21

Simulation

Simulation is a very powerful and widely used management science technique for the analysis and study of complex systems. In previous chapters, we were concerned with the formulation of models that could be solved analytically. In almost all of those models, our goal was to determine optimal solutions. However, because of complexity, stochastic relations, and so on, not all real-world problems can be represented adequately in the model forms of the previous chapters. Attempts to use analytical models for such systems usually require so many simplifying assumptions that the solutions are likely to be inferior or inadequate for implementation. Often, in such instances, the only alternative form of modeling and analysis available to the decision maker is simulation.

Simulation may be defined as a technique that imitates the operation of a real-world system as it evolves over time. This is normally done by developing a simulation model. A *simulation model* usually takes the form of a set of assumptions about the operation of the system, expressed as mathematical or logical relations between the objects of interest in the system. In contrast to the exact mathematical solutions available with most analytical models, the simulation process involves executing or running the model through time, usually on a computer, to generate representative samples of the measures of performance. In this respect, simulation may be seen as a sampling experiment on the real system, with the results being sample points. For example, to obtain the best estimate of the mean of the measure of performance, we average the sample results. Clearly, the more sample points we generate, the better our estimate will be. However, other factors, such as the starting conditions of the simulation, the length of the period being simulated, and the accuracy of the model itself, all have a bearing on how good our final estimate will be. We discuss such issues later in the chapter.

As with most other techniques, simulation has its advantages and disadvantages. The major advantage of simulation is that simulation theory is relatively straightforward. In general, simulation methods are easier to apply than analytical methods. Whereas analytical models may require us to make many simplifying assumptions, simulation models have few such restrictions, thereby allowing much greater flexibility in representing the real system. Once a model is built, it can be used repeatedly to analyze different policies, parameters, or designs. For example, if a business firm has a simulation model of its inventory system, various inventory policies can be tried on the model rather than taking the chance of experimenting on the real-world system. However, it must be emphasized that simulation is not an optimizing technique. It is most often used to analyze "what if" types of questions. Optimization with simulation is possible, but it is usually a slow process. Simulation can also be costly. However, with the development of special-purpose simulation languages, decreasing computational cost, and advances in simulation methodologies, the problem of cost is becoming less important.

In this chapter, we focus our attention on simulation models and the simulation technique. We present several examples of simulation models and explore such concepts as random numbers, time flow mechanisms, Monte Carlo sampling, simulation languages, and statistical issues in simulation.

21.1 Basic Terminology

We begin our discussion by presenting some of the terminology used in simulation. In most simulation studies, we are concerned with the simulation of some system. Thus, in order to model a system, we must understand the concept of the system. Among the many different ways of defining a system, the most appropriate definition for simulation problems is the one proposed by Schmidt and Taylor (1970).

DEFINITION ■ A **system** is a collection of entities that act and interact toward the accomplishment of some logical end. ■

In practice, however, this definition generally tends to be more flexible. The exact description of the system usually depends on the objectives of the simulation study. For example, what may be a system for a particular study may be only a subset of the overall system for another.

Systems generally tend to be dynamic—their status changes over time. To describe this status, we use the concept of the state of a system.

DEFINITION ■ The **state** of a system is the collection of variables necessary to describe the status of the system at any given time. ■

As an example of a system, let us consider a bank. Here, the system consists of the servers and the customers waiting in line or being served. As customers arrive or depart, the status of the system changes. To describe these changes in status, we require a set of variables called the **state variables.** For example, the number of busy servers, the number of customers in the bank, the arrival time of the next customer, and the departure time of the customers in service together describe every possible change in the status of the bank. Thus, these variables could be used as the state variables for this system. In a system, an object of interest is called an **entity,** and any properties of an entity are called **attributes.** For example, the bank's customers may be described as the entities, and the characteristics of the customers (such as the occupation of a customer) may be defined as the attributes.

Systems may be classified as discrete or continuous.

DEFINITION ■ A **discrete system** is one in which the state variables change only at discrete or countable points in time. ■

A bank is an example of a discrete system, since the state variables change only when a customer arrives or when a customer finishes being served and departs. These changes take place at discrete points in time.

DEFINITION ■ A **continuous system** is one in which the state variables change continuously over time. ■

A chemical process is an example of a continuous system. Here, the status of the system is changing continuously over time. Such systems are usually modeled using differential equations. We do not discuss any continuous systems in this chapter.

There are two types of simulation models: static and dynamic.

DEFINITION ■ A **static simulation model** is a representation of a system at a particular point in time. ■

We usually refer to a static simulation as a **Monte Carlo simulation.**

DEFINITION ■ A **dynamic simulation** is a representation of a system as it evolves over time. ■

Within these two classifications, a simulation may be deterministic or stochastic. A **deterministic simulation model** is one that contains no random variables; a **stochastic simulation model** contains one or more random variables. Discrete and continuous simulation models are similar to discrete and continuous systems. In this chapter, we concentrate mainly on discrete stochastic models. Such models are called *discrete-event* simulation models. Discrete-event simulation concerns the modeling of a stochastic system as it evolves over time by a representation in which state variables change only at discrete points in time.

21.2 An Example of a Discrete-Event Simulation

Before we proceed to the details of simulation modeling, it will be useful to work through a simple simulation example to illustrate some of the basic concepts in discrete-event simulation. The model we have chosen as our initial example is a single-server queuing system. Customers arrive into this system from some population and either go into service immediately if the server is idle or join a waiting line (queue) if the server is busy. Examples of this kind of a system are a one-person barber shop, a small grocery store with only one checkout counter, and a single ticket counter at an airline terminal.

The same model was studied in Chapter 20 in connection with queuing theory. In that chapter, we used an analytical model to determine the various operating characteristics of the system. However, we had to make several restrictive assumptions to use queuing theory. In particular, when we studied an $M/M/1$ system, we had to assume that both interarrival times and service times were exponentially distributed. In many situations, these assumptions may not be appropriate. For example, arrivals at an airline counter generally tend to occur in bunches, because of such factors as the arrivals of shuttle buses and connecting flights. For such a system, an empirical distribution of arrival times must be used, which implies that the analytical model from queuing theory is no longer feasible. With simulation, any distribution of interarrival times and service times may be used, thereby giving much more flexibility to the solution process.

To simulate a queuing system, we first have to describe it. For this single-server system, we assume that arrivals are drawn from an infinite calling population. There is unlimited waiting room capacity, and customers will be served in the order of their arrival—that is, on a first come, first served (FCFS) basis. We further assume that arrivals occur one at a time in a random fashion, with the distribution of interarrival times as specified in Table 1. All arrivals are eventually served, with the distribution of service times shown in Table 2. Service times are also assumed to be random. After service, all customers return to the calling population. This queuing system can be represented as shown in Figure 1.

Before dealing with the details of the simulation itself, we must define the state of this system and understand the concepts of events and clock time within a simulation. For this