce. The inputs might then be simulated to the real system; in doing this, one would be simulating part of the system environment.

(7) When, as often occurs in operations research problems, it is desirable to plan policy for operating a large system without actually trying out various alternative policies.

(8) When there exists the problem of training operators of a large system and it is not practical to do this on the real system (it may not offer a predictable variety of possible conditions, or may offer some of them only rarely).

If a digital computer simulation is indicated, there is a need to examine the feasibility of such a study. A simulation is possible only if the system can be broken up into a set of interrelated elements having operating rules which can be specified (if only in a probabilistic sense). Certainly, the feasibility depends upon the availability of a computer, programmers, and problem analysts. A knowledge of the capabilities of machines and the time and cost of producing and running programs clearly is needed to assess feasibility and decide on the proper level of detail in the model. Generally, the larger the data storage (memory) of a computer and the more flexible its logic, the better suited it is for simulation of large systems.

The time and dollar costs for simulations vary as much as the models used. A typical "small" simulation might require on the order of 1500 instructions and 500 words of memory. A particular example of this type, involving the air-traffic application mentioned earlier, required up to 55 minutes of computer time to follow 35 to 75 aircraft for three hours of real-time flight activity. With computer time costing hundreds of dollars per hour, even a small simulation is not trivially expensive, especially when many runs are contemplated. By way of contrast, a "large" simulation may require on the order of 25,000 words of program and, roughly, the same number of words of fast-access memory, plus considerable secondary storage. This could entail as many as eighteen months of programming effort. In some cases, a limitation on the amount of computer storage available can be overcome by doing the simulation in separate "blocks," with the outputs of certain blocks being used as inputs to others. The various parts of the simulation then can be run at different times, and, in fact, on different computers, if desirable.

The cost of simulation for large systems is generally between that of analysis and experimental testing, being more expensive than the former and less expensive than the latter (there are exceptions, of course). A simulation could take from a few weeks to several years and cost from a few thousand to several hundred thousand dollars, whereas building and testing a real system could take years and cost millions. The real measure of cost is the cost of doing without simulation, and this could be as great as building a useless system.

SIMULATION TECHNIQUES

Various lists of steps in performing a simulation have been suggested. A typical one is:

(1) The system is studied by one or more analysts who roughly fix the size and scope of the model and its nature, as determined by the type of system, the questions to be answered, and the support available.

(2) The system is divided into subsystems whose relationships are known (input/output relationship, for example). A logical flow diagram showing these elements and their interconnection is drawn up. Various questions, such as how to simulate the passage of time and what system functions or elements can be safely left out of the model, must be answered here.

(3) A full-scale test of the model may be carried out manually at this point, by going step-by-step through the simulation using a desk calculator, followed by revision and/or modification.

(4) A computer having adequate storage and logical flexibility is chosen. A program flow chart is developed and the problem is coded for the machine, using one of the available programming languages.

(5) The program is run and debugged, and, if possible, the model is calibrated against the results of tests on actual equipment.

(6) A detailed test program is designed for using the simulation, including such things as number of values to be used for different variables, number of cases to be treated, and (if the model contains random elements) number of runs required to give statistically valid results.

(7) A number of runs are made, as called for by the test plan, and the resulting data are analyzed and used to predict the performance or quality of the system being simulated.

In the foregoing list, items are seen to fall into one of two classes: those concerned with the design of the model, and those concerned with its use or application. If different groups are concerned with these two categories, it is important that they coordinate their activities, since decisions in either area have an effect on the other.

Of the specific techniques useful in simulation, the Monte Carlo technique is one of the more important. To "Monte Carlo" a particular thing means, very roughly, to "randomize" it. If the time-of-arrival of an aircraft is known to be somewhat variable, or unpredictable, it may be called a "random variable" or a "chance variable." There are degrees of randomness, however, and the time of arrival may be known, from experience, to vary no more than 45 minutes either way, in a particular instance. In simulating the aircraft's flight, the number representing arrival time would not be fixed (this would be unrealistic) but would be "Monte Carloed," allowing it to take on random values (within the 90-minute limitation). Each simulation run would result in a different, and unpredictable, arrival time, but after many Monte Carlo runs, a pattern would begin to emerge. In this case, the pattern would be the one imposed by the simulation designer when he specified a probability distribution for the random number, and might be described, after 100 runs, by the statement that 10 arrivals (or 10 percent of the total) were within 5 minutes of the scheduled time, 40 arrivals (or 40 percent) were within 20 minutes of schedule, and so on. An alternative description would be that the aircraft could be expected to arrive no more than 5 minutes early or late with a probability of 0.1, or that the probability of a 20-minute deviation from schedule was 0.4, and so on. Such a description is called a probability distribution.

A practical application of Monte Carlo would be one in which certain variables were chosen at random during each run, again according to some distribution imposed by the designer, but in which these variables react in the simulation to produce a new variable (such as the number of aircraft waiting in a "stack") which is, accordingly, also random. After many Monte Carlo runs, one would be able to obtain the pattern or distribution (unknown at the start) for the new variable and to make such statements as, "There will be at least eight aircraft in the stack 75 percent of the time," or, "With probability 0.75, there will be eight or more aircraft waiting to land." Clearly, any system or process containing randomness can only be specified by such statements, and can only be evaluated, for a given exercise of the system, by a Monte Carlo technique. On the other hand, since randomness does not always play an important part, Monte Carlo is not necessarily a part of a simulation.

A second specific technique useful in digital simulation is the writing of computer programs in special simulation languages, such as the Gordon GPS Simulator, SIMPAC and SIMSCRIPT. The latter is similar in general configuration to FORTRAN, but is specifically designed to make it easy to write certain types of simulation programs. In the vocabulary of SIMSCRIPT, a system consists of a set of entities, each of which can have several numerically described attributes. There are two types of entities: temporary (those which enter into only part of the simulation) and permanent (those which exist throughout the entire simulation). Examples of these types might be, respectively, a single aircraft flight which lasts for only two minutes out of a 20-minute simulation run, and an airport facility which exists throughout the run. The attributes of a permanent entity may take on several values during a run (the number of runways available at a given time, for example) and one might well be interested in the probabilities associated with these different "states" of the system.

ADVANTAGES AND DANGERS IN SIMULATING

At the risk of repetition, a summary of some advantages of simulation follows:

(1) Simulation provides experience and permits experimenting without the risks and costs involved in dealing with the real thing.

(2) With respect to systems, simulation and modelling permit the demonstration of system operation before hardware is built, the evaluation of already constructed systems which cannot be adequately tested otherwise, aid in setting up field test procedures for system checkout, and evaluation of a system's capability for assuming missions or performing functions not originally specified.

(3) Simulation can be considerably faster than making comparable operational tests on equipment.

(4) Environmental conditions, system parameters, and subsystem operating characteristics may be varied almost at will in many simulation models. "Changes" in the system involve only changes in programs or substitution of new programs. Such changes can, of course, be difficult in certain cases.

(5) Simulation of a complex process may provide an indication of which variables are especially important, and may reveal unforeseen difficulties resulting from apparently minor changes in the system or its environment. Such indications may lead to evolution of new policies and ideas, or the realization of simple but hidden truths.

(6) Simulation gives control over time. In a dynamic model, one may either compress or expand time from its real-world pace.

(7) Simulation generally has beneficial "fallout." The data collected may turn out to be useful in answering questions other than the ones which led to the simulation, since these data may be analyzed and reassembled in a variety of ways.

As with any technique, simulation has its dangers as well as its advantages. Simulation is not always a faster, cheaper way of doing the job; there are certainly real-world processes which are better handled by the analytical methods of mathematics or by prototype testing. Generally speaking, the larger and more complex the process, the more advantages are offered by simulation. Hence, one of the primary dangers is the poor use of simulation; i.e., its use in cases where other methods are indicated. Even when indicated, the technique can be carried to extremes such that the diminishing returns are not worth the additional costs. Fortunately, there exist, at least for Monte Carlo schemes, statistical criteria which indicate a reasonable maximum number of runs. Techniques also exist for designing the test plan, i.e., choosing the combinations of parameter values to be used for simulations having a large number of variables.

Another class of dangers can be grouped under the label of poor design. In some areas, at least, simulation design is more art than science, and there often must be strong reliance on common sense and experience. Some specific pitfalls are given in the (by no means complete) list below:

(1) Sometimes the designer cannot foresee all of the variables needed; this results either in important omissions or in such a conservatively large number of variables that test design and/or data evaluation become hopeless tasks.

(2) There is danger that the simulation will be designed with too much emphasis on imitating the real world and too little emphasis on the questions to be answered or on the problems to be solved.

(3) In simulations of man/machine systems, there is the danger of inadequately representing the human being. When simple motor tasks are to be performed, a fixed or Monte Carlo time delay may suffice. When the human's response is functional, the problem becomes difficult because of the non-linear stimulus response, non-zero memory time, and highly complicated pattern-recognition capabilities of this "device." When decision-making functions are also to be considered, the only recourse is to make the human being part of the simulation. This has many ramifications, such as the requirement for carrying out the simulation in real time and the problem of choosing a set of humans for participation which is representative of the class occurring in the real system.

(4) The simulation design can be too narrowly conceived and thus limited in its application. A rigid adherence to only the specific requirements envisioned at the start may result in a simulation program so inflexible and incapable of expansion or modification as to be of little value.

CLASSIFICATION OF SIMULATIONS

One division of general simulation techniques that can be made is manual versus automated. Automated techniques can be further divided into simulations utilizing analog computers and those utilizing digital computers, the latter being particularly appropriate in the area of systems simulation. As a final overview of this latter area, one might attempt a classification, by various (and somewhat overlapping) categories, of digital computer simulations, as follows:

(1) Simulation of the environment for a real system, as opposed to simulation of the system in a real environment, or to simulation of both.

(2) Models of noncomputer-based systems constructed <u>on</u> a computer as opposed to models which <u>use</u> a computer in the way which it is used in the real (computer-based) system. Here, the actual system computer may be used, or its program may be simulated by a different type of computer.

(3) Analytic versus real-time simulations. The latter is self-explanatory. the former is based on a model in which time-pacing or sequencing of events is not important. A third class includes time-based models which are run on the computer in scaled time.

(4) Classification by use or purpose, such as design, evaluation, training, or the study of human reactions in a simulated environment (operational gaming).

(5) Deterministic versus Monte Carlo.

(6) Man/machine simulations as opposed to pure machine simulations. In the former, the emphasis is usually upon real-world processes in which the human decision-making role is of paramount importance, such as in military command and control systems. In the latter, the system being modelled includes human elements almost inconsequentially, such as in automated sensing and reporting systems.

SUMMARY

The basic ideas involved in experimentation with models are both familiar and possessed of a long history of application. A relatively recent culmination of endeavors in one particular application, the prediction of large system performance, is the use of large digital computers in performing simulations. When certain conditions of feasibility are met by the system processes, such computers often constitute a tremendous aid in modelling these processes.

As with any technical tool, simulations pose problems in addition to providing helpful answers. These dangers and advantages are, to some extent, reflections of the basic tradeoff that must be made between realistic detail and cost (or time). The final choice is one between accuracy of prediction and ease of prediction.

BIBLIOGRAPHY

Adams, R. H., and Jenkins, J. L., "Simulation of Air Operations with the Air-Battle Model," Operations Research, 8:5, Sept.-Oct. 1960.

Buffa, E. S., Models for Production and Operations Management, John Wiley and Sons, Inc., New York, 1963.

Conway, R. W., "Some Tactical Problems in Digital Simulation," Management Science, 10:1, October 1963.

Goode, H. H., "Simulation--Its Place in System Design," Proceedings of the IRE, 39:12, December 1951.

Goode, H. H., and Machol, R. E., System Engineering, McGraw-Hill Book Company, New York, 1957.

Greenberger, M., "Simulation and the Problems of Air Traffic Control," Industrial Management Review, May 1961.

Harling, J., "Simulation Techniques in Operations Research--A Review," Operations Research, 6:3, May-June 1958.

Harris, B., "Some Problems in the Theory of Intra-Urban Location," Operations Research, 9:5, Sept.-Oct. 1961.

Haythorne, W. E., "Information Systems Simulation and Modelling," First Congress on the Information System Sciences, Session 7, November 1962.

LeGrande, E., "The Development of a Factory Simulation System Using Actual Operating Data," Management Technology, 3:1, May 1963.

Malcolm, D. G., "Bibliography of the Use of Simulation in Management Analysis," Operations Research, 8:2, March-April 1960.

Markowitz, H. M., Hausner, B., and Karr, H. W., SIMSCRIPT, A Simulation Programming Language, Prentice-Hall, Englewood Cliffs, New Jersey, 1963.

McKenney, J. L., "An Evaluation of a Decision Simulation as a Learning Environment," Management Technology, 3:1, May 1963.

Morgenthaler, G. W., "The Theory and Application of Simulation in Operations Research," Chapter 9 of Progress in Operations Research, Vol. 1, John Wiley and Sons, Inc., New York, 1961.

Shubik, M., "Bibliography on Simulation Gaming, Artificial Intelligence and Allied Topics," Journal of American Statistical Association, Vol. 55, December 1960.