Proceedings of the 2003 Winter Simulation Conference
S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice, eds.

DESIGNING SIMULATION PROJECTS

Melanie L. Hatch
Donald L. Byrkett

Computer Science and Systems Analysis Department
Miami University
Oxford, OH 45056, U.S.A.

ABSTRACT

Simulation projects provide a useful way to tie together the expected learning outcomes in a simulation class. Designing a good project is a challenging task and simulation instructors are always on the lookout for interesting ideas. This paper provides simulation instructors with two interesting simulation applications, a supply chain project and a web site design project. The supply chain problem is based on a real company problem and the web site design problem is based on a textbook problem. Each application is described in detail and links are provided so that other simulation instructors can access handouts and possible solutions. The authors also provide some insights about simulation project design.

1 INTRODUCTION

Teaching a simulation class is quite challenging. A typical syllabus for a discrete event simulation class includes problem formulation, modeling, simulation programming, input data analysis, random number generation, random variate generation, verification and validation, output analysis, and experimentation. Though it is obvious to us, students often have difficulty connecting these topics together. Providing a simulation project as a part of the course requirements is a very useful way to make the necessary connections and to insure that students understand how each part of the course can be used in a simulation study.

Designing a good simulation project is one of the more difficult aspects of teaching a simulation class, especially when the class size is 30 or more students. With the larger class size it is almost impossible to give a totally open-ended problem in which the students go through the problem formulation step. This means that as part of the simulation assignment, students must be given the problem formulation including the controllable variables and a description of the system. If the project has a real client, this information can be presented by the client, otherwise the instructor will need to prepare a write-up of the problem and a description of the system. Most of the other aspects of the problem can be left for the student to do.

The ideal simulation project will have as many of the following attributes as possible to give the student an opportunity to use their creativity in creating a model and reaching a solution.

- *Unspecified objective* – It is important for students to think through the problem and determine a response variable or variables to appropriately solve the problem. This would allow the student to also make decisions regarding the type of data to be collected.
- *Complexity* – The system being modeled must be complex enough that significant effort will be required for students both to visualize and understand the system and to create a model to accurately represent the system. If possible, the student should be required to make some decisions about the level of detail included in the model.
- *Creativity* – Some aspect of the problem should require some creative modeling. Not all queuing systems should be simple SEIZE-DELAY-RELEASE. Some aspects of the system should require some custom random variate generation, such as a non-homogeneous arrival pattern. Some modeling assumptions should be required to complete the model.
- *Raw data* – Not all modeling distributions should be given in the problem statement. Ideally, the student should determine what data is needed, though this is likely not possible. At the least, raw data should be provided (or collected) for some activities, so that students need to decide how to model the activity using some of the input data analysis methodologies.
- *Unnecessary or missing data* – To help students understand the relationship between modeling and data, it can be useful to omit some needed data from the problem or to provide some unneeded data.

- *Language features* – Normally only a small set of the features of a simulation language are covered in class. Students can learn a lot and gain self confidence in their programming ability by figuring out how to use new features of a simulation language.
- *Model verification* – Validation of the model is usually not possible unless there is a real client and an existing system. However, verification steps can be performed with any model. It is important that students are required to explicitly go through some steps to insure that their model is working as intended.
- *Experimentation* – Experimenting with a properly verified model is needed to link the final topics of the course to the project. The experimental step should include different cases, replications, and statistical testing to reach a statistically significant conclusion. If there are many controllable variables, experimental design may be involved.

Designing a project to meet all of these objectives is challenging, if not impossible. Over time, the projects should involve a wide variety of systems such as manufacturing, computing, military, service, and/or transportation. This variety also adds to the difficulty for the simulation instructor, since the instructor must gain a detailed understanding of many applications.

The main purpose of this paper is to share two project ideas. One problem involves a supply chain system and a real client and the other involves a web service system based on a textbook problem. Sufficient details are provided for each problem to make them useable by other simulation instructors. A web site provides detailed write-ups, data files, solutions, and student reports. In addition, the authors provide a few insights on the advantages and disadvantages of using a real client problem.

2 SUPPLY CHAIN PROBLEM

The Computer Science and Systems Analysis (CSA) department at Miami University enjoys an excellent relationship with Eli Lilly and Company. Eli Lilly has hired many CSA graduates and is currently a member of our Industrial Advisory Council. Because of this relationship, the first author approached Eli Lilly with the idea of collaborating on a case study for the senior-level computer simulation course. Eli Lilly offered to provide background information and raw data from the supply chain for one of their drug product families which we will call LLY123.

Two employees of Eli Lilly and Company visited the simulation class twice during the Fall 2000 semester. They presented an overview of the company and the pharmaceutical supply chain on their first visit. The students then had six weeks to work on the project either individually or in groups. The top three project groups (or individuals) were selected to present their findings to Eli Lilly during a pizza party at the end of the semester.

2.1 Problem Description

At the macro level, the typical stages of Eli Lilly's supply chain consist of: a source of raw material, intermediate and final bulk manufacturing, form/fill and finish processing, distribution centers and wholesale outlets. A network representation is shown in Figure 1. In actuality, any of these high level stages, or nodes in the network can consist of several physical locations. Different facilities in the supply chain may specialize in processing or distributing to specific areas of the world.

The bulk manufacturing site is located in Kinsale, Ireland. It is a large chemical plant that manufactures the active ingredient for LLY123. The facility operates 24 hours per day, 7 days per week, 52 weeks per year. The equipment in the plant is also used to produce two other products for Eli Lilly. The same equipment in the plant is used for intermediate and final processing. Due to a significant cleanup/setup process between intermediate and final processing, LLY123 is produced twice a year.

The process flow for the Kinsale operation consists of intermediate processing, inspection, rework and retest as necessary, and cleanup. These steps are then repeated for final processing. Any product that does not pass inspection, after being reworked, is destroyed.

There are two form/fill and finish facilities that combine the active ingredient from Kinsale with other raw materials to produce the LLY123 tablets. The Puerto Rico facility primarily packages the tablets into bottles and cartons for Canada and the United States but is also certified for other markets if necessary. The Basingstoke, England facility packages the tablets into cards and cartons for all other markets outside the U.S. and Canada.

One unique feature of the pharmaceutical supply chain is the lifesaving nature of the product. Eli Lilly can not allow product backorders to occur because the potential harm of not having a medication available when needed outweighs the cost of stocking inventory. Eli Lilly will also conduct an expedited production run or hold stock at intermediate locations along the chain before permitting any backorders.

The students were asked to compute the total flow time to manufacture, ship, and package the product at the two finishing facilities. Eli Lilly also wanted to know how many production runs were needed to meet three years of future demand. The students also computed the expected number of retests during inspection and how much raw materials and final product should be stocked to prevent back orders from occurring.

2.2 Problem Attributes

This problem compares well to the "ideal" case outlined in Section 1 of the paper.

- Unspecified *objective* – Since this project was being conducted for an actual client, the objective

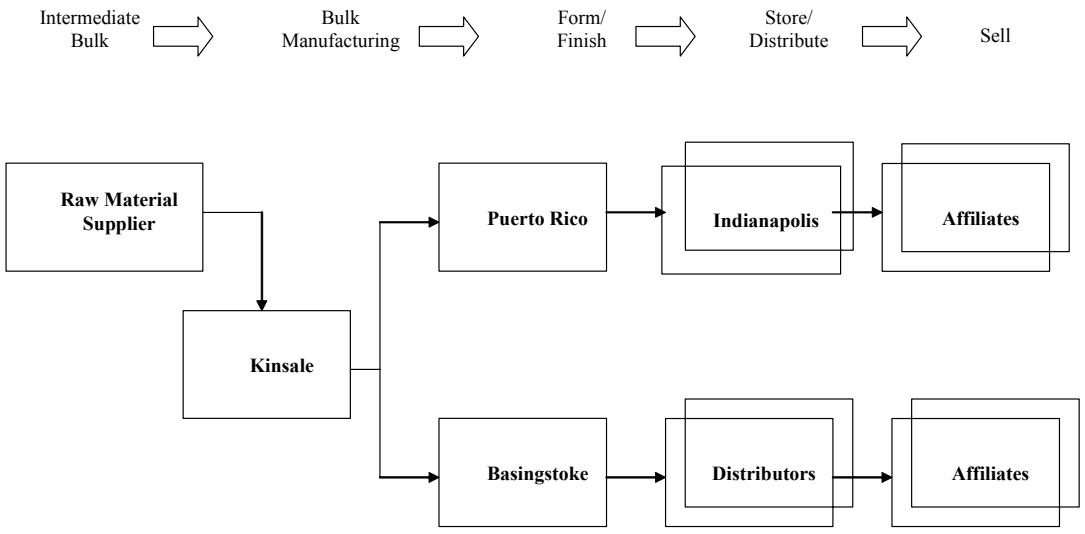

Figure 1: Supply Chain Network

and response variable was determined by Eli Lilly. The project goals of minimizing flow time and backorders were defined by Eli Lilly.

- *Complexity* – The true supply chain operation for LLY123 is very complex. Since we only had approximately six weeks for the students to complete the project, we had to make some simplifications. All simplifications were approved by Eli Lilly before being included in the case write-up.
- *Creativity* – Eli Lilly specifically did not include cost information to allow the students to pursue creative solutions to the flow time problem. Initially, this might make the problem appear less realistic, but Eli Lilly was willing to consider all reasonable solutions to prevent backorders and did not want to influence the students with cost information. Some project teams became very creative in considering alternative job sequencing at the Kinsale facility, devising various stocking policies, or purchasing additional equipment.
- *Raw data* – Eli Lilly provided a sample of approval times for the inspection operation at the Kinsale plant. The students had to analyze a sample of 35 observations from the approval process to determine the underlying distribution. Eli Lilly offered other information such as process times and ship times in terms of minimum and maximum values. These times were assumed to be uniformly distributed between the upper and lower values.
- *Unnecessary or missing data* – The students were not given any information regarding the other products being produced at the Kinsale plant. They had to make assumptions about where the LLY123 product would fit into the overall schedule for the plant. Since the client presented the problem to the students, they were not given any unnecessary data for the problem.
- *Language features* – The students were studying the Arena simulation language (Kelton, Sadowski, and Sadowski 2002) during the semester. Many features of the language were discussed in class and the students were given six lab assignments that covered a variety of Arena modules. The students did not have to learn new syntax to complete the project, but they were told the model would be part of the presentation to the client. The students had to learn Arena's animation capabilities to provide the client with a realistic looking system on the screen as the model ran.
- *Model verification* – Although this project had a real client, model validation against the real system was not done. Since we made several simplifications for the project, we felt this would negate the possibility for validation. The students did perform verification of their models. This was required for the model write-up to the client. The students were told to document any assumptions they made in building the model and to show any test scenarios conducted for verification purposes.
- *Experimentation* – The students were not given any guidance in this area. They were told the project's objectives and had to decide what alternatives to test to minimize flow time and product backorders. The students were told to conduct statistical output analysis to verify their recommendations to the client. Formal use of the optimization package that is included in the Arena simulation package was not required for this project. Optimization is an elective topic for the course and was not discussed.

2.3 Solution Notes

This supply chain problem has no unique correct solution. Several student teams recommended purchasing additional equipment, while other teams explored different inventory holding policies. My solution included a different sequencing of the intermediate and final processing operations at the Kinsale plant. It was possible to run several lots through intermediate processing then incur only one cleanup operation before performing the final processing. This was a solution that Eli Lilly had not previously considered.

2.4 Student Feedback

Twenty-four students formed 12 project teams for this case study. Three teams were chosen to present their findings to Eli Lilly at the end of the semester. A survey was given to the students using a five-point scale where 5 = strongly agree and 1 = strongly disagree. The average response for three of the survey questions are presented below.

- I understand the relevance of using simulation to make business decisions – average response = 4.3.
- I have an appreciation of real-world simulation projects – average response = 4.1.
- I would feel comfortable assisting with a simulation project in my job – average response = 3.7.

3 WEB SITE SERVER PROBLEM

In the recent edition of Banks, Carson, Nelson, and Nicol (2001), a new chapter was added on the Simulation of Computer Systems. This chapter provides a rich source of computer applications. The second author was particularly interested in the web site server system described on pages 553-557. Unfortunately, after many careful readings of this section, it became clear that a better understanding of the technical operation of a web site is needed before this application can be used in a simulation class.

Since the book was published two years ago, considerable time was spent trying to understand the operation of web server systems and to develop a reasonable simulation project for use in a simulation course. This background was obtained by attending classes in database and e-Technology, listening to a lecture from a systems architect at UBID.COM, a variety of readings, and discussions with knowledgeable faculty colleagues. This project is being shared so that other simulation instructors can be brought up to speed more rapidly.

3.1 Problem Description

The problem is to design a web site server system for an auction site similar to <www.ubid.com>. Like UBID, this auction site will buy closeout merchandise from manufacturers and auction it to the general public. The goal of the study is to determine the appropriate number of servers to handle a peak load of 50 document requests per second.

The initial site concept is illustrated in Figure 2. Internet visitors will enter through a firewall/router to a 100 MBPS (million bits per second) LAN containing a bank of web servers. UBID plans to divide the web servers into two groups, one to serve normal document requests and static images and the other to serve dynamic images of the products currently being auctioned. The application servers and data servers are on a separate 100 MBPS LAN and protected from the web servers by an additional router/firewall. The application servers are divided into servers supporting client requests and servers supporting employee updates to the database. The web server routes dynamic pages to one of the client application servers to run an application to build the page. The client application server will access the customer data servers for customer information and/or the product data servers for auction information to obtain the appropriate information for the document requested. The populated document is returned via the internal router to the web server, which then forwards the document to the client.

Traffic also comes from the Company LAN to the employee application servers to maintain the customer databases, the product databases, and the product images on the web server. The internal traffic from the Company LAN is estimated at 2 document requests per second.

3.2 Problem Attributes

This problem contains some of the attributes of an ideal simulation project, but certainly not all of them.

- *Unspecified objective* – The actual project write-up specifies that the client would like to use the minimum amount of resources such that (1) all resources will be no more than 70% utilized and (2) the web site response time for client requests will be less than ½ second. Though this is a fairly clear statement of the objective, it turns out that computing the response time is not as easy as one might expect.
- *Complexity* – This problem is a little on the complex side. The project description contains detailed write-ups about the process of fetching documents which must be understood by the students. A couple of tricky aspects involve the implementation of round-robin server selection, the maintenance of state on a particular server while a document is created, measuring response time, and modeling the network lines as a resource. Most of the level of detail decisions have been made in the problem description document.
- Creativity – There are a number of challenging problems to solve to create a model of this system. New document requests are assigned a server

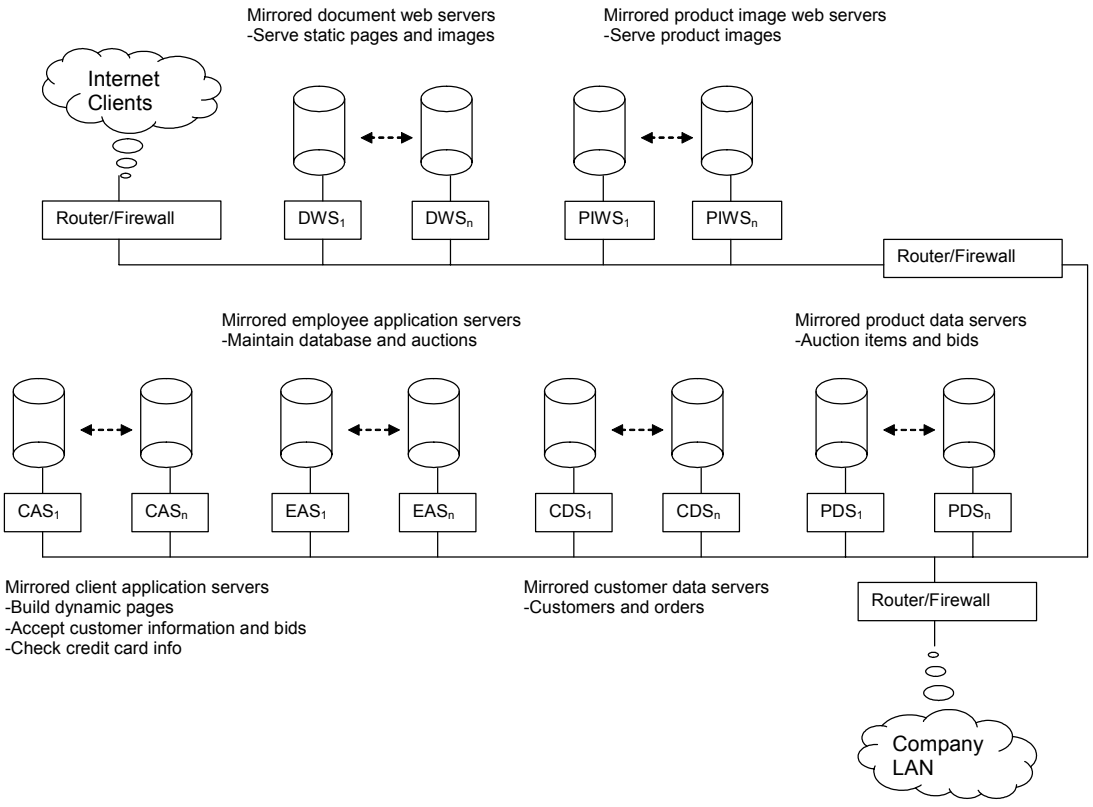

Figure 2: Web Site Server System

in a round robin fashion requiring each server to have an identity (that is, it is not a standard multiple server queuing system). As document requests are passed from web server to application server to data server and back, state must be maintained so the information is returned to the same server. This is a modeling situation that will be new to most students. One of the most interesting opportunities for creativity is measuring response time. Often web pages contain text and graphic images. Initially, a web page is returned without the images and additional requests are generated to obtain the images specified in the web page. Thus, the response time should be the total time to retrieve the initial web page plus all of the images. Students will need to make some modeling assumptions to calculate this response variable.

- *Raw data* – Many of the data items for this problem were specified as ranges for the uniform distribution or as three points for the triangular distribution. Caching percentages and modeling assumptions about caching were provided. The most interesting data modeling task was to model the frequency a particular document is requested. The problem statement provided a data file containing characteristics of all possible external and internal document requests and a link to Ziph's law `<http://linkage.rockefeller.edu/wli/zipf/>` The students were expected to figure out how to use Ziph's law to model the document requests.
- *Unnecessary or missing data* – There was no unnecessary or missing data in this problem. Students may have thought it would be unnecessary to model the network lines, but it turns out they are a bottleneck. In reality, this is no longer an issue since most networks have moved from the bus architecture to a switched architecture.
- *Language features* – Most of the language features needed to model this system were covered in class. The only tricky problem was to measure response time for a complete web page. The entities moving through the system are documents. This includes text documents and images. An initial document request may generate requests for many image documents. Holding an initial document until all of the image documents are returned requires some advanced modeling features to allow entities to communicate with each other.
- *Model verification* – Validation of this model is not possible. Students were asked to walkthrough a working model with the instructor and to describe the steps they used to verify their model. Suggestions were to display as many statistics as

possible, use a trace to verify tricky modeling situations, and create an analytic model for comparison. One student was able to verify some resource utilizations using an analytic model.

- *Experimentation* – At first glance experimentation seems very difficult. There are multiple performance measures and decisions are needed for the number of resources of many types. Despite the large number of variables, there is little interaction between the variables. For example, the number of routers needed is almost independent of the number of employee application servers. It is fairly easy to locate a pretty good solution. Students were asked to make some preliminary runs with 10 replications to locate several near optimal solutions and then to perform a thorough statistical comparison of this small number of cases. Experimental design techniques were not covered in class.

3.3 Solution Notes

This system was successfully modeled using CSIM (Schwetman 1996). An initial model of Figure 2 indicated that at least two routers are required for the internet clients and the LAN containing the document web servers and the product image web servers does not have sufficient capacity. It was necessary to reconfigure the LAN. The instructor's solution was to place the document web servers and the product image web servers on separate LANs, though other solutions are possible. This complexity can be removed by providing an appropriate LAN configuration from the start.

3.4 Student Feedback

This project was given to a class of six students. Four students choose to work independently and two as a team. Three individual students and the team were able to successfully complete the project. The other student did not put sufficient time into the project to reach a satisfactory solution. Enthusiasm for the project was high. Many interesting discussions about modeling were generated and the students learned a lot about the way web site server systems worked. One of the students was so enthused that she began exploring opportunities for graduate study in modeling.

4 REAL VS. TEXTBOOK PROJECTS

As can be seen from the previous two sections, both textbook and real-world problems exhibit many of the ideal attributes that were mentioned Section 1. In this section, we present a comparison of the two types of problems based on the ideal attributes along with the advantages and disadvantages of each.

In both cases, the project objective was clearly defined and performance measures were given to the students. The students working on the supply chain project had an opportunity to meet with the client to ask questions and receive additional information while the students working on the web server problem consulted with the instructor.

Both problems represented complex systems to be studied. Each one had to be modified to allow the students the time to solve the problem within the course of one semester. When using a real-world problem, any changes to the problem must be approved by the client. One needs to be careful not to simplify or eliminate portions of the problem that are important to the client. In contrast, a textbook problem may be modified in any fashion to fit the needs of a particular offering of a course.

Both problems provided opportunities for the students to show creativity in finding a solution. The students were not directed toward any particular solution alternatives. The students also had to conduct some input data analysis in the course of solving the problem. The supply chain problem contained actual observations from the manufacturing plant while the web-server problem required the students to determine how to model document requests.

The supply chain problem referred to issues that might impact the chain, like other products produced in the Kinsale plant, but they were not given specific information about the other products. The students needed to make some assumptions about the production sequence.

Most of the language features needed to model the problem were presented to the students for each project. Some additional study was needed to make the model presentable through graphics and animation to the client for the supply chain problem. The students had to discover the means for modeling response time for the web-server problem.

The students were expected to verify the model for each project and conduct some experiments to solve the problem. The instructor required the students to perform a walkthrough of their working model and describe how it was verified as part of the overall project. The students who worked on the supply chain problem turned in their verification procedures with the final report.

An interesting by-product of the supply chain problem resulted from the final presentations to the client at the end of the semester. Due to time limitations, all the project teams were not provided with the opportunity to present their findings. Three project teams were chosen based on the number of alternatives explored, the statistical output analysis and the final written report. This created an environment of competition between the project teams for the privilege of presenting. This competitive atmosphere would be difficult to duplicate in a textbook problem.

An advantage of using a textbook problem is the level of control over problem design. The problem can be manipulated to any degree of complexity, raw data can be

generated according to any distribution for input analysis and the project can be changed at any time if necessary to eliminate unexpected events. This is not true with real-world problems. The client must be consulted regarding any changes to the problem and input data observations must be presented in their raw form.

Finally, the students had to be concerned with the appearance of the model for the supply chain problem. They were asked to provide a model that was aesthetically pleasing as well as producing the correct solution. This is not as important when there is not a real client viewing the results. The students were also asked to carefully document any assumptions they made to build a working model.

5 SUMMARY

This paper has provided some insights about designing simulation projects for teaching discrete event simulation. Two projects have been described in this paper and details are available at an accompanying website <http://unixgen.muohio.edu/~byrketdl/wsc2003/>.

While it appears that a real-world problem offers some advantages not available with a textbook problem, it is not always possible to gain access to real problems. It can require a great deal of time and effort for a real client to visit the class or produce the necessary background information and raw data. Textbook problems are easier to modify to illustrate specific points or achieve a desired level of complexity.

REFERENCES

Banks, J., J. Carson, B. Nelson, and D. Nicol. 2001. *Discrete-Event System Simulation*. Prentice-Hall.

Kelton, W., R. Sadowski, and D. Sadowski. 2002. *Simulation With Arena*. McGraw-Hill.

Schwetman, H. 1996. *CSIM Users Guide*, Mesquite Software.

AUTHOR BIOGRAPHIES

MELANIE L. HATCH is an Assistant Professor of Computer Science and Systems Analysis at Miami University in Oxford, OH. Her teaching and research interests include math programming, simulation, and service organization design and operation. She can be contacted by e-mail at <hatchml@muohio.edu>

DONALD L. BYRKETT is a Professor of Computer Science and Systems Analysis at Miami University in Oxford, OH. His teaching and research interests include stochastic modeling and simulation. He has been designing classroom simulation projects for over 25 years. He can be contacted by e-mail at <byrketdl@muohio.edu>.